

Bridging the Narrative Divide: Cross-Platform Discourse Networks in Fragmented Ecosystems

Patrick Gerard¹, Hans W. A. Hanley², Luca Luceri¹, Emilio Ferrara³

¹Information Sciences Institute, University of Southern California

²Stanford University

³Thomas Lord Department of Computer Science, University of Southern California

pgerard@isi.edu, hhanley@cs.stanford.edu, lluceri@isi.edu, emiliofe@usc.edu

Abstract

Political discourse has grown increasingly fragmented across different social platforms, making it challenging to trace how narratives spread and evolve within such a fragmented information ecosystem. Reconstructing social graphs and information diffusion networks is challenging, and available strategies typically depend on platform-specific features and behavioral signals which are often incompatible across systems and increasingly restricted. To address these challenges, we present a platform-agnostic framework that allows to accurately and efficiently reconstruct the underlying social graph of users' cross-platform interactions, based on discovering latent narratives and users' participation therein. Our method achieves state-of-the-art performance in key network-based tasks: *information operation detection*, *ideological stance prediction*, and *cross-platform engagement prediction*—while requiring significantly less data than existing alternatives and capturing a broader set of users. When applied to cross-platform information dynamics between Truth Social and X (formerly Twitter), our framework reveals a small, mixed-platform group of *bridge users*, comprising just 0.33% of users and 2.14% of posts, who introduce nearly 70% of *migrating narratives* to the receiving platform. These findings offer a structural lens for anticipating how narratives traverse fragmented information ecosystems, with implications for cross-platform governance, content moderation, and policy interventions.

Code — <https://tinyurl.com/CANE2025>

Data — <https://tinyurl.com/CANE2025>

Introduction

Political communication today unfolds across a fragmented digital landscape, shaped by the dynamics of ideologically distinct platforms (Zhang et al. 2025). During the January 6th, 2021, Capitol insurrection, participants engaged in parallel but intersecting conversations across Parler, X (formerly Twitter), and Gab (Sipka, Hannak, and Urman 2022; Vishnuprasad et al. 2024), which were later amplified on mainstream platforms via news reporting (Zulli, Coe, and Isaacs 2023; Luceri, Cresci, and Giordano 2021). More recently, the emergence of Truth Social, a platform created by U.S. President Donald Trump, has introduced a new locus

of politically salient discourse (Zhang et al. 2025; Gerard, Botzer, and Weninger 2023), with narratives frequently originating there before diffusing outward (Shah et al. 2024). This fragmentation became particularly salient during the 2024 U.S. presidential election, when narratives, unverified claims, and emotionally charged rhetoric moved rapidly across ecosystem boundaries (Minici et al. 2024). Yet despite growing concern, we still lack systematic tools for tracing how discourse migrates between these siloed communities, or for identifying the structures that enable it.

Most traditional network approaches rely on platform-specific behaviors such as reposts, mentions, or follower ties (Cinelli et al. 2021), which are often incompatible across ecosystems and increasingly inaccessible due to API restrictions (Tromble 2021). More recent content-based methods infer user relationships through linguistic similarity (Ng, Cruickshank, and Carley 2022; Luceri et al. 2024a), but typically depend on pairwise message comparisons or token-level co-occurrence (Cinus et al. 2025). As we show in this work, such approaches fail to capture persistent alignment, generate unstable user communities, and generalize poorly across platforms. These limitations obscure the structural architecture of discourse, particularly the aspects that allow narratives to move between fragmented online communities.

To overcome these limitations and understand cross-platform dynamics during the 2024 U.S. Presidential election, we model users not through interaction patterns or surface-level text similarity, but as distributions over latent narratives. Our discourse-centered framework links users through shared participation in semantic clusters of ideas, capturing persistent alignment even without direct ties. This approach disentangles user similarity from platform-specific behavior, allowing for robust and scalable network construction across sparse and fragmented ecosystems.

Empirically, we find that this method offers several advantages. We test its effectiveness across both standard user representation tasks and a new benchmark for cross-platform engagement prediction. In both intra- and inter-platform settings, our method matches or exceeds the performance of existing approaches while requiring significantly less data. It captures a broader set of users, identifies stable communities across platforms, and supports large-scale modeling of discourse. Our method, without relying on platform-specific metadata, not only improves generalization and diffusion

modeling across siloed ecosystems but also performs competitively within single-platform contexts.

Applied to posts from Truth Social and X during the 2024 U.S. Presidential election, our framework reveals not only when and where political narratives migrated between ecosystems but also how and through whom this diffusion occurred. Utilizing our framework, we concretely ask:

1. How do political narratives migrate across fragmented platforms like X and Truth Social during the 2024 U.S. Presidential election, and what structural pathways enable their movement?
2. Do shared discourse communities emerge across X and Truth Social during the 2024 election cycle, and if so, how are they positioned within the broader structure of political communication?
3. Are certain users disproportionately associated with cross-platform narrative diffusion during the 2024 U.S. Presidential election, and what structural roles do they occupy within the discourse network?

Our analysis reveals that political narratives migrated between X and Truth Social in a highly structured manner. Many of these narratives migrated between X and Truth Social through a small but structurally distinct set of users, comprising just 0.33% of the total. These users accounted for only 2.14% of all posts, yet served as the initial cross-platform carriers for nearly 70% of narratives that successfully migrated. These *bridge users* were embedded at the intersection of fragmented discourse communities and consistently appeared early in the life cycle of narratives that later gained traction across ecosystems.

Critically, these insights do not rely on retrospective interaction patterns or explicit behavioral ties. Instead, they emerge from a structural perspective made possible by our modeling approach. By linking users through shared participation in latent discourse clusters, our platform-agnostic framework uncovers connective structures that remain hidden in traditional interaction-based or semantic similarity-based graphs. This shift opens new avenues for studying how influence operates across media environments and positions cross-platform discourse networks as a scalable and generalizable lens for understanding narrative diffusion.

Related Work

Network-Based User Representations. A central challenge in social media analysis is how to represent users in a way that reflects their relationships, ideological alignment, and interactions. Traditional methods rely heavily on platform-specific signals, including follower networks (Bollen et al. 2011), reposts and mentions (Cinelli et al. 2021), hashtags (Alieva, Ng, and Carley 2022; Burghardt et al. 2024), or URLs (Tardelli et al. 2024a). These signals, however, face increasing limitations due to restrictive data policies (Tromble 2021) and fragmentation of user activities across platforms (Minici et al. 2024). Semantic approaches, which construct graphs from linguistic similarity of user content (Ng, Cruickshank, and Carley 2023; Luceri et al. 2024a), attempt to bypass such lim-

itations. However, these methods typically rely on pairwise comparisons and often suffer from computational complexity and sensitivity to minor linguistic variations (Luceri et al. 2024a). Thus, current methods rarely provide platform-agnostic and scalable solutions.

Behavioral Modeling with Social Graphs. Once constructed, social graphs are typically evaluated by their utility in downstream tasks, including ideological stance prediction and coordination detection. Ideological stance prediction methods often combine content and network signals (Xiao et al. 2020; Jiang, Ren, and Ferrara 2023), yet struggle when metadata is sparse or inaccessible (Jiang, Ren, and Ferrara 2023). Coordination detection generally relies on identifying synchronized behaviors that often depend on behavioral traces that may be limited or inconsistent across platforms (Luceri et al. 2024a; Ng and Carley 2022; Magelinski, Ng, and Carley 2022). Overall, existing models are frequently constrained by platform-specific assumptions and data availability, limiting their adaptability in fragmented or low-signal environments.

Cross-Platform Dynamics and Behavior. As online discourse becomes increasingly fragmented across platforms, users and narratives move fluidly between ecosystems, disrupting traditional modeling assumptions (Ribeiro et al. 2021; Russo et al. 2023). Understanding these dynamics has become increasingly important, yet modeling them remains methodologically difficult. Cross-platform behavior is hard to capture due to the lack of consistent identifiers and semantic variation across ecosystems (Ng, Cruickshank, and Carley 2022, 2023; Luceri et al. 2024a; Cinus et al. 2025; Magelinski, Ng, and Carley 2022). For example, common strategies such as user-matching heuristics (Iofciu et al. 2011) and link-tracking (Cinus et al. 2025) often struggle with generalizability and scale. Despite growing interest in semantic similarity-based alignment methods, prior approaches remain limited in scope or too narrow to generalize across diverse ecosystems (Luceri et al. 2024a; Ng, Cruickshank, and Carley 2022). Meanwhile, influence operations exploit this fragmentation (Minici et al. 2024; Ng, Cruickshank, and Carley 2022), underscoring the need for user graph construction methods that are content-driven, platform agnostic, and robust to missing metadata.

Structural Influence and Network Theory. We connect our concept of *bridge users* to classic theories of structural influence. In sociology and organizational science, *boundary spanners* and *brokers* are actors who enable information flow across otherwise disconnected groups (Cross and Prusak 2002; Gould and Fernandez 1989). These roles emphasize that influence stems not only from activity or status, but from an actor’s position within a network. This framing aligns with Granovetter’s theory of weak ties (Granovetter 1983), which highlights how loosely connected individuals often facilitate the spread of novel information.

Methods: Platform-Agnostic Social Graphs

Illustrated in Figure 1, we introduce Cluster Affiliation Network Embedding (CANE): a data-efficient and platform-agnostic framework for constructing user-user networks di-

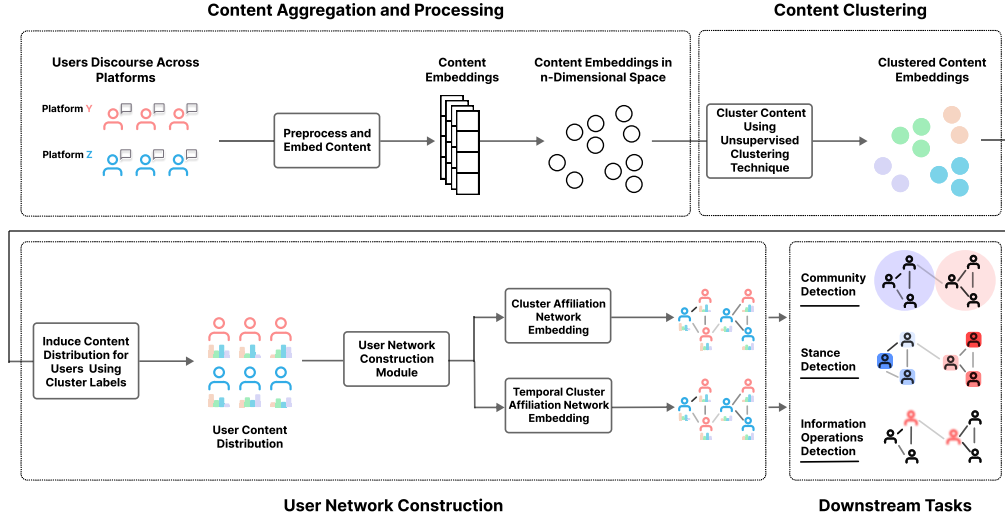


Figure 1: Overview of our cross-platform user network inference framework. Content is embedded and clustered into semantically coherent narratives, which form the basis for constructing user-user networks. These networks support a range of downstream tasks, including community detection, stance prediction, and content analysis.

rectly from content. Unlike traditional interaction-based approaches, our method does not depend on reposts, mentions, or other platform-specific signals. Instead, it links users through shared participation in latent narratives, supporting robust graph construction in fragmented, diverse, and sparse data environments. This reframes network modeling around discourse, replacing reliance on platform-specific interactions with the underlying structure of narrative alignment.

Given a set of users $U = u_1, \dots, u_n$ and their posts $C = c_1, \dots, c_m$, we construct a weighted, undirected graph $G = (V, E, w)$, where edges capture inferred behavioral alignment via shared participation in semantically coherent content clusters. Our pipeline includes three steps: (1) embedding user posts into a shared semantic space, (2) clustering posts into latent topics, and (3) connecting users based on shared cluster participation. We introduce two variants: **CANE**, a static model based on aggregated content, and **t-CANE**, a temporal extension. Full benchmarks and complexity analysis are provided in the Appendix.

Content Aggregation and Embedding

We begin by preprocessing post content for semantic modeling, removing platform-specific artifacts, such as URLs, hashtags, and user mentions. Standard normalization procedures, including whitespace trimming and character encoding cleanup, are applied. We embed each post using MP-Net (Song et al. 2020), a transformer model chosen for its strong performance on semantic similarity tasks. This yields a 768-dimensional vector for each post, which captures contextual meaning while remaining agnostic to platform-specific syntax. We note that our framework is embedding-agnostic and supports substitution with alternative monolingual, multilingual, or multimodal models (Wang et al. 2024).

Content Clustering

DP-Means. To identify coherent content clusters, we apply DP-Means to embed representations of user posts. Unlike k-means, which requires specifying the number of clusters in advance, DP-Means adaptively creates new clusters based on a distance threshold λ . This threshold, which we set using cosine distance and empirical semantic equivalence cut-offs (typically around 0.65 (Hanley, Kumar, and Durumeric 2023)), governs when new clusters are formed—allowing us to strike a balance between topical specificity and generality. Full algorithmic details, justification, and sensitivity for different λ values are provided in Appendix A.

User Network Construction

Once the content on particular platforms has been clustered (Figure 1), we utilize two unique strategies for constructing user-user edges, designed to balance expressiveness and computational efficiency. Both methods aim to robustly connect users based on shared narrative engagement, with one capturing static behavioral alignment and the other modeling temporal dynamics.

Cluster Affiliation Network Embedding (CANE). The first method we utilize for constructing user-user edges amongst social media after contracting their participation in different content clusters is based on CANE. CANE models user similarity by capturing how individuals align along content clusters, providing a proxy for their narrative preferences. CANE treats each user’s cluster participation as a document-term-like matrix (e.g. akin to methods like TF-IDF (Aizawa 2003)), where clusters are analogous to terms and user participation corresponds to frequency. This formulation allows us to assess not just whether users engage with the same clusters, but how distinctive those shared affiliations are relative to all content clusters.

This is such that for each user u , we define its full cluster

affiliation across all clusters C vector:

$$\mathbf{v}_u = [w_{u,1}, w_{u,2}, \dots, w_{u,|C|}] \quad (1)$$

where each weight $w_{u,c}$ for $c \in C$ is computed as:

$$w_{u,c} = t f_{u,c} \cdot \log \left(\frac{|U|}{|\{u : u \in c\}|} \right) \quad (2)$$

Here, $t f_{u,c}$ is the normalized frequency of user u 's participation in cluster c (i.e., the number of times that the user posted about the content in the cluster c), and the second log term penalizes broadly popular clusters, thereby emphasizing more distinctive affiliations.

Once cluster affiliation vectors are computed, we measure user similarity using cosine similarity, a standard metric for high-dimensional, sparse representations (Song et al. 2020; Luceri et al. 2024a; Pacheco et al. 2021; Wang et al. 2024). This captures the proportional overlap in narrative engagement, ensuring that users with similarly distributed content preferences are more strongly connected, even if they do not share identical cluster memberships. We evaluated alternative weighting strategies, including raw counts and softmax normalization, and found that a TF-IDF-inspired scheme yielded the best performance in downstream tasks. This supports the intuition that TF-IDF weighting better emphasizes distinctive cluster participation while down-weighting generic, viral, and ubiquitous content (Appendix B; Table 8).

To improve scalability, we compute user-user similarity using the GPU-supported FAISS¹ library (Malkov and Yashunin 2018). This approximate nearest neighbor method replaces brute-force comparisons, supports efficient edge construction, and produces sparse, high-quality user graphs. Empirical evaluation on multiple datasets shows that using FAISS-HNSW yields performance comparable to brute-force similarity computation, with no statistically significant degradation in F1 or AUC (Appendix B; Table 9).

Temporal Cluster Affiliation Network Embedding (t-CANE). While the static CANE model effectively captures user alignment through shared narrative engagement, it collapses time, obscuring the dynamics of when and how relationships form and shift. Yet many critical processes in political discourse, such as coordination, radicalization, and narrative seeding, are inherently temporal (Tardelli et al. 2024b). To capture these dynamics, we introduce **t-CANE**, a discrete temporal extension. This allows us to trace how user affinities evolve and to distinguish fleeting overlaps from sustained ideological alignment.

At each timestep t , we compute a TF-IDF-weighted cluster affiliation vector \mathbf{v}_u^t for each user u , summarizing their narrative participation during that window. Pairwise similarities are measured via cosine similarity. To construct a dynamic graph, we apply a memory-based update rule inspired by Hawkes processes (Laub, Taimre, and Pollett 2015), which reinforces repeated user alignment across timesteps while gradually weakening connections that lapse. This mechanism allows the graph to capture both persistent relationships and newly emerging discourse patterns. Neighborhoods at each timestep are computed using the GPU-supported FAISS library to ensure scalable similarity search.

¹<https://github.com/facebookresearch/faiss>

Network Type	Construction Description
Co-Repost	Connect users who repost the same piece of content.
Co-URL	Connect users who share the same URLs in their posts.
Fast Repost	Connect users who repost identical content within a short time window.
Hashtag Sequence	Connect users based on ordered sequences of shared hashtags.
Text Similarity	Connect users if they post at least one highly similar post; edge weight reflects average text similarity across matches (Pacheco et al. 2021).
k-NN Embedding Graph	Build a full text-to-text kNN graph first, then induce a user-to-user graph reflecting overall proximity across all posts (Ng, Cruickshank, and Carley 2023).
Fused Graph	Construct a unified network where users are linked if they are connected in any underlying similarity network (Co-Repost, Co-URL, Fast Repost, Hashtag Sequence, or Text Similarity) (Luceri et al. 2024a).

Table 1: Baseline similarity network construction methods.

Method	Macro-F ₁	AUC
Co-Repost	0.71 ± 0.02	0.77 ± 0.03
Co-URL	0.50 ± 0.02	0.56 ± 0.06
Fast Repost	0.46 ± 0.01	0.54 ± 0.02
Hashtag Sequence	0.51 ± 0.04	0.49 ± 0.03
Text Similarity	0.53 ± 0.01	0.33 ± 0.00
k-NN Embedding Graph	0.58 ± 0.02	0.39 ± 0.01
Fused Graph	0.74 ± 0.02	0.81 ± 0.02
CANE	0.72 ± 0.02	0.91 ± 0.01
t-CANE	0.83 ± 0.01	0.98 ± 0.01

Table 2: Performance on the China IO dataset.

We validate this design through downstream task performance and targeted ablation studies (Appendix C).

Benchmark Tasks and Evaluation

We evaluate our framework on three tasks: information operation detection, ideological stance prediction, and a new cross-platform engagement prediction benchmark. The first two tasks test intra-platform utility, while the third assesses whether user-user networks capture behavioral alignment across ecosystems. For each task, we compare the user-user network construction methods in Table 1 against our own while keeping the downstream prediction model fixed. This setup thus isolates the input graph’s role on performance.

To assess robustness under sparse data conditions, we simulate increasing data availability by incrementally sampling $n\%$ of each user’s posts in 5% steps, retaining previously included posts for consistency. As part of our evaluation framework, we define the point of data efficiency of a method as the point at which it reaches 95% of its maximum AUC. This threshold serves as a standard for all methods and offers a principled balance between performance and efficiency. We validate the appropriateness of this choice by confirming that performance increases smoothly with additional data, and that the 95% point is not an outlier but reflects consistent trends (see Figure 4).

Method	Macro-F ₁	AUC
Co-Repost	0.80 ± 0.01	0.73 ± 0.01
Co-URL	0.49 ± 0.00	0.52 ± 0.01
Fast Repost	0.59 ± 0.01	0.60 ± 0.01
Hashtag Sequence	0.54 ± 0.03	0.48 ± 0.01
Text Similarity	0.62 ± 0.03	0.52 ± 0.02
k-NN Embedding Graph	0.65 ± 0.01	0.57 ± 0.00
Fused Graph	0.81 ± 0.01	0.76 ± 0.01
CANE	0.82 ± 0.01	0.94 ± 0.00
t-CANE	0.90 ± 0.01	0.94 ± 0.02

Table 3: Performance on the Iran IO dataset.

Information Operations

Information operations (IOs) represent a critical and high-stakes challenge for online platform governance and political communication research. Detecting such operations often requires robust inference of latent user alignment in the absence of explicit coordination signals, making them a strong test case for evaluating the effectiveness of user-user network construction methods. We evaluate our framework on the task of IO detection using the dataset from Seckin et al. (2024), which includes labeled state-backed campaigns identified by Twitter and released to the research community to enable research on IO detection. We focus on two operations: one attributed to Iran and one to China, each comprising users labeled as either IO drivers or organic accounts. See Appendix D for details.

Following the methodology of Luceri et al. (2024a), we embed user similarity networks (Table 1) where edges reflect inferred alignment between users based on shared content or behavioral patterns using node2vec (Grover and Leskovec 2016) and train a random forest classifier to distinguish drivers from organic users.

Overall Results. Shown in Tables 2 and 3, our method achieves the highest performance across both campaigns, with t-CANE reaching 0.85 Macro-F₁ and 0.92 AUC on the China IO dataset, and 0.90 Macro-F₁ and 0.94 AUC on the Iran IO dataset. The temporal variant (t-CANE) consistently outperforms static methods, highlighting the importance of modeling evolving behavior.

Sparse Data Setting. To evaluate data efficiency, we incrementally sample user content in 5% intervals and report how much data each method requires to reach 95% of its peak AUC. Our method achieves near-optimal performance with just 5–10% of the training data, significantly outperforming baselines. Although semantic similarity-based methods (e.g., k-NN Embedding Graph) also show data efficiency, they plateau at much lower overall performance. Full results are available in Appendix D (Table 15).

Ideological Mapping on X and TikTok

We apply our framework to an ideological classification task using 2024 U.S. election data from X and TikTok, sampled via a shared keyword query (Pinto et al. 2024; Balasubramanian et al. 2024). The task involves predicting each user’s political leaning as either Liberal or Conservative, a common binary framing in prior work on ideology pre-

diction (Jiang, Ren, and Ferrara 2023). X serves as a standard benchmark for this task; TikTok presents a more challenging and increasingly common environment, where traditional graph construction is constrained by sparse metadata and platform limitations, making it a useful testbed for evaluating structure-based approaches.

Many ideological stance prediction models rely on the structure of user-user graphs, often constructed from follower links or interaction data that reflect homophily and alignment (Xiao et al. 2020; Jiang, Ren, and Ferrara 2023). To isolate the role of structure, we exclude all content features and apply a fixed downstream model across all graph variants. This setup allows us to evaluate how well each network captures latent ideological patterns, independent of content or platform-specific behaviors.

As in previous sections, we compare against the baseline network construction methods listed in Table 1. On TikTok, where available signals are limited, we evaluate Hashtag Sequence, Text Similarity (from video descriptions), k-NN embedding Graph, and a partial fused graph.

We evaluate classification performance using graph-based models. While multiple architectures were tested, final results are reported using node2vec embeddings combined with a feedforward neural network classifier. This setup emphasizes structural information, aligning with our goal of isolating graph quality without relying on content features; performance was comparable across alternatives. See Appendix F for details.

US-2024 X Dataset. We use the large-scale dataset of political discourse on X during the 2024 U.S. election introduced by Balasubramanian et al. (2024). From this, we sample 24,000 users who posted at least five times. This threshold filters out the bottom 5% of users, who are often associated with low-quality or noise-prone content producers, and aligns with standard practices for reducing volatility in downstream modeling (Luceri et al. 2024b). A manually labeled subset of 2,100 users was annotated as liberal, conservative, or other/NA. Labeling was conducted by trained annotators with political science and media studies expertise, following standardized coding guidelines 19. Each user was independently labeled by two raters; disagreements were excluded from the final dataset. See Appendix E for full label distributions and inter-annotator agreement statistics.

US-2024 TikTok Dataset. We use the corresponding 2024 U.S. election dataset on TikTok from Pinto et al. (2024), sampling 52,859 active users (≥ 5 posts). A subset of 1,758 users was manually labeled using both video descriptions and Whisper-transcribed audio², following the same schema as for the X dataset. Labeling metrics for both platforms appear in Appendix E (Table 16).

US-2024 X Results. As shown in Table 4, our method achieves the highest Macro-F₁ (0.83) and AUC (0.76) scores across all network construction approaches, demonstrating strong downstream performance. Traditional networks underperform largely due to platform-specific signal sparsity. For instance, only 0.003% of X posts contain URLs, and just

²<https://github.com/openai/whisper>

0.65% of users share them—rendering URL-based graphs nearly empty. Repost and hashtag graphs perform similarly, covering only 5.2% and 5.4% of users, respectively. Even the Fused Graph struggles with limited behavioral coverage.

To isolate performance from coverage effects, we evaluate each method using only the users included in its graph. As shown in Table 20, our method remains competitive under these constraints, suggesting its advantage stems from structural quality, not just breadth. Class distributions and coverage-adjusted subsets appear in Table 18.

US-2024 TikTok Results. As shown in Table 5, our method again achieves the highest performance (Macro-F₁: 0.83, AUC: 0.83). Unlike the X-based evaluation, no user filtering is needed here, as the graph covers 100% of labeled users, likely due to the dataset design by Pinto et al. (2024), where hashtags appear in 92.6% of video descriptions.

Sparse Data Setting: US-2024 X and TikTok Results. We observe similar trends in the US-2024 X and TikTok datasets as in the IO experiments: our networks require significantly less data to reach near-optimal performance. As shown in Table 21, our methods reach 95% of their peak AUC with minimal user content (5-10%), consistently outperforming alternatives in both efficiency and final performance. Likely due to their operation on individual posts, semantic similarity-based methods also exhibit relatively low data requirements. However, consistent with our IO setting findings, these methods plateau at lower performance levels compared to our method.

Method	Similarity Network	Macro-F ₁	AUC
Baseline	Co-Repost	0.43 ± 0.30	0.56 ± 0.22
	Co-URL	0.51 ± 0.26	0.53 ± 0.05
	Fast Repost	0.48 ± 0.04	0.51 ± 0.03
	Hashtag Sequence	0.65 ± 0.22	0.63 ± 0.14
	Text Similarity	0.39 ± 0.12	0.32 ± 0.09
	k-NN Embedding Graph	0.41 ± 0.03	0.34 ± 0.11
	Fused Graph	0.69 ± 0.15	0.70 ± 0.05
Ours	CANE	0.81 ± 0.01	0.73 ± 0.02
	t-CANE	0.83 ± 0.01	0.76 ± 0.02

Table 4: Ideological classification on X.

Method	Similarity Network	Macro-F ₁	AUC
Baseline	Hashtag Sequence	0.71 ± 0.03	0.62 ± 0.01
	Text Similarity	0.38 ± 0.14	0.42 ± 0.07
	k-NN Embedding Graph	0.45 ± 0.10	0.45 ± 0.08
	(Partial) Fused Graph	0.71 ± 0.03	0.64 ± 0.02
Ours	CANE	0.83 ± 0.02	0.83 ± 0.01
	t-CANE	0.82 ± 0.02	0.83 ± 0.03

Table 5: Ideological classification on TikTok.

Cross-Platform Engagement Prediction

We introduce a novel benchmark to test whether user-user networks can predict future topic engagement across platforms. This evaluates whether structural proximity in the graph corresponds to behavioral alignment in fragmented media ecosystems. Formulated as a user-topic prediction task, the model receives a graph at time t and predicts if user u will engage with topic k in the future using only structural

features. No content-based features are used during prediction, isolating the representational value of the graph.

We apply this to a dataset of U.S. election-related posts from X and Truth Social (May–Nov 2024), clustered into 321 cross-platform narrative themes using multilingual MP-Net embeddings and DP-means clustering. Each theme is labeled using top TF-IDF terms. Full preprocessing details, cluster labeling procedures, and representative narrative examples are provided in Appendix G. As in prior tasks, we train GCNs (Zhang et al. 2019) on each graph, using binary user-topic features to predict future engagement. We report performance at multiple time windows using Macro-F₁ and AUC. A randomly rewired graph is included to verify that performance reflects real structure, not model capacity.

Method	Similarity Network	Macro-F ₁	AUC
Baseline	Random GCN	0.00 ± 0.00	0.56 ± 0.03
	Co-URL	0.01 ± 0.00	0.43 ± 0.04
	Hashtag Sequence	0.11 ± 0.04	0.48 ± 0.08
	Text Similarity	0.02 ± 0.00	0.58 ± 0.05
	k-NN Embedding Graph	0.02 ± 0.01	0.61 ± 0.06
	(Partial) Fused Graph	0.05 ± 0.01	0.64 ± 0.06
Ours	CANE	0.30 ± 0.02	0.89 ± 0.02
	t-CANE	0.35 ± 0.06	0.94 ± 0.02

Table 6: Performance of similarity networks for cross-platform narrative engagement prediction ($t=7$).

Results. As shown in Table 6 and corroborated across all timesteps in Appendix Table 25, our proposed approach, t-CANE, substantially outperforms all baseline methods. The strongest baseline achieves a Macro-F₁ score of 0.11 (Hashtag Sequence), while t-CANE reaches 0.35, more than tripling predictive accuracy. Even using our non-temporal model CANE, our model reached a Macro-F₁ of 0.30. Similarly, the highest baseline AUC is 0.64 (Fused Graph), compared to 0.94 for t-CANE.

Connective Narratives: X and Truth Social

To analyze how political narratives migrated between Truth Social and X during the 2024 U.S. Presidential election, we use the CANE framework to construct a unified discourse network that connects users through shared narrative engagement. This representation allows us to investigate not only whether narratives moved between platforms, but also how, when, and through whom they traveled. Our analysis proceeds in three stages: first, we characterize overall patterns of narrative migration; second, we model the structure of narrative flow to assess whether it follows concentrated and repeatable routes; and third, we examine whether a small set of users embedded across discourse communities consistently enables cross-platform diffusion.

We find that narrative diffusion between X and Truth Social followed structured and repeatable patterns. It was shaped by asymmetries in platform roles and by repeated involvement from users positioned at the boundaries of discourse communities. By viewing the ecosystem through a discourse network, we are able to uncover concrete pathways through which narratives consistently migrated, pathways that remain invisible to interaction- or semantic

similarity-based. In the sections that follow, we use this structural lens to characterize patterns of narrative movement, identify the users who drive it, and assess the broader architecture of the cross-platform discourse that occurred.

Cross-Platform Narrative Migration

To understand how political discourse flows across platforms, we begin by modeling narratives, seeking to trace how coherent narratives emerge, evolve, and migrate between Truth Social and X. Following prior work on narrative tracking in multilingual information ecosystems (Hanley, Kumar, and Durumeric 2024), we define a narrative as a collection of posts that focus on the same issue or event. This definition treats narratives as coherent units of discourse that evolve over time and platform, rather than as isolated posts or terms. For example, in the context of the 2024 U.S. election, a narrative might involve claims that mail-in ballots are being used to rig the outcome: potentially unfolding over multiple posts that cite anecdotal evidence, inaccurate information, and emotionally charged appeals.

To identify narratives, we embed posts using the multilingual MPNet model fine-tuned for semantic similarity and apply DP-means clustering with a cosine distance cutoff of 0.30 (i.e., minimum similarity of 0.70 within clusters). Prior work in narrative tracking (Hanley, Kumar, and Durumeric 2024) typically uses thresholds between 0.60 and 0.80. Given the greater linguistic variability across Truth Social and X, we adopt a 0.65 threshold to balance semantic coherence with topical specificity, minimizing the risk of merging distinct narratives. We validate this choice via human evaluation of 50 post pairs at various thresholds (Table 11) and find this threshold to be comparable to benchmarks in similar work (Hanley and Durumeric 2024).

To establish a baseline for narrative movement, we first identify cases of *simple migration*: narratives that cross platform boundaries, regardless of sustained dynamics. A narrative is considered to originate on a platform if it appears there at least 24 hours before surfacing on the other, filtering out near-simultaneous mentions triggered by shared external events. Migration is confirmed when the narrative surpasses a minimum engagement threshold on the receiving platform: defined as 10 posts, the 35th percentile across all narrative-platform pairs. This ensures focus on substantive diffusion events and follows best practices in prior coordination studies (Magelinski, Ng, and Carley 2022; Nizzoli et al. 2021). We validate this threshold via sensitivity analysis (Table 11).

Next, to assess structured and sustained diffusion, we identify cases of *significant migration*: narratives showing statistically directional flow. We compute Transfer Entropy (TE) to assess directional information flow. TE is a non-linear, asymmetric measure that detects whether past activity on one platform reliably reduces uncertainty about future activity on the other (Schreiber 2000), allowing us to move beyond simple symmetric association to capture predictive, time-structured diffusion dynamics. To validate directionality, we perform a permutation test by randomly shuffling the source platform’s time series 1,000 times to generate a null TE distribution. Narratives are considered significant if their observed TE exceeds the 95th percentile of this distribution

($p < 0.05$). This procedure isolates narratives with robust, predictive diffusion dynamics.

Results. From this process, we identify 1,552 narratives exhibiting at least *simple migration* and 238 narratives exhibiting directional, *significant migration* (that is, 15.3% of narratives exhibiting at least simple migration also exhibit significant migration). We enumerate example representative narratives along with their TE in Table 7.

At the level of *simple migration*, we find that Truth Social plays a disproportionately large role in initiating cross-platform diffusion. Despite accounting for just 1.7% of total post volume across migrating narratives, Truth Social is the origin point for 18.9% of them, making its narratives more than eleven times more likely to initiate cross platform spread than expected by chance. These findings suggest that Truth Social, although small in scale, may serve as an important incubator for narratives that later gain broader traction.

Qualitatively, narratives originating on Truth Social tend to exhibit more overt themes of fear or conspiratorial framing, often revolving around election integrity, government control, or existential threat. This contrasts with X originating narratives, which more commonly emphasize partisan conflict or elite criticism without the same fear-centric framing. To test this distinction systematically, we apply a model fine-tuned on human-annotated fear speech data (Saha et al. 2023) to all posts within each migrating narrative (see Appendix H for details). Narratives seeded from Truth Social exhibit significantly higher rates of fear-laden language than those originating on X ($p < 0.01$), with a log-odds ratio of +0.22 and a 22.5% relative increase.

Turning to *significant migration*, our stricter criteria yield 238 narratives exhibiting statistically significant, directional cross-platform flow. These narratives form the empirical foundation for the rest of our analysis. Within this set, we again find that Truth Social plays an outsized role in shaping subsequent activity on X. Despite contributing only 1.1% of total post volume within these clusters, Truth Social leads 15.6% of significant migration cases, with the remainder originating from X, representing a $14.2\times$ overrepresentation relative to its share of total post volume (1.1%), calculated as the ratio between observed and expected initiation rates. This suggests that Truth Social frequently functions as an upstream source within the temporal structure of cross-platform discourse. Overall, these results indicate that the observed cross-platform migration is common, structurally directional, and disproportionately seeded by Truth Social.

Discourse as Structures: Bridge Zones

To examine the structural dynamics of the narrative flow between X and Truth Social, we adopt a discourse-centered approach. Rather than treating each platform as an isolated system, we use CANE to model a unified information ecosystem, where users are linked through shared narrative participation. To uncover its internal organization, we apply the Louvain algorithm (Traag, Waltman, and Van Eck 2019) to detect discourse communities: groups of users who consistently engage with similar narratives over time. We then assess the cross-platform composition of each community using Shannon entropy, a standard measure of distributional

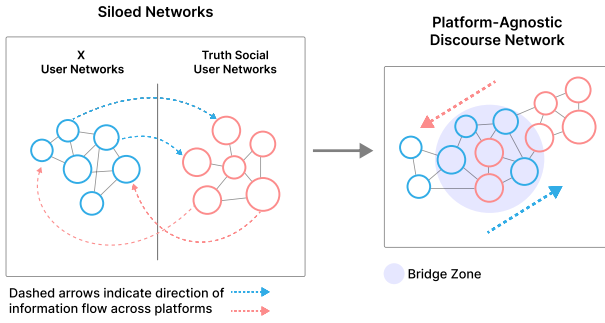


Figure 2: A conceptual illustration of cross-platform narrative diffusion. While most communities are siloed within platforms, some users form *bridge zones*: structural overlaps where narrative transfer is more likely.

diversity. Here, entropy reflects how evenly users are distributed across platforms: values near 0 indicate platform homogeneity, while values near 1 suggest a balanced mix.

Bridge Communities. We identify a single high-entropy community (entropy = 0.72) comprising just 0.33% of users and 2.14% of posts (visualized in Figure 5). Notably, this community’s users do not stand out in terms of volume, but through their distinctive position within the discourse network inferred from content alignment. They span discursive boundaries between platforms, occupying a region through which narratives are especially likely to migrate.

We refer to this region as a *bridge zone*: a conceptual area in the discourse network where users from different platforms are densely connected through shared narrative engagement (illustrated in Figure 2) (Mendelsohn et al. 2023). To validate that this zone reflects meaningful structural alignment rather than artifacts of duplicated accounts across platforms, we apply a two-stage de-duplication process across the full user set. First, we compute character-level string similarity using the Ratcliff/Obershelp algorithm (Hasugian and Padang 2023), removing 1,644 high-similarity username pairs (0.19% of users) based on a liberal threshold of 0.7. Second, to assess potential residual duplication, we manually audit 200 randomly sampled users from the *bridge zone*. For each, we searched the other platform’s user interface for accounts with matching or near-matching usernames and names where applicable. In cases of near matches, we cross-referenced bios and profile images to assess the likelihood of duplication. Only six users (3%) appeared to maintain accounts on both platforms, and none of them were retained in the final data set. This suggests that account duplication is unlikely to explain the distinct structural role of the observed *bridge zone*.

Importantly, these structural dynamics only become visible when discourse is modeled as a unified, cross-platform network. While some baseline graph construction methods surface large or mildly cross-platform communities, none isolate a region with such substantial platform overlap and strong association with narrative migration. See Appendix I for empirical comparisons.

Bridge Users. Building on the concept of bridge zones (regions of structural overlap in the discourse network where

narratives migrate across platforms), we define *bridge users* as members of these high-entropy, cross-platform communities. Rather than standing out by volume or engagement, they are structurally embedded at key intersections between Truth Social and X. Notably, their activity metrics are near the global median (Appendix I, Table 26), suggesting their influence stems from position, not popularity.

These 2,864 *bridge users* comprise just 0.33% of all users and are responsible for only 2.14% of posts, yet they play a disproportionately large role in cross-platform diffusion. They serve as the first cross-platform carriers for 68% of simple migrating narratives and 69% of significant migrating narratives: an overrepresentation of more than 200× their population share. In addition, they are not only conduits, but also initiators, seeding 26% of all narratives, including 22% of simple migrating narratives and 19% of significant migrating narratives.

Interestingly, even within this small subgroup, influence is highly concentrated. Just 122 bridge users account for all earliest introductions of Truth Social-originating narratives into X, with four users alone responsible for roughly a quarter of these events. Although we claim no evidence of explicit coordination, two of these users independently pinned an identical inflammatory image (“F*** this Government”) nearly a year apart, pointing to convergent symbolic signaling despite the absence of direct ties. These findings suggest that bridge users occupy structurally important points in the discourse network: locations where cross-platform narrative movement is significantly more likely to occur.

