Joint Source Localization in Different Platforms via Implicit Propagation Characteristics of Similar Topics

Zhen Wang\(^{1,2}\), Dongpeng Hou\(^{1,2}\), Shu Yin\(^{4,2}\), Chao Gao\(^{2}\) and Xianghua Li\(^{2*}\)

\(^{1}\)School of Mechanical Engineering, Northwestern Polytechnical University
\(^{2}\)School of Artificial Intelligence, OPtics and ElectroNics (iOPEN), Northwestern Polytechnical University
\(^{3}\)School of Cybersecurity, Northwestern Polytechnical University
\(^{4}\)School of Computer Science, Northwestern Polytechnical University

li_xianghua@nwpu.edu.cn

Abstract

Different social media are widely used in our daily lives. Inspired by the fact that similar topics have similar propagation characteristics, we mine the implicit knowledge of cascades with similar topics from different platforms to enhance the localization performance for scenarios where limited propagation data leads to the weak learning ability of existing localization models. In this work, we first construct a multiple platform propagation cascade dataset, aligning similar topics from both Twitter and Weibo, and enriching it with user profiles. Leveraging this dataset, we propose a Dual-channel Source Localization Framework (DSLF) for the joint cascades with similar topics. Specifically, a self-loop attention based graph convolutional network is designed to adaptively adjust the neighborhood aggregation scheme of different users with heterogeneous features in the message-passing process. Additionally, a dual-structure based Kullback-Leibler (KL) regularization module is proposed to constrain the latent distribution space of the source probabilities of similar characteristic-level users for a similar topic, enhancing the robustness of the model. Extensive experiments across Twitter and Weibo platforms demonstrate the superiority of the proposed DSLF over the SOTA methods. The code is available at https://github.com/cgao-comp/DSLF.

1 Introduction

In today’s digital age, the widespread usage of social media platforms like Weibo and Twitter has dramatically changed the spreading way of information [Ji et al., 2017]. However, the appearance of malicious fake news or sensationalized topics on multiple platforms presents major challenges to the economy and society [Yin et al., 2024; Islam et al., 2020; Wang et al., 2022b]. Focusing on their features and locating the sources are key steps in controlling abnormal spread.

Snapshot based and sensor based source identification methods are developed to locate the sources in the different scenarios [Jiang et al., 2016; Jin and Wu, 2021]. Since the sensor deployment process requires time and space overhead [Paluch et al., 2020a], more and more studies focus on the snapshot branch due to the accessibility and low cost of snapshot acquiring. The captured snapshot provides a set of users who participated in a topic. Then the inference strategy is proposed to identify the first participant in the snapshot.

Considering the complexity of propagation dynamics, most existing snapshot-based methods require abundant cascades to achieve an acceptable performance. A network with fewer than 200 nodes typically needs thousands of cascades to provide reasonable localization performance [Dong et al., 2019]. In practical applications, however, the amount of data is usually insufficient. For example, only a few thousand observed cascades are captured in a million-scale Weibo network [Ma et al., 2017]. Consequently, accurately locating the source using limited data in real-world scenarios becomes a significant challenge. Fortunately, we have identified that there exist potential similar features between cascades with similar topics in different platforms [Hou et al., 2024b]. And the localization performance can be enhanced by leveraging this experience in source inference. However, most snapshot-based methods focus on a single-layer network localization [Hou et al., 2023; Ling et al., 2022], thereby limiting their applicabil-
ity in joint localization scenarios. Although several localization methods for multiplex networks are proposed [Paluch et al., 2020b], their required prior information, such as spreading rates between two layers, is challenging to obtain.

To address these issues, our work seeks to bridge the ‘simulation-to-reality’ gap by employing real-world data collected from Twitter and Weibo for joint source localization. As can be seen from Fig. 1, first, different real propagation cascade data of similar topics from Weibo and Twitter is aligned based on the comments, and the user-specific profiles are injected and aligned with the nodes within the cascades based on the unique user identification (UID) of each node in the datasets. Subsequently, a localization method tailored for dual-platform similar topics is developed. It’s noteworthy that the propagation way of information is not confined to online platforms alone. The offline behaviors and external knowledge can also greatly affect intra-layer and inter-layer spreading rates. However, the prior information is costly to make the statistics. Therefore, to ensure the scalability of our proposed localization framework, we do not leverage external experience or prior knowledge and only use observed snapshot information in each platform to infer the source.

Based on two sets of snapshots at available timestamps observed from Weibo and Twitter, respectively, we propose a Dual-channel Source Localization Framework (DSLF) based on the regularized constraint for cascades with similar topics in different platforms. Specifically, eight essential propagation features is constructed for each user at every timestamp. These are further integrated with seven unique user profile features, such as tweet count and fan numbers, to create a comprehensive representation for each user. Then a self-loop attention based GCN is designed to dynamically assess and adapt to the varying importance of heterogeneous users in the information aggregation process of GCN instead of fixed coefficients. Moreover, a variational autoencoder is employed for each user at every timestamp to learn the latent distribution space of becoming a source, allowing us to develop a dual-structure based Kullback-Leibler (KL) regularized constraint module, which realizes the practical experience that the behavior and determination of users with analogous behavioral characteristics regarding similar topics are likely to stem from approximation distribution spaces. To further refine our model, a composite loss function is incorporated to enable DSLF to capture and learn the propagation dynamics in dual platforms more accurately. The major contributions of DSLF are as follows.

- A topic-aligned propagation cascade dataset incorporating user profiles for Twitter and Weibo is developed to bridge the ‘simulation-to-reality’ gap in source localization research. And a dual-channel source localization framework (DSLF) for cascades with similar topics in different platforms is proposed, validating its practicality in real-world multiplex network scenarios. Extensive experiments further demonstrate that DSLF can be flexibly extendable to a single or more layer network.

- A self-loop attention-based GCN is designed to refine the neighborhood aggregation scheme, learning a higher-quality representation of graph structures and user profiles, which improves the source detection performance in real-world scenarios.

- A dual-structure based KL regularization is designed to focus on implicit joint features in the latent distribution space of the source probability between platforms. By realizing the alignment of behavioral distribution characteristics for users with similar profile levels about similar topics, the robustness of the model can be enhanced.

2 Related Work

2.1 Diffusion Models

In the last few years, many diffusion models have been proposed to capture the diffusion feature and simulate the propagation data which is applied for evaluating the performance of localization methods, such as the Susceptible-Infect (SI) model [Yang et al., 2020; Paluch et al., 2021; Zang et al., 2015; Zhu and Ying, 2014; Tang et al., 2018]. Further, recognizing that in reality, every individual in social networks has unique attributes, heterogeneous diffusion models like the heterogeneous SI (HSI) and heterogeneous SIR (HSIR) are proposed to consider varied infