

Predicting Information Pathways Across Online Communities

Yiqiao Jin

Georgia Institute of Technology
Atlanta, GA, United States
yjin328@gatech.edu

Meng Ye

SRI International
Princeton, NJ, United States
meng.ye@sri.com

Yeon-Chang Lee

Georgia Institute of Technology
Atlanta, GA, United States
yeonchang@gatech.edu

Kartik Sharma

Georgia Institute of Technology
Atlanta, GA, United States
ksartik@gatech.edu

Karan Sikka

SRI International
Princeton, NJ, United States
karan.sikka@sri.com

Ajay Divakaran

SRI International
Princeton, NJ, United States
ajay.divakaran@sri.com

Srijan Kumar

Georgia Institute of Technology
Atlanta, GA, United States
srijan@gatech.edu

ABSTRACT

The problem of *community-level information pathway prediction* (CLIPP) aims at predicting the transmission trajectory of content across online communities. A successful solution to CLIPP holds significance as it facilitates the distribution of valuable information to a larger audience and prevents the proliferation of misinformation. Notably, solving CLIPP is non-trivial as inter-community relationships and influence are unknown, information spread is multi-modal, and new content and new communities appear over time. In this work, we address CLIPP by collecting large-scale, multi-modal datasets to examine the diffusion of online YouTube videos on Reddit. We analyze these datasets to construct community influence graphs (CIGs) and develop a novel dynamic graph framework, INPAC (Information Pathway Across Online Communities), which incorporates CIGs to capture the temporal variability and multi-modal nature of video propagation across communities. Experimental results in both warm-start and cold-start scenarios show that INPAC outperforms seven baselines in CLIPP. Our code and datasets are available at <https://github.com/claws-lab/INPAC>

CCS CONCEPTS

- **Information systems** → **Content ranking; Data mining; Collaborative and social computing systems and tools.**

KEYWORDS

graph neural networks, information pathway, information diffusion

ACM Reference Format:

Yiqiao Jin, Yeon-Chang Lee, Kartik Sharma, Meng Ye, Karan Sikka, Ajay Divakaran, and Srijan Kumar. 2023. Predicting Information Pathways Across Online Communities. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*, August 6–10, 2023, Long



This work is licensed under a Creative Commons Attribution International 4.0 License.

Beach, CA, USA. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3580305.3599470>

1 INTRODUCTION

Background. Social media users form communities based on their interests, beliefs, ethnicity, and geographical location [82, 85]. These communities are prevalent on popular social platforms such as Reddit, WhatsApp, and Telegram, enabling users to connect with like-minded individuals as well as consume and disseminate information in an interactive manner. As communities grow in size, they become hubs of information flow, facilitating the exchange of information across communities. Existing research has shown that online communities interact with and influence one another [21, 50, 55, 101, 105].

As information spreads from one community to the other, it can rapidly reach all members in the new community. While individual posts and hyperlinks may propagate in varying patterns, the underlying pathways on which information propagates remain relatively stable [25, 90]. Their stability is partially due to the behavior of common users who repeatedly spread information among the same communities, creating a reinforcing effect of the underlying information pathways.

The fast-paced evolution of social media has accelerated the spread of information, including a variety of content types ranging from news articles, commercial advertisements, to harmful content such as online rumors [5, 74, 107, 123], fake news [5, 107, 115, 121], hate speech [75, 133], and political bias [47]. The unmoderated spread of these contents can cause adverse social impacts. For example, the COVID-19 pandemic has led to the formation and growth of multiple online communities, such as subreddits r/CoronavirusUS, r/COVID-19Positive, and r/COVID19, where users discuss various topics related to the pandemic. These communities are interconnected, with similar topics and user groups, thus having a significant influence on each other. Sometimes misinformation proliferates in online communities, such as the unfounded claim that 5G technology can spread the virus [1, 100]. Despite a lack of scientific evidence, this conspiracy theory gained traction in several online communities, including r/conspiracy, r/5G, r/CoronavirusUS,

and r/COVID19, causing unwarranted fear and concern among the public.

The Community-Level Information Pathway Prediction (CLIPP) problem seeks to predict the transmission trajectory of information among online communities. CLIPP is of significant importance as it enables prediction of communities where information, including problematic content, is likely to emerge and spread. Such capability can provide numerous benefits across a wide range of applications. Efficient prediction of misinformation spread with CLIPP can guide intervention strategies, while for advertising, CLIPP can refine strategies and maximize the efficacy of marketing campaigns, increasing the visibility of information and providing insights into the communities where their target audience is most active.

Challenges. Solving CLIPP is challenging. First, community-to-community influence is usually unknown [25, 93], and the mechanisms of interactions between communities and how they impact users remains hidden [55]. Different communities may have different norms, values, and communication patterns that influence the temporal patterns of information diffusion [117]. In this case, we only observe where new content is propagated to the new communities and when it takes place. The underlying community influence, *i.e.*, who influences the propagation, remains unknown. Most existing works focus on predicting information diffusion at the user level (*i.e.*, microscopic influence) [61, 90]. Meanwhile, existing datasets [79, 80, 103, 133] only contain limited information about community structures, making it difficult to study cross-community information spread.

Second, the spread of information is characterized by a complex and dynamic diffusion environment [66]. Posts contain multi-modal signals, such as text, images, and videos [4, 10, 40]. Diffusion patterns vary across content types. For example, misleading news and inflammatory microblogs spread faster and wider than true information [30, 41, 104]. Niche content are usually shared within a few narrow-interest communities, while broad-interest contents create far-reaching cascades and reach several disparate communities [89, 105, 106]. Understanding these propagation patterns is essential to predicting information spread across communities.

Our Work. In this work, we investigate the dynamics of community-level information flow while jointly addressing the challenges of complex diffusion environment and the continuously evolving information ecosystem.

We choose Reddit as the platform for studying community-level information diffusion since it provides numerous communities, named “subreddits,” that are dedicated to specific topics or interests. Towards this goal, we collect two large-scale and multi-modal datasets that enable us to study the community-level diffusion of visual contents for information pathway prediction. Based on that, we identify distinct temporal patterns of information sharing using inter-activity time distribution, infer macroscopic community-to-community influence, and construct novel community influence graphs (CIGs).

We design INPAC, or Information Pathway Across Online Communities, a dynamic graph-based method to predict community-level information pathways using CIGs and content’s multi-modal information (visual features and channel metadata). INPAC integrates structure, content semantics, and temporal information by utilizing

Table 1: Statistics of our datasets.

| | Large | Small |
|--------------------|-----------|----------|
| #Videos / URLs | 183,596 | 6,802 |
| #Subreddits | 57,894 | 7,319 |
| #Users | 291,047 | 8,752 |
| #Shares | 1,323,714 | 36,118 |
| Density | 7.96E-05 | 6.11E-04 |
| #Cold-start Videos | 3,042,068 | 68,095 |

Continuous-Time Dynamic Graphs (CTDGs) to model the time-aware propagation patterns of videos. In INPAC, nodes and edges are continuously introduced to the graph, incorporating both visual features and channel metadata of the content.

Contributions. Our main contributions are as follows:

- **Novel Multi-modal Datasets and Analysis:** We collect two large-scale, multi-modal datasets to study community-level diffusion of visual contents for information pathway prediction. We identify distinct temporal content sharing patterns that are used to infer community-to-community influence graphs.
- **Information Pathway Prediction Framework:** To solve CLIPP, we propose INPAC, a dynamic graph framework based on CIGs that learns from multimodal data and the dynamics of the interactions between users and communities.
- **Experimental Evaluation:** We demonstrate the effectiveness of INPAC framework and its design choices through experiments in various scenarios, *e.g.*, prediction of cold/warm-start videos on communities with various popularity. INPAC reaches performance improvements of up to 18.8% on MRR, 13.8% on NDCG@5, and 6.2% on Rec@5.

2 DATASET AND PROBLEM

2.1 Dataset Description

In this study, we aim to study the spread of visual content across communities on social media. To this end, we collect massive visual contents on YouTube and long-term community activity on Reddit. The reasons for selecting these two platforms in this study as follows:

- **YouTube** is one of the most widely used video-sharing platforms that contains over 2.56 billion users¹ and provides a venue for users to upload, share, and view videos.
- **Reddit** is one of the largest social platforms for content creation, rating, and sharing. It allows users to interact in a variety of communities (*i.e.*, subreddits). Reddit is an ideal platform for studying the propagation of online visual contents such as YouTube videos because of its vast and diverse user base as well as its open-source nature and community structures.

As the first step, we collected 54 months of historical Reddit posts from January 2018 to June 2022 via PushShift². We removed any posts that did not contain valid URLs and retained URLs associated with valid YouTube videos, resulting in 5,723,910 posts and 3,737,191

¹<https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

²<https://pushshift.io/>

Table 2: Examples of cross-community information flow in our datasets. A video is usually shared on a set of semantically similar subreddits. “→” indicates the temporal order of the sharing.

| Title of the Video | Subreddits on Which the Video Appears |
|---|--|
| Canadian Trudeau Investigation | Liberate_Canada → conspiracy → TheNewRight → PeoplesPartyofCanada → Canada_First |
| Reviews: Super Dragon Ball Heroes Episode 19 | promote → AnimeReviews → anime_manga → YouTubeAnimeCommunity → Anime_and_Manga |
| Warcraft 3 Reforged Cutscene Only | WC3 → pcgaming → warcraft3 → gaming → legaladviceofftopic |
| Practical Greeting Phrases for Chinese New Year | learnchinese → learnmandarin → learnmandarinchinese |
| Accepting what is. (Realize Instant Freedom) | AnxietyDepression → Soulexus → SpiritualAwakening → Meditation → spirituality → awakened → inspiration |
| Covid-19 Explained with Data Science | Python → CoronavirusUS → CanadaCoronavirus → CoronaVirus_2019_nCoV → CoronavirusUK |
| Implement RNN-LSTM for Music Genre Classification | learnmachinelearning → Python → tensorflow → musictheory |

associated videos. Finally, following previous works [9, 33], we retained videos shared within at least 3 communities. Table 1 shows the statistics of the two datasets we construct. The large dataset covers 54 months of video propagation history from January 2018 to June 2022, while the small dataset covers a 3-month period from January to March 2020. Table 1 reveals that both datasets contain a considerable number of cold-start videos with only one interaction in a subreddit, which reflects the real-world distribution and the challenges associated with information pathway prediction.

2.2 Problem Formulation

We formulate the CLIPP problem as follows: Given a video and a sequence of subreddits in which it has been posted, predict the next community the video will be posted in at a given time. Formally, we define a posting of a video as a video link appearing on a subreddit, either as a standalone post or as part of a longer post. A *posting instance* is represented as a 4-tuple $p_{ij} = (v_i, s_j, u_j, t_j)$, where v_i is a video posted by a user u_j in an online community s_j at time t_j . The *posting sequence* for v_i is defined as a list of posting instances $P_i = \{(v_i, s_j, u_j, t_j)\}_{j=1}^N$, which indicates the dissemination trajectory with length N across communities for the video v_i . Then, our problem can be defined as follows:

PROBLEM 1 (INFORMATION PATHWAY PREDICTION). Given a video v_i , its posting sequence $P_i = \{(v_i, s_j, u_j, t_j)\}_{j=1}^N$ with length N , and a target timestamp t_j' , our model outputs a ranked list of communities $\{s_k\}$ indicating the most likely communities that v_i will appear at time t_j' .

Table 3 summarizes a list of notations used in this paper.

3 THE PROPOSED FRAMEWORK: INPAC

3.1 Overview

In this work, we aim to study the propagation of online visual content on social media. To this end, we propose a dynamic graph framework INPAC based on Community Influence Graphs (CIGs) that learns the dynamics of cross-community information flow and accurately predicts information pathways. As shown in Figure 1,

Table 3: Notations used in this paper.

| Notation | Description |
|--|---|
| V, S | Set of videos and communities |
| v_i, s_j, u_k | Video, community, user |
| S^u | Historical interaction sequence for user u |
| P_i | Posting sequence of video v_i |
| G_i^S | Community-community influence graph for v_i |
| G^D | Dynamic graph |
| n | Maximum sequence length |
| e_{jk} | Edge weights |
| α | Teleport probability for APPNP |
| λ_1, λ_2 | Hyperparameters |
| $\Delta_t^{\text{Same}}, \Delta_t^{\text{Diff}}$ | Time intervals for same / different users |
| $f_\theta(\cdot, \cdot)$ | Message function for dynamic modeling |

INPAC consists of three key modules: (1) community influence modeling; (2) video content modeling; and (3) dynamic modeling.

3.2 Community Influence Modeling

Given a community (e.g., a subreddit), INPAC learns its embedding such that the embedding preserves its influence on other communities during information propagation. We infer the influence relationships between communities using content sharing patterns in those communities. Specifically, a video is usually shared in communities that have similar topics. For example, in Table 2, the video “Practical Greeting Phrases for Chinese New Year” is shared within a set of subreddits related to language learning, such as $r/learnchinese$ and $r/learnmandarin$. To model this, we create a novel influence network by leveraging the video’s temporal interaction patterns.

Influence Graph Construction. In the context of CLIPP, community-level influence is defined as the presence of causal relationships between posting of a video in two different communities. This can happen when two communities share a common group of users. To infer the influence exerted by one community on another, we employ a sequence of communities $\{s_1, s_2, \dots\}$, in which a video v_i is posted. Assuming that users require a certain amount of time to engage in online content, the interval between the appearance of a video v_i in two communities s_1 and s_2 serves as an indicator of the influence of s_1 on the appearance of the video in s_2 . If a video

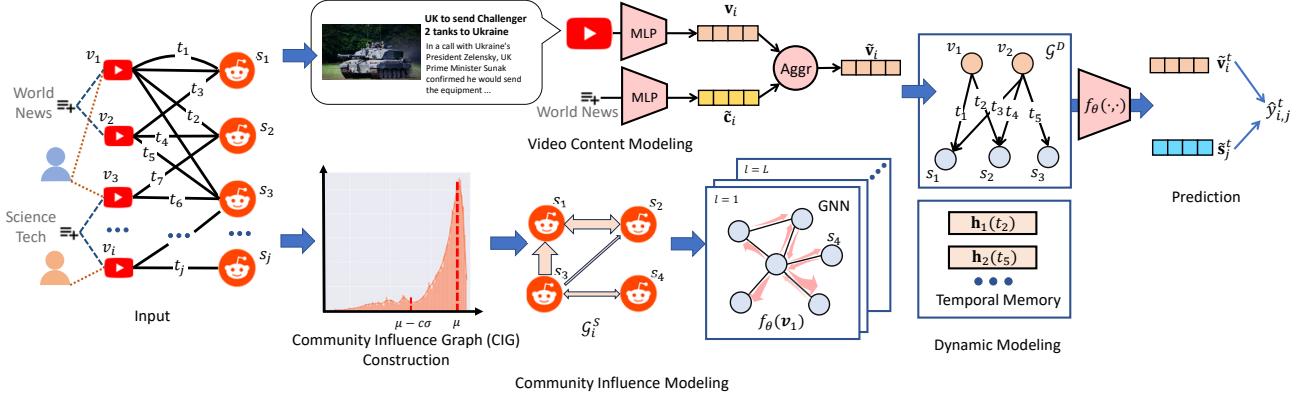


Figure 1: The overview of our proposed INPAC framework, which consists of static modeling, including video content and community influence modeling, as well as dynamic modeling.

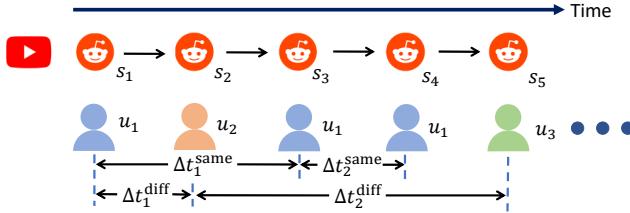


Figure 2: Illustration of how Δt^{Same} and Δt^{Diff} are calculated.

is shared by two users within a very short time interval, it suggests that the shares occur simultaneously and are not influenced by one another. Based on this assumption, we model a posting sequence P_i of v_i among communities as a directed graph \mathcal{G}_i^S consisting of community nodes s_j involved in the propagation of a video v_i .

To model the propagation sequence of a video, we first identify its concurrent sharing events, where the propagation of the video takes place within a brief time period, referred to as a session, in the same or different communities. To this end, one needs to decide whether two shares are within the same session. A straightforward approach is to set a threshold time limit, such as one hour or one day, as is common in session-based recommender systems [8, 29, 64, 73, 116]. However, this ad-hoc use of the time limit is insufficient as it can vary across datasets, videos, and platforms [39, 84].

We note that consecutive sharing of a video can occur due to the same user or different users, resulting in differing sharing patterns and motivations. Therefore, we create two distributions of time difference between consecutive shares of each video v_i : (1) Δt^{Same} , representing the time intervals between consecutive shares of v_i by the *same user*; (2) Δt^{Diff} , representing the time intervals between the first share of v_i by *different users*. Figure 2 illustrates an example of a video's consecutive sharing on several communities over time by three users. From Figure 2, we can observe how the two time intervals $\Delta t_1^{\text{same}}, \Delta t_2^{\text{same}}$ for the same user u_1 as well as the two intervals $\Delta t_1^{\text{diff}}, \Delta t_2^{\text{diff}}$ for different users u_1, u_2, u_3 are computed.

For Δt^{Same} , it is important to consider that a user's multiple postings of the same video in different communities should not be viewed as one community influencing another. This is because users usually post the same content in various venues to enhance its

visibility and attract more "likes" [22, 42, 122]. This is not indicative of natural flow of content from one community to another.

Thus, we only utilize Δt^{Diff} to identify community-level influence. Specifically, we plot the distribution of Δt^{Diff} across sharing events of all videos, as shown in Figure 3, where the x -axis represents the time interval in seconds with a logarithmic scale of base 10, and the y -axis indicates the percentage. Then, we fit a Gaussian distribution to Δt^{Diff} and found that the distribution has a mean of 6.844 and a standard deviation of 0.823 on the logarithmic scale. Based on this finding, we determine the cutoff time for partitioning sessions using $\Delta t^{\text{Thres}} = \mu - c\sigma$, where c is a hyperparameter that represents the confidence level for determining concurrent shares. When the time difference between two postings exceeds Δt^{Thres} , the later posting is considered to be influenced by the former.

Now, we construct the community influence graph (CIG) \mathcal{G}_i^S with respect to v_i based on the threshold Δt^{Thres} . Each node in \mathcal{G}_i^S indicates a community s_j and a directed edge from s_j to s_k indicates s_k is influenced by s_j . Specifically, if two shares of v_i from different users occur within Δt^{Thres} , they are considered concurrent postings in the same session and not influenced by each other. Otherwise, a directed edge is added from s_j to s_k for $t_j < t_k$ in \mathcal{G}_i^S . Furthermore, when v_i is simultaneously shared by the same user in two different communities s_j and s_k , a bi-directional edge is added between these communities to reflect their mutual influence as a result of overlapping users.

Message Aggregation. After the construction of \mathcal{G}_i^S , the graph is transformed from a multigraph to a weighted graph by merging multiple edges with the same source and destination nodes. Let \mathcal{E}_{jk} denote the set of edges between s_j and s_k . The new edge weight e_{jk} is calculated using the logarithmic value

$$e_{jk} = \ln(1 + |\mathcal{E}_{jk}|). \quad (1)$$

As \mathcal{G}_i^S consists of a number of periphery nodes, such as inactive online communities with few propagations, long-range dependencies should be considered to learn distinct node representation. To this end, we leverage the propagation scheme of APPNP [23] based on the personalized PageRank algorithm [2]. APPNP adds a probability of teleporting back to the root node, which ensures

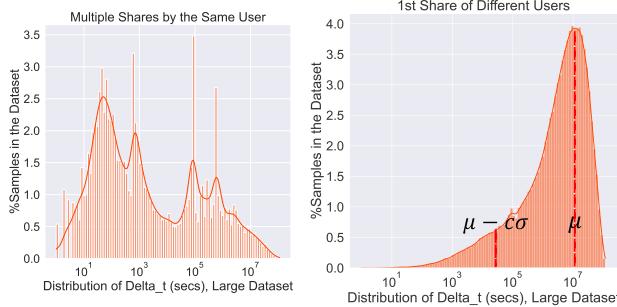


Figure 3: Distribution of Δt^{Same} (Left) and Δt^{Diff} (Right) for videos on the Large dataset.

that the PageRank score encodes the local neighborhood for every node and mitigates the oversmoothing issues.

Then, we obtain the embedding matrix $S_i^{(l)}$ at layer l for communities involved in the i -th propagation sequence P_i :

$$S_i^{(l)} = (1 - \alpha)\hat{D}_i^{-1/2}\hat{A}_i\hat{D}_i^{-1/2}S_i^{(l-1)} + \alpha S_i^{(0)}, \quad (2)$$

where $S_i^{(0)} = [s_1 || \dots || s_{|\mathcal{G}_i^S|}]$ is the initial embedding matrix for all $s_i \in \mathcal{G}_i^S$. \hat{A}_i and \hat{D}_i are the adjacency matrix and the diagonal degree matrix, respectively. $\alpha \in [0, 1]$ is the teleport probability.

During training, we derive a probability distribution over all communities $\mathbb{P}(s_{N+1}|v_i, \mathcal{G}_i^S)$, which indicates the most likely community for the next share of v_i . This requires both the current status of the sharing and the global information about \mathcal{G}_i^S . The current status can be represented using the latest posting event encoded in $s_{|\mathcal{G}_i^S|}$. For global information, we leverage soft-attention to derive β_i , the importance of each community in the posting sequence

$$\beta_i = \mathbf{w}_1^\top \sigma \left(\text{Linear} \left(s_{|\mathcal{G}_i^S|} \right) + (\text{Linear} (s_i)) \right), \quad (3)$$

where $\mathbf{w}_1 \in \mathbb{R}^d$ is trainable parameter. $\sigma(\cdot)$ is the sigmoid activation function.

Finally, we compute the probability by taking linear transformation over the concatenation:

$$\mathbb{P}(s_{N+1}|v_i, \mathcal{G}_i^S) = \text{Softmax}(\text{MLP} \left(\left[s_{|\mathcal{G}_i^S|} || \sum_{i=1}^n \beta_i s_i \right] \right) S_i), \quad (4)$$

where $||$ is the concatenation operand. $S_i = [s_1 || s_2 || \dots || s_{|\mathcal{G}_i^S|}]$ is the concatenation of all community embeddings in the sessions.