Bridge Users and Narrative Diffusion. As shown above, *bridge users* played a disproportionately large role in the narrative flow. However, visibility alone does not explain their significance. Structurally embedded actors who engage early can play a decisive role in shaping how information spreads, particularly in fragmented systems where direct ties are sparse (Granovetter 1983; Gould and Fernandez 1989). Engagement at the early stages of a narrative’s lifecycle reflects not just activity but also a form of narrative buy-in which may determine which ideas take root and diffuse. We therefore test whether narratives in which bridge users appear during the earliest stages of formation (defined as within the first 5% of participants; we validate the robustness of this in Appendix I) exhibit higher downstream engagement than those initiated by comparable users.

To ensure a fair comparison, each bridge user is matched to a non-bridge user from a single-platform community using k-nearest neighbors based on z-normalized behavioral metrics: post volume, likes, replies, and reposts. We then evaluate engagement across likes, replies, and reposts. To quantify the strength of the observed differences, we report the rank-biserial correlation, which measures the probability that a randomly selected value from one group will exceed one from the other, providing an interpretable effect size for non-parametric comparisons. Narratives involving early bridge user participation receive significantly more interaction across all metrics (Mann–Whitney U , $p < 0.01$), with large effect sizes for likes (0.60) and reposts (0.46), and a moderate effect for replies (0.34). These results are

Narrative Theme	TE (TS → X)
Criticism of Stormy Daniels’ testimony against Trump.	0.95
Discussions around Bud Light’s sales slump after boycott fallout.	0.58
Claims about FBI suppression regarding Seth Rich.	0.50
Allegations of FBI suppression of the Hunter Biden laptop story.	0.49
Posts suggesting Biden mental decline.	0.49
Valorization of El Salvador President, Nayib Bukele.	0.36
Accusations of Kamala Harris being racial “chameleon.”	0.33
Calls to “cancel” people, events, or shows.	0.31

Table 7: Representative narratives exhibiting significant directional influence from Truth Social to X, as identified via Transfer Entropy. Themes were assigned by a human expert based on posts and top TF-IDF terms within each narrative.

robust across a range of early engagement thresholds (see Appendix I), and hold even when restricting to low-virality narratives, indicating that bridge users are not simply reacting to already popular content.

These patterns may bolster the idea that these *bridge users* partially function as early validators structurally positioned at the seams of the discourse network. Their engagement aligns with classic network theory, where actors such as brokers (Gould and Fernandez 1989) and weak ties (Granovetter 1983) help facilitate early information flow across community boundaries. Rather than reflecting visibility or volume, their engagement patterns appear to coincide with structural embedding in positions that are more likely to be associated with cross-platform diffusion.

Discussion and Conclusion

Political narratives during the 2024 U.S. Presidential election did not diffuse haphazardly across X and Truth Social. Rather, they followed structured and asymmetric paths, enabled by a small subset of users embedded at the intersection of fragmented discourse communities. These *bridge users*, who comprise just 0.33% of users and account for only 2.14% of posts, nonetheless served as initial cross-platform conduits for approximately 70% of narratives that migrated between platforms. Their centrality to cross-platform narrative flow stemmed not from visibility or volume, but from their structural embedding within a unified discourse network that spanned both ecosystems. Crucially, this structure only becomes visible when discourse is modeled directly. While some baseline graph methods surface modest cross-platform communities, none recover the tightly embedded bridge zone that emerges from our discourse network. This not only allows us to recover patterns of cross-platform diffusion that remain hidden in interaction-based graphs, but also makes our method scalable and robust in environments where behavioral data is limited or siloed.

These results expand our understanding of influence in fragmented media. Rather than treating platforms as silos, our approach surfaces the connective tissue through which narratives moved across X and Truth Social. This highlights the unique value of discourse-centered modeling: it not only improves prediction but also reveals latent structures of narrative diffusion that remain hidden in traditional, behavior- or metadata-based graphs. Moreover, this structural view aligns with classic theories in sociology and organizational science. Our notion of “bridge users” paral-

els boundary spanners in organizational networks (Cross and Prusak 2002) and brokers in Gould and Fernandez’s typology (Gould and Fernandez 1989). Most directly, it aligns with Granovetter’s theory of weak ties (Granovetter 1983), which highlights how those positioned across structural holes, rather than central figures, often drive information flow. In our case, bridge users span platform divides, occupying boundary positions in the discourse network that are consistently associated with cross-platform diffusion.

As such, our contributions are both empirical and methodological. Empirically, we show that narrative diffusion across X and Truth Social exhibits highly structured patterns, with a small set of users consistently positioned at key points of cross-platform narrative movement. Methodologically, we offer a content-driven, platform-agnostic framework for constructing discourse networks: scalable, unsupervised, and robust to missing behavioral metadata. More broadly, we argue that understanding online discourse requires shifting focus from individual platforms to the connective structures that link them. Ideas flow across ecosystems, and capturing this movement demands tools attuned to the networked nature of modern communication: not just in terms of behavior, but in the deeper discursive alignments that sustain influence. Although prior work has shown that cross-platform contagion is often driven by high-attention or elite accounts (Wilson and Starbird 2020; Ribeiro et al. 2021), few have examined the structural patterns formed through shared engagement with ideas rather than direct interaction. By representing users as distributions over shared narratives, our framework identifies these latent pathways and the actors who traverse them, linking fragmented publics through structurally embedded roles akin to boundary spanners and brokers. This discourse-centered approach offers a scalable and theory-informed lens to model influence in a fractured information environment.

Limitations. While our study reveals important patterns in cross-platform narrative migration, it is not without limitations. First, our analysis is limited to two platforms, X and Truth Social, and does not capture the full complexity of the broader information ecosystem. Second, while our semantic clustering approach enables scalable narrative tracking, it may miss highly subtle variations or instances of near-synonymous discursive drift. Finally, although we identify structural convergence and symbolic signaling among key *bridge users*, we cannot definitively infer intentional coordination from observational data alone.

Future Work. Future research can extend this framework in several directions. First, incorporating additional platforms such as Telegram, TikTok, Reddit, and Gab would allow for a more comprehensive analysis of cross-ecosystem narrative flow. Second, while this study focuses on user-level discourse, the method is readily adaptable to other entities such as media outlets or organizations, making it possible to trace narrative migration across actors, genres, or communities. Finally, integrating dynamic temporal modeling could uncover how bridge zones form, dissolve, and shift in response to evolving narrative and platform-specific conditions.

Ethical Statement. This study uses publicly available

data from X and Truth Social, collected in accordance with platform terms. Because our analysis examines patterns of influence and discourse across platforms, all usernames are anonymized in the released dataset, and no identifying information is shared. Our focus is on aggregate narrative dynamics and structural features of discourse networks, rather than individual-level behavior. We do not make claims about user intent or coordination. The dataset may contain offensive content given the nature of political discourse, but our analysis focuses on aggregate patterns rather than individual behavior. Data will be released with clear documentation to support open access and reuse for public-interest research on political discourse, narrative diffusion, and cross-platform communication dynamics.

Conclusion. Unlike traditional interaction-based or semantic similarity-based graphs that rely on surface level overlap or direct behavioral ties, CANE models users as distributions over shared narratives. This captures not just what they say or how they interact, but what ideas they inhabit. It re-frames network construction from tracing explicit signals to uncovering the conceptual structure that underlies discourse across fragmented ecosystems.

By modeling discourse in this way, we find that cross-platform diffusion between X and Truth Social during the 2024 U.S. election was not only structured, but also predictable: revealing pathways that behavioral and similarity-based methods fail to detect. These patterns point to deeper structures that shape how narratives moved between these fragmented communities. As political discourse continues to fracture, identifying such structures is critical: not just for understanding how information spreads, but for actively anticipating its movement across fragmented publics.

References

- Aizawa, A. 2003. An information-theoretic perspective of tf-idf measures. *Information Processing & Management*, 39(1): 45–65.
- Alieva, I.; Ng, L. H. X.; and Carley, K. M. 2022. Investigating the spread of Russian disinformation about biolabs in Ukraine on Twitter using social network analysis. In *2022 IEEE international conference on big data (big data)*, 1770–1775. IEEE.
- Balasubramanian, A.; Zou, V.; Narayana, H.; You, C.; Luceri, L.; and Ferrara, E. 2024. A public dataset tracking social media discourse about the 2024 us presidential election on twitter/x. *arXiv preprint arXiv:2411.00376*.
- Bollen, J.; Gonçalves, B.; Ruan, G.; and Mao, H. 2011. Happiness is assortative in online social networks. *Artificial life*, 17(3): 237–251.
- Burghardt, K.; Rao, A.; Chochlakis, G.; Sabyasachee, B.; Guo, S.; He, Z.; Rojecki, A.; Narayanan, S.; and Lerman, K. 2024. Socio-linguistic characteristics of coordinated inauthentic accounts. In *Proceedings of icwsm*, volume 18, 164–176.
- Cinelli, M.; De Francisci Morales, G.; Galeazzi, A.; Quattrociocchi, W.; and Starnini, M. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9): e2023301118.
- Cinus, F.; Minici, M.; Luceri, L.; and Ferrara, E. 2025. Exposing cross-platform coordinated inauthentic activity in the run-up to the 2024 us election. In *Proceedings of the ACM on Web Conference 2025*, 541–559.
- Cross, R.; and Prusak, L. 2002. The people who make organizations go—or stop. *Networks in the Knowledge Economy*, 80(6): 248–260.
- Gerard, P.; Botzer, N.; and Weninger, T. 2023. Truth social dataset. In *Proceedings of icwsm*, volume 17, 1034–1040.
- Gerard, P.; Theisen, W.; Weninger, T.; and Lerman, K. 2024. Fear and Loathing on the Frontline: Decoding the Language of Othering by Russia-Ukraine War Bloggers. *arXiv:2409.13064*.
- Gould, R. V.; and Fernandez, R. M. 1989. Structures of mediation: A formal approach to brokerage in transaction networks. *Sociological methodology*, 89–126.
- Granovetter, M. 1983. The strength of weak ties: A network theory revisited. *Sociological theory*, 201–233.
- Grover, A.; and Leskovec, J. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 855–864.
- Hanley, H. W.; and Durumeric, Z. 2024. Partial mobilization: Tracking multilingual information flows amongst russian media outlets and telegram. In *Proceedings of icwsm*, volume 18, 528–541.
- Hanley, H. W.; Kumar, D.; and Durumeric, Z. 2023. Happenstance: utilizing semantic search to track Russian state media narratives about the Russo-Ukrainian war on Reddit. In *Proceedings of icwsm*, volume 17, 327–338.
- Hanley, H. W.; Kumar, D.; and Durumeric, Z. 2024. Specious sites: Tracking the spread and sway of spurious news stories at scale. In *2024 IEEE Symposium on Security and Privacy (SP)*, 1609–1627. IEEE.
- Hasugian, A. H.; and Padang, S. 2023. Image Similarity Detection Application Using The Ratcliff/Obershelp Method And Levenshtein Distance. *Instal: Jurnal Komputer*, 15(01): 80–88.
- Iofciu, T.; Fankhauser, P.; Abel, F.; and Bischoff, K. 2011. Identifying users across social tagging systems. In *Proceedings of icwsm*, volume 5, 522–525.
- Jiang, J.; Ren, X.; and Ferrara, E. 2023. Retweet-bert: political leaning detection using language features and information diffusion on social networks. In *Proceedings of icwsm*, volume 17, 459–469.
- Laub, P. J.; Taimre, T.; and Pollett, P. K. 2015. Hawkes processes. *arXiv preprint arXiv:1507.02822*.
- Luceri, L.; Cresci, S.; and Giordano, S. 2021. Social media against society. *The Internet and the 2020 Campaign*, 1.
- Luceri, L.; Pantè, V.; Burghardt, K.; and Ferrara, E. 2024a. Unmasking the web of deceit: Uncovering coordinated activity to expose information operations on twitter. In *Proceedings of the ACM on Web Conference 2024*, 2530–2541.
- Luceri, L.; Ye, J.; Jiang, J.; and Ferrara, E. 2024b. The Susceptibility Paradox in Online Social Influence. *arXiv preprint arXiv:2406.11553*.

- Magelinski, T.; Ng, L.; and Carley, K. 2022. A synchronized action framework for detection of coordination on social media. *Journal of Online Trust and Safety*, 1(2).
- Malkov, Y. A.; and Yashunin, D. A. 2018. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE transactions on pattern analysis and machine intelligence*, 42(4): 824–836.
- Mendelsohn, J.; Ghosh, S.; Jurgens, D.; and Budak, C. 2023. Bridging nations: quantifying the role of multilinguals in communication on social media. In *Proceedings of icwsm*, volume 17, 626–637.
- Minici, M.; Cinus, F.; Luceri, L.; and Ferrara, E. 2024. Uncovering coordinated cross-platform information operations: Threatening the integrity of the 2024 US presidential election. *First Monday*.
- Ng, L. H. X.; and Carley, K. M. 2022. A Combined Synchronization Index for Grassroots Activism on Social Media. *arXiv preprint arXiv:2212.13221*.
- Ng, L. H. X.; Cruickshank, I. J.; and Carley, K. M. 2022. Cross-platform information spread during the January 6th capitol riots. *Social Network Analysis and Mining*, 12(1): 133.
- Ng, L. H. X.; Cruickshank, I. J.; and Carley, K. M. 2023. Coordinating Narratives Framework for cross-platform analysis in the 2021 US Capitol riots. *Computational and Mathematical Organization Theory*, 29(3): 470–486.
- Nizzoli, L.; Tardelli, S.; Avvenuti, M.; Cresci, S.; and Tesconi, M. 2021. Coordinated behavior on social media in 2019 UK general election. In *Proceedings of icwsm*, volume 15, 443–454.
- Pacheco, D.; Hui, P.-M.; Torres-Lugo, C.; Truong, B. T.; Flammini, A.; and Menczer, F. 2021. Uncovering coordinated networks on social media: methods and case studies. In *Proceedings of ICWSM*, volume 15, 455–466.
- Pinto, G.; Bickham, C.; Salkar, T.; Luceri, L.; and Ferrara, E. 2024. Tracking the 2024 US Presidential Election Chatter on TikTok: A Public Multimodal Dataset. *arXiv:2407.01471*.
- Ribeiro, M. H.; Blackburn, J.; Bradlyn, B.; De Cristofaro, E.; Stringhini, G.; Long, S.; Greenberg, S.; and Zannettou, S. 2021. The evolution of the manosphere across the web. In *Proceedings of icwsm*, volume 15, 196–207.
- Russo, G.; Verginer, L.; Ribeiro, M. H.; and Casiraghi, G. 2023. Spillover of antisocial behavior from fringe platforms: The unintended consequences of community banning. In *Proceedings of icwsm*, volume 17, 742–753.
- Saha, P.; Garimella, K.; Kalyan, N. K.; Pandey, S. K.; Meher, P. M.; Mathew, B.; and Mukherjee, A. 2023. On the rise of fear speech in online social media. *Proceedings of the National Academy of Sciences*, 120(11): e2212270120.
- Schreiber, T. 2000. Measuring information transfer. *Physical review letters*, 85(2): 461.
- Seckin, O. C.; Pote, M.; Nwala, A.; Yin, L.; Luceri, L.; Flammini, A.; and Menczer, F. 2024. Labeled Datasets for Research on Information Operations. *arXiv preprint arXiv:2411.10609*.
- Shah, K.; Gerard, P.; Luceri, L.; and Ferrara, E. 2024. Unfiltered Conversations: A Dataset of 2024 US Presidential Election Discourse on Truth Social. *arXiv preprint arXiv:2411.01330*.
- Sipka, A.; Hannak, A.; and Urman, A. 2022. Comparing the Language of QAnon-related content on Parler, Gab, and Twitter. In *Proceedings of the 14th ACM Web Science Conference 2022*, 411–421.
- Song, K.; Tan, X.; Qin, T.; Lu, J.; and Liu, T.-Y. 2020. MpNet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*, 33: 16857–16867.
- Tardelli, S.; Nizzoli, L.; Avvenuti, M.; Cresci, S.; and Tesconi, M. 2024a. Multifaceted online coordinated behavior in the 2020 US presidential election. *EPJ Data Science*, 13(1): 33.
- Tardelli, S.; Nizzoli, L.; Tesconi, M.; Conti, M.; Nakov, P.; Da San Martino, G.; and Cresci, S. 2024b. Temporal dynamics of coordinated online behavior: Stability, archetypes, and influence. *Proceedings of the National Academy of Sciences*, 121(20): e2307038121.
- Traag, V. A.; Waltman, L.; and Van Eck, N. J. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific reports*, 9(1): 1–12.
- Tromble, R. 2021. Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media+ Society*, 7(1): 2056305121988929.
- Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; and Bengio, Y. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903*.
- Vishnuprasad, P. S.; Nogara, G.; Cardoso, F.; Cresci, S.; Giordano, S.; and Luceri, L. 2024. Tracking fringe and coordinated activity on Twitter leading up to the US Capitol attack. In *Proceedings of ICWSM*, volume 18, 1557–1570.
- Wang, L.; Yang, N.; Huang, X.; Yang, L.; Majumder, R.; and Wei, F. 2024. Multilingual e5 text embeddings: A technical report. *arXiv preprint arXiv:2402.05672*.
- Wilson, T.; and Starbird, K. 2020. Cross-platform disinformation campaigns: lessons learned and next steps. *Harvard Kennedy School Misinformation Review*, 1(1).
- Xiao, Z.; Song, W.; Xu, H.; Ren, Z.; and Sun, Y. 2020. TIMME: Twitter ideology-detection via multi-task multi-relational embedding. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, 2258–2268.
- Zhang, S.; Tong, H.; Xu, J.; and Maciejewski, R. 2019. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1): 1–23.
- Zhang, Y.; Lukito, J.; Suk, J.; and McGrady, R. 2025. Trump, Twitter, and Truth Social: how Trump used both mainstream and alt-tech social media to drive news media attention. *Journal of Information Technology & Politics*, 22(2): 229–242.
- Zulli, D.; Coe, K.; and Isaacs, Z. 2023. News framing in the aftermath of the January 6 attacks on the US Capitol: An analysis of labels, definitional uncertainty, and contextualization. *American behavioral scientist*, 67(6): 702–720.

Paper Checklist to be included in your paper

1. For most authors...

- (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? [Our study advances computational methods for modeling cross-platform discourse without infringing on user privacy or reinforcing harmful social dynamics. We anonymize all released data and focus on aggregate narrative structures, not individual behavior.](#)
- (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? [We state our contributions in narrative tracking, platform-agnostic user graph construction, and identification of bridge users.](#)
- (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? [Our method is designed to detect cross-platform narrative flow through latent discourse clusters. This aligns with the claims about discourse structure and user roles across platforms.](#)
- (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? [See Limitations: We acknowledge that platform differences and linguistic variability may introduce artifacts; also, we we apply clustering thresholds and human validation to mitigate potential artifacts.](#)
- (e) Did you describe the limitations of your work? [Yes, see the Limitations section.](#)
- (f) Did you discuss any potential negative societal impacts of your work? [We caution against misinterpreting structural diffusion patterns as intentional coordination. Additionally, we warn about overreliance on structural roles.](#)
- (g) Did you discuss any potential misuse of your work? [See Ethical Statement: We note that identifying bridge users could be misused to target or surveil individuals. We mitigate this risk through full anonymization.](#)
- (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? [We anonymized all data, did not release any user-identifiable information, and documented our methodology in the appendix and supplemental code.](#)
- (i) Have you read the ethics review guidelines and ensured that your paper conforms to them?

2. Additionally, if your study involves hypotheses testing...

- (a) Did you clearly state the assumptions underlying all theoretical results? [NA](#)
- (b) Have you provided justifications for all theoretical results? [NA](#)

- (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? [NA](#)
- (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? [See Bridge Users and Narrative Diffusion: We discuss alternative explanations such as visibility-based influence versus structural positioning.](#)
- (e) Did you address potential biases or limitations in your theoretical framework? [See Methodology: We acknowledge that structure may not capture all aspects of influence, and that clustering thresholds may miss subtle narrative drift.](#)
- (f) Have you related your theoretical results to the existing literature in social science? [See Related Work, Discussion: We connect our findings to theories of weak ties, boundary spanning, and brokerage.](#)
- (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? [We discuss implications for incorrect information tracking, platform governance, and cross-platform content moderation.](#)

3. Additionally, if you are including theoretical proofs...

- (a) Did you state the full set of assumptions of all theoretical results? [NA](#)
- (b) Did you include complete proofs of all theoretical results? [NA](#)

4. Additionally, if you ran machine learning experiments...

- (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [We provide anonymized code and datasets through an anonymized GitHub link.](#)
- (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [See Appendix.](#)
- (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [We report standard deviations across runs for all evaluations.](#)
- (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [See Appendix: We include compute details \(e.g., use of FAISS with GPU support\).](#)
- (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? [See Experiments: Our evaluation includes multiple downstream tasks across three real-world datasets.](#)
- (f) Do you discuss what is “the cost” of misclassification and fault (in)tolerance? [See Ethical Statement.](#)

5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...

- (a) If your work uses existing assets, did you cite the creators? [All datasets and models are appropriately cited.](#)
 - (b) Did you mention the license of the assets? [See Appendix.](#)
 - (c) Did you include any new assets in the supplemental material or as a URL? [An anonymized dataset and replication code.](#)
 - (d) Did you discuss whether and how consent was obtained from people whose data you’re using/curating? [All data were collected from publicly available sources in accordance with platform terms of service, and derived from existing datasets released in prior research. We use these datasets with appropriate citations.](#)
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [See Ethical Statement: We anonymized all usernames and note that offensive content may be present, but is analyzed only in aggregate.](#)
 - (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see ?)? [We describe plans for public access, clear documentation, and support for reuse. We include more information in the supplemental material/link.](#)
 - (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see ?)? [We include a datasheet in the supplemental material/link.](#)
6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity...**
- (a) Did you include the full text of instructions given to participants and screenshots? [NA](#)
 - (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? [NA](#)
 - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [NA](#)
 - (d) Did you discuss how data is stored, shared, and deidentified? [NA](#)

Appendix A: DP-Means Details

Why Use DP-Means? We adopt DP-Means over alternatives like DBSCAN or agglomerative clustering for its linear time complexity ($O(nKT)$), low memory overhead ($O(K)$), and conceptual clarity. In contrast, DBSCAN has worst-case complexity $O(n^2)$ and often struggles with high-dimensional, sparse embeddings; agglomerative clustering requires $O(n^2 \log n)$ time and is impractical at our scale. Moreover, DP-Means’ use of a single interpretable parameter contrasts with the more opaque controls of DBSCAN (e.g., minimum cluster size) and agglomerative methods (e.g., linkage criteria). Finally, DP-Means supports online updates, making it suitable for large-scale or streaming analyses. Full algorithmic details are provided in Appendix A.

DP-Means Objective. Given a set of embedded posts $X = \{x_1, \dots, x_n\}$ where $x_i \in R^d$, DP-Means minimizes the following objective:

$$\min_{z, \mu} \sum_{i=1}^n \min_{k=1}^K d(x_i, \mu_k), \lambda \quad (3)$$

where z is the cluster assignment, μ_k are cluster centers, and $d(\cdot, \cdot)$ is cosine distance. A new cluster is created when no existing cluster is within λ distance of a point.

Efficiency and Interpretability. DP-Means runs in $O(nKT)$ time and $O(K)$ space, supporting online streaming. We favor it over DBSCAN ($O(n^2 \log n)$) and agglomerative clustering ($O(n^2 d)$) due to its speed, memory efficiency, and direct interpretability via λ . Unlike opaque hyperparameters (e.g., ‘minPts’ in DBSCAN), λ has a direct semantic meaning.

Appendix B: User Network Construction Pipeline Evaluation Details

Weighting Strategy Ablation for CANE. To understand how different edge weighting schemes affect performance, we evaluate three variants of our CANE graph construction method on the China and Iran information operation datasets: (1) *Raw Count*, where edge weights are simple co-participation counts; (2) *Softmax-Normalized*, which applies a softmax over per-user neighbors; and (3) *TF-IDF-Inspired*, our proposed method which weighs user-cluster participation based on global distinctiveness.

Table 8 reports the comparative results. Across both datasets, the TF-IDF-inspired weighting strategy consistently outperforms the other two, illustrating the value of incorporating global cluster salience when modeling user similarity.

Approximate Nearest Neighbor Evaluation (FAISS). To scale our CANE graph construction, we replace brute-force similarity comparisons with approximate nearest neighbor search using FAISS³. Specifically, we use the Hierarchical Navigable Small World (HNSW) index (Malkov and Yashunin 2018), which enables sublinear retrieval while

³<https://github.com/facebookresearch/faiss>

Weighting Strategy	China		Iran	
	F1	AUC	F1	AUC
Raw Count	0.70 \pm 0.01	0.90 \pm 0.01	0.80 \pm 0.01	0.93 \pm 0.01
Softmax-Normalized	0.54 \pm 0.02	0.87 \pm 0.02	0.67 \pm 0.01	0.86 \pm 0.00
TF-IDF-Inspired	0.72 \pm 0.02	0.91 \pm 0.01	0.82 \pm 0.01	0.94 \pm 0.00

Table 8: Performance comparison of CANE weighting strategies on the China and Iran information operation datasets. The TF-IDF-inspired scheme outperforms both raw count and softmax-normalized weighting approaches.

Similarity Backend	China		Iran	
	F1	AUC	F1	AUC
Brute-force	0.72 \pm 0.01	0.92 \pm 0.02	0.83 \pm 0.01	0.94 \pm 0.00
FAISS-HNSW	0.72 \pm 0.02	0.91 \pm 0.01	0.82 \pm 0.01	0.94 \pm 0.00

Table 9: Comparison of exact (brute-force) vs. approximate (FAISS-HNSW) nearest neighbor computation in CANE. FAISS-HNSW yields virtually identical performance while theoretically enabling approximately $\sim 10\times$ faster construction for large graphs (e.g., 77k users, $k = 800$, 48M edges).

maintaining high recall. This substitution accelerates user-user similarity computation without compromising downstream performance.

To confirm that approximation does not degrade model quality, we compare the original brute-force CANE construction with its FAISS-based counterpart on the China and Iran IO datasets. As shown in Table 9, performance remains statistically indistinguishable across both campaigns, validating the use of FAISS-HNSW as a practical and scalable alternative.

Computational and Time Complexity. We outline the computational complexity of our framework in three key stages: (1) post-level embedding, (2) clustering via DP-Means, and (3) user-user network construction.

All methods, including baselines, require the full embedding of N posts into d -dimensional vectors, yielding a shared base cost of $O(Nd)$. Our framework adds a clustering stage, using DP-Means, to form a latent topic structure which improves network expressivity and scalability.

We compare our method against platform-agnostic baselines from previous works that also attempt to connect users based on their embedded content, representing fair baselines for content/platform-agnostic methods.

- **Embedding Averaging:** Each user’s messages are embedded, and pairwise user similarity is computed as the average cosine similarity between all messages authored by each user.
- **k-NN Embedding Graph:** All messages are embedded and used to construct a k-NN graph; user-user edges are then inferred based on how many of their messages are neighbors.

We enumerate this in Table 10.

We then use dataset statistics from our experimental settings to empirically model how cluster count, user count, and

Method	Embedding	Clustering	Graph Construction
<i>Our Methods</i>			
CANE (FAISS)	$O(Nd)$	$O(NKT)$	$O(U \log U + Uk)$
t-CANE (FAISS)	$O(Nd)$	$O(NKT)$	$O(T \cdot (U \log U + Uk))$
<i>Baselines</i>			
Embedding Averaging	$O(Nd)$	—	$O(U^2d)$
k-NN Embedding Graph	$O(Nd)$	—	$O(N \log N + kN)$

Table 10: Breakdown of computational costs for different network construction methods. All methods require embedding N posts into d -dimensional space. CANE and t-CANE include clustering via DP-Means (K clusters), and use approximate user-user graph construction via FAISS. Baseline methods compute similarity at the post or user level without intermediate clustering. **Notation:** N - posts, U - users, d - embedding dim, K - clusters, T - time bins, k - neighbors.

post volume scale in real-world cases. Based on these trends, we plot the estimated time complexity of each method in Figure 3. While both of our methods are slightly less efficient than the k-NN Embedding Graph baseline, they are significantly more efficient than the Embedding Averaging approach. As shown in subsequent experiments, we find this minor efficiency tradeoff acceptable given the substantial gains in performance and network expressivity achieved by our methods.

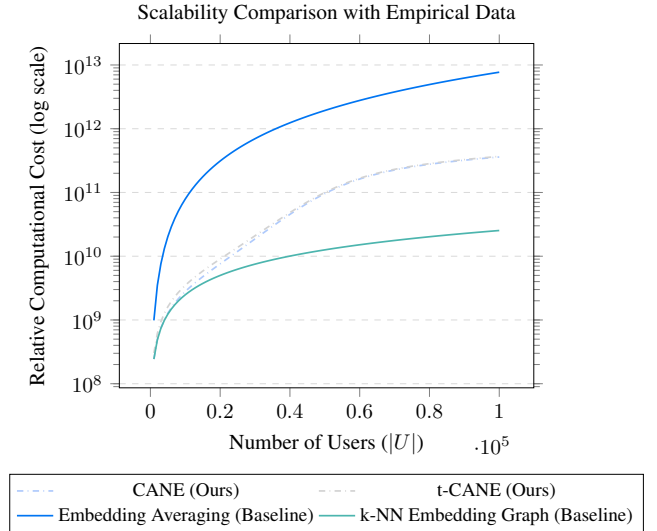


Figure 3: Comparison of computational complexity across network construction methods using empirically-informed scaling trends. CANE and t-CANE scale more efficiently with users due to FAISS and clustering, while baselines exhibit steep cost increases tied to post or user volume.

Setting λ : Cosine Similarity Cutoff. To determine an appropriate cosine similarity threshold for clustering narratives, we conduct a human evaluation of semantic similarity across five thresholds. For each threshold (e.g., 0.60), we sample 50 random paragraph pairs whose cosine similarities fall within a narrow window (e.g., between 0.59 and 0.61).

Threshold	Human Agreement
0.60	74%
0.65	86%
0.70	90%
0.75	94%
0.80	98%

Table 11: Human evaluation accuracy for narrative clustering at different cosine similarity thresholds.

These pairs are annotated by an expert to assess whether they refer to the same underlying topic. The evaluation is conducted exclusively in English, the dominant language in our dataset. Following the protocol outlined in prior work (Hanley and Durumeric 2024), we compute human agreement as the proportion of pairs correctly identified as referring to the same or different topics. Table 11 reports human agreement with model similarity scores across thresholds, providing an empirical basis for selecting a clustering cutoff.

We adopt a threshold of 0.70 as it offers a strong balance between semantic precision and narrative coverage. As shown in Table 11, human agreement increases sharply between 0.60 and 0.65, indicating that the model’s similarity scores begin to meaningfully capture topic-level coherence at this point. While stricter thresholds such as 0.75 and 0.80 yield marginally higher agreement, they risk over-segmenting the narrative space by excluding topically related posts that fall just below the cutoff.

Clustering Parameter Choice. To evaluate the sensitivity of our clustering threshold λ , we conduct a human annotation task across multiple cosine similarity cutoffs. At each threshold, we sample 50 random pairs of posts from different clusters and ask annotators to judge whether they express semantically similar narratives. We report the percentage of agreement with the clustering decision as a proxy for semantic coherence at that threshold. Results in Table 11 show how human-labeled accuracy varies across similarity cutoffs.

Appendix C: CANE Temporal Extension (t-CANE) and Edge Update Equations

Edge Update Rule. At each time step t , we compute similarity scores s_{uv}^t from user vectors and update edges as:

$$e_{uv}^t = \begin{cases} \alpha \cdot \tilde{s}_{uv}^t + (1 - \alpha) \cdot e_{uv}^{t-1}, & \text{if } s_{uv}^t \text{ is defined} \\ (1 - \beta) \cdot e_{uv}^{t-1}, & \text{otherwise} \end{cases} \quad (4)$$

Here, $\alpha \in [0, 1]$ controls how strongly new similarities influence edge weights, while $\beta \in [0, 1]$ controls the rate at which inactive edges decay. This rewards users who consistently co-engage with similar clusters across timesteps, allowing their edge to strengthen over time, while transient alignments naturally diminish.

Hyperparameter Settings. We set default values of $\alpha = 0.8$ and $\beta = 0.2$, based on downstream validation for information operation detection via grid a targeted grid search. We show example values and corresponding Macro-F₁ and

AUC in Table 12. While these settings generalize well across tasks, we leave more exhaustive hyperparameter optimization to future work.

Component Contribution and Ablation. To understand the contribution of each mechanism in **t-CANE**, we ablate the model’s core components. Specifically, we evaluate performance when removing temporal memory ($\alpha = 0$) and decay weighting ($\beta = 0$), while keeping other elements fixed. We also compare against simpler aggregation baselines, including unweighted similarity summation and naive averaging over time.

As shown in Table 13, the full model outperforms ablated variants and baselines, underscoring the importance of both memory integration and decay control. Averaging performs reasonably but lacks the capacity to distinguish stable alignment from noise.

Appendix D: Information Operations Results

Dataset Composition. We evaluate on Twitter’s publicly released Information Operations archive (Seckin et al. 2024), using labeled campaigns attributed to China and Iran. Each dataset includes state-backed IO driver accounts and a large set of matched organic control users. Table 14 summarizes the number of accounts and total posts per group.

Network Construction Details. Following Luceri et al. (2024a), we generate user-user similarity graphs for each network construction method. For our methods, CANE and t-CANE, we compute participation profiles over clustered topics and connect users based on cosine similarity. Temporal variants further aggregate user behavior over time windows.

Evaluation Setup. Each graph is embedded using node2vec (Grover and Leskovec 2016), and a random forest classifier is trained to predict whether a user is part of the IO campaign. Models are evaluated using 5-fold cross-validation, with Macro-F₁ and AUC reported. Full results are shown in Tables 2 and 3.

Sparse Data Robustness. To evaluate data efficiency, we simulate limited post availability by sampling $n\%$ of each user’s content in 5% increments. At each step, we measure when a method reaches 95% of its best AUC. Table 15 reports these thresholds, demonstrating that CANE t-CANE achieves strong performance with minimal data.

Interpreting Peak AUC Thresholds. Figure 4 visualizes the proportion of peak AUC achieved as a function of available training data. Each method’s curve is truncated at the earliest point where 95% of its maximum AUC is reached. This threshold balances interpretability and robustness: the trendlines shows smooth monotonic increases, and most methods saturate well before full data availability. For instance, both CANE and t-CANE reach 95% of peak performance with just 5% of user data. These saturation points validate the 95% threshold as a conservative but effective marker of data efficiency.

Configuration	ROC-AUC	Precision	Recall	F ₁
$\alpha = 0.2, \beta = 0.0$	0.98 ± 0.00	1.00 ± 0.00	0.74 ± 0.02	0.82 ± 0.02
$\alpha = 0.2, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.83 ± 0.01
$\alpha = 0.2, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.81 ± 0.02
$\alpha = 0.2, \beta = 0.15$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.01	0.81 ± 0.01
$\alpha = 0.2, \beta = 0.2$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.82 ± 0.02
$\alpha = 0.3, \beta = 0.0$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.3, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.82 ± 0.02
$\alpha = 0.3, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.3, \beta = 0.15$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.3, \beta = 0.2$	0.98 ± 0.00	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.4, \beta = 0.0$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.4, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.83 ± 0.01
$\alpha = 0.4, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.82 ± 0.02
$\alpha = 0.4, \beta = 0.15$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.4, \beta = 0.2$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.83 ± 0.01
<hr/>				
$\alpha = 0.5, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.82 ± 0.02
$\alpha = 0.5, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.01	0.81 ± 0.01
$\alpha = 0.6, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.81 ± 0.01
$\alpha = 0.6, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.01	0.82 ± 0.01
$\alpha = 0.7, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.82 ± 0.02
$\alpha = 0.7, \beta = 0.1$	0.97 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.81 ± 0.02
$\alpha = 0.7, \beta = 0.2$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.83 ± 0.02
$\alpha = 0.8, \beta = 0.05$	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.82 ± 0.02
$\alpha = 0.8, \beta = 0.1$	0.98 ± 0.01	1.00 ± 0.00	0.74 ± 0.02	0.82 ± 0.02
$\alpha = 0.8, \beta = 0.15$	0.98 ± 0.00	1.00 ± 0.00	0.73 ± 0.02	0.82 ± 0.01
$\alpha = 0.8, \beta = 0.2$	0.98 ± 0.01	1.00 ± 0.00	0.75 ± 0.01	0.83 ± 0.01

Table 12: Expanded grid search results on the China Information Operations detection task. All metrics are macro-averaged and reported as mean \pm standard deviation across cross-validation folds.

Appendix E: Annotation Statistics for Ideological Mapping

Initial Annotation and Agreement. Table 16 presents the initial ideological annotations of users from X and TikTok. Each user was labeled as *Liberal*, *Conservative*, or *Other/NA*. Inter annotator agreement, measured by Cohen’s κ , was high across both platforms, with values of 0.84 for X and 0.87 for TikTok, indicating strong reliability.

Filtered High-Confidence Labels. For downstream modeling, we retain only users where both annotators agreed. Table 17 shows the final label distributions after filtering. This filtered set serves as the basis for all supervised experiments and classification tasks.

Controlled Evaluation for Network Comparison. In Section *Ideological Mapping on X and TikTok*, we compare network construction methods by training classifiers to predict user ideology from node embeddings. However, raw comparisons are complicated by differing graph coverage (e.g., co-URL and co-Reply) graphs include only small subsets of users due to sparse signals on X.

To isolate model performance from coverage effects, we

conduct a controlled evaluation using only the users included in each graph’s subgraph. Table 18 reports the number of Liberal, Conservative, and Other/NA users in each graph’s subset. For classification experiments, we exclude Other/NA users and balance the liberal and conservative classes to prevent skewed results.

Appendix F: Ideological Mapping Results

Testing Procedure for Ideological Mapping. We frame the ideology classification task as a purely network-based prediction problem, using only the structure of user-user graphs to predict ideological leaning. While recent studies often combine network and content-based signals (Jiang, Ren, and Ferrara 2023), we deliberately exclude all text features to isolate the structural contribution of different graph construction methods.

We evaluate our approach on annotated data from X and TikTok, as shown in Section . The underlying label distributions are summarized in Appendix E (Tables 16 through 17). For model training, we focus only on users labeled as *Liberal* or *Conservative*, discarding all *Other/NA* labels. To ensure that results are not driven by class imbalance, we

Model Variant	ROC-AUC	Precision	Recall	F ₁
Full t-CANE	0.98 ± 0.01	1.00 ± 0.00	0.75 ± 0.01	0.83 ± 0.01
<i>w/o Memory</i> ($\alpha=0$)	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
<i>w/o Decay</i> ($\beta=0$)	0.98 ± 0.01	1.00 ± 0.00	0.73 ± 0.02	0.81 ± 0.02
<i>No Memory or Decay</i> (Average Similarity)	0.96 ± 0.01	1.00 ± 0.00	0.72 ± 0.02	0.79 ± 0.02
<i>No Memory or Decay</i> (Naive Aggregation)	0.97 ± 0.00	0.99 ± 0.00	0.71 ± 0.02	0.80 ± 0.01

Table 13: Ablation study on key components of **t-CANE** based on results from the China Information Operations detection task. All metrics are macro-averaged. The full model incorporates temporal memory (α) and decay smoothing (β). Removing memory ($\alpha=0$) leads to a collapse in performance, as user similarity cannot meaningfully propagate across time. The final two rows reflect ablations of both components, where user similarity is computed independently at each timestep and aggregated via averaging or summation.

Country	IO Drivers [Posts]	Control Users [Posts]
China	5,191 [13.8M]	76,286 [3.5M]
Iran	209 [9.9M]	16,885 [2.5M]

Table 14: IO datasets used in this study. For each campaign, we report the number of state-backed accounts (IO drivers) and control users, along with post volumes.

Category	Method	China	Iran
Baseline	Co-Repot	15%	20%
	Co-URL	20%	25%
	Fast Repost	20%	20%
	Hashtag Sequence	80%	60%
	Text Similarity	5%	10%
	k-NN Embedding Graph	10%	5%
	Fused Graph	15%	30%
Ours	CANE	5%	10%
	t-CANE	5%	5%

Table 15: Percentage of training data required to reach 95% of peak AUC for each method on the China and Iran IO campaigns. Lower values reflect higher data efficiency.

undersample the majority class within each graph subset to achieve balanced training sets.

We compare our method against several traditional and hybrid network construction baselines: Co-Repot, Co-URL, Fast Repost, Hashtag Sequence, Text Similarity, the k-NN Embedding Graph, and a Fused Graph combining available signals. Due to metadata constraints, the TikTok dataset includes only Hashtag Sequence, Text Similarity (from video descriptions), k-NN Embedding Graph, and a partial Fused Graph constructed from these sources.

To evaluate classification performance, we train three types of graph-based models on each network: (1) Graph Attention Networks (GAT) (Veličković et al. 2017), (2) Graph Convolutional Networks (GCN) (Zhang et al. 2019), and (3) node2vec (Grover and Leskovec 2016) embeddings combined with a feedforward neural network (NN) classifier. We initially tested node2vec with a random forest (RF) classi-

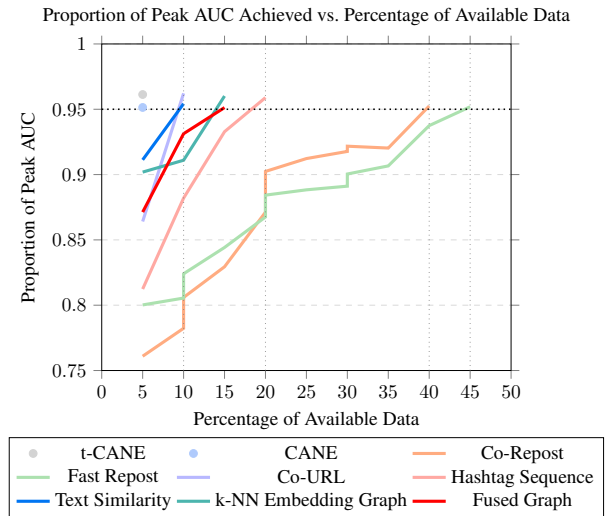


Figure 4: Proportion of peak AUC achieved as a function of training data percentage. Each curve is truncated at the first point where the method reaches 95% or more of its peak performance, illustrating the data required to achieve near-optimal results.

fier as well, but found it consistently underperformed and therefore excluded it from subsequent comparisons. Overall, performance across GAT, GCN, and node2vec+NN was broadly comparable; for consistency and ease of interpretation, we report main results using node2vec+NN.

All models are trained using 5-fold cross-validation and evaluated on held-out validation sets using Macro-F₁ and AUC. Statistical comparisons between models are conducted using unpaired t-tests on the fold-level results. Performance results (mean and standard deviation) are reported in Table 4. Full model hyperparameters and training code are available in our public repository.

Annotation Guidelines. Table 19 reproduces the full labeling instructions shown to annotators for the ideology classification task. These were presented in advance of annotation and were standardized across annotators.

Controlled Evaluation Across Network Types. To isolate

Label	X		TikTok	
	Count (%)	Cohen’s κ	Count (%)	Cohen’s κ
Liberal	503 (24%)	0.86	408 (23%)	0.88
Conservative	1,161 (55%)	0.88	681 (39%)	0.89
Other / NA	436 (21%)	0.79	670 (38%)	0.85
Total / Avg	2,100	0.84	1,759	0.87

Table 16: Initial annotation statistics for X and TikTok users. Counts and inter-annotator agreement are shown by label.

Label	X		TikTok	
	Count	%	Count	%
Liberal	433	22.8%	359	22.2%
Conservative	1,057	55.6%	620	38.4%
Other / NA	410	21.6%	636	39.4%
Total	1,900	100%	1,615	100%

Table 17: Final label distributions after removing all users with annotator disagreement. Only users with consensus labels are retained.

the effect of graph structure from node coverage, we rerun our ideological classification experiments using only the intersection of users shared across all graph types. This controlled setup allows us to evaluate structural quality independent of user inclusion, ensuring a fair comparison. As shown in Table 20, our method performs on par with or better than established graphs, validating that performance gains are not simply a result of broader user coverage.

Data Efficiency Analysis. To evaluate how robust each method is under limited data conditions, we progressively subsample user posts and measure the minimum percentage of data needed to reach 95% of each method’s peak AUC. Table 21 summarizes these thresholds for both X and TikTok. Our methods consistently achieve near-optimal performance with substantially less data, demonstrating strong data efficiency and robustness to real-world sparsity.

Appendix G: Cross-Platform Engagement Prediction Task

Narrative Identification. Narratives are identified via multilingual MPNet embeddings fine-tuned for semantic similarity, followed by DP-means clustering with a cosine distance threshold of 0.35. This threshold was validated through manual evaluation for coherence and specificity. Each cluster was assigned a topic label via top TF-IDF terms. To prevent circularity, narrative text was excluded from graph features.

Dataset Summary. Table 22 provides summary statistics for the full set of narrative themes evaluated in the engagement prediction task. We report the number of users and posts associated with each theme, as well as platform composition and distributional metrics. This context helps characterize the underlying structure and balance of the prediction space

Similarity Network	Liberal	Conservative	Other / NA
Co-Repost	182	428	220
Co-URL	160	269	109
Fast Repost	270	161	72
Hashtag Sequence	331	421	108
Fused Graph	681	483	217

Table 18: Label distribution in each similarity network’s test subset. Controlled evaluation (Section §) uses only users with Liberal or Conservative labels in each subset.

Table 19: Annotation Instructions Presented to Annotators

Guidelines

You will be assigned a list of 150 users. For each user, you will be shown a sample of five posts. Based on these posts, assign one of the following ideological labels to the user: *Liberal*, *Conservative*, or *Other/NA*.

These posts are drawn from a dataset of political discourse related to the 2024 U.S. Presidential Election on X (formerly Twitter). In many cases, a user’s political leaning will be evident through partisan language, endorsements, or other clear cues. However, if a user’s stance is unclear, ambiguous, or unrelated to politics, label them as *Other/NA*.

We acknowledge that this labeling framework is inherently limited. The liberal/conservative binary does not capture political ideology’s full complexity. Given the constraints of the task, however, please do your best to assign the most appropriate label based on available evidence.

across X and Truth Social.

Representative Narrative Themes. Table 23 presents a sample of narrative themes with substantial activity across both platforms. These examples illustrate the mix of episodic and persistent content our model handles: ranging from emergent protest events to recurring partisan issues. Platform distribution is also shown to highlight asymmetries in narrative engagement.

Model Architecture and Training Setup. Table 24 outlines the architecture and training parameters used in our engagement prediction model. We adopt a two-layer Graph Convolutional Network (GCN) with binary input features capturing user-topic interaction histories. Evaluation relies on Micro-F₁, AUC, and accuracy, with thresholds selected via validation sweep. A randomly rewired GCN serves as a structural baseline.

We use a two-layer Graph Convolutional Network (GCN) (Zhang et al. 2019) to predict future engagement with trending topics. Each user is represented by a binary feature vector indicating past engagement with a consistent set of clusters, and edges reflect user similarity based on the selected network (e.g., Co-URL, Fused Graph).

We initially tested multiple architectures, including Graph Attention Networks (GAT), but observed no meaningful per-

Method	Similarity Network	F1	AUC
Baseline	Co-Repost	0.82 ± 0.02	0.68 ± 0.05
CANE	Co-Repost	0.79 ± 0.04	0.70 ± 0.09
t-CANE	Co-Repost	0.86 ± 0.02	0.83 ± 0.03
Baseline	Co-URL	0.77 ± 0.01	0.54 ± 0.03
CANE	Co-URL	0.78 ± 0.04	0.72 ± 0.05
t-CANE	Co-URL	0.82 ± 0.02	0.79 ± 0.05
Baseline	Fast Repost	0.88 ± 0.02	0.81 ± 0.05
CANE	Fast Repost	0.83 ± 0.02	0.77 ± 0.04
t-CANE	Fast Repost	0.89 ± 0.03	0.87 ± 0.02
Baseline	Hashtag Sequence	0.81 ± 0.02	0.83 ± 0.03
CANE	Hashtag Sequence	0.77 ± 0.04	0.78 ± 0.05
t-CANE	Hashtag Sequence	0.83 ± 0.02	0.84 ± 0.03
Baseline	Fused Graph	0.86 ± 0.01	0.84 ± 0.02
CANE	Fused Graph	0.77 ± 0.04	0.78 ± 0.05
t-CANE	Fused Graph	0.83 ± 0.02	0.84 ± 0.03

Table 20: Controlled evaluation of ideological classification performance on X, using only the subset of users present in each similarity network’s graph. This setup isolates structural quality by holding user coverage constant across methods. Under these conditions, our method performs comparably to established approaches like Repost and Fused Graphs, indicating that it offers similar structural effectiveness while providing broader user coverage. We omit Text Similarity and k-NN Embedding Graph methods from this comparison, as they naturally include the same users as our approach and require no such control.

formance gains over GCN. For clarity and reproducibility, we report final results using GCN.

We train the model using binary cross-entropy loss with dynamic class weighting to address label imbalance, optimizing over 100 epochs using the Adam optimizer. Evaluation is conducted every 5 epochs, with micro- F_1 optimized via threshold sweeping on the validation set. To contextualize performance, we include, as a baseline, a randomly initialized GCN trained on a degree-preserving rewiring of the graph. Performance metrics (F_1 , AUC, and accuracy) are reported as the mean and standard deviation across snapshots.

Engagement Prediction Across Time Windows. Table 25 presents the performance of different similarity networks in predicting future user engagement with specific narratives across four time horizons: 3, 5, 7, and 14 days. Performance is measured using F_1 and AUC.

Appendix H: Fear Speech Detection and Examples

Definition of Fear Speech. Following Saha et al. (2023), we define *fear speech* as language that expresses apprehension or alarm about a group, portraying it as a threat to the in-group’s safety, values, or way of life. Unlike hate speech, which tends to be overtly derogatory or dehumanizing, fear speech is often framed as reasoned or cautionary discourse, designed to evoke anxiety or urgency. It may implicitly jus-

Category	Method	X	TikTok
Baseline	Co-Repost	15%	–
	Co-URL	20%	–
	Fast Repost	20%	–
	Hashtag Sequence	80%	60%
	Text Similarity	5%	10%
	k-NN Embedding Graph	10%	5%
Ours	Fused Graph	15%	30%
	CANE	5%	10%
	t-CANE	5%	5%

Table 21: Percentage of training data required to reach 95% of peak AUC for each method on the X and TikTok datasets. Lower values indicate faster convergence toward near-optimal performance.

Metric	Value
Total Narrative Themes	321
Total Posts	374,148
Total Users	261,398
Median Posts per Theme	154
Median Users per Theme	116
Mean X Post %	72.3
Mean Truth Social Post %	27.7
Median X User Count	89
Median Truth Social User Count	15

Table 22: Summary statistics for all tested narrative themes.

tify exclusionary or punitive responses while maintaining a rhetorically restrained tone (Gerard et al. 2024).

Examples From the Dataset.

“Besides, they [immigrants] don’t believe our laws apply to them even when they come here. They’re already practicing p***philia in their enclaves.”

“Keep this one going, help the ones still blinded by communist regime. The media while Venezuelan gangs torture, kill citizens in 5 states! Biden/Kamala open borders—no one saw this on CNN, etc. Corrupt media. This was in Merit Street News!”

Fear Speech Detection Model. We follow the training setup and hyperparameter choices from Saha et al. (2023), fine-tuning a RoBERTa-based classifier ⁴ on their annotated dataset of over 22,000 social media posts labeled as hate, fear, or benign. The model is trained to identify group-directed alarmist rhetoric: *e.g.*, language invoking cultural threat, physical danger, or national decline. Saha et al. (2023) report a macro- F_1 of 0.62 for their model on this task; our implementation yields comparable performance.

Each post in our dataset receives a continuous score between 0 and 1 indicating the likelihood of expressing fear speech. For our main analysis, we define fear speech as any post with a model score exceeding 0.75. This threshold, slightly more conservative than the default 0.5, prioritizes precision over recall to reduce false positives and en-

⁴<https://huggingface.co/FacebookAI/xlm-roberta-large>

Narrative Theme	Posts	Users	X (%)	TS (%)	X Users	TS Users
Columbia University and Protest Suppression	688	382	60.9	39.1	325	57
Promotion of Conservative Media	518	271	50.0	50.0	243	29
FBI Investigations and Federal Overreach	302	199	57.0	43.0	162	37
Chemical Leaks and Environmental Disasters	289	154	55.7	44.3	125	29
Condemnation of Pro-Palestinian Protest Tactics	329	140	50.5	49.5	102	39
Mar-a-Lago Probe and GOP Investigations	247	159	56.3	43.7	123	38
Domestic Shootings and Public Safety	239	118	43.1	56.9	85	33
Nuclear Energy Conspiracy Theories	89	73	61.8	38.2	51	22
Criticism of Jan 6 Committee Chair	87	68	60.9	39.1	52	18
Global Unrest and War Escalations	81	57	49.4	50.6	36	21
Comparing Jan 6 and Tiananmen Square	66	47	57.6	42.4	33	14
Transgender Policy Debate	53	42	52.8	47.2	27	15

Table 23: Examples of narrative themes with presence on both X and Truth Social (TS), shown to illustrate our inclusion of both emerging events and broader, persistent discourse. Narrative themes were manually assigned by a human annotator based on the five posts closest to each narrative cluster centroid and the top TF-IDF terms. Full metadata is available in the accompanying repository.

Setting	Value
Model Type	2-layer Graph Convolutional Network (GCN)
Hidden Dimension	64
Input Features	Prior engagement (binary per topic)
Output	Future engagement (binary per topic)
Loss Function	Binary Cross-Entropy with class weights
Optimizer	Adam
Learning Rate	0.01
Epochs	100
Evaluation Metrics	Micro-F ₁ , AUC, Accuracy
Baselines	Random GCN
Threshold Selection	F ₁ -optimized sweep (0.05–0.95)

Table 24: GCN architecture and training settings for future engagement prediction.

sure that identified posts reflect clearer expressions of fear-laden rhetoric. We verify that the same trend holds even under stricter criteria (up to thresholds of 0.95) and find that platform-origin differences persist. For example, when applying a stricter 0.9 cutoff, Truth Social-seeded narratives still exhibit a +0.23 log-odds ratio and a 23.5% relative increase in fear-laden content compared to those originating from X ($p < 0.01$).

Appendix I: Bridge Users

Bridge Zone Visualization. Figure 5 displays the bridge zone as an extracted subgraph from the full discourse network. Node positions are determined using a force-directed layout computed in Gephi⁵ with the Force Atlas 2 algorithm. Each node is colored by platform (Truth Social in red, X in blue) and sized by degree. This view isolates a structurally integrated region where users from both platforms are densely connected, reflecting cross-platform alignment within the discourse network. We visualize node size using degree to highlight structural embeddedness: an indicator of how centrally users are positioned within the network topol-

ogy.

Bridge User Activity Characteristics. Table 26 reports median engagement metrics and within-platform percentile ranks for this subset. Given substantial differences in engagement norms and distributional skew between X and Truth Social, we use rank-based percentiles computed within each platform and then aggregate across them. We use rank-based percentiles to capture users’ relative position in a way that is robust to outliers and platform-specific scale (because engagement distributions vary widely across platforms like X and Truth Social, and are often heavy-tailed, raw comparisons or z-scores can be misleading). Notably, bridge users are not uniformly high-volume or high-engagement accounts. Their engagement levels consistently hover near the platform medians, reinforcing that their influence stems not from visibility but from their structural positioning across fragmented discourse communities.

Notably, bridge users are not uniformly high-volume or high-engagement accounts. Their engagement levels (likes, replies, and reposts) consistently fall near the platform-level medians, with percentile ranks between the 47th and 58th percentiles. These results reinforce the idea that bridge users are not defined by activity level but by structural embeddedness: they serve as connectors across fragmented platform communities rather than high-visibility influencers.

Robustness Check on Low-Prominence Narratives. Although our matched schema likely already accounts for differences in engagement with viral or high-visibility content (confirmed via Mann–Whitney U tests), we further restrict our analysis to test the robustness of this pattern. Specifically, we isolate a stricter subset of low-prominence narratives that, within the first 24 hours, attract no more than five users, span at most two communities, and appear on only one platform. In other words, narratives that exhibit minimal early signs of virality.

Even within this conservative subset, we observe large and statistically significant engagement differences. Clusters seeded early by bridge users receive substantially more likes ($p < 0.001$, rank-biserial = 0.503), replies ($p < 0.001$,

⁵<https://gephi.org/>

Time (days)	Method	F1	AUC
t = 3	Random GCN	0.00 ± 0.00	0.00 ± 0.00
	Co-URL	0.00 ± 0.00	0.00 ± 0.00
	Hashtag Sequence	0.05 ± 0.01	0.42 ± 0.06
	Text Similarity	0.01 ± 0.00	0.49 ± 0.04
	k-NN Embedding Graph	0.01 ± 0.01	0.52 ± 0.05
	(Partial) Fused Graph	0.03 ± 0.01	0.55 ± 0.05
	CANE	0.29 ± 0.02	0.86 ± 0.03
	t-CANE	0.34 ± 0.02	0.88 ± 0.03
t = 5	Random GCN	0.00 ± 0.00	0.00 ± 0.00
	Co-URL	0.00 ± 0.00	0.00 ± 0.00
	Hashtag Sequence	0.08 ± 0.02	0.45 ± 0.07
	Text Similarity	0.02 ± 0.01	0.53 ± 0.05
	k-NN Embedding Graph	0.02 ± 0.01	0.57 ± 0.06
	(Partial) Fused Graph	0.04 ± 0.01	0.60 ± 0.06
	CANE	0.30 ± 0.02	0.87 ± 0.03
	t-CANE	0.34 ± 0.02	0.88 ± 0.03
t = 7	Random GCN	0.00 ± 0.00	0.56 ± 0.03
	Co-URL	0.01 ± 0.00	0.43 ± 0.04
	Hashtag Sequence	0.11 ± 0.04	0.48 ± 0.08
	Text Similarity	0.02 ± 0.00	0.58 ± 0.05
	k-NN Embedding Graph	0.02 ± 0.01	0.61 ± 0.06
	(Partial) Fused Graph	0.05 ± 0.01	0.64 ± 0.06
	CANE	0.30 ± 0.02	0.89 ± 0.02
	t-CANE	0.35 ± 0.06	0.94 ± 0.02
t = 14	Random GCN	0.00 ± 0.00	0.00 ± 0.00
	Co-URL	0.00 ± 0.00	0.00 ± 0.00
	Hashtag Sequence	0.06 ± 0.01	0.40 ± 0.06
	Text Similarity	0.01 ± 0.00	0.47 ± 0.05
	k-NN Embedding Graph	0.01 ± 0.01	0.51 ± 0.05
	(Partial) Fused Graph	0.02 ± 0.01	0.53 ± 0.05
	CANE	0.29 ± 0.02	0.86 ± 0.03
	t-CANE	0.34 ± 0.02	0.88 ± 0.03

Table 25: Performance of similarity networks for cross-platform narrative engagement prediction across time windows. Values reflect mean F₁ and AUC (± standard deviation) across validation folds. t-CANE consistently outperforms all baselines.

$r = 0.432$), and reposts ($p < 0.001$, $r = 0.373$). These results confirm that bridge users are not simply interacting with content already trending toward virality. Rather, they are consistently associated with elevated engagement even when narratives begin with minimal visibility, suggesting that their influence operates independently of early popularity signals.

Robustness Checks on Early Seeding Thresholds. To assess whether our findings regarding the impact of bridge users on engagement are sensitive to threshold selection, we replicate our main analysis across a range of early-seeding definitions. Specifically, we vary the user participation threshold from 5% to 30% (in increments of 5%), while keeping the 24-hour temporal window fixed.

Across all thresholds, clusters seeded by bridge users consistently receive more likes, replies, and reposts than those seeded by matched non-bridge users. Effect sizes remain statistically significant and moderately large throughout, with rank-biserial correlations ranging from 0.27 to 0.60 for likes,

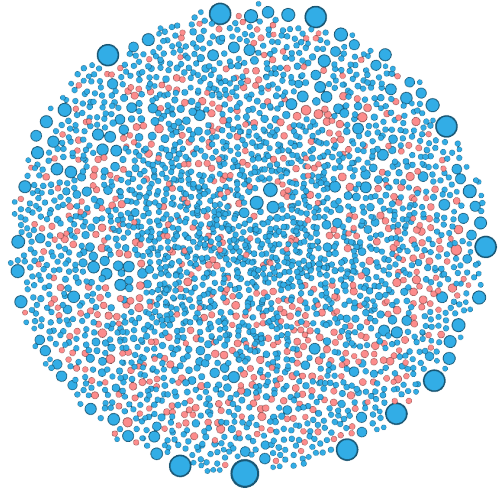


Figure 5: Visualization of the discourse network colored by platform (blue: X, red: Truth Social), with node size scaled by degree. This layout reveals a structurally embedded *bridge zone*: a dense, mixed-platform region near the center where users from both platforms are highly interconnected. These users serve as key conduits for narrative migration across fragmented media environments. The concentration of red nodes within the blue-majority core illustrates cross-platform entanglement not evident in interaction-based graphs.

0.32 to 0.43 for replies, and 0.27 to 0.46 for reposts. These results demonstrate that our engagement-based findings are not an artifact of a particular threshold choice. Rather, they underscore the robustness of bridge users’ outsized role in driving interaction, even under conservative definitions of early seeding.

Baseline Graph Evaluation. To assess whether other graph construction methods can uncover the same cross-platform alignment surfaced by our discourse-based approach, we construct several alternative user-user graphs: co-hashtag, co-URL, fused lexical similarity, and sentence-level embedding kNN graphs. Each graph is subjected to Louvain community detection, followed by platform entropy and narrative migration analysis.

Table 27 compares the proportion of users and posts identified as bridge users, along with the percentage of simple migrating narratives they introduced across platforms (we note that the trend remains the same when limited to significant migrating narratives):

Our method identifies 67.7% of cross-platform narrative introductions using only 0.33% of users: over 21× more efficient than the best-performing baseline (Fused Graph), which captures just 27.0% using 2.68% of users. Measured as the ratio of narratives introduced to users involved, our model achieves an efficiency score of **205**, far exceeding all

Metric	Median	Bridge Percentile (Platform-Normalized)
Total Posts	14.00	58.49%
Reply Count	0.16	47.06%
Like Count	0.58	53.26%
Repost Count	0.02	55.19%

Table 26: Median engagement metrics and corresponding platform-normalized percentile ranks for bridge users. Percentiles are calculated within each platform and then averaged.

Graph Type	% Users	% Posts	% Narratives Introduced
Co-URL	0.00%	0.00%	0.00%
Hashtag Sequence	0.04%	0.37%	4.21%
Text Similarity	1.42%	2.21%	15.70%
kNN Embedding Graph	1.73%	2.41%	25.90%
Fused Graph	2.68%	3.19%	27.00%
Ours (Discourse)	0.33%	2.14%	67.72%

Table 27: Comparison of bridge user detection and narrative introduction across graph construction methods.

other methods. Notably, the Hashtag Sequence graph performs poorly despite benefiting from feature leakage: hashtags were partially baked into the upstream data collection process by Balasubramanian et al. (2024) and (Shah et al. 2024). Both it and the co-URL graph fail to surface meaningful cross-platform bridges.

These findings underscore the distinctiveness of our bridge community and validate our clustering strategy, which surfaces discourse-level alignments rather than relying on shallow behavioral or lexical overlap. Our method reveals structurally coherent zones of narrative migration that remain invisible to traditional similarity-based approaches.

Bridge User Role in Cross-Platform Narrative Introduction. To better understand the temporal role bridge users play in narrative diffusion, Table 28 reports the proportion of clusters for which a bridge user appeared within the first n posts on the receiving platform, across different cluster types. For example, among *significant migrating clusters* (as identified by Pearson correlation), bridge users were the first to introduce the narrative 69.3% of the time and were among the first three posters in 79.4% of cases. These results suggest that bridge users are not simply eventual adopters, but consistent early initiators.

Cluster Type	Top 1	Top 2	Top 3	Top 4	Top 5
Significant Migration	69.33%	77.73%	79.41%	79.83%	80.67%
Simple Migration	67.72%	76.61%	78.29%	78.99%	79.45%

Table 28: Percentage of cross-platform narrative clusters where a bridge user appeared within the first n posts on the receiving platform.