Blowing Seeds across Gardens: Visualizing Implicit Propagation of Cross-Platform Social Media Posts

Jianing Yin, Hanze Jia, Buwei Zhou, Tan Tang, Lu Ying, Shuainan Ye, Tai-Quan Peng, and Yingcai Wu

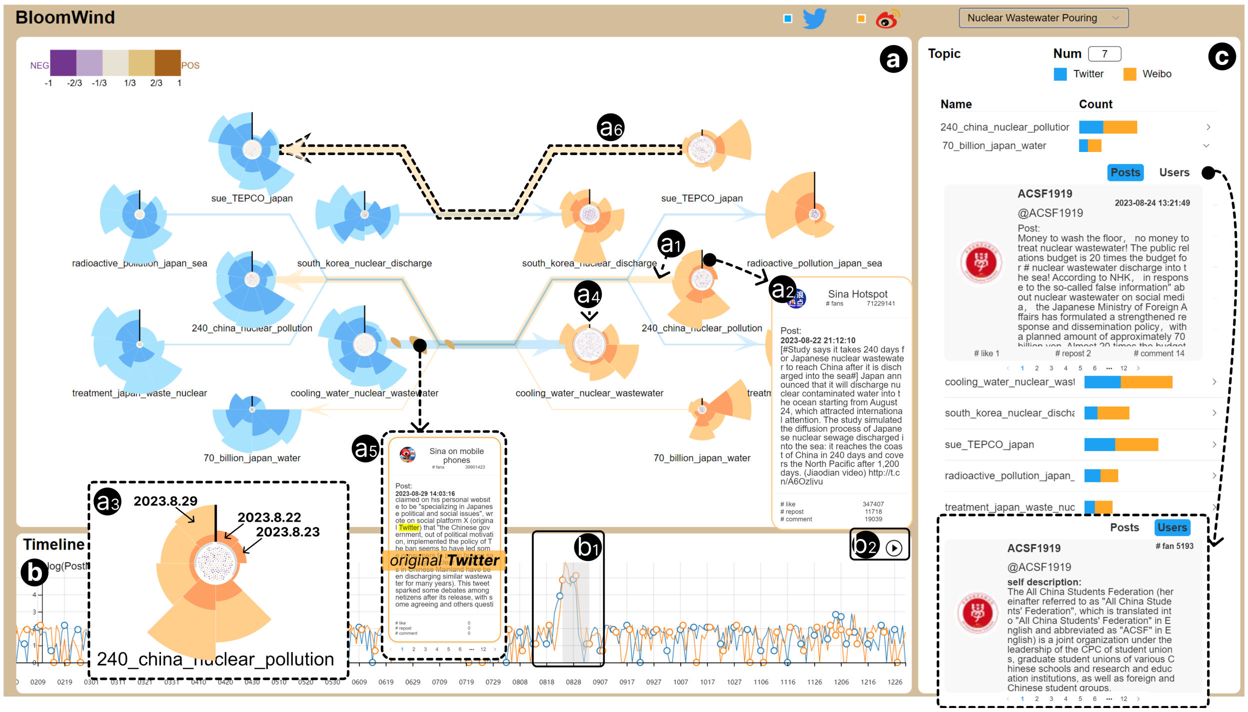

Fig. 1: Interface Overview (Cluster-level): (a) Cluster-level Propagation View, demonstrating the diffusion process of topics among platforms; (b) Timeline View, for selecting a time frame and controlling the animation process of propagation; (c) Cluster-level Detail View, listing the post and user information by topic and platform.

Abstract—Propagation analysis refers to studying how information spreads on social media, a pivotal endeavor for understanding social sentiment and public opinions. Numerous studies contribute to visualizing information spread, but few have considered the implicit and complex diffusion patterns among multiple platforms. To bridge the gap, we summarize cross-platform diffusion patterns with experts and identify significant factors that dissect the mechanisms of cross-platform information spread. Based on that, we propose an information diffusion model that estimates the likelihood of a topic/post spreading among different social media platforms. Moreover, we propose a novel visual metaphor that encapsulates cross-platform propagation in a manner analogous to the spread of seeds across gardens. Specifically, we visualize platforms, posts, implicit cross-platform routes, and salient instances as elements of a virtual ecosystem — gardens, flowers, winds, and seeds, respectively. We further develop a visual analytic system, namely BloomWind, that enables users to quickly identify the cross-platform diffusion patterns and investigate the relevant social media posts. Ultimately, we demonstrate the usage of BloomWind through two case studies and validate its effectiveness using expert interviews.

Index Terms—Propagation analysis, social media visualization, cross-platform propagation, metaphor design

1 Introduction

Social media has become a part of daily life where users would share life moments [50], express personal emotions [3], and discuss public events [18]. Owing to diverse social media platforms, information could easily spread among them, leading to an emergent research problem [29, 47, 67]. Understanding how information propagates across different platforms could assist analysts in analyzing complex user behaviors [43], designing effective marketing strategies [6, 19], and preventing harmful content [20, 25, 46, 57]. Despite the existence of numerous studies [7, 12, 56, 63, 74] that visualize information propagation on social media, few have considered cross-platform propagation due to its complex and implicit mechanisms.

To visualize information propagation, existing methods mainly employ metaphor-based designs to illustrate the diffusion process of spatial-temporal information on a social media platform (e.g.,

- *J. Yin, H. Jia, B. Zhou, L. Ying, S. Ye, and Y. Wu are with the State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China. E-mail: {yinjianing, hzjia, zhoubuwei, yingluu, sn_ye, ycwu}@zju.edu.cn.*
- *T. Tang is with the Laboratory of Art and Archaeology Image, Zhejiang University, Hangzhou, China. E-mail: tangtan@zju.edu.cn.*
- *T. Peng is with the Department of Communication, Michigan State University, East Lansing, United States. E-mail: pengtaiq@msu.edu.*
- *Tan Tang is the corresponding author.*

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

Weibo [12] or Twitter [7] [1]). However, these visualizations may not be able to reveal underlying data patterns without a deep understanding of the cross-platform propagation mechanisms. We also notice that several studies [4, 36, 39] have considered multiple social media platforms. For example, TopicPanorama [36] investigates how to merge topics extracted from different platforms into a unified graph. Others [4, 39] merely leverage information from multiple platforms to gather comprehensive insights related to their topic of interest. To our knowledge, visually uncovering propagation patterns across social media platforms using the cross-platform mechanisms remains an unresolved issue.

In this study, our target users are experts in social media cross-platform propagation analysis. We aim to empower them to analyze the propagation situation and uncover meaningful insights interactively. Accomplishing such a goal encounters three challenges:

C1. Deep understanding of cross-platform mechanism. The driving forces behind cross-platform propagation can vary significantly across different cases, resulting in diverse and complex propagation patterns. Identifying and analyzing such patterns demands significant effort and extensive professional knowledge, posing a challenging and labor-intensive task.

C2. Probabilistic inference of implicit propagation. Information spread within a platform is usually visible due to user interactions such as reposting, commenting, and liking. However, these connections may not be readily accessible in the context of cross-platform propagation, complicating the inference of implicit propagation likelihood.

C3. Scalable visualization of multi-faceted information. Given the huge volume of posts and the intricate dynamics of information spread, a scalable visual design is required to effectively demonstrate diffusion across various platforms. Furthermore, the visual design must be comprehensive enough to support multi-faceted analysis, given the complexity of the social media data.

To address the above challenges, we develop a novel visual analytic system, namely BloomWind, to explore cross-platform information diffusions on social media. To understand complex mechanisms (**C1**), we collaborate with domain experts to systematically summarize cross-platform diffusion patterns, analyze their underlying mechanisms, and extract significant factors. Based on that, we introduce a cross-platform influence model (**C2**) that incorporates content relevance and other significant factors to estimate the likelihood of cross-platform information diffusion at the topic/post level. Moreover, we propose a novel visual metaphor to provide a vivid and multi-faceted representation of cross-platform diffusion (**C3**). Specifically, we employ "overview + detail" lenses and a hierarchical design based on text clustering. After doing text clustering on all posts, each topic cluster is further split based on the different platforms to which the posts belong. Subsequently, the propagation likelihood is calculated among the clusters belonging to the same topic. A social media platform is visualized as a garden and each cluster is represented by a flower that encodes the basic information of the cluster and propagation factors. Upon concentrating on a specific propagation path across platforms, the two corresponding clusters are magnified into detailed circular gardens, where each post is visualized as a flower or a dot. The cross-platform propagation is thereby visualized as a breeze blowing seeds across different gardens. We finally demonstrate the usage of BloomWind through two real-world case studies and validate its effectiveness by conducting semi-structured interviews with domain experts.

Our contributions are summarized as follows:

- We systematically summarize the **cross-platform patterns** with professionals and propose a novel **influence model** that estimates the potential for information to spread across platforms.
- We propose a novel **visual metaphor** for vividly illustrating cross-platform information diffusion and develop a **visual analytic system** to facilitate cross-platform propagation analysis.
- We hold two **case studies** on real-world datasets to indicate how to discover effective propagation insights and introduce **expert interviews** to validate the effectiveness of our system.

[1]Currently known as X, but for the sake of familiarity, we still refer to it as Twitter throughout the paper.

2 Related Work

This section introduces a detailed literature review of information diffusion on social media and relevant visual analytic studies.

2.1 Visual Analytics on Information Diffusion

Various visualization methods reveal information diffusion from different perspectives, including messages' diffusion network, people's reposting network, and geographic information diffusion [13]. Studies on the messages' diffusion network explore methods to visualize information propagation efficiently and understandably. Some focus on depicting a post's reposting tree with enhanced node-link diagrams [1, 31, 48, 62, 69]. R-Map [12] and E-Map [10] improve these ideas through map-like metaphors to present Weibo posts and repostings during an event. Ripples [58] employs a balloon graph layout to show the sharing process of all posts containing a specific link. Liu et al. [35] combined an uncertainty glyph with a node-link diagram to depict hashtag propagation. River-like visualizations are also widely utilized to present temporal information diffusion [56, 59, 63, 66, 74]. Some other works focus on visualizing the performance of diffusion algorithms and models [5, 52, 65].

Studies on people's reposting networks visualize user interactions based on reposting and commenting, using improved node-link diagrams and river-like visualizations. Yuan et al. [70] used node-link diagrams to uncover repost paths among users, helping identify key players and their roles. D-Map [11] leverages a map-based ego-centric visualization to show people's involvement in a message diffusion process. Shareflow [23] employs a circle layout with edge bundling to illustrate diffusion paths from opinion leaders to followers, as well as the activities of active users around a specific topic. To assess the potential threats of an anomalous user, TargetVue [8] combines user activity threads into a volume circle view, which helps illustrate how the user's messages are reacted to by others. Liu et al. [37] proposed a fluid dynamics-inspired visualization technique to facilitate the comprehension of users' retweeting behaviors for different topic categories. To analyze the overall dynamic patterns of Twitter accounts in a specific event and identify pivotal moments, GraphFlow [16] enhances the comprehensive flow view with a filtered graph view. Additionally, studies on geographic information diffusion integrate maps to visualize spatial patterns of diffusion [7, 22, 42, 55, 62].

However, the aforementioned works merely focus on diffusion within a single platform, leaving the problem of effectively visualizing cross-platform information propagation unresolved. These studies primarily rely on reposting or commenting relationships between users or posts, while cross-platform propagation relationships require additional multi-faceted information to substantiate existence and demonstrate patterns. Although several multi-platform visualizations [4, 36, 39] can be used to infer cross-platform diffusion indirectly, they lack a focus on the direct, systematic, and comprehensive revelation of cross-platform relations. In our work, we provide a visual analytics system with a novel visual metaphor that depicts cross-platform information propagation as the spread of seeds across gardens. Each platform is visualized as a garden, implicit propagation routes as winds, and salient instances as seeds. With this metaphor, we consistently reveal cross-platform diffusion patterns and mechanisms in a visually intuitive manner.

2.2 Propagation Analysis of Social Media

Social media platforms offer a huge space for information diffusion. Extensive research has been undertaken on three primary objectives: modeling and explaining this process [14, 20, 22, 25, 26, 28, 38, 41, 43, 46, 55, 56, 60, 61, 63], predicting and simulating the diffusion process [9, 15, 17, 24, 27, 30, 32–34, 37, 44, 49, 51, 53, 68, 71–73], and detecting anomalous diffusion patterns [54, 74]. Gravity model [55] and influence model [63] have been utilized to approximate the propagation process. Sun et al. [56] proposed a co-opetition model to quantify the competition and cooperation between topics and the influence of topic leaders in this dynamic. Infectious disease models [17, 73], agent-based models [44], regression-based models [34], and hydrodynamic models [24, 37] are well-established models for simulating information propagation, taking into account the influence of individual behavior and interaction within the simulation environment. Researchers have also identified several key factors crucial for diffusion to expand and improve these foundational models, such as salience of information [13], recency effect [63], and opinion leader's influence [24, 28].

Several explorations have delved into cross-platform information diffusion. Some choose to employ multi-layer networks [53, 72] for simulation to help understand how different platforms interact and facilitate information dissemination. In situations requiring detailed case analysis, the previously mentioned approach encounters limitations in accuracy and interpretability, leading researchers to primarily adopt alternative methods. Lu et al. [38] introduced a semi-automated method for the quantitative analysis of pandemic-related information flow from Twitter to Weibo. Additionally, other research mainly relies on the detection of common URLs or identification of cross-platform user relationships [14, 20, 25, 43, 46, 61, 64], to unveil specific cross-platform dissemination scenarios. However, due to the absence of a systematic summary of cross-platform diffusion patterns, existing efforts are not enough to achieve comprehensive analyses despite the precise identification of cases under some specific patterns. Moreover, a lack of effective human-machine collaboration results in a considerable human workload. We build our cross-platform diffusion model based on the significant factors extracted from our comprehensive analysis of cross-platform diffusion patterns and their underlying mechanisms.

Table 1: Cross-platform diffusion patterns.

Type of pattern	Types of mechanisms
Among same-entity accounts	Operational needs
	Experimental needs
Driven by account positioning	Relaying / News / Media account
	Fan account
	Subsidiary work / Functional account
Driven by content	Aroused with the desire to share or express
	Promotion task
	Content extension

3 Cross-Platform Diffusion Patterns

In this section, we worked closely with three domain experts: a professor specializing in media and communication, and two senior researchers with over a decade of experience in propagation analysis. To understand cross-platform diffusion patterns and their underlying mechanisms, we first presented online questionnaires for our collaborators, investigating the general workflow of propagation analysis. We followed the professor's suggestions to introduce patterns and mechanisms of cross-platform diffusion. The professor recommended a detailed analysis of patterns in real-world examples. We collected cross-platform diffusion instances based on the data supplied by our collaborators, sourced from popular platforms, including Weibo, Douban [2], Twitter, Facebook, and Reddit. We organized bi-weekly meetings with the senior researchers to discuss our ongoing instance collection and analysis. In our regular meetings with the professor, we focused on the characteristics of these instances. We shared preliminary findings, and the professor provided insights into potential mechanisms and judgment criteria. Consequently, we reach the following conclusions shown in Tab. 1. There are three primary patterns of cross-platform diffusion: diffusion among same-entity accounts, diffusion driven by account positioning, and diffusion driven by content.

3.1 Diffusion among Same-entity Accounts

Users who maintain accounts on multiple social media platforms often share information across these platforms. There are two primary needs associated with this type of diffusion: operational needs and experimental needs. For operational needs, many individuals and organizations nowadays maintain accounts on multiple platforms as a strategy to broaden their influence. Mostly, they share similar information across platforms, but occasionally adjust their expression strategies to fit each platform's context. For the latter one, users may initially test various message diffusion effects on a less significant platform before implementing the successful ones on the target platform. For example, during the 2016 U.S. presidential election, the Internet Research Agency appeared to test content on Reddit to determine the most effective messages for propagation on Twitter, as investigated by Lukito et al. [40]. To detect diffusion among accounts from the same entity, it is crucial to identify accounts with similarities in user information, such as usernames, self-descriptions, and validation details. Furthermore, this diffusion primarily occurs within influential accounts, which are distinguishable by follower count and engagement metrics on posts.

3.2 Diffusion Driven by Account Positioning

Some accounts pay close attention to certain information on other platforms, resulting in frequent cross-platform propagation activities. These accounts fall into three categories: relaying/media/news accounts, fan accounts, and subsidiary work/functional accounts. Relaying/media/news accounts monitor and amplify specific information from various platforms. Relaying accounts copy viral news across platforms to maximize audience reach. Media accounts, managed by professional media outlets or individuals, often share reliable and well-articulated content from other platforms. News accounts focus on rapidly spreading breaking news, prioritizing speed over source reliability and detailed expression. Fan accounts typically center on a specific individual or subject, frequently sharing updates about their idol sourced from other platforms. Fan accounts not only show admiration for their idols but also play a crucial role in cultivating and engaging the fan community on the platform. Subsidiary work/functional accounts often convey messages on behalf of their organizations. For example, the Weibo account of the U.S. Embassy in China serves as a channel for important communications, regularly conveying messages posted by the U.S. President and the Ambassador on Twitter. To detect this pattern, user information like usernames, self-descriptions, and validation details is useful for account positioning. Behaviors like sharing message sources to enhance credibility are also typical. Additionally, these accounts are often influential on their platforms, reflected by follower count or engagement metric.

3.3 Diffusion Driven by Content

Content-driven cross-platform diffusion follows a more irregular pattern, as individuals spread content across platforms due to its appeal. Generally, it can be divided into three categories: the desire to share or express, promotion tasks, and content extension. People usually share or express their ideas when they are captivated by the content. Upon finding something interesting on other platforms, they would be likely to share similar information on a familiar platform. They may also integrate their own thoughts into the original content or simply add words to stimulate discussions among their audience. For instance, some people on Twitter expressed their ridicule towards the notion circulating on Reddit that the offline sales of singer Lucas' album are struggling. Expressing a positive attitude is also common, which can be explained by the selective exposure effect. People spread information across platforms to endorse it or cite original information as evidence supporting their viewpoints. Some people share content from another platform as a promotion task, aiming to direct attention or garner support for the original content. For example, the Weibo user 'Light years drunk with tea' posted a message recommending the content of a Douban post and provided its URL, encouraging everyone to come and support it. Similarly, the Twitter user 'RSK XFM Remastered' wrote a post, hoping the audience support a Reddit post where a URL was also attached. Additionally, embedding another post's URL is also a common practice to enrich detail and extend content. When a post on another platform articulates a point clearly, users often prefer to provide its URL rather than replicate its content entirely. Varied expression styles on different social platforms further motivate this mode. For instance, Twitter posts are typically brief, whereas posts on Reddit can be lengthier and more formal. It is common for a Twitter post to direct its readers towards a Reddit post, as observed in many cases. To detect this pattern, specific URLs or indicative words within the post content can be helpful. Additionally, the poster's level of activity on their platform and interest in the original content can be inferred from their historical behavior information, which may include instances of cross-platform content sharing. Generally, the original content capable of catalyzing cross-platform propagation already possesses an established audience base on the target platform. This means that on the target platform, the relevant topic not only has a certain discussion heat but may also reveal a relatively clear attitude tendency.

4 Cross-Platform Influence Model

This section introduces a novel influence model that estimates the likelihood of cross-platform propagation at the cluster and post levels.

4.1 Significant Factors

According to Sec. 3 and literature reviews, we identify several significant factors for estimating cross-platform propagation likelihood.

Information Salience. Diffusion capacity increases with the number of posts. A hot topic is more likely to spread between platforms [55].

KOL Influence. KOL means Key Opinion Leader. As confirmed by our collaborators, influential accounts and posts within the platform are more likely to be shared outside [38]. Katz et al. [28] proposed a "two-step flow" propagation theory, showing the ability of opinion leaders to spread information to broad audiences. The effect of influential ones has been utilized and proved in many propagation studies [24,56,60].

Recency Effect. Recently encountered information can have more influence on audiences than the old ones. The massive amount of information on social media makes this effect particularly evident during information diffusion [30,41].

Content Relevance. When posts contain content that exhibits some degree of relevance, the likelihood of cross-platform diffusion between them increases. For instance, many URL-based re-posting behaviors can reveal the cross-platform connection between posts, as demonstrated by Murdock et al. [43] and further supported in Section 3. This relevance can also manifest through similar linguistic patterns, as users tend to assimilate information that displays linguistic similarities to their prior exposures [55,56].

User Relevance. Our results in Sec. 3 indicate that cross-platform diffusion, whether motivated by account positioning or occurring among same-entity accounts, consistently shows a strong correlation between the source account, which initially posts the content, and the target account, which propagates the content on a new platform.

History Relevance. Historical behavior information can reflect account habits. When two accounts share similar interests, consistently discuss some topics, or have established cross-platform propagation relationships, the probability of ongoing interactions increases.

4.2 Cluster-level Cross-platform Propagation Likelihood

To track cross-platform information diffusion over time for a specific event, Bertopic [21] is adopted to divide posts into topic clusters $\{C(h)\}$, where h represents the topic. Each topic cluster $C(h)$ is divided into $\{C^x(h)\}$ by platform x, and $C^j(h)$ is further divided into $\{C^{j,t}(h)\}$ by time period t. Considering content relevance in diffusion, we assume that cross-platform diffusion occurs between clusters of the same topic on different platforms. Consequently, we calculate the ability of topic h to propagate from platform i to platform j during period t, denoted as $P^t_{i \xrightarrow{h} j}$. We define the cluster-level cross-platform propagation likelihood as follows:

$$P^t_{i \xrightarrow{h} j} = 1 - e^{-S-K*E} \tag{1}$$

where S, K, and E represent information salience, KOL influence, and content relevance, respectively. This formula is inspired by the principles in infectious disease models. S corresponds to the basic transmission rate, reflected by the number of posts. K corresponds to source strength, reflected by engagement metrics and follower count. E corresponds to the contact rate, reflected by content similarity and content associations implying diffusion. These factors are computed from various variables (see Tab. 3) arranged by their positive or negative

Table 2: Explanations of notations used in the model.

Notation	Explanation
i , j	platform
h	topic
t , t-1	time period t , time period $[t\text{-}\Delta t,\ t]$
$C^{j,t}(h)$	cluster of posts related to h on j published during t
$\lvert C^{j,t}(h)\rvert$, $\lvert C^{j,t}\rvert$	post count of $C^{j,t}(h)$, post count of $C^{j,t}$
$C^{j,t}_{\text{infposts}}(h)$, $C^{j,t}_{\text{infaccts}}(h)$	influential post/account count of $C^{j,t}(h)$
$C^{j,t}_{\text{reprword}}(h)$	the representative words of $C^{j,t}(h)$
$C^{j,t}_{\text{hashtag}}(h)$	hashtags in $C^{j,t}(h)$
$C^{j,t}_{\text{textURL}}(h)$	URLs in the text of posts belonging to $C^{j,t}(h)$
$C^{j,t}_{\text{postURL}}(h)$	URLs of the posts belonging to $C^{j,t}(h)$
$C^{j,t}_{\text{text}}(h)$	the texts of posts belonging to $C^{j,t}(h)$
$C^{j,t}_{\text{propword}}(h)$	specific propagation words in $C^{j,t}(h)$
$\mathrm{Sim}(\cdot,\cdot)$, $\mathrm{Same}(\cdot,\cdot)$	calculate the similarity , calculate the overlap
$\mathrm{Direct}(\cdot,\cdot)$	identify typical cross-platform diffusion cases
T , W	post T , post W
$U(T)$	the account that posted T
T_{text} , T_{hashtag}	text of T , hashtags in T
T_{postURL} , T_{textURL}	the URL of T , URLs in T 's text
$\mathrm{Mention}(T,W)$	check if W's text has specific propagation words
$\mathrm{Score}(\cdot)$	the influence score
$U_{\text{userinfo}}(T)$	user information of $U(T)$
$U_{\text{interest}}(T)$	interest distribution of $U(T)$'s historical posts
$\mathrm{PastSpread}(U(T),U(W))$	examine the propagation relationships between $U(T)$ and $U(W)$'s historical posts
$\mathrm{PastNum}(h,U(T),U(W))$	quantify $U(T)$ and $U(W)$'s posting frequency under topic h

Table 3: Computing methods for the significant factors.

Factor	Computing Method
S	$\dfrac{\lvert C^{i,t-1}(h)\rvert\lvert C^{j,t}(h)\rvert}{\lvert C^{i,t-1}\rvert\lvert C^{j,t}\rvert}$
K	$\dfrac{C^{i,t-1}_{\text{infposts}}(h)}{\max_h\left(C^{i,t-1}_{\text{infposts}}(h)\right)} + \dfrac{C^{i,t-1}_{\text{infaccts}}(h)}{\max_h\left(C^{i,t-1}_{\text{infaccts}}(h)\right)}$
E	$\omega_1 \cdot \mathrm{Sim}(C^{i,t-1}_{\text{reprword}}(h), C^{j,t}_{\text{reprword}}(h)) + \omega_2 \cdot \mathrm{Sim}(C^{i,t-1}_{\text{hashtag}}(h), C^{j,t}_{\text{hashtag}}(h))$ $+ \omega_3 \cdot \mathrm{Same}(C^{i,t-1}_{\text{textURL}}(h), C^{j,t}_{\text{textURL}}(h)) + \omega_4 \cdot \mathrm{Direct}(C^{i,t-1}_{\text{postURL}}(h), C^{j,t}_{\text{text}}(h))$ $+ \omega_5 \cdot \mathrm{Direct}(i, C^{j,t}_{\text{text}}(h))$
CR	$\frac{1}{2}\Big(\mathrm{Mention}(T,W) + \mathrm{Score}(T) \cdot \max(\mathrm{Sim}(T_{\text{text}}, W_{\text{text}}), \mathrm{Sim}(T_{\text{hashtag}}, W_{\text{hashtag}}), \mathrm{Same}(T_{\text{textURL}}, W_{\text{textURL}}))\Big)$
UR	$\mathrm{Score}(U(T)) \cdot \mathrm{Sim}(U_{\text{userinfo}}(T), U_{\text{userinfo}}(W))$
HR	$\dfrac{2(\omega_6 \cdot \mathrm{Sim}(U_{\text{interest}}(T), U_{\text{interest}}(W)) + \omega_7 \cdot \mathrm{PastSpread}(U(T),U(W)))}{1+\exp(-\mathrm{PastNum}(h,U(T),U(W)))}$

correlation with the factors. The use of +/- is to avoid non-linear relationships and to enhance interpretability. The weights of the variables are determined based on empirical observations from small-scale data. The explanations of notations are in Tab. 2. Considering the recency effect and propagation time cost, we assume posts published during period t could be propagated from posts published during $[t-\Delta t, t]$, simplified as $t-1$. For detailed weights and specific calculations of each variable, please refer to the supplementary material.

4.3 Post-level Cross-platform Propagation Likelihood

The cluster-level propagation likelihood provides an overview of a topic's cross-platform diffusion. Subsequently, within the same topic, the propagation likelihood between two posts on different platforms is calculated. To approximate the cross-platform propagation likelihood from one post T to another post W within the same topic, several factors are utilized to define the following formula:

$$P_{T \to W} = \max\left(\mathrm{Direct}(T_{\text{postURL}}, W_{\text{text}}),\ 1 - e^{-\frac{CR+UR+HR}{1+0.2(t(W)-t(T))}}\right) \tag{2}$$

In Eq. 2, $\mathrm{Direct}(T_{\text{postURL}}, W_{\text{text}})$ checks if the content of W contains the URL of T, which is regarded as a direct sign of propagation. The combination of other factors is inspired by the principles in infectious disease models. *CR* is the content relevance between T and W. $U(T)$ is the account that posted T. *UR* and *HR* represent the user relevance and history relevance between $U(T)$ and $U(W)$, respectively. $t(W)$ is the publication time of W, and $t(W) - t(T)$ quantifies the recency effect. These factors are computed from various variables (see Tab. 3) arranged by their positive or negative correlation with the factors. The weights of the variables are determined based on empirical observations from small-scale data. For detailed weights and specific calculations of each variable, please also refer to the supplementary material.

Estimate the type of diffusion. Our model is also able to identify the type of cross-platform diffusion pattern with the following logic: if $\mathrm{Sim}(U_{\mathrm{userinfo}}(T), U_{\mathrm{userinfo}}(W))$ is very large (0.5 in our case), then it is a diffusion among same-entity accounts; if $\mathrm{Sim}(U_{\mathrm{userinfo}}(T), U_{\mathrm{userinfo}}(W))$ is relatively large ([0.3, 0.5) in our case), or it is not an ordinary account, or the source platform-related words appear in the user information, then it is a diffusion driven by account positioning; Otherwise, it is a diffusion driven by content.

5 Visualization Design

This section introduces the visualization design of BloomWind. To clarify requirements, we conducted semi-structured interviews with our collaborated experts. They helped us understand their routines, primary concerns, and practical needs while doing cross-platform diffusion analysis. Based on the interview, we summarized the design requirements to design our visual analytics system as follows.

R1. Support hierarchical exploration. Considering the large amounts of posts on the internet, experts prefer to see the overall diffusion situation before delving into specific details.

R2. Organize posts into topics. Multiple topics emerge within a single event. Experts are interested in analyzing the differences between topics during the same period and answering questions like which kind of topics are more likely to be spread out.

R3. Demonstrate the temporal dynamics of propagation. Besides vertical comparisons within the same time frame, experts are also keen on analyzing how propagation patterns evolve, through which they can dig out the propagation story line by line.

R4. Support detailed analysis of significant factors. Besides the observation of the cross-platform information diffusion phenomenon, the experts are more interested in analyzing the driving mechanisms behind it. Therefore, they would like to see the encoding of significant factors used in the calculation of propagation likelihood.

R5. Visualize bi-directional cross-platform propagation. The experts are particularly interested in the information exchange between platforms. They seek to determine whether, within a specific event and time frame, a dominant platform emerges or if a multi-polar pattern develops, with several platforms simultaneously importing and exporting information. Their focus is on understanding the intricacies of the information ecosystem created by these inter-platform exchanges.

R6. Demonstrate propagation in an intuitive manner. The experts expect the system to illustrate the where, when, how, and why of cross-platform propagation vividly and clearly.

5.1 Interface Overview

BloomWind is designed to illustrate and facilitate the analysis of cross-platform diffusion among posts related to an event. Fig. 1 illustrates the interface of BloomWind, composing three views: Propagation View (see Fig. 1a), Timeline View (see Fig. 1b), and Detail View (see Fig. 1c). Users first select the desired event and platforms, then Timeline View (see Fig. 1b) shows the variation in posting volume on each platform over time. According to **R3**, this view also supports the selection of time frames (see Fig. 1(b1)) and animation control (Fig. 1(b2)). Following the selection of a time frame, Detail View (see Fig. 1c) and Propagation View (see Fig. 1a) present the topic clustering results and propagation abilities at the cluster level (**R1**). Specifically, Detail View lists the posts' and users' detailed information organized by topic clusters (**R2**). The Propagation View presents the cluster-level propagation across platforms with a metaphor-based visualization (**R4, R5, R6**). After grasping the overall diffusion situation between

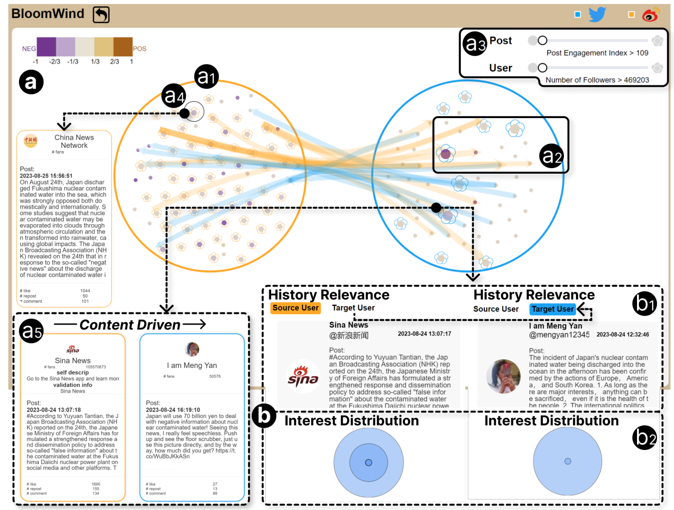

Fig. 2: Interface Overview (Post-level). (a) Post-level Propagation View: the diffusion of posts related to a topic between platforms; (b) Post-level Detail View: topic-related posts and user information by platform, along with historical relevance data for the focused post pair.

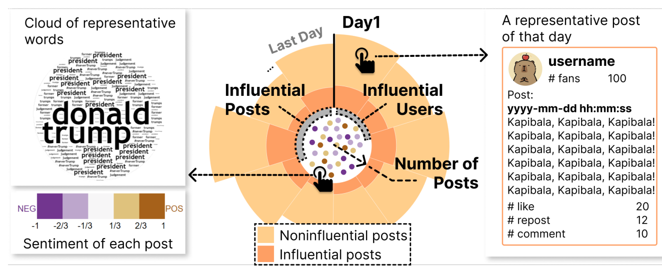

Fig. 3: Cluster-level Flower Glyph Design. A cluster-level flower glyph represents a cluster of posts related to a topic published on a platform during a period. The encodings and interactions are designed to learn about the cluster quickly. Petals' color (i.e., sectors) matches the platform's color. The inner circle's radius encodes the number of posts in this cluster, with two semicircular rings around it whose angles encode the number of influential users and number of influential posts in cluster.

topic clusters across different platforms, users can right-click on the propagation path between a cluster pair to explore their cross-platform propagation at the post level (see Fig. 2). Once switching into the post level, Detail View (see Fig. 2b) would adjust contents in coordination with Propagation View (see Fig. 2a). Additionally, our system applies uniform color coding, where information about a platform is encoded in the same color scheme: Twitter is denoted by the color blue, Weibo by orange, and Facebook by green.

5.2 Propagation View

To better demonstrate and explain the cross-platform propagation (**R4, R5, R6**), Propagation View visualizes social platforms, posts, implicit cross-platform routes, and salient propagation instances as elements of a virtual ecosystem, including gardens, flowers, winds, and seeds.

5.2.1 Cluster-level Glyph Design

In the cluster-level analysis, a flower-shaped glyph represents a cluster of posts related to a specific topic published on a platform within a certain time frame. As shown in Fig. 3, the cluster-level flower glyph is designed to encapsulate the overall situation of the cluster and visually encode factors that influence its propagation potential.

Encoding. The flower-based glyph is adapted from the Nightingale rose chart([45]), and the outer radius of these glyphs is uniform. The color of the petals (i.e., sectors) corresponds to the platform's color, and their number equals the days with topic-related posts in the selected time frame. To clarify the start and end date, we use a vertical black line between the petal representing the earliest date and the one for the latest date. Users can understand the chronological changes in posting activity by following a clockwise direction. The length of each petal encodes the daily posting volume about this topic on the platform. We also highlight the influential posts by layering them with a darker color stack. The radius of the inner circle encodes the total number of posts in this cluster, with two semicircular rings around it whose angles encode the total number of influential posts and the total number of influential users in the cluster. One dot inside the inner circle represents one post, with its color reflecting the sentiment of the post.

Interaction and Layout. Clicking on a petal reveals a representative post on that day. Clicking on the inner circle reveals the representative word cloud of this cluster. We use hexagonal grids to place the glyphs, with each glyph placed in the center of a hexagon. To enhance the proximity of glyphs within the same platform, glyphs from different platforms are arranged in distinct directions, with each platform's glyphs organized in a honeycomb clustering style.

5.2.2 Cluster-level Cross-platform Propagation Path Design

We visualize propagation paths (see Fig. 1(a6)) to demonstrate the calculation results of cluster-level cross-platform propagation likelihood.

Layout. We use a hexagonal grid-based algorithm to lay out the propagation path from the source glyph to the target one. To determine the paths between the flower glyphs under the same topic, we use a modified Breadth-First Search (BFS) algorithm to make edges along the points in the hexagonal grids. Starting from the source hexagon, we first find its vertex closest to the destination hexagon. The subsequent path-finding process is as follows: we aim to maximize segments that share the same line of hexagons when pursuing the shortest path, considering aesthetics and neatness. Consequently, for each computed path, we increase the incentive value of each hexagonal grid vertex it passes through, ensuring that vertices with higher incentive values are given priority in subsequent path calculations.

Encoding. We use the width of the path to encode the magnitude of propagation likelihood. Since our propagation likelihood is directional, we display the direction with the higher propagation likelihood. The color of the path is encoded using the source platform's color.

Interaction. Clicking on the animation control button (Fig. 1(b2)) can show the dynamic process of all the paths sequentially. Hovering over a propagation path will illuminate it, enhancing its visibility. Left-clicking on the path reveals some salient propagation instances used to calculate this propagation likelihood. As illustrated in Fig. 1(a5), upon left-clicking, three seeds will emerge on the path, each representing a type of notable cross-platform cases: the source and target post including a common URL, the target post including the source post's URL, and the target post containing some specific reference to the source post. After checking the cluster-level propagation situation, users can right-click on a path to explore the propagation relationship between these two clusters at the post level.

5.2.3 Post-level Design

The post level concentrates on the post-to-post propagation between two clusters of the same topic on different platforms (see Fig. 2).

Encoding and Layout. Each cluster is visualized as a circle, with posts in it using bubble chart layouts (Fig. 2(a1)). The encoding of a post is based on its influential metric (Fig. 2(a2)). If its post engagement index or number of followers exceeds a threshold, then the post is visualized as a flower whose size is directly proportional to the influential metric. Its inner circle's color represents the post's sentiment. Otherwise, a post is represented as a dot, whose color represents the post's sentiment. Users can adjust the thresholds through the control strips (Fig. 2(a3)). Between clusters, cross-platform paths extend from one post to another, with their width encoded by post-level cross-platform likelihood. The system only displays paths that have a considerable width with edge bundling. As users adjust thresholds, less significant paths are filtered out. When dragging the control strips to the rightmost position, only the most significant paths are visible on the screen.

Interaction. Clicking on the post's symbol (a flower or a dot) shows its detailed content (Fig. 2(a4)). Clicking on the animation control button (Fig. 1(b2)) shows the dynamic process of all the paths sequentially. Clicking on a propagation path shows the details of this post pair (Fig. 2(a5)). The two posts' information is displayed in parallel for easy comparison, with specific propagation words and similar contents like hashtags and URLs highlighted. The cross-platform diffusion pattern of this case identified by our model is also directly displayed. Meanwhile, the history information of these two posts' posting accounts is shown in the 'History Relevance' module in Detail View, including their history posts (see Fig. 2(b1)) and interest distribution (see Fig. 2(b2)).

6 Evaluation

We validate the effectiveness of BloomWind through two case studies and expert interviews. As the party that raised the requirements without engaging in solution development, our domain experts could help conclude whether the requirements were satisfied. We also interviewed a computer scientist and an industry engineer to collect their feedback on usability and effectiveness.

6.1 Case Studies

To evaluate BloomWind, we invited our domain expert to conduct two case studies about two events, namely Nuclear Wastewater Pouring and the U.S. Presidential Election. The Nuclear Wastewater Pouring dataset was collected from Twitter and Weibo, with 1445 tweets and 2948 Weibo posts from April 20, 2021 to December 10, 2023. The Presidential Election dataset was collected from Twitter, Weibo, and Facebook, with 783 tweets, 1875 Weibo posts, and 172 Facebook posts from August 10, 2023 to March 25, 2024. The data was collected using keywords like "japan wastewater", "primaries trump", etc.

6.1.1 Nuclear Wastewater Pouring

In this case, the expert explored the information diffusion between Twitter and Weibo related to the nuclear wastewater pouring event (a Japanese government's plan to release treated nuclear wastewater from the Fukushima nuclear plant into the sea). BloomWind provides an overview of cross-platform propagation whose topics and patterns evolve along the timeline (see Fig. 1).