3.3 Video Content Modeling

Given a video, INPAC encodes its visual content into a low-dimensional feature vector. The content modeling component of INPAC can utilize a diverse range of encoders. Here, we note that online visual content is highly diverse in terms of topics, languages, and subject matter. Therefore, the titles, descriptions, and metadata of these videos such as channel information, can provide valuable insights into their content. These additional data sources can be leveraged to better categorize and understand the content of videos. We thus utilize the titles, descriptions, and channel information as the static features for each video. Specifically, inspired by the

success of pre-trained language models in natural language understanding [6, 12, 46, 108, 109], we encode the title and descriptions of each video v_i into a feature vector $\mathbf{v}_i \in \mathbb{R}^D$ based on a multilingual version of MiniLM [110]. Similarly, we encode each video's channel $c_{\rho(i)}$ into a feature vector $\mathbf{c}_{\rho(i)}$, where $\rho(\cdot) : V \rightarrow C$ maps each video to the channel that posts v_i . Then, the two feature vectors are aggregated into a joint representation

$$\tilde{\mathbf{v}}_i = \text{Aggr}(\mathbf{v}_i, \mathbf{c}_{\rho(i)}). \quad (5)$$

Here, a wide variety of aggregation schemes can be applied, including addition, concatenation, and element-wise multiplication, to obtain the joint representation. In Section 4.3, we investigate the impact of using different aggregation schemes for \mathbf{v}_i and $\mathbf{c}_{\rho(i)}$.

3.4 Dynamic Modeling

In the dynamic modeling component, INPAC models the temporal variability of each video's propagation on communities, obtaining temporal embedding of videos and communities. Here, we note that a video can be shared multiple times within a short amount of time [63]. Inspired by continuous-time dynamic graph (CTDG) [68, 71, 94, 129], we design a dynamic modeling module to provide a robust representation of the video sharing process and better handles the bursty nature of information sharing.

First, we leverage temporal graph network (TGN) [94] and represent our dynamic network as a pair (\mathcal{G}_0^D, E) where \mathcal{G}_0^D is the initial state of the dynamic network represented as a static graph. E is a set of graph events with timestamps. In INPAC, we consider two types of graph events, including node additions (*i.e.*, the emergence of new videos and communities) and edge additions (*i.e.*, a video is posted in an online community).

Input Encoding. The input embeddings $\mathbf{x}_i(t)$ and $\mathbf{x}_j(t)$ are raw feature representations for each video v_i and community s_j , respectively. We leverage the embeddings derived from Section 3.2-3.3 as the raw node embeddings. Namely, $\mathbf{x}_i(t) = \tilde{\mathbf{v}}_i$ for video v_i and $\mathbf{x}_j(t) = \mathbf{s}_j^{(L)}$ for community s_j , where $\mathbf{s}_j^{(L)}$ is the representation of s_j at the final layer in Equation 2.

Time Encoding. Similar to [94, 97, 118], the time encoding function $\phi(\cdot) : \mathbb{R} \rightarrow \mathbb{R}^d$ maps a continuous timestamp to the d -dimensional vector space:

$$\phi(t) = \cos(t\mathbf{w}_2 + \mathbf{b}_1), \quad (6)$$

where $\mathbf{w}_2, \mathbf{b}_1 \in \mathbb{R}^d$ are learnable parameters.

Temporal Memory. As in [94], to track the propagation state for each node, v_i or s_j , at any timestamp, there exists a memory vector, $\mathbf{h}_i(t)$ or $\mathbf{h}_j(t)$, to store history interactive memory in a compressed format. The memory of each node is initialized to zero and updated after each graph event. Given a node addition event of v_i , v_i 's message $\mathbf{m}_i^{\text{node}}(t)$ at time t is computed as the concatenation of i 's raw features and memory:

$$\mathbf{m}_i^{\text{node}}(t) = \text{MLP} \left([\mathbf{h}_i(t') || \mathbf{x}_i(t) || \phi(t)] \right), \quad (7)$$

where $\mathbf{h}_i(t')$ is v_i 's memory from time t' , *i.e.*, the time of the previous interaction involving v_i . In the same manner, we obtain each community s_j 's message $\mathbf{m}_j(t)$ at t given s_j 's event.

For an edge addition event involving v_i and s_j , the edge’s message $\mathbf{m}_i^{\text{edge}}(t)$ with respect to v_i at t is computed as:

$$\mathbf{m}_i^{\text{edge}}(t) = \text{MLP}([\mathbf{h}_i(t') || \mathbf{h}_j(t') || \mathbf{x}_i(t) || \mathbf{x}_j(t) || \phi(t)]). \quad (8)$$

Similarly, we can obtain the edge’s message $\mathbf{m}_j^{\text{edge}}(t)$ with respect to s_j at t .

During batch training, multiple events in the same batch can be associated with the same nodes. Therefore, we aggregate multiple messages of video v_i and community s_j from t_1 to t_B through mean pooling, thus obtaining $\overline{\mathbf{m}}_i(t)$ and $\overline{\mathbf{m}}_j(t)$ as in [94].

Based on these messages, the memory embeddings of v_i and s_j are updated upon each event involving v_i and s_j , respectively:

$$\mathbf{h}_i(t) = \text{GRU}(\overline{\mathbf{m}}_i(t), \mathbf{h}_i(t')), \quad (9)$$

$$\mathbf{h}_j(t) = \text{GRU}(\overline{\mathbf{m}}_j(t), \mathbf{h}_j(t')). \quad (10)$$

During prediction, we pass the representation $\mathbf{h}_i(t), \mathbf{h}_j(t)$ through multiple GNN layers to aggregate the features of each node from its neighbors on G^D

$$\tilde{\mathbf{v}}_i^t = f_\theta(\mathbf{h}_i(t), \mathcal{G}^D), \quad \tilde{\mathbf{s}}_j^t = f_\theta(\mathbf{h}_j(t), \mathcal{G}^D), \quad (11)$$

where $\tilde{\mathbf{v}}_i^t, \tilde{\mathbf{s}}_j^t$ are the transformed representation of v_i, s_j . The aggregation function $f_\theta(\cdot, \cdot)$ can be chosen from a wide range of GNN operators, such as GCN [53], GraphSAGE [31], Transformer-Conv [95], and GIN [119]. In practice, we employ a 2-layer Graph Attention Network (GAT) [102].

3.5 Training

We employ element-wise multiplication to calculate the score between each video v_i and each community s_j at time t :

$$\hat{y}_{ij}^t = \text{MLP}(\tilde{\mathbf{v}}_i^t \odot \text{MLP}(\tilde{\mathbf{s}}_j^t)), \quad (12)$$

where \hat{y}_{ij}^t is the predicted score between v_i and s_j . We train our model using the Bayesian Personalized Ranking (BPR) [92] loss, which encourages the prediction of an observed interaction to be greater than an unobserved one:

$$\mathcal{L}_{\text{BPR}} = \sum_{(i, j^+, j^-, t)} -\ln(\text{sigmoid}(\hat{y}_{ij^+}^t - \hat{y}_{ij^-}^t)), \quad (13)$$

where (i, j^+, j^-, t) denotes an example in the pairwise training data. j^+ indicates that one sharing of v_i is observed in community s_{j^+} , and j^- indicates an unobserved one.

Furthermore, for the training of the community influence graph, we use the next item prediction objective. Given each \mathcal{G}_i^S , the loss function $\mathcal{L}_{\text{CE}}^i$ is defined as the cross-entropy of the predicted and ground-truth community that will propagate v_i at the next timestamp:

$$\mathcal{L}_{\text{CE}}^i = \text{CrossEntropy}(\mathbb{P}(s_{N+1}|v_i, P_i), \mathbf{y}_{N+1}), \quad (14)$$

where $\mathbf{y}_{N+1} \in \mathbb{R}^{|S|}$ is a one-hot vector that encodes the ground-truth community interacted at the next timestamp.

The overall optimization objective is defined as follows:

$$\mathcal{L} = \mathcal{L}_{\text{BPR}} + \lambda_1 \sum_{i \in \mathcal{V}} \mathcal{L}_{\text{CE}}^i + \lambda_2 \|\Theta\|_2, \quad (15)$$

where Θ denotes all trainable model parameters. λ_1 and λ_2 are hyperparameters in INPAC.

4 EVALUATION

In this section, we conduct experiments to answer the following evaluation questions (EQs):

- **(EQ1)** Does INPAC outperform the baseline models for the task of community-level information pathway prediction (Section 4.2.1)?
- **(EQ2)** Does INPAC provide excellent inductive reasoning for cold-start videos (Section 4.2.2)?
- **(EQ3)** What is the contribution of each component in INPAC (Section 4.3)?
- **(EQ4)** Do community influence graphs (CIGs) constructed by INPAC manifest macroscopic influences (Section 4.4)?

4.1 Experimental Setup

4.1.1 Datasets. We construct 2 multi-modal datasets that provide the diffusion of YouTube videos on Reddit. Details can be found in Section 2.1. Table 1 provides an overview of their statistics. To partition the datasets into train/validation/test sets, we used a 70/15/15 ratio based on the timestamps in a sequential manner. To ensure validity, we constructed the community influence graphs (CIGs) exclusively using the interactions from the training set to prevent any potential information leakage.

4.1.2 Baselines. To evaluate the effectiveness of INPAC, we compare INPAC with seven baselines. We categorize them into four folds: (1) *Matrix Factorization*, including MF [92]; (2) *Graph-based Recommendation*, including NGCF [111], LightGCN [35], and SVD-GCN [88]; (3) *Sequential Recommendation*, including TiSASRec [65]; (4) *Representation Learning on Temporal Graphs*, including TGAT [118], and TGN [94].

4.1.3 Metrics. We measure the models’ performances using three widely adopted metrics in the field of ranking systems: (1) recall@ K , which measures the proportion of relevant items (*i.e.*, ground truth) that are retrieved among the top- K items; (2) normalized discounted cumulative gain (NDCG@ K), which evaluates the ranking quality of the top- K items, with a score of 1 assigned to the ideal ranking; (3) mean reciprocal rank (MRR), which computes the average reciprocal rank of the top-ranked relevant item. In this paper, we set K to 5 and 10. Our evaluation procedure follows the established method [19, 36, 65] by randomly selecting 100 communities with no observed propagations of the video and ranking the test item among the 100 items. Additionally, we exclude any existing interactions in the training set from the test set.

4.1.4 Implementation Details. We implemented INPAC in PyTorch [87] and PyG [20]. For a fair comparison, we set the embedding size to 64 in all methods including INPAC and perform Xavier initialization [24] on the model parameters. We use Adam optimizer [52] with a batch size of 256. For the baseline models, the hyperparameters are set to the optimal values as reported in the original paper. For all models, we search the learning rate within the range of $[1e-4, 3e-4, 1e-3, 3e-3, 1e-2]$ and select the best setting. We set $\alpha = 0.1$, $c = 3$, $\lambda_1 = 1$, and $\lambda_2 = 1e-3$, respectively. L , the number of GNN layers in Community Influence Modeling (Section 3.2), is set to 4.

Table 4: Performances of INPAC and 7 competitors for warm-start videos on the Large dataset. Values in bold and underline represent the best and the second best performances in each row, respectively. “Impr.” denotes the performance improvement of INPAC compared to the best baseline.

| | (a) Popular Subreddits | | | | | | | | |
|---------|----------------------------|--------------|--------------|--------------|--------------|--------------|---------------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 52.26 ± 0.52 | 52.97 ± 0.40 | 54.74 ± 0.55 | 56.85 ± 0.32 | 57.02 ± 0.44 | 56.85 ± 0.41 | 57.28 ± 0.57 | 60.19 ± 0.38 | 5.1% |
| Rec@5 | 71.98 ± 0.33 | 73.14 ± 0.25 | 73.56 ± 0.45 | 75.78 ± 0.40 | 76.00 ± 0.51 | 76.02 ± 0.58 | <u>76.08 ± 0.43</u> | 78.32 ± 0.47 | 3.0% |
| NDCG@10 | 60.15 ± 0.44 | 55.99 ± 0.54 | 56.81 ± 0.58 | 60.22 ± 0.49 | 60.34 ± 0.33 | 61.35 ± 0.31 | <u>61.46 ± 0.40</u> | 63.89 ± 0.39 | 4.0% |
| Rec@10 | 84.76 ± 0.23 | 84.82 ± 0.26 | 85.12 ± 0.49 | 85.19 ± 0.43 | 85.39 ± 0.38 | 85.34 ± 0.39 | <u>85.56 ± 0.35</u> | 87.98 ± 0.50 | 2.8% |
| MRR | 47.38 ± 0.51 | 52.52 ± 0.36 | 52.20 ± 0.35 | 53.85 ± 0.29 | 53.60 ± 0.46 | 53.58 ± 0.47 | <u>55.82 ± 0.53</u> | 58.27 ± 0.36 | 4.4% |
| | (b) Non-popular Subreddits | | | | | | | | |
| NDCG@5 | 13.54 ± 0.31 | 14.05 ± 0.48 | 15.42 ± 0.44 | 16.15 ± 0.55 | 16.72 ± 0.35 | 16.82 ± 0.43 | <u>16.89 ± 0.52</u> | 18.04 ± 0.56 | 6.8% |
| Rec@5 | 22.08 ± 0.38 | 22.19 ± 0.51 | 24.32 ± 0.38 | 25.44 ± 0.53 | 25.90 ± 0.36 | 25.95 ± 0.45 | <u>25.99 ± 0.44</u> | 27.46 ± 0.40 | 5.7% |
| NDCG@10 | 18.28 ± 0.59 | 18.50 ± 0.45 | 19.93 ± 0.50 | 20.70 ± 0.42 | 20.82 ± 0.39 | 21.26 ± 0.46 | <u>21.42 ± 0.57</u> | 22.67 ± 0.37 | 5.8% |
| Rec@10 | 36.03 ± 0.27 | 36.88 ± 0.38 | 38.39 ± 0.55 | 39.68 ± 0.46 | 39.62 ± 0.30 | 39.75 ± 0.31 | <u>39.76 ± 0.38</u> | 41.88 ± 0.52 | 5.3% |
| MRR | 15.17 ± 0.55 | 15.87 ± 0.57 | 16.96 ± 0.51 | 17.45 ± 0.58 | 17.75 ± 0.48 | 17.79 ± 0.49 | <u>18.22 ± 0.41</u> | 19.27 ± 0.55 | 5.8% |

Table 5: Performances of INPAC and 7 competitors for cold-start videos on the Large dataset.

| | (a) Popular Subreddits | | | | | | | | |
|---------|----------------------------|--------------|--------------|--------------|--------------|---------------------|---------------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 52.88 ± 0.43 | 56.28 ± 0.33 | 57.73 ± 0.45 | 58.07 ± 0.37 | 58.52 ± 0.39 | <u>58.95 ± 0.53</u> | 58.79 ± 0.38 | 61.85 ± 0.38 | 4.9% |
| Rec@5 | 73.55 ± 0.42 | 74.83 ± 0.34 | 75.41 ± 0.28 | 76.31 ± 0.46 | 76.02 ± 0.55 | <u>76.38 ± 0.57</u> | 76.35 ± 0.47 | 78.54 ± 0.53 | 2.8% |
| NDCG@10 | 56.73 ± 0.38 | 58.51 ± 0.52 | 60.13 ± 0.54 | 60.32 ± 0.37 | 60.22 ± 0.51 | <u>61.03 ± 0.57</u> | 60.92 ± 0.39 | 64.08 ± 0.44 | 5.0% |
| Rec@10 | 84.12 ± 0.56 | 83.95 ± 0.58 | 83.84 ± 0.34 | 84.06 ± 0.37 | 83.85 ± 0.59 | <u>84.23 ± 0.53</u> | 84.12 ± 0.49 | 86.89 ± 0.52 | 3.2% |
| MRR | 48.25 ± 0.45 | 51.17 ± 0.50 | 52.57 ± 0.47 | 53.33 ± 0.35 | 53.28 ± 0.32 | <u>54.88 ± 0.49</u> | 54.64 ± 0.54 | 57.58 ± 0.36 | 4.9% |
| | (b) Non-Popular Subreddits | | | | | | | | |
| NDCG@5 | 10.64 ± 0.36 | 13.79 ± 0.39 | 14.34 ± 0.37 | 14.83 ± 0.33 | 15.13 ± 0.37 | 15.84 ± 0.52 | 16.23 ± 0.32 | 17.61 ± 0.58 | 8.5% |
| Rec@5 | 15.92 ± 0.54 | 22.91 ± 0.57 | 25.24 ± 0.38 | 25.31 ± 0.37 | 25.38 ± 0.30 | <u>25.49 ± 0.45</u> | <u>25.46 ± 0.59</u> | 27.06 ± 0.40 | 6.2% |
| NDCG@10 | 13.98 ± 0.51 | 18.33 ± 0.38 | 19.51 ± 0.52 | 19.44 ± 0.55 | 19.89 ± 0.50 | <u>20.66 ± 0.44</u> | <u>20.79 ± 0.37</u> | 22.04 ± 0.45 | 6.0% |
| Rec@10 | 25.88 ± 0.41 | 37.30 ± 0.49 | 39.33 ± 0.41 | 39.70 ± 0.36 | 39.72 ± 0.54 | 40.06 ± 0.39 | <u>40.17 ± 0.57</u> | 42.37 ± 0.32 | 5.5% |
| MRR | 12.42 ± 0.55 | 15.06 ± 0.38 | 15.83 ± 0.55 | 16.94 ± 0.59 | 17.03 ± 0.38 | <u>17.58 ± 0.58</u> | 17.30 ± 0.52 | 18.64 ± 0.48 | 6.0% |

4.2 Overall Performances

We conducted comparative experiments on 2 datasets to demonstrate the superiority of INPAC over the 7 baselines. To this end, we grouped the videos into warm-start and cold-start videos. We define warm-start and cold-start videos as videos with ≥ 2 postings and 1 posting in the training phase, respectively. Furthermore, the number of videos posted in communities creates an imbalanced distribution. For instance, in the small dataset, more than 20% of videos were posted on the two most popular subreddits. Since it is relatively trivial to make predictions for such popular subreddits, we split subreddits into popular (*i.e.*, top 25 percentile subreddits where YouTube videos are posted most frequently) and non-popular (*i.e.*, the rest of the subreddits). The results are partitioned with respect to whether the target community is a popular subreddit or a non-popular subreddit.

4.2.1 Warm-Start Prediction. Tables 4(a)-(b) show the results for warm-start prediction for popular and non-popular subreddits, respectively, on the Large dataset. The results for the Small dataset can be found in Appendix B. We observe that INPAC consistently and significantly outperforms all baselines on both datasets for

both groups of subreddits. On the Large dataset, INPAC outperforms the best baseline by 5.1% on NDCG@5 and 4.4% on MRR for the popular communities, as well as 6.8% on NDCG@5 and 5.8% on MRR for non-popular communities, respectively. On the Small dataset, INPAC outperforms the best competitor by 8.6% and 7.5% on the two metrics for popular communities, and 12.9% and 18.8% for non-popular communities, respectively. Our results demonstrate the effectiveness of INPAC in the task of CLIPP. Moreover, we observe that representation learning methods on temporal graphs (*i.e.*, TGAT and TGN) outperform all other baselines. This observation underscores the importance of considering temporal information in predicting information pathways.

4.2.2 Cold-Start Prediction. As the content sharing network evolves, the emergence and spread of new content to a diverse range of communities presents considerable challenges for CLIPP, particularly in cold-start scenarios where historical propagation of videos is absent. Thus, the prediction problem becomes: *given a video that has only 1 propagation, how can we predict its second propagation?* Tables 5(a)-(b) show the performances of seven baselines and INPAC for popular and non-popular subreddits, respectively, on the Large dataset. The results for the Small dataset can be found in Appendix B. We observe that INPAC is able to achieve even greater

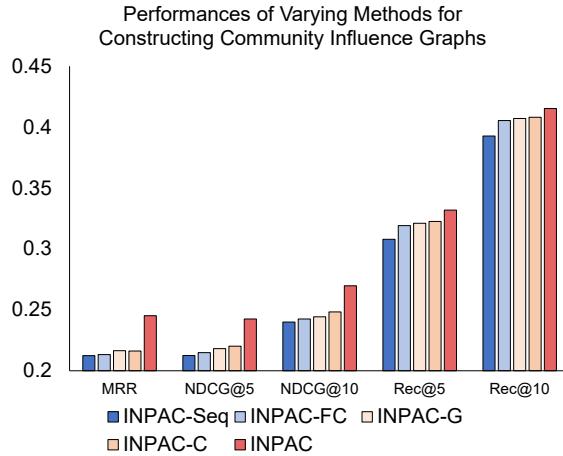


Figure 4: Performances of different methods for constructing the community influence graph (CIG) on the Small dataset.

performance improvements in the cold-start scenario through its inductive reasoning capability, consistently outperforming all competitors on both datasets for both groups of subreddits. Moreover, from Table 5(a), we observed that when the cold-start videos are propagated to popular communities, predicting these flows is relatively straightforward for all the models, including INPAC. On the other hand, the results in Table 5(b) show that predicting the flow of cold-start videos to less popular communities is a more challenging task. Despite this, INPAC still shows the best performance. These results encourage further investigation into such flows, which we consider to be a potential area of future work.

4.3 Ablation Studies

We validate the effectiveness of the design choices in INPAC. In Section 3.2, we designed a way to construct community influence graphs (CIGs) by considering the time that videos were propagated in communities. To evaluate this design, we made 4 variants of INPAC: INPAC-Seq connects the community nodes sequentially, *i.e.*, we create a directed edge from s_j to s_k if they are adjacent in the corresponding propagation sequence P_i . INPAC-FC establishes connections in a fully-connected manner, meaning that an edge is created between s_j and s_k if s_j precedes s_k in P_i . INPAC-G adopts the graph construction method of CIG as suggested by the GAINRec model [77]. INPAC-C omits any content information about the video and its channel. Specifically, the video embeddings v_i and the channel embedding $c_{p(i)}$ in Eq. (5) are randomly initialized.

From Figure 4, we observe that INPAC-Seq exhibits the lowest performances. This result can be attributed to the limitations of the sequential connection method, which fails to capture the underlying influencing relationships between communities as manifested by the sharing events. On the other hand, INPAC-FC performs better than INPAC-Seq in terms of Rec@5 and Rec@10. However, the fully-connected method can potentially lead to spurious correlations. Perhaps surprisingly, INPAC-C outperforms both INPAC-Seq, INPAC-FC, and INPAC-G on most metrics, suggesting that the model can still achieve remarkable performance in the absence of content features, given that the Community Influence Graphs

(CIGs) are properly constructed and modeled. This has broader implications for its applicability to other types of information with less available content, such as short online posts or URLs to misinformative websites. The superior performance of INPAC-C highlights the importance of the construction of the Community Influence Graph (CIG) in our approach. The CIG captures the interactions and influence patterns among communities, which is a crucial aspect when modeling the spread of information in online social networks. By focusing on the underlying social structures, our method is able to identify and predict the propagation of information more effectively than solely relying on content features. Overall, the method employed by INPAC achieves the best performances, demonstrating the effectiveness of our graph construction approach.

4.4 Analysis of CIG

Figure 5 presents the visualization of Community Influence Graphs (CIGs) for 4 videos with different topics (Section 3.2). Each video was propagated in exactly 20 communities. The node colors and sizes in the graphs represent the node degrees, while edge colors indicate the edge weights. We observe that CIGs generated from different videos demonstrate diverse connectivities and structures. We categorized the CIGs into two groups: (1) CIGs with multiple clusters, such as Figures 5(a)(c); and (2) CIGs with a single cluster, such as Figures 5(b)(d).

Regarding the CIGs with multiple clusters, we analyzed the differences between the clusters and the factors that contributed to the video spreading across different clusters. In Figures 5(a)(c), the videos were first posted in highly active communities. As the videos gained visibility over time, they spread to different clusters of communities. For instance, in Figure 5(a), the video was initially shared in r/AskScienceDiscussion, a community focused on in-depth scientific discussions, which aligned with the video's original purpose. Subsequently, as the video gained popularity, it was shared by distinct users in highly active COVID-19 related communities such as r/CoronavirusUS and r/China_Flu. Furthermore, the video also sheds light on the poor living conditions of animals in produce markets, where animals are confined in stacked cages and subjected to unsanitary conditions, evoking sympathy among viewers regarding animal welfare. As a result, the video was shared in 5 topically similar communities related to vegetarianism and animal welfare, including r/Vegan, r/VeganActivism, r/PlantBasedDiet, r/AnimalRights, and r/animalwelfare. In fact, the same group of users spread the video to multiple semantically similar communities potentially due to overlapping interests. Our INPAC model successfully models these correlated sharing behaviors as a 5-clique.

On the other hand, the CIGs in Figures 5(b)(d) exhibit a single cluster. We manually examined how these videos spread to communities with less obvious topical similarities. For instance, in Figure 5(d), the video first appeared in subreddits like r/WorshipTaylorSwift, a popular subreddits centered around the famous singer Taylor Swift, which directly relates to the posted video. Subsequently, the video propagated to multiple semantically distinct communities at different time periods. These communities included r/terracehouse, a subreddit about the reality TV show Terrace House, where users compared Taylor Swift's songs with the show's theme song and other famous singers' songs. Another example is r/NoStupidQuestions,

a subreddit for discussing a wide range of curious questions, where a user shared this video and questioned people’s obsession with Taylor Swift. Our key findings are as follows:

- Initially, online content tends to be shared within communities that closely match its topic. As the content gains popularity, it gradually spreads to multiple communities with a broader range of topics.
- Content is shared within topically similar communities in a short period, regardless of whether it is shared by the same user or different users. This observation aligns with previous studies [96] that found faster/slower information diffusion among topically similar/distant communities, respectively.
- There exist “Super spreaders” on online platforms who actively engage in and disseminate content across multiple topically diverse communities. For example, we identified a user who played a significant role in spreading the video in Figure 5(a) across vegetarian-related subreddits. This user has posted a total of 118 YouTube videos, with 67 shared in vegetarian-related communities. Another similar observation from Figure 5(c) is a user who actively contributed to communities about emotions, philosophy, Marvel Comics, and anime before eventually spreading the video among depression-related subreddits.

5 RELATED WORKS

5.1 Information Diffusion

Modeling the spread of information in online social networks has been a challenging task. Previous works have investigated information diffusion on social media [18, 25], prediction of popularity [3], social influence [61, 90], and topological analysis of follower networks [57, 98] for information sharing. While these studies cover a broad spectrum of social interactions in online communities, they generally focus on user-level influence and interactions. Research has shown that the dissemination of information within a community is different from that at the individual level [7, 81, 86, 130]. In this sense, diffusion models have been used to understand the spread of ideas, information and influence on social and information networks [62, 83]. Our study differs from the prior studies in its methodology as it endeavors to delve into the intricacies of community-level interactions.

5.2 Graph Neural Networks

Graph Neural Networks (GNNs) [11, 13–15, 53, 131, 132] have received increased attention in recent years due to their exceptional capacity to model complex, non-Euclidean graph structures. Recently, GNNs have achieved state-of-the-art performances in various applications, including recommendation [17, 32, 51, 54, 76, 112, 128], user modeling [28, 37, 127], and social network analysis [45, 61, 90, 126]. These methods typically structure events into interaction graphs and leverage high-order relationships to derive node/edge attributes [38, 43, 48, 59, 88]. Recently, dynamic graph models [49, 56, 69, 113, 118, 125] have emerged as powerful tools for various tasks, e.g., node classification, link prediction, and representation learning. The CLIPP problem can be modeled using dynamic networks [68, 70, 71] in which time-dependent representations of videos and communities are learned to infer future interactions.

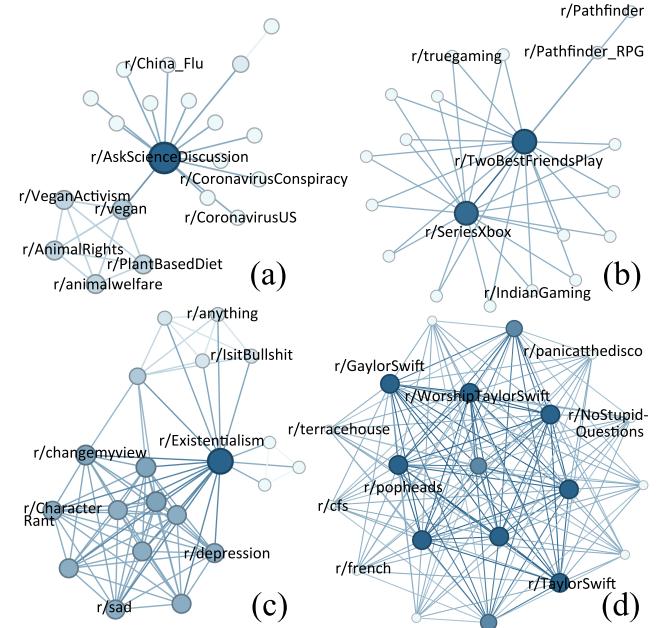


Figure 5: Community Influence Graphs (CIGs) of 4 different videos, all of which were propagated in exactly 20 communities. (a) How Wildlife Trade is Linked to Coronavirus; (b) Black Myth: Wukong - Official 13 Minutes Gameplay Trailer; (c) Thought experiment “BRAIN IN A VAT”; (d) Taylor Swift - ME! Node sizes and colors indicate the node degrees. Edge colors indicate the edge weights.

6 DISCUSSION AND CONCLUSION

Inference of community influence pathways can provide important information about the structure and dynamics of online platforms and the resulting information flow in the platform. This work created and utilized this influence graph in a dynamic graph framework INPAC to predict the flow of YouTube videos across Reddit communities (subreddits). Some shortcomings of this work include: (i) studying only YouTube-Reddit data and (ii) difficulty in the validation of the inferred influence graph. Future work includes alternate approaches to generate and validate influence graphs, creation of new dynamic graph models to predict information flow, and using multi-platform data.

ACKNOWLEDGMENTS

This research/material is based upon work supported in part by NSF grants CNS-2154118, IIS-2027689, ITE-2137724, ITE-2230692, CNS-2239879, Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00112290102 (subcontract No. PO70745), and funding from Microsoft, Google, and Adobe Inc. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the position or policy of DARPA, DoD, SRI International, NSF and no official endorsement should be inferred. We thank the reviewers for their comments.

REFERENCES

- [1] Wasim Ahmed, Josep Vidal-Alaball, Joseph Downing, Francesc López Seguí, et al. 2020. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. *Journal of Medical Internet Research* 22, 5 (2020), e19458.
- [2] Sergey Brin and Lawrence Page. 1998. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems* 30, 1-7 (1998), 107–117.
- [3] Qi Cao, Huawei Shen, Jinhua Gao, Bingzheng Wei, and Xueqi Cheng. 2020. Popularity prediction on social platforms with coupled graph neural networks. In *WSDM*. 70–78.
- [4] Shiyu Chang, Yang Zhang, Jiliang Tang, Dawei Yin, Yi Chang, Mark A Hasegawa-Johnson, and Thomas S Huang. 2016. Positive-Unlabeled Learning in Streaming Networks.. In *KDD*. 755–764.
- [5] Canyu Chen, Haoran Wang, Matthew Shapiro, Yunyu Xiao, Fei Wang, and Kai Shu. 2022. Combating Health Misinformation in Social Media: Characterization, Detection, Intervention, and Open Issues. *arXiv preprint arXiv:2211.05289* (2022).
- [6] Changyu Chen, Xiting Wang, Yiqiao Jin, Victor Ye Dong, Li Dong, Jie Cao, Yi Liu, and Rui Yan. 2023. Semi-Offline Reinforcement Learning for Optimized Text Generation. In *ICML*.
- [7] Jinyin Chen, Xiaodong Xu, Lihong Chen, Zhongyuan Ruan, Zhaoyan Ming, and Yi Liu. 2022. CTL-DIFF: Control Information Diffusion in Social Network by Structure Optimization. *IEEE Transactions on Computational Social Systems* (2022).
- [8] Tianwen Chen and Raymond Chi-Wing Wong. 2020. Handling information loss of graph neural networks for session-based recommendation. In *KDD*. 1172–1180.
- [9] Wanyu Chen, Fei Cai, Honghui Chen, and Maarten De Rijke. 2019. Joint neural collaborative filtering for recommender systems. *TOIS* 37, 4 (2019), 1–30.
- [10] Yixuan Chen, Dongsheng Li, Peng Zhang, Jie Sui, Qin Lv, Lu Tun, and Li Shang. 2022. Cross-modal ambiguity learning for multimodal fake news detection. In *TheWebConf*. 2897–2905.
- [11] Hejie Cui, Zijie Lu, Pan Li, and Carl Yang. 2022. On positional and structural node features for graph neural networks on non-attributed graphs. In *CIKM*. 3898–3902.
- [12] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *NAACL*. 4171–4186.
- [13] Yushun Dong, Jian Kang, Hanghang Tong, and Jundong Li. 2021. Individual fairness for graph neural networks: A ranking based approach. In *KDD*. 300–310.
- [14] Yushun Dong, Ninghao Liu, Brian Jalaian, and Jundong Li. 2022. Edits: Modeling and mitigating data bias for graph neural networks. In *TheWebConf*. 1259–1269.
- [15] Yushun Dong, Bilichi Zhang, Yiling Yuan, Na Zou, Qi Wang, and Jundong Li. 2023. Reliant: Fair knowledge distillation for graph neural networks. In *SDM*. 154–162.
- [16] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, and Sylvain Gelly. 2021. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In *ICLR*.
- [17] Jane Dwivedi-Yu, Yi-Chia Wang, Lijing Qin, Cristian Canton-Ferrer, and Alon Y Halevy. 2022. Affective Signals in a Social Media Recommender System. In *KDD*. 2831–2841.
- [18] Ahmed El-Kishky, Thomas Markovich, Serim Park, Chetan Verma, Baekjin Kim, Ramy Eskander, Yury Malkov, Frank Portman, Sofia Samaniego, Ying Xiao, et al. 2022. Twihin: Embedding the twitter heterogeneous information network for personalized recommendation. In *KDD*. 2842–2850.
- [19] Ali Mamdouh Elkahky, Yang Song, and Xiaodong He. 2015. A multi-view deep learning approach for cross domain user modeling in recommendation systems. In *TheWebConf*. 278–288.
- [20] Matthias Fey and Jan E. Lenssen. 2019. Fast Graph Representation Learning with PyTorch Geometric. In *ICLR Workshop on Representation Learning on Graphs and Manifolds*.
- [21] Casey Fiesler, Joshua McCann, Kyle Frye, Jed R Brubaker, et al. 2018. Reddit rules! characterizing an ecosystem of governance. In *ICWSM*.
- [22] Syeda Nadia Firdaus, Chen Ding, and Alireza Sadeghian. 2018. Retweet: A popular information diffusion mechanism—A survey paper. *Online Social Networks and Media* 6 (2018), 26–40.
- [23] Johannes Gasteiger, Aleksandar Bojchevski, and Stephan Günnemann. 2018. Predict then Propagate: Graph Neural Networks meet Personalized PageRank. In *ICLR*.
- [24] Xavier Glorot and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. In *AISTATS*. 249–256.
- [25] Manuel Gomez-Rodriguez, Jure Leskovec, and Andreas Krause. 2012. Inferring networks of diffusion and influence. *TKDD* 5, 4 (2012), 1–37.
- [26] Alex Graves and Alex Graves. 2012. Long short-term memory. *Supervised sequence labelling with recurrent neural networks* (2012), 37–45.
- [27] Anmol Gulati, James Qin, Chung-Cheng Chiu, Niki Parmar, Yu Zhang, Jiahui Yu, Wei Han, Shibo Wang, Zhengdong Zhang, Yonghui Wu, et al. 2020. Conformer: Convolution-augmented Transformer for Speech Recognition. *Interspeech* (2020), 5036–5040.
- [28] Jiayan Guo, Peiyan Zhang, Chaozhuo Li, Xing Xie, Yan Zhang, and Sunghun Kim. 2022. Evolutionary Preference Learning via Graph Nested GRU ODE for Session-based Recommendation. In *CIKM*. 624–634.
- [29] Lei Guo, Hongzhi Yin, Qinyong Wang, Tong Chen, Alexander Zhou, and Nguyen Quoc Viet Hung. 2019. Streaming session-based recommendation. In *KDD*. 1569–1577.
- [30] Nils Gustafsson. 2010. This time it's personal: Social networks, viral politics and identity management. In *Emerging practices in cyberspace and social networking*. Brill, 1–23.
- [31] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *NIPS* 30 (2017).
- [32] Junheng Hao, Tong Zhao, Jin Li, Xin Luna Dong, Christos Faloutsos, Yizhou Sun, and Wei Wang. 2020. P-companion: A principled framework for diversified complementary product recommendation. In *CIKM*. 2517–2524.
- [33] F Maxwell Harper and Joseph A Konstan. 2015. The movieLens datasets: History and context. *TIS* 5, 4 (2015), 1–19.
- [34] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *CVPR*. 770–778.
- [35] Xiangnan He, Kuai Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 2020. Lightgen: Simplifying and powering graph convolution network for recommendation. In *SIGIR*. 639–648.
- [36] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural collaborative filtering. In *TheWebConf*. 173–182.
- [37] Zijie Huang, Yizhou Sun, and Wei Wang. 2021. Coupled graph ode for learning interacting system dynamics. In *KDD*.
- [38] Bo Hui, Da Yan, Haiquan Chen, and Wei-Shinn Ku. 2021. Trajnet: A trajectory-based deep learning model for traffic prediction. In *KDD*. 716–724.
- [39] Dietmar Jannach and Malte Ludewig. 2017. When recurrent neural networks meet the neighborhood for session-based recommendation. In *RecSys*. 306–310.
- [40] Akshay Java, Xiaodan Song, Tim Finin, and Belle Tseng. 2009. Why we twitter: An analysis of a microblogging community. In *Advances in Web Mining and Web Usage Analysis: 9th International Workshop on Knowledge Discovery on the Web, WebKDD 2007, and 1st International Workshop on Social Networks Analysis, SNA-KDD 2007, San Jose, CA, USA, August 12–15, 2007. Revised Papers*. 118–138.
- [41] Maximilian Jenders, Gjergji Kasneci, and Felix Naumann. 2013. Analyzing and predicting viral tweets. In *TheWebConf*. 657–664.
- [42] Guoyin Jiang, Xiaodong Feng, Wenping Liu, and Xingjun Liu. 2020. Clicking position and user posting behavior in online review systems: A data-driven agent-based modeling approach. *Information Sciences* 512 (2020), 161–174.
- [43] Yiqiao Jin, Yunsheng Bai, Yanqiao Zhu, Yizhou Sun, and Wei Wang. 2023. Code Recommendation for Open Source Software Developers. In *WWW*.
- [44] Yiqiao Jin, Yunsheng Bai, Yanqiao Zhu, Yizhou Sun, and Wei Wang. 2023. Code Recommendation for Open Source Software Developers. In *TheWebConf*. 1324–1333.
- [45] Yiqiao Jin, Yeon-Chang Lee, Kartik Sharma, Meng Ye, Karan Sikka, Ajay Dikavaran, and Srijan Kumar. 2023. Predicting Information Pathways Across Online Communities. In *KDD*.
- [46] Yiqiao Jin, Xiting Wang, Yaru Hao, Yizhou Sun, and Xing Xie. 2023. Prototypical Fine-tuning: Towards Robust Performance Under Varying Data Sizes. *AAAI* (2023).
- [47] Yiqiao Jin, Xiting Wang, Ruichao Yang, Yizhou Sun, Wei Wang, Hao Liao, and Xing Xie. 2022. Towards fine-grained reasoning for fake news detection. In *AAAI*. Vol. 36. 5746–5754.
- [48] Jaehun Jung, Jinhong Jung, and U Kang. 2021. Learning to walk across time for interpretable temporal knowledge graph completion. In *KDD*. 786–795.
- [49] Seyed Mehran Kazemi, Rishabh Goel, Kshitij Jain, Ivan Kobyzev, Akshay Sethi, Peter Forsyth, and Pascal Poupart. 2020. Representation learning for dynamic graphs: A survey. *JMLR* 21, 1 (2020), 2648–2720.
- [50] Brian C Keegan. 2019. The Dynamics of Peer-Produced Political Information During the 2016 US Presidential Campaign. *HCI* 3, CSCW (2019), 1–20.
- [51] Taeri Kim, Yeon-Chang Lee, Kijung Shin, and Sang-Wook Kim. 2022. MARIO: Modality-Aware Attention and Modality-Preserving Decoders for Multimedia Recommendation. In *CIKM*. 993–1002.
- [52] Diederik P Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR*.
- [53] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. In *ICLR*.
- [54] Taeyong Kong, Taeri Kim, Jinsung Jeon, Jeongwan Choi, Yeon-Chang Lee, Noseong Park, and Sang-Wook Kim. 2022. Linear, or Non-Linear, That is the Question!. In *WSDM*. 517–525.
- [55] Srijan Kumar, William L Hamilton, Jure Leskovec, and Dan Jurafsky. 2018. Community interaction and conflict on the web. In *TheWebConf*. 933–943.
- [56] Srijan Kumar, Xikun Zhang, and Jure Leskovec. 2019. Predicting dynamic embedding trajectory in temporal interaction networks. In *KDD*. 1269–1278.
- [57] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. 2010. What is Twitter, a social network or a news media?. In *TheWebConf*. 591–600.
- [58] Yann LeCun, Yoshua Bengio, et al. 1995. Convolutional networks for images, speech, and time series. *The Handbook of Brain Theory and Neural Networks*

- 3361, 10 (1995), 1995.
- [59] Yeon-Chang Lee, JaeHyun Lee, Dongwon Lee, and Sang-Wook Kim. 2022. THOR: Self-Supervised Temporal Knowledge Graph Embedding via Three-Tower Graph Convolutional Networks. In *ICDM*. 1035–1040.
- [60] Jure Leskovec, Lars Backstrom, and Jon Kleinberg. 2009. Meme-tracking and the dynamics of the news cycle. In *KDD*. 497–506.
- [61] Carson K Leung, Alfredo Cuzzocrea, Jiaxing Jason Mai, Deyu Deng, and Fan Jiang. 2019. Personalized DeepInf: enhanced social influence prediction with deep learning and transfer learning. In *BigData*. 2871–2880.
- [62] Cheng Li, Jiaqi Ma, Xiaoxiao Guo, and Qiaozhu Mei. 2017. Deepcas: An end-to-end predictor of information cascades. In *TheWebConf*. 577–586.
- [63] Haitao Li, Xiaoqiang Ma, Feng Wang, Jiangchuan Liu, and Ke Xu. 2013. On popularity prediction of videos shared in online social networks. In *CIKM*. 169–178.
- [64] Jing Li, Pengjie Ren, Zhumin Chen, Zhaochun Ren, Tao Lian, and Jun Ma. 2017. Neural attentive session-based recommendation. In *CIKM*. 1419–1428.
- [65] Jiacheng Li, Yujie Wang, and Julian McAuley. 2020. Time interval aware self-attention for sequential recommendation. In *WSDM*. 322–330.
- [66] Zhiyuan Lin, Niloufar Salehi, Bowen Yao, Yiqi Chen, and Michael S Bernstein. 2017. Better when it was smaller? community content and behavior after massive growth. In *ICWSM*.
- [67] Chen Ling, Ihab AbuHilal, Jeremy Blackburn, Emiliano De Cristofaro, Savvas Zannettou, and Gianluca Stringhini. 2021. Dissecting the meme magic: Understanding indicators of virality in image memes. *ACM HCI 5, CSCW1* (2021), 1–24.
- [68] Meng Liu, Ke Liang, Bin Xiao, Sihang Zhou, Wenxuan Tu, Yue Liu, Xihong Yang, and Xinwang Liu. 2023. Self-Supervised Temporal Graph learning with Temporal and Structural Intensity Alignment. *arXiv:2302.07491* (2023).
- [69] Meng Liu, Ke Liang, Bin Xiao, Sihang Zhou, Wenxuan Tu, Yue Liu, Xihong Yang, and Xinwang Liu. 2023. Self-supervised temporal graph learning with temporal and structural intensity alignment. *arXiv:2302.07491* (2023).
- [70] Meng Liu and Yong Liu. 2021. Inductive representation learning in temporal networks via mining neighborhood and community influences. In *SIGIR*. 2202–2206.
- [71] Meng Liu, Yue Liu, Ke Liang, Siwei Wang, Sihang Zhou, and Xinwang Liu. 2023. Deep Temporal Graph Clustering. *arXiv:2305.10738* (2023).
- [72] Beth Logan et al. 2000. Mel frequency cepstral coefficients for music modeling. In *ISMIR*, Vol. 270. 11.
- [73] Malte Ludewig and Dietmar Jannach. 2018. Evaluation of session-based recommendation algorithms. *User Modeling and User-Adapted Interaction* 28 (2018), 331–390.
- [74] Jiachen Ma, Yong Liu, Meng Liu, and Meng Han. 2022. Curriculum Contrastive Learning for Fake News Detection. In *CIKM*. 4309–4313.
- [75] Yingchen Ma, Bing He, Nathan Subrahmanian, and Srijan Kumar. 2023. Characterizing and Predicting Social Correction on Twitter. In *WebSci*. 86–95.
- [76] Haitao Mao, Lixin Zou, Yujia Zheng, Jiliang Tang, Xiaokai Chu, Jiashu Zhao, and Dawei Yin. 2022. Whole Page Unbiased Learning to Rank. *arXiv preprint arXiv:2210.10718* (2022).
- [77] Qing Meng, Hui Yan, Bo Liu, Xiangguo Sun, Mingrui Hu, and Jiaxin Cao. 2023. Recognize News Transition from Collective Behavior for News Recommendation. *TOIS* 41, 4 (2023), 1–30.
- [78] Nicholas Micallef, Bing He, Srijan Kumar, Mustaque Ahamed, and Nasir Memon. 2020. The role of the crowd in countering misinformation: A case study of the COVID-19 infodemic. In *BigData*. 748–757.
- [79] Nicholas Micallef, Bing He, Srijan Kumar, Mustaque Ahamed, and Nasir D. Memon. 2020. The Role of the Crowd in Countering Misinformation: A Case Study of the COVID-19 Infodemic. *BigData* (2020), 748–757.
- [80] Nicholas Micallef, Marcelo Sandoval-Castañeda, Adir Cohen, Mustaque Ahamed, Srijan, Kumar, and Nasir D. Memon. 2022. Cross-Platform Multimodal Misinformation: Taxonomy, Characteristics and Detection for Textual Posts and Videos. In *ICWSM*.
- [81] Seth A Myers and Jure Leskovec. 2014. The bursty dynamics of the twitter information network. In *TheWebConf*. 913–924.
- [82] Usman Naseem, Adam G Dunn, Jinman Kim, and Matloob Khushki. 2022. Early identification of depression severity levels on reddit using ordinal classification. In *TheWebConf*. 2563–2572.
- [83] Lynnette Hui Xian Ng, Iain J Cruickshank, and Kathleen M Carley. 2022. Cross-platform information spread during the january 6th capitol riots. *Social Network Analysis and Mining* 12, 1 (2022), 133.
- [84] Sejoon Oh, Ankur Bhardwaj, Jongseok Han, Sungchul Kim, Ryan A Rossi, and Srijan Kumar. 2022. Implicit Session Contexts for Next-Item Recommendations. In *CIKM*. 4364–4368.
- [85] Maya Okawa and Tomoharu Iwata. 2022. Predicting Opinion Dynamics via Sociologically-Informed Neural Networks. *KDD* (2022).
- [86] Diogo Pacheco, Pik-Mai Hui, Christopher Torres-Lugo, Bao Tran Truong, Alessandro Flammini, and Filippo Menczer. 2021. Uncovering Coordinated Networks on Social Media: Methods and Case Studies. *ICWSM* 21 (2021), 455–466.
- [87] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *NIPS* 32.
- [88] Shaowen Peng, Kazunari Sugiyama, and Tsunenori Mine. 2022. SVD-GCN: A Simplified Graph Convolution Paradigm for Recommendation. In *CIKM*. 1625–1634.
- [89] Shruti Phadke, Mattia Samory, and Tanushree Mitra. 2021. What makes people join conspiracy communities? Role of social factors in conspiracy engagement. *HCI* 4, CSCW3 (2021), 1–30.
- [90] Jiezhang Qiu, Jian Tang, Hao Ma, Yuxiao Dong, Kuansan Wang, and Jie Tang. 2018. Deepinf: Social influence prediction with deep learning. In *CIKM*. 2110–2119.
- [91] Yiting Qu, Xinlei He, Shannon Pierson, Michael Backes, Yang Zhang, and Savvas Zannettou. 2022. On the Evolution of (Hateful) Memes by Means of Multimodal Contrastive Learning. In *2023 IEEE Symposium on Security and Privacy (SP)*. 1348–1365.
- [92] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *UAI*. 452–461.
- [93] Yu Rong, Qiankun Zhu, and Hong Cheng. 2016. A model-free approach to infer the diffusion network from event cascade. In *CIKM*. 1653–1662.
- [94] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. 2020. Temporal graph networks for deep learning on dynamic graphs. *arXiv:2006.10637* (2020).
- [95] Yunsheng Shi, Zhengji Huang, Shikun Feng, Hui Zhong, Wenjing Wang, and Yu Sun. 2021. Masked Label Prediction: Unified Message Passing Model for Semi-Supervised Classification. In *IJCAI*.
- [96] Jieun Shin, Lian Jian, Kevin Driscoll, and François Bar. 2018. The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior* 83 (2018), 278–287.
- [97] Chenguang Song, Kai Shu, and Bin Wu. 2021. Temporally evolving graph neural network for fake news detection. *Information Processing & Management* 58, 6 (2021), 102712.
- [98] Xiran Song, Jianxun Lian, Hong Huang, Mingqi Wu, Hai Jin, and Xing Xie. 2022. Friend Recommendations with Self-Rescaling Graph Neural Networks. In *KDD*. 3909–3919.
- [99] Statista. 2021. *Most popular social networks worldwide as of January 2023, ranked by number of monthly active users*. <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>
- [100] Tristan Sturm and Tom Albrecht. 2021. Constituent Covid-19 apocalypses: contagious conspiracism, 5G, and viral vaccinations. *Anthropology & medicine* 28, 1 (2021), 122–139.
- [101] Chenhao Tan. 2018. Tracing community genealogy: how new communities emerge from the old. In *ICWSM*.
- [102] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. In *ICLR*.
- [103] Gaurav Verma, Ankur Bhardwaj, Talayeh Ale davood, Munmun De Choudhury, and Srijan Kumar. 2022. Examining the impact of sharing COVID-19 misinformation online on mental health. *Scientific Reports* 12, 1 (2022), 1–9.
- [104] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018), 1146–1151.
- [105] Isaac Waller and Ashton Anderson. 2019. Generalists and specialists: Using community embeddings to quantify activity diversity in online platforms. In *TheWebConf*. 1954–1964.
- [106] Isaac Waller and Ashton Anderson. 2021. Quantifying social organization and political polarization in online platforms. *Nature* 600, 7888 (2021), 264–268.
- [107] Haoran Wang, Yingtong Dou, Canyu Chen, Lichao Sun, Philip S Yu, and Kai Shu. 2023. Attacking Fake News Detectors via Manipulating News Social Engagement. In *TheWebConf*. 3978–3986.
- [108] Weiqi Wang, Tianqing Fang, Wenxuan Ding, Baixuan Xu, Xin Liu, Yangqiu Song, and Antoine Bosselut. 2023. CAR: Conceptualization-Augmented Reasoner for Zero-Shot Commonsense Question Answering. *arXiv preprint arXiv:2305.14869* (2023).
- [109] Weiqi Wang, Tianqing Fang, Baixuan Xu, Chun Yi Louis Bo, Yangqiu Song, and Lei Chen. 2023. CAT: A contextualized conceptualization and instantiation framework for commonsense reasoning. In *ACL*.
- [110] Wenhui Wang, Furu Wei, Li Dong, Hangbo Bao, Nan Yang, and Ming Zhou. 2020. MiniLM: Deep self-attention distillation for task-agnostic compression of pre-trained transformers. *NIPS* 33 (2020), 5776–5788.
- [111] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In *SIGIR*. 165–174.
- [112] Xiting Wang, Kunpeng Liu, Dongjie Wang, Le Wu, Yanjie Fu, and Xing Xie. 2022. Multi-level recommendation reasoning over knowledge graphs with reinforcement learning. In *TheWebConf*. 2098–2108.
- [113] Yanbang Wang, Yen-Yu Chang, Yunyu Liu, Jure Leskovec, and Pan Li. 2021. Inductive representation learning in temporal networks via causal anonymous walks. *ICLR* (2021).

- [114] Wikipedia. 2023. *List of Most Visited Websites*. https://en.wikipedia.org/wiki/List_of_most_visited_websites
- [115] Junfei Wu, Qiang Liu, Weizhi Xu, and Shu Wu. 2022. Bias mitigation for evidence-aware fake news detection by causal intervention. In *SIGIR*. 2308–2313.
- [116] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. 2019. Session-based recommendation with graph neural networks. In *AAAI*, Vol. 33. 346–353.
- [117] Wenwen Xia, Yuchen Li, Jun Wu, and Shenghong Li. 2021. Deepis: Susceptibility estimation on social networks. In *WSDM*. 761–769.
- [118] Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, and Kannan Achan. 2020. Inductive representation learning on temporal graphs. *ICLR* (2020).
- [119] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2018. How Powerful are Graph Neural Networks?. In *ICLR*.
- [120] Sheng Xu, Yanjing Li, Teli Ma, Bohan Zeng, Baochang Zhang, Peng Gao, and Jinhui Lv. 2022. TerViT: An efficient ternary vision transformer. *arXiv:2201.08050* (2022).
- [121] Weizhi Xu, Junfei Wu, Qiang Liu, Shu Wu, and Liang Wang. 2022. Evidence-aware fake news detection with graph neural networks. In *TheWebConf*. 2501–2510.
- [122] Zhiheng Xu, Yang Zhang, Yao Wu, and Qing Yang. 2012. Modeling user posting behavior on social media. In *SIGIR*. 545–554.
- [123] Ruichao Yang, Jing Ma, Hongzhan Lin, and Wei Gao. 2022. A weakly supervised propagation model for rumor verification and stance detection with multiple instance learning. In *SIGIR*. 1761–1772.
- [124] Hongzhi Yin, Qinyong Wang, Kai Zheng, Zhixu Li, Jiali Yang, and Xiaofang Zhou. 2019. Social influence-based group representation learning for group recommendation. In *ICDE*. 566–577.
- [125] Hyunsik Yoo, Yeon-Chang Lee, Kijung Shin, and Sang-Wook Kim. 2023. Disentangling Degree-related Biases and Interest for Out-of-Distribution Generalized Directed Network Embedding. In *TheWebConf*. 231–239.
- [126] Jinghao Zhang, Yaqiao Zhu, Qiang Liu, Shu Wu, Shuhui Wang, and Liang Wang. 2021. Mining Latent Structures for Multimedia Recommendation. In *ACM MM*. 3872–3880.
- [127] Peiyuan Zhang, Jiayan Guo, Chaozhuo Li, Yueqi Xie, Jae Boum Kim, Yan Zhang, Xing Xie, Haohan Wang, and Sung hun Kim. 2023. Efficiently leveraging multi-level user intent for session-based recommendation via atten-mixer network. In *WSDM*. 168–176.
- [128] Peiyuan Zhang and Sung hun Kim. 2023. A Survey on Incremental Update for Neural Recommender Systems. *arXiv:2303.02851* (2023).
- [129] Peiyuan Zhang, Yuchen Yan, Chaozhuo Li, Senzhang Wang, Xing Xie, Guojie Song, and Sung hun Kim. 2023. Continual Learning on Dynamic Graphs via Parameter Isolation. *arXiv:2305.13825* (2023).
- [130] Fan Zhou, Xovee Xu, Kunpeng Zhang, Goce Trajcevski, and Ting Zhong. 2020. Variational information diffusion for probabilistic cascades prediction. In *IEEE INFOCOM*. 1618–1627.
- [131] Yaqiao Zhu, Weizhi Xu, Jinghao Zhang, Qiang Liu, Shu Wu, and Liang Wang. 2021. Deep graph structure learning for robust representations: A survey. *arXiv:2103.03036* (2021).
- [132] Yaqiao Zhu, Yichen Xu, Hejie Cui, Carl Yang, Qiang Liu, and Shu Wu. 2022. Structure-enhanced heterogeneous graph contrastive learning. In *SDM*. 82–90.
- [133] Caleb Ziems, Bing He, Sandeep Soni, and Srijan Kumar. 2020. Racism is a virus: anti-asian hate and counterspeech in social media during the COVID-19 crisis. *ASONAM* (2020).

A DISCUSSION

A.1 Difference between CLIPP and Recommendation Problems

A.1.1 Distinct underlying dynamics. In recommendation problems, the focus is on user behavior, as it largely reflects their interests, making it crucial to model user preferences accurately for precise recommendations. Group recommendation models [124] suggest items based on the combined preferences of users in a group, whereas sequential recommendation [44, 116] models concentrate on individual users' preferences and the extent to which item attributes align with those preferences.

On the other hand, the CLIPP problem encompasses a combination of factors that influence a user's decision to post a video within a community, where different users can share the video on different communities. An information-sharing event within a community is subject to factors such as user interests, community characteristics,

and the relationship between the community and the information being shared. For example, a piece of information can be posted in some online community due to the following reasons:

- Community members find the information valuable and wish to share it with other, driven by internal factors such as interest or altruism;
- Some users who originally do not belong to the community want to promote their product or service to a wider audience;
- Users with malicious intent seek to spread false or misleading information

A.1.2 The User Behaviors to be Modeled are Different. The goal of the proposed CLIPP problem requires simultaneous understanding of multiple users' behavior. One video can be shared by different users on different communities with completely different motivations. For example, a video v_1 can be shared in community s_1 by user u_1 with positive intent (e.g., promoting the video) while by another user u_2 in community s_2 with negative intent (e.g., criticizing the video). Yet, the goal is still to predict the next community on which the video will appear.

A.1.3 The Goals of the Problems are Different. The primary objective of the CLIPP problem is to model information flow across online communities rather than creating a recommender system. Although our proposed INPAC approach can be adapted for sequential recommendation, its primary focus is on capturing the complex interactions between users, communities, and information. Our experimental results (Tables 4–5) demonstrate that existing recommendation models were not designed to address the CLIPP problem and have inherent limitations when applied to it.

A.1.4 The Datasets are not Directly Transferable. As the first three points suggest, existing recommender system datasets, such as LastFM³, MovieLens⁴, and Goodreads⁵, are not directly applicable to solving the CLIPP problem, as they lack information about clearly defined online communities and the sharing of information across those online communities. This discrepancy highlights the need for distinct datasets that capture the complex dynamics specific to the CLIPP problem.

In summary, although there may be some overlap between the methods used in recommender systems and the CLIPP problem, they are fundamentally different problems that require distinct approaches to model the unique interactions between users, communities, and information sharing.

A.2 Extension to Other Types of Features

Our proposed framework can be extended to handle other complex types of information, such as images and audio. We outline the simple modifications required to accommodate these data formats. Specifically, in Section 3.3, we can add the appropriate encoders for image or audio instead of using the encoder for video content. Below are the potential encoders to use image and audio content

Image: To handle images, we can incorporate a variety of image encoders, such as CNN [58], ResNet [34], or Vision Transformer [16, 120], which will convert each input image into a D -dimensional

³<http://millionsongdataset.com/lastfm/>

⁴<https://grouplens.org/datasets/movielens/>

⁵<https://sites.google.com/eng.ucsd.edu/ucsdbookgraph/home>

Table 6: Performances on the Small dataset for warm-start videos

| (a) Popular Subreddits | | | | | | | | | |
|------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 35.92 ± 0.30 | 36.39 ± 0.56 | 37.95 ± 0.53 | 38.89 ± 0.33 | 39.17 ± 0.50 | 39.22 ± 0.46 | 40.33 ± 0.22 | 43.81 ± 0.29 | 8.6% |
| Rec@5 | 52.21 ± 0.48 | 53.01 ± 0.41 | 55.05 ± 0.52 | 56.50 ± 0.37 | 56.34 ± 0.50 | 56.79 ± 0.17 | 57.33 ± 0.30 | 60.89 ± 0.40 | 6.2% |
| NDCG@10 | 40.25 ± 0.79 | 40.91 ± 0.25 | 41.77 ± 0.36 | 42.45 ± 0.24 | 42.95 ± 0.39 | 42.65 ± 0.16 | 43.27 ± 0.25 | 46.11 ± 0.26 | 6.6% |
| Rec@10 | 65.83 ± 0.52 | 66.30 ± 0.66 | 68.21 ± 0.61 | 68.48 ± 0.37 | 68.39 ± 0.33 | 68.53 ± 0.31 | 68.60 ± 0.20 | 70.31 ± 0.59 | 2.5% |
| MRR | 33.47 ± 0.61 | 34.12 ± 0.31 | 34.71 ± 0.38 | 36.22 ± 0.29 | 36.55 ± 0.46 | 36.75 ± 0.29 | 37.45 ± 0.30 | 40.24 ± 0.44 | 7.5% |

| (b) Non-Popular Subreddits | | | | | | | | | |
|----------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 7.71 ± 0.31 | 8.14 ± 0.07 | 8.58 ± 0.23 | 9.50 ± 0.17 | 9.65 ± 0.22 | 9.64 ± 0.17 | 9.87 ± 0.03 | 11.14 ± 0.18 | 12.9% |
| Rec@5 | 12.00 ± 0.33 | 12.52 ± 0.18 | 13.16 ± 0.22 | 14.17 ± 0.11 | 14.38 ± 0.23 | 14.47 ± 0.18 | 14.58 ± 0.12 | 15.31 ± 0.17 | 5.0% |
| NDCG@10 | 9.93 ± 1.08 | 10.09 ± 0.13 | 11.96 ± 0.21 | 12.05 ± 0.13 | 12.37 ± 0.31 | 12.61 ± 0.32 | 12.98 ± 0.08 | 14.25 ± 0.31 | 9.8% |
| Rec@10 | 19.61 ± 0.46 | 18.21 ± 0.11 | 22.56 ± 0.31 | 23.16 ± 0.32 | 23.51 ± 0.31 | 23.27 ± 0.20 | 23.63 ± 0.10 | 25.28 ± 0.44 | 7.0% |
| MRR | 8.08 ± 0.24 | 9.09 ± 0.35 | 9.74 ± 0.20 | 10.20 ± 0.17 | 10.60 ± 0.15 | 11.11 ± 0.20 | 11.57 ± 0.16 | 13.75 ± 0.21 | 18.8% |

Table 7: Performances on the Small dataset for cold-start videos.

| (a) Popular Subreddits | | | | | | | | | |
|------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 34.57 ± 1.56 | 36.46 ± 0.17 | 39.24 ± 0.97 | 40.27 ± 1.29 | 41.33 ± 0.94 | 42.45 ± 1.16 | 42.66 ± 1.16 | 46.44 ± 1.28 | 8.9% |
| Rec@5 | 58.19 ± 1.85 | 58.51 ± 0.51 | 57.58 ± 0.17 | 60.60 ± 0.95 | 64.46 ± 0.90 | 67.07 ± 1.30 | 67.51 ± 1.27 | 71.53 ± 1.17 | 6.0% |
| NDCG@10 | 41.02 ± 0.82 | 43.63 ± 0.16 | 43.75 ± 0.99 | 44.74 ± 0.64 | 46.11 ± 1.10 | 47.77 ± 1.17 | 47.90 ± 1.08 | 50.81 ± 1.33 | 6.1% |
| Rec@10 | 84.98 ± 0.86 | 83.64 ± 0.41 | 85.53 ± 0.26 | 87.52 ± 0.89 | 88.13 ± 0.99 | 88.13 ± 1.14 | 88.19 ± 1.13 | 91.17 ± 1.41 | 3.4% |
| MRR | 29.65 ± 0.76 | 31.95 ± 0.46 | 31.93 ± 0.76 | 32.56 ± 0.76 | 36.55 ± 1.27 | 36.64 ± 1.24 | 36.89 ± 1.14 | 38.44 ± 1.27 | 4.2% |

| (b) Non-Popular Subreddits | | | | | | | | | |
|----------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------------|-------|
| | MF | NGCF | LightGCN | SVD-GCN | TiSASRec | TGAT | TGN | INPAC | Impr. |
| NDCG@5 | 7.23 ± 1.18 | 8.16 ± 0.38 | 8.04 ± 1.09 | 7.86 ± 0.95 | 8.45 ± 1.36 | 8.39 ± 1.15 | 8.83 ± 1.12 | 10.05 ± 1.20 | 13.8% |
| Rec@5 | 10.53 ± 1.02 | 12.30 ± 0.50 | 11.72 ± 0.38 | 13.41 ± 0.80 | 13.23 ± 1.28 | 13.92 ± 1.29 | 14.40 ± 1.14 | 15.23 ± 0.81 | 5.8% |
| NDCG@10 | 11.59 ± 1.76 | 11.12 ± 0.30 | 10.55 ± 0.90 | 11.27 ± 1.06 | 10.13 ± 0.32 | 11.50 ± 1.18 | 11.62 ± 1.16 | 12.75 ± 0.21 | 9.7% |
| Rec@10 | 20.40 ± 0.86 | 23.07 ± 0.34 | 22.78 ± 0.76 | 22.72 ± 1.42 | 23.38 ± 1.32 | 23.71 ± 1.21 | 23.90 ± 1.09 | 25.09 ± 1.20 | 5.0% |
| MRR | 8.61 ± 1.34 | 9.78 ± 0.24 | 10.98 ± 0.38 | 10.00 ± 0.43 | 9.30 ± 0.16 | 10.68 ± 1.21 | 10.97 ± 1.18 | 11.83 ± 1.12 | 7.8% |

feature vector. This vector can then be fed into our current model architecture as an input for community prediction.

Audio To accommodate audio data, there are several options for encoding audio into a D -dimensional representation, including MFCC-based models [72], LSTM [26], or Transformer-based models such as Conformer [27]. The choice of encoder would depend on the specific characteristics of the audio data, the acceptable level of computational costs, and the desired level of representation.

A.3 Rationale Behind Using Reddit Data

The rationale for focusing on YouTube videos on Reddit is:

- **Reddit:** We chose to study Reddit because Reddit is one of the largest global social platforms. It is ranked among the top 10 visited websites worldwide [114]. Most importantly, unlike many other social platforms, Reddit users form clearly defined community structures. These communities, also known as subreddits, are typically centered around specific topics or interests, such as music, politics, science, or gaming. The community-centric nature of Reddit makes it easy to analyze user group behavior and identify patterns of information sharing across different communities.
- **Cross-posting of YouTube videos on Reddit:** YouTube videos have previously been shown to be a major means of spreading misinformation on other platforms including Reddit [78].

YouTube is the second most popular platform in the world with 2.51 billion monthly active users [99], and is one of the most popular ways for users to consume online information.

- **Rich semantic information:** YouTube videos contain a wealth of textual and visual information that can help us develop a more comprehensive understanding of the content and its potential for spreading misinformation. This multimodal nature allows us to extract features from both visual and textual data.
- **Traceability of sharing patterns:** The sharing of YouTube videos on Reddit can be easily traced, enabling us to study the dissemination of misinformation across communities. In contrast, other types of information such as quotes and online memes [60, 67, 91], can be more difficult to track due to the evolution and modification of their content [60, 91].

B EVALUATION

B.1 Performances on the Small Dataset

Tables 6 and 7 present the results on the Small dataset for warm-start and cold-start videos, respectively. Remarkably, our INPAC model consistently surpasses all baseline methods with statistically significant improvements.