Browsing Timeline View, the expert can see a distinct "peak" (see Fig. 1(b1)) during late August 2023, indicating that the platforms had a very hot discussion about this event during this period. Selecting this time frame, according to the Detail View (see Fig. 1c), seven main topics were discussed, namely, "240_china_nuclear_pollution", "cooling_water_nuclear_wastewater", "70_billion_japan_water", "sue_TEPCO_japan", "south_korea_nuclear_discharge", "radioactive_pollution_japan_sea", and "treatment_japan_waste_nuclear". The Propagation View (see Fig. 1a) shows the mutual diffusion pattern. It showcases how certain topics propagated from Twitter to Weibo, whereas others followed the reverse path.

Wider pathways reflected the topics with greater potential for cross-platform propagation, which caught the expert's attention first. The topic "240_china_nuclear_pollution" (see Fig. 1(a1)) had a strong propagation likelihood from Weibo to Twitter. BloomWind is helpful in quickly understanding the main content of the topic cluster. By reading the representative word cloud, the representative posts (see Fig. 1(a2)), and conducting a cursory scan of the post list, the expert identified that the discussions within the "240_china_nuclear_pollution" cluster primarily revolved around the estimated 240-day journey of Japan's nuclear wastewater to reach China. The Weibo glyph "240_china_nuclear_pollution" (see Fig. 1(a3)) represented an interesting diffusion pattern, with only influential posts written by influential accounts involved at first. Subsequently, an increasing number of influential accounts joined the discussion, attracting many ordinary accounts. As the event's popularity waned, the presence of influential accounts significantly decreased, and the activity level of ordinary accounts also declined. This seemed like a topic that could spread between Twitter and Weibo, therefore the expert right-clicked on this propagation line to access more detailed information. He observed that the most significant post-level propagation paths demonstrated a clear trend from Weibo to

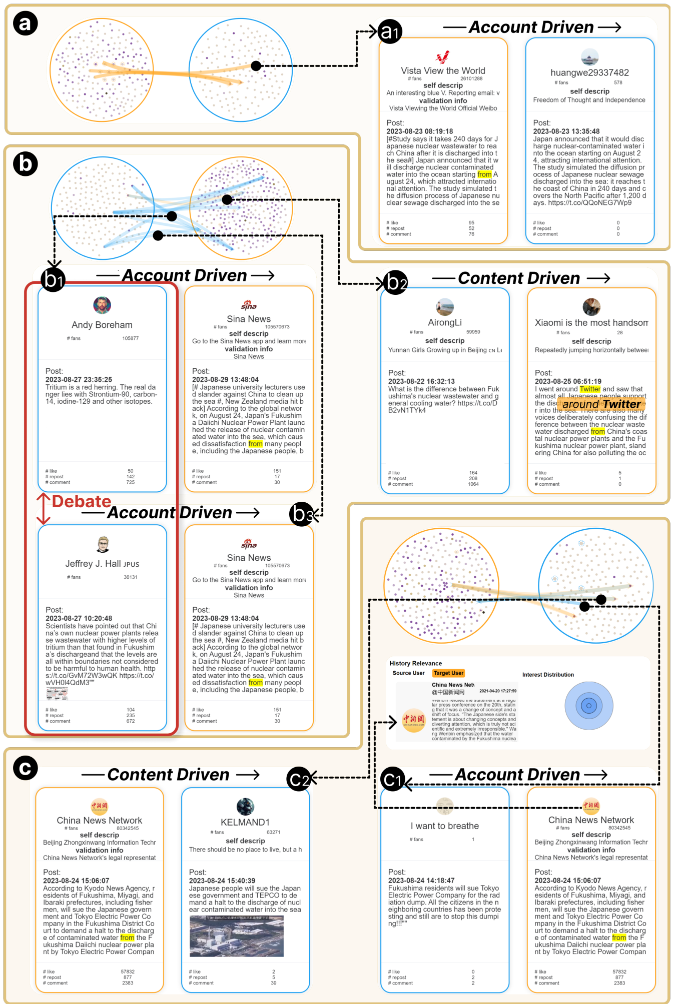

Fig. 4: Post-level Discoveries of Nuclear Wastewater Pouring (I). (a) Topic “240_china_nuclear_pollution”: About the estimated 240-day journey of Japan’s nuclear wastewater to reach China. It had a clear trend from Weibo to Twitter driven by relaying accounts. (b) Topic “cooling_water_nuclear_wastewater”: About the difference between cooling water and nuclear wastewater. It mainly diffused from Twitter to Weibo driven by news accounts and content. (c) Topic “sue_TEPCO_japan”: About Japanese people suing Tepco over nuclear wastewater. It showcased bidirectional propagation driven by media accounts and content.

Twitter (see Fig. 4(a)). Clicking on some significant propagation paths, he found a clear diffusion pattern driven by the positioning of relaying accounts based on the high degree of textual similarity (see Fig. 4(a1)). The topic “cooling_water_nuclear_wastewater” (see Fig. 1(a4)) was about the difference between cooling water and nuclear wastewater. As the largest topic cluster during this period, with many influential users and posts involved, the expert assessed that this topic possessed strong cross-platform propagation potential. The seeds on its cluster-level propagation path unveiled salient instances of cross-platform propagation, with many Weibo posts mentioning ‘Twitter’ (see Fig. 1(a5)). Its post-level propagation paths, mainly from Twitter to Weibo, revealed some interesting patterns (Fig. 4(b)). Certain news accounts on Weibo reported the debate between two Twitter users about this topic, attracting some attention within Weibo (see Fig. 4(b1)). Meanwhile, certain Weibo users expressed their disapproval of the tendency among many Twitter users to conflate two distinct types of water (see Fig. 4(b2)). Based on their wording and account information, the expert inferred that they were ordinary users. Their behavior exemplifies a typical instance of cross-platform diffusion driven by content, where people chose to express their thoughts on Weibo instead of on Twitter

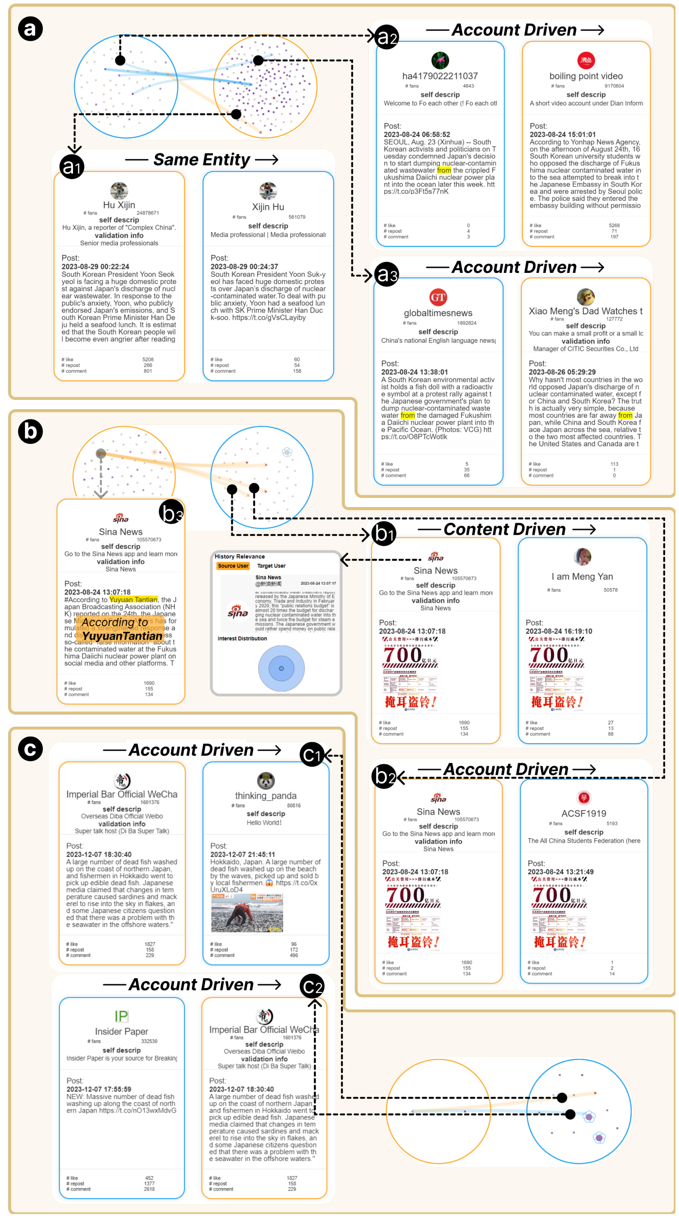

Fig. 5: Post-level Discoveries of Nuclear Wastewater Pouring (II). (a) Topic “south_korea_nuclear_discharge”: About Korea’s reactions towards the nuclear wastewater pouring event. It revealed bidirectional propagation driven by same-entity and media accounts. (b) Topic “70_billion_japan_water”: About Japan’s 70 billion yen public relations budget for nuclear sewage discharge. It showed a clear diffusion from one Weibo account to multiple Twitter accounts driven by content and media accounts. (c) Topic “japan_dead_fish”: About a mass of dead fish off the coast of Japan. Its cross-platform diffusion showed an interesting export-to-domestic sales mode driven by news accounts.

since more users on Weibo supported their standpoint. Similar account positioning-driven patterns occurred in the topic “sue_TEPCO_japan” (see Fig. 4(c)), where some media accounts enthusiastic about this topic (Japanese people sued TEPCO) continuously reported related information across platforms (see Fig. 4(c1)). Content-driven patterns were also detected automatically and validated by the expert (see Fig. 4(c2)).

The cluster-level propagation paths of other topics were relatively narrow, so the expert chose to focus on the two topics that appear to have clear semantics. The topic “south_korea_nuclear_discharge” was about Korea’s attitude and actions towards the event. Various patterns were detected among posts on this topic. Similar user information

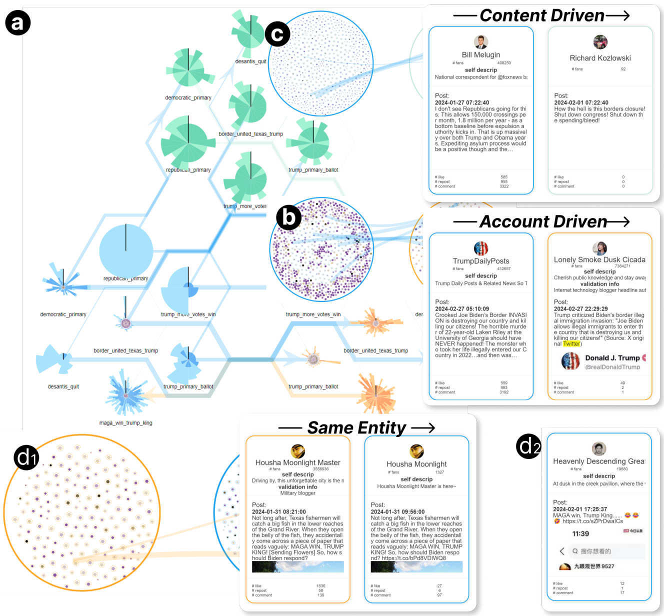

Fig. 6: Discoveries of U.S. 2024 Presidential Election. The diffusion was primarily from Twitter to Facebook and Weibo, mainly driven by news accounts for Weibo and by content for Facebook. Some interesting non-linear diffusion modes were also found.

assisted in locating a case of propagation between the same entity's accounts. The expert discovered that Xijin Hu spread the news of Yoon Suk-yeol's seafood lunch from his Weibo account to his Twitter account (see Fig. 5(a1)). It was interesting to find that Hu's words were more moderate on Twitter. The expert speculated that the variation of the audience's characteristics led him to adopt different expression strategies. Additionally, many media accounts on Weibo disseminated news about public protests in South Korea, because a significant portion of Weibo users shared a similar viewpoint and expressed concerns regarding protests taking place in foreign countries (see Fig. 5(a2-a3)). In the topic "70_billion_japan_water", BloomWind showed a clear cross-platform propagation from one Weibo account to multiple Twitter accounts for different reasons (see Fig. 5(b1-b2)). Users attracted by the content and some media accounts spread the information to Twitter. By checking the watermark on the images carefully, the expert found that the source post pointed out by BloomWind was not the image's real origin. However, the source post detected by our model was a very influential one posted by an active user under the topic. Its content explicitly revealed the real information source (see Fig. 5(b3)). Therefore, the expert assumed our model accurately pointed out the direction in identifying cross-platform propagation. Moreover, the expert observed that during early December 2023, there was an interesting export-to-domestic sales pattern in the topic "japan_dead_fish". A Weibo news account 'Imperial Bar Official WeChat' spread the news about dead fish occurrence in Japan from a Twitter account 'Insider Paper' (see Fig. 5(c2)). Then Twitter user 'thinking_panda' saw this Weibo post and shared the information on its tweet (see Fig. 5(c1)).

This case study demonstrates the effectiveness of BloomWind to explore and understand cross-platform propagation. The visualization design is meaningful, and the model can successfully infer the propagation path as well as pattern types. The most frequent pattern in this case is diffusion driven by account positioning, and content-driven propagation is also common. Many news media and the general public play a role in facilitating cross-platform information exchange.

6.1.2 U.S. Presidential Election

BloomWind also supports simultaneous research on the propagation relationship between more than two platforms. During the U.S. presidential election, there was a clear trend from Twitter to other platforms like Weibo and Facebook (see Fig. 6a). The expert speculated that these phenomena could be attributed to two factors. First, the majority of Twitter's users are located in the United States, enabling faster access to relevant news. Second, in recent years, Twitter has shown higher levels of activity compared to Facebook. The right part of Fig. 6 shows the post-level propagation paths in the topic "border_united_texas_trump", which was about the U.S.-Mexico border issue and two presidential candidates' different attitudes. Checking the paths, the expert found some evident cross-platform diffusion since many users kept the habit of containing the source while bringing some news onto their platform. However, the primary diffusion pattern from Twitter to Weibo (see Fig. 6b) differed from that of Twitter to Facebook (see Fig. 6c). The former was account-driven diffusion by media accounts, while the latter was content-driven diffusion. The expert assumed this could be attributed to the higher overlap of users between Twitter and Facebook. Consequently, much propagation from Twitter to Facebook was content-driven, incorporating more personal attitudes and comments.

Additionally, the expert found an interesting propagation case with the assistance of BloomWind. Our model detected a same-entity diffusion from Weibo to Twitter (see Fig. 6(d1)). However, although he spread this interesting idea to Twitter, another Twitter user encountered this meme elsewhere and propagated it to Twitter again (see Fig. 6(d2)). This example demonstrated that information propagation between platforms is not a linear process and can involve more complex situations. This case study showcases the scalability of BloomWind to support the analysis of multiple platforms.

6.2 Expert Interview

We conducted semi-structured interviews with three experts: our collaborator in media and communication (P1), a computer science professor (P2), and an industry engineer (P3). P2 has studied information diffusion for over 5 years with a focus on propagation algorithms. P3 is developing a commercial product for public opinion analysis and has sufficient expertise for information propagation. Each interview took about an hour. We first introduced BloomWind and let the expert explore our system freely (10 minutes). After that, P1 conducted case studies, and we demonstrated these case studies to the other two experts (30 minutes). Then we held a post-interview discussion (20 minutes). All experts were required to comment on the usability of BloomWind and the cases. Moreover, we interviewed P1 to determine if all requirements were satisfied and evaluate the effectiveness of BloomWind. P2 and P3 were not involved in our study before the interviews. Our discussion with P2 was mainly around the effectiveness of the cross-platform influence model. We discussed with P3 about how this system could be applied in real-world applications.

After a short introduction, all experts learned to use BloomWind quickly and provided positive feedback on its usability. They appreciated the interface design and the layered exploration approach. They found these vivid metaphors very intuitive and effective. P1 appreciated the cluster-level flower glyph for its clarity and comprehensive information presentation. He could learn about the variation trend of a topic on a platform within a time frame simply by reading the petals along the circle. Moreover, the stacked color design and gray half-rings gave him insights into the influential posts and users, which he regarded as very important in social media analysis. He also liked the idea of using the inner circle size to encode the total post number, which gave him a straight comparison between different topics to find the dominant ones. P1 opined that the color coding and the animation of the propagation paths offered meaningful insights into the overall input-output relationships between platforms, providing a valuable reference for understanding the dynamics of information flow. He thought the cluster-level layout was visually attractive but deficient in information density. He suggested enhancing the layout's meaningfulness by organizing it based on semantic relevance or propagation likelihood. Moreover, the post-level Propagation View provided an overview of the overall diffusion situation through paths, but he also wanted to see specific statistics on the portion of diffusion in different directions. He also commented, "Given a topic, what are the proportions of these cross-platform diffusion patterns: same-entity, account-driven, and content-driven? And what is the trend of these proportions over time?" He emphasized its importance for defining the inherent driving force behind a topic, which is valuable for his further studies.

P1 and P2 praised the detection ability of our model and found it helpful for their cross-platform propagation analysis. Through the post-level propagation paths, they could find many interesting cross-platform propagation cases and see their details. The pattern type detected by our model also helped them a lot in exploring the driving force of specific diffusion cases. P2 appreciated the approach of summarizing patterns, extracting significant factors, and subsequently constructing the model based on these factors. He commented that the calculations for each factor took into account various aspects affecting propagation judgments, offering clear and understandable interpretations. He also mentioned that our pattern summary and model design were inspirational for his future research. P2 noted that our model holds potential for further development, suggesting the incorporation of multimodal information processing. P3 confirmed the practical value of BloomWind in correlating multi-platform public opinions from an industry standpoint. P3 appreciated our model's ability to correlate public opinion across various platforms, emphasizing its significance in obtaining a holistic perspective of an event. He stated, "We can use BloomWind to identify the most discussed topics, determine which platforms are the main information sources, and capture significant cross-platform propagation cases. It's very helpful for us in grasping public opinion."

7 Discussion

This section discusses the implications, limitations, and, future work of BloomWind, while also introducing the design lessons learned.

7.1 Implications

Our study has several important implications. First, to the best of our knowledge, BloomWind is the first trial to visualize the cross-platform propagation among posts. Implicit as cross-platform propagation is, we introduce a novel visual metaphor to reveal this process as flowers in different gardens exchange their seeds. It is appropriate to compare the information spread between platforms to species dispersal in an ecosystem. The dispersal of species and the structure of niches in ecosystems offer explicit explanations for comprehending both macro and micro cross-platform propagation dynamics. By disassembling and integrating the factors in cross-platform propagation, we present our model's speculation and relevant basis for diffusion with this appealing and effective visual metaphor. Second, we systematically summarize the cross-platform diffusion patterns and their various mechanisms, which also instructs the design of our cross-platform influence model. The summary offers insights into the common driving forces of cross-platform diffusion, offering a fresh perspective for creating new information diffusion models and visualizations for future research in cross-platform propagation analysis. Third, we develop BloomWind to facilitate cross-platform propagation analysis and assist users in discovering cross-platform patterns and investigating the underlying mechanisms. This tool presents a workflow paradigm for studying cross-platform diffusion through its hierarchical design and "overview + detail" lenses. Given an event appears on multiple platforms, the experts first need to figure out what is discussed on these platforms. Then, they need to find out what topics have the potential and characteristics of cross-platform propagation by reading the details summarized at the topic cluster level. Posts on these topics can be further studied. After learning about the overall picture of posts, significant propagation paths between posts should be checked in detail. After identifying cross-platform diffusion instances, it is meaningful to analyze the inherent driving force, which is useful for the experts to figure out the evolution of public opinions.

7.2 Design Lessons Learned

Balancing aesthetics and effectiveness appropriately is crucial in metaphor design. Several discoveries from our iterative design process can be used to illustrate this point. Initially, to visually represent the patterns of information input and output between platforms, we contemplated encoding the path color using the distinctive color associated with the source platform. In this approach, the information exchange between platforms could be immediately discernible through color intensity, while also conveying path directions. However, user feedback suggested that employing colors to denote path direction lacked sufficient intuitiveness, so we opted to incorporate arrows along the paths to indicate the directions directly. Our cluster-level flower glyph is praised by experts due to its intuitiveness and expressiveness, allowing for understanding the chronological story along the circle. However, using a looping circle to represent the linear time flow could bias users. Therefore, we ultimately drew a vertical line between the petals representing the start and end date for clarity. We designed an icon for each cross-platform diffusion pattern type to align with our metaphor, which would appear after clicking on a post-level propagation path. However, it was hard to understand their meaning without enough annotation. Consequently, to avoid unnecessary over-designing, we replaced them with direct words. Sometimes, a metaphor design may seem great in its own context but can be confusing in practical use. Therefore, iterative design and careful balance are required when using metaphor design.

7.3 Limitations and Future Work

There are several limitations to be solved and shed light on the potential future work. First, data scarcity and heterogeneity are common challenges in social media propagation analysis, especially in cross-platform situations. To address such issues, we implement a strategy whereby certain discretionary indicators are computed solely when their corresponding fields are present. Concerning data heterogeneity, more precise localization of diverse fields is essential to harness valuable insights and portray distinct patterns in an enhanced manner. These practices are usually labor-intensive and require further improvements. One promising direction is to incorporate Large Language Models (LLMs) for data preprocessing. However, integrating LLMs raises concerns about data privacy and availability. Ensuring data anonymization and adherence to privacy regulations, as well as managing data quality and access, will be significant for system development. Second, we plan to consider multimodal analysis in our system. Users on Platforms like TikTok and Instagram mainly share photos and videos. Even on more traditional platforms like Twitter and Weibo, there exists a trend towards visual communication. In our case studies, we have observed some common photos across multiple platforms, indicating that media resources, often marked by distinct watermarks, could serve as a robust indicator of cross-platform diffusion. To enhance multimodal analysis, we plan to integrate Siamese networks to compare image similarity and employ Optical Character Recognition technologies to extract textual information from visual content. Ultimately, expert interviews have indicated that there is a desire for the overview, such as a temporal view, to display more information regarding diffusion patterns, which remains an objective for future developments.

8 Conclusion

In this paper, we present a visualization system BloomWind for exploring cross-platform information propagation. We first systematically summarize the cross-platform diffusion patterns and their underlying mechanisms. Based on the above summary and related literature, we extract crucial factors to build our cross-platform influence model. The model estimates the cross-platform propagation likelihood of a topic/post and identifies the type of its diffusion pattern. Then we provide a novel visual metaphor to demonstrate the model's results as the interaction between different gardens. Hierarchical design and "overview + detail" lenses are presented to facilitate understanding and exploration. Two case studies and expert interviews have validated the effectiveness and usefulness of BloomWind. In the future, we intend to further improve BloomWind by solving the limitations.

Acknowledgments

The work was supported by National Key R&D Program of China (2023YFB3107100), Key "Pioneer" R&D Projects of Zhejiang Province (2023C01120), NSFC (U22A2032, 62302440), and the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU).

REFERENCES

[1] Revisit. Website, 2013. `http://truth-and-beauty.net/projects/revisit/`. 2

[2] Douban. Website, 2023. `https://www.douban.com`. 3

[3] N. Andalibi and J. Buss. The human in emotion recognition on social media: Attitudes, outcomes, risks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 551:1–551:16. ACM, New York, 2020. doi: 10.1145/3313831.3376680 1

[4] J. R. Arie Croitoru, Andrew Crooks and A. Stefanidis. Geosocial gauge: a system prototype for knowledge discovery from social media. *International Journal of Geographical Information Science*, 27(12):2483–2508, 2013. doi: 10.1080/13658816.2013.825724 2

[5] A. Arleo, W. Didimo, G. Liotta, S. Miksch, and F. Montecchiani. Influence maximization with visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 28(10):3428–3440, 2022. doi: 10.1109/TVCG.2022.3190623 2

[6] M. Bossetta. The digital architectures of social media: Comparing political campaigning on facebook, twitter, instagram, and snapchat in the 2016 u.s. election. *Journalism & Mass Communication Quarterly*, 95(2):471–496, 2018. doi: 10.1177/1077699018763307 1

[7] 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 1, 2

[8] N. Cao, C. Shi, S. Lin, J. Lu, Y.-R. Lin, and C.-Y. Lin. Targetvue: Visual analysis of anomalous user behaviors in online communication systems. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):280–289, 2016. doi: 10.1109/TVCG.2015.2467196 2

[9] J. Caverlee, Z. Cheng, D. Z. Sui, and K. Y. Kamath. Towards geo-social intelligence: Mining, analyzing, and leveraging geospatial footprints in social media. *IEEE Data Engineering Bulletin*, 36(3):33–41, 2013. 2

[10] 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. IEEE, Piscataway, 2017. doi: 10.1109/VAST.2017.8585638 2

[11] 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. IEEE, Piscataway, 2016. doi: 10.1109/VAST.2016.7883510 2

[12] 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, 2020. doi: 10.1109/TVCG.2019.2934263 1, 2

[13] S. Chen, L. Lin, and X. Yuan. Social media visual analytics. *Computer Graphics Forum*, 36(3):563–587, 2017. doi: 10.1111/cgf.13211 2, 3

[14] Z. Chen, K. Aslett, J. Reynolds, J. Freire, J. Nagler, J. Tucker, and R. Bonneau. An automatic framework to continuously monitor multi-platform information spread. In *Proceedings of the Workshop on Misinformation Integrity in Social Networks*, pp. 49–68. CEUR-WS, Aachen, 2021. 2, 3

[15] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec. Can cascades be predicted? In *Proceedings of the International Conference on World Wide Web*, pp. 925–936. ACM, New York, 2014. doi: 10.1145/2566486.2567997 2

[16] W. Cui, X. Wang, S. Liu, N. H. Riche, T. M. Madhyastha, K. L. Ma, and B. Guo. Let it flow: A static method for exploring dynamic graphs. In *Proceedings of IEEE Pacific Visualization*, pp. 121–128. IEEE, Piscataway, 2014. doi: 10.1109/PacificVis.2014.48 2

[17] L. Dinh and N. Parulian. Covid-19 pandemic and information diffusion analysis on twitter. *Proceedings of the Association for Information Science and Technology*, 57(1):252–261, 2020. doi: 10.1002/pra2.252 2

[18] T. Dong, C. Liang, and X. He. Social media and internet public events. *Telematics and Informatics*, 34(3):726–739, 2017. doi: 10.1016/j.tele.2016.05.024 1

[19] C. Gao, T.-H. Lin, N. Li, D. Jin, and Y. Li. Cross-platform item recommendation for online social e-commerce. *IEEE Transactions on Knowledge and Data Engineering*, 35(2):1351–1364, 2023. doi: 10.1109/TKDE.2021.3098702 1

[20] V. L. Gatta, L. Luceri, F. Fabbri, and E. Ferrara. The interconnected nature of online harm and moderation: Investigating the cross-platform spread of harmful content between youtube and twitter. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media*, pp. 39:1–39:10. ACM, New York, 2023. doi: 10.1145/3603163.3609058 1, 2, 3

[21] M. Grootendorst. BERTopic: Neural topic modeling with a class-based tf-idf procedure. arXiv preprint, 2023. To appear. doi: 10.48550/arXiv.2203.05794 4

[22] C.-T. Ho, C.-T. Li, and S.-D. Lin. Modeling and visualizing information propagation in a micro-blogging platform. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 328–335. IEEE, Piscataway, 2011. doi: 10.1109/ASONAM.2011.37 2

[23] C.-C. Hu, H.-X. Wei, and M.-T. Chi. Shareflow: A visualization tool for information diffusion in social media. In *Proceedings of the International Conference on Ubiquitous Information Management and Communication*, pp. 563–581. Springer, Cham, 2019. doi: 10.1007/978-3-030-19063-7_45 2

[24] Y. Hu, R. J. Song, and M. Chen. Modeling for information diffusion in online social networks via hydrodynamics. *IEEE Access*, 5:128–135, 2017. doi: 10.1109/ACCESS.2016.2605009 2, 3, 4

[25] K. Hunt, B. Wang, and J. Zhuang. Misinformation debunking and cross-platform information sharing through twitter during hurricanes harvey and irma: a case study on shelters and id checks. *Natural Hazards*, 103(1):861–883, 2020. doi: 10.1007/s11069-020-04016-6 1, 2, 3

[26] S. Jain and A. Sinha. Identification of influential users on twitter: A novel weighted correlated influence measure for covid-19. *Chaos, Solitons Fractals*, 139:110037:1–110037:8, 2020. doi: 10.1016/j.chaos.2020.110037 2

[27] K. Y. Kamath and J. Caverlee. Spatio-temporal meme prediction: learning what hashtags will be popular where. In *Proceedings of the ACM International Conference on Information & Knowledge Management*, pp. 1341–1350. ACM, New York, 2013. doi: 10.1145/2505515.2505579 2

[28] E. Katz, P. F. Lazarsfeld, and E. Roper. *Personal influence: The part played by people in the flow of mass communications*. Routledge, 2017. doi: 10.4324/9781315126234 2, 3, 4

[29] S. Lee, M. S. Cho, and T.-Q. Peng. Understanding sentiment toward racial unrest through temporal and geographic lenses: a multilevel-analysis across metropolitan areas in the united states. *Frontiers in Communication*, 9:1259046:1–1259046:14, 2024. doi: 10.3389/fcomm.2024.1259046 1

[30] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 497–506. ACM, New York, 2009. doi: 10.1145/1557019.1557077 2, 4

[31] Q. Li, H. Qu, L. Chen, R. Wang, J. Yong, and D. Si. Visual analysis of retweeting propagation network in a microblogging platform. In *Proceedings of the International Symposium on Visual Information Communication and Interaction*, pp. 44–53. ACM, New York, 2013. doi: 10.1145/2493102.2493108 2

[32] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences*, 102(33):11623–11628, 2005. doi: 10.1073/pnas.0503018102 2

[33] A.-A. Liu, X. Wang, N. Xu, J. Guo, G. Jin, Q. Zhang, Y. Tang, and S. Zhang. A review of feature fusion-based media popularity prediction methods. *Visual Informatics*, 6(4):78–89, 2022. doi: 10.1016/j.visinf.2022.07.003 2

[34] J. Liu, W. Chung, Y. Huang, and C. Toraman. CrossSimON: A novel probabilistic approach to cross-platform online social network simulation. In *IEEE International Conference on Intelligence and Security Informatics*, pp. 7–12. IEEE, Piscataway, 2019. doi: 10.1109/ISI.2019.8823276 2

[35] M. Liu, S. Liu, X. Zhu, Q. Liao, F. Wei, and S. Pan. An uncertainty-aware approach for exploratory microblog retrieval. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):250–259, 2016. doi: 10.1109/TVCG.2015.2467554 2

[36] S. Liu, X. Wang, J. Chen, J. Zhu, and B. Guo. Topicpanorama: A full picture of relevant topics. *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2508–2521, 2016. doi: 10.1109/TVCG.2016.2515592 2

[37] Y. Liu, C. Wang, P. Ye, and K. Zhang. Analysis of micro-blog diffusion using a dynamic fluid model. *Journal of Visualization*, 18(2):201–219, 2015. doi: 10.1007/s12650-015-0277-y 2

[38] Y. Lu, J. Schaefer, K. Park, J. Joo, and J. Pan. How information flows from the world to china. *The International Journal of Press/Politics*, 29(2):305–327, 2024. doi: 10.1177/19401612221117470 2, 3, 4

[39] Y. Lu, H. Wang, S. Landis, and R. Maciejewski. A visual analytics framework for identifying topic drivers in media events. *IEEE Transactions on*

Visualization and Computer Graphics, 24(9):2501–2515, 2018. doi: 10.1109/TVCG.2017.2752166 2

[40] J. Lukito. Coordinating a multi-platform disinformation campaign: Internet research agency activity on three u.s. social media platforms, 2015 to 2017. *Political Communication*, 37(4):238–255, 2019. doi: 10.1080/10584609.2019.1661889 3

[41] S. Macskassy and M. Michelson. Why do people retweet? anti-homophily wins the day! In *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 209–216. AAAI Press, Palo Alto, 2021. doi: 10.1609/icwsm.v5i1.14110 2, 4

[42] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 227–236. ACM, New York, 2011. doi: 10.1145/1978942.197897 2

[43] I. Murdock, K. M. Carley, and O. Yağan. Identifying cross-platform user relationships in 2020 u.s. election fraud and protest discussions. *Online Social Networks and Media*, 33:100245:1–100245:16, 2023. doi: 10.1016/j.osnem.2023.100245 1, 2, 3, 4

[44] G. Murić, A. Tregubov, J. Blythe, A. Abeliuk, D. Choudhary, K. Lerman, and E. Ferrara. Large-scale agent-based simulations of online social networks. *Autonomous Agents and Multi-Agent Systems*, 36(2):38:1–38:21, 2022. doi: 10.1007/s10458-022-09565-7 2

[45] F. Nightingale. *Notes on matters affecting the health, efficiency, and hospital administration of the British army: founded chiefly on the experience of the late war*. Harrison and Sons, 1987. 5

[46] O. Papakyriakopoulos, J. C. M. Serrano, and S. Hegelich. The spread of covid-19 conspiracy theories on social media and the effect of content moderation. *Harvard Kennedy School Misinformation Review*, 08(3):34:1–34:19, 2020. doi: 10.37016/mr-2020-034 1, 2, 3

[47] T.-Q. Peng and J. Zhu. Competition, cooperation, and coexistence: An ecological approach to public agenda dynamics in the united states (1958-2020). *Communication Research (OnlineFirst)*, 2023. Published online ahead of print. doi: 10.1177/00936502221118156 1

[48] D. Ren, X. Zhang, Z. Wang, J. Li, and X. Yuan. Weiboevents: A crowd sourcing weibo visual analytic system. In *Proceedings of IEEE Pacific Visualization*, pp. 330–334. IEEE, Piscataway, 2014. doi: 10.1109/PacificVis.2014.38 2

[49] D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *Proceedings of the International Conference on World Wide Web*, pp. 695–704. ACM, New York, 2011. doi: 10.1145/1963405.1963503 2

[50] K. Saha, J. Seybolt, S. M. Mattingly, T. Aledavood, C. Konjeti, G. J. Martinez, T. Grover, G. Mark, and M. De Choudhury. What life events are disclosed on social media, how, when, and by whom? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 335:1–335:22. ACM, New York, 2021. doi: 10.1145/3411764.3445405 1

[51] K. Saito, R. Nakano, and M. Kimura. Prediction of information diffusion probabilities for independent cascade model. In *Proceedings of the International Conference on Knowledge-based and Intelligent Information and Engineering Systems*, pp. 67–75. Springer, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-85567-5_9 2

[52] D. Sehnan, V. Goel, S. Masud, C. Jain, V. Goyal, and T. Chakraborty. Diva: A scalable, interactive and customizable visual analytics platform for information diffusion on large networks. *ACM Transactions on Knowledge Discovery from Data*, 17(4):47:1–47:33, 2023. doi: 10.1145/3558771 2

[53] H. Shen, L. Tu, Y. Guo, and J. Chen. The influence of cross-platform and spread sources on emotional information spreading in the 2e-sir two-layer network. *Chaos, Solitons Fractals*, 165:112801:1–112801:11, 2022. doi: 10.1016/j.chaos.2022.112801 2, 3

[54] Y. Song, Z. Wen, C.-Y. Lin, and R. Davis. One-class conditional random fields for sequential anomaly detection. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1685–1691. AAAI Press, Palo Alto, 2013. doi: 10.1145/2505515.2505579 2

[55] G. Sun, T. Tang, T.-Q. Peng, R. Liang, and Y. Wu. SocialWave: Visual analysis of spatio-temporal diffusion of information on social media. *ACM Transactions on Intelligent Systems and Technology*, 9(2):15:1–15:23, 2017. doi: 10.1145/3106775 2, 4

[56] 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 1, 2, 4

[57] N. Velasquez, R. Leahy, N. J. Restrepo, Y. Lupu, R. Sear, N. Gabriel, O. Jha, B. Goldberg, and N. Johnson. Online hate network spreads malicious covid-19 content outside the control of individual social media platforms. *Scientific Reports*, 11(1):11549:1–11549:8, 2021. doi: 10.1038/s41598-021-89467-y 1

[58] F. Viégas, M. Wattenberg, J. Hebert, G. Borggaard, A. Cichowlas, J. Feinberg, J. Orwant, and C. Wren. Google+ripples: a native visualization of information flow. In *Proceedings of the International Conference on World Wide Web*, pp. 1389–1398. ACM, New York, 2013. doi: 10.1145/2488388.2488504 2

[59] 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. IEEE, Piscataway, 2016. doi: 10.1109/VAST.2016.7883511 2

[60] Z. Wang, H. Liu, W. Liu, and S. Wang. Understanding the power of opinion leaders' influence on the diffusion process of popular mobile games: Travel frog on sina weibo. *Computers in Human Behavior*, 109:106354:1–106354:11, 2020. doi: 10.1016/j.chb.2020.106354 2, 4

[61] T. Wilson and K. Starbird. Cross-platform disinformation campaigns: lessons learned and next steps. *Harvard Kennedy School Misinformation Review*, 1(1):2:1–2:11, 2020. doi: 10.37016/mr-2020-002 2, 3

[62] Y. Wu, Z. Chen, G. Sun, X. Xie, N. Cao, S. Liu, and W. Cui. Streamexplorer: a multi-stage system for visually exploring events in social streams. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2758–2772, 2018. doi: 10.1109/TVCG.2017.2764459 2

[63] Y. Wu, S. Liu, K. Yan, M. Liu, and F. Wu. Opinionflow: Visual analysis of opinion diffusion on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1763–1772, 2014. doi: 10.1109/TVCG.2014.2346920 1, 2, 3

[64] L. Xing, K. Deng, H. Wu, P. Xie, H. V. Zhao, and F. Gao. A survey of across social networks user identification. *IEEE Access*, 7:137472–137488, 2019. doi: 10.1109/ACCESS.2019.2942840 3

[65] J. Xu, C. Zhang, M. Xie, X. Zhan, L. Yan, Y. Tao, and Z. Pan. Imvis: Visual analytics for influence maximization algorithm evaluation in hypergraphs. *Visual Informatics*, 8(2):13–26, 2024. doi: 10.1016/j.visinf.2024.04.006 2

[66] P. Xu, Y. Wu, E. Wei, T.-Q. Peng, S. Liu, J. J. Zhu, and H. Qu. Visual analysis of topic competition on social media. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2012–2021, 2013. doi: 10.1109/TVCG.2013.221 2

[67] Y. Xu and T.-Q. Peng. Ecological constraints on audience size in the digital media system: evidence from the longitudinal tracking data from 2019 to 2022. *Human Communication Research*, 50(1):14–26, 2024. doi: 10.1093/hcr/hqad028 1

[68] J. Yang and S. Counts. Predicting the speed, scale, and range of information diffusion in twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, pp. 355–358. AAAI Press, Palo Alto, 2010. doi: 10.1609/icwsm.v4i1.14039 2

[69] P. Ye, C. Wang, Y. Liu, Q. Zhu, and K. Zhang. Visual analysis of microblog retweeting using an information diffusion function. *Journal of Visualization*, 19(4):823–838, 2016. doi: 10.1007/s12650-016-0347-9 2

[70] X. Yuan, Z. Wang, Z. Liu, C. Guo, H. Ai, and D. Ren. Visualization of social media flows with interactively identified key players. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 291–292. IEEE, Piscataway, 2014. doi: 10.1109/VAST.2014.7042535 2

[71] J. Zhang, J. Tang, J. Li, Y. Liu, and C. Xing. Who influenced you? predicting retweet via social influence locality. *ACM Transactions on Knowledge Discovery from Data*, 9(3):25:1–25:26, 2015. doi: 10.1145/2700398 2

[72] L. Zhang, C. Su, Y. Jin, M. Goh, and Z. Wu. Cross-network dissemination model of public opinion in coupled networks. *Information Sciences*, 451–452:240–252, 2018. doi: 10.1016/j.ins.2018.04.037 2, 3

[73] M. Zhang, S. Qin, and X. Zhu. Information diffusion under public crisis in ba scale-free network based on seir model — taking covid-19 as an example. *Physica A: Statistical Mechanics and its Applications*, 571:125848:1–125848:19, 2021. doi: 10.1016/j.physa.2021.125848 2

[74] J. Zhao, N. Cao, Z. Wen, Y. Song, Y.-R. Lin, and C. Collins. # fluxflow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1773–1782, 2014. doi: 10.1109/TVCG.2014.2346922 1, 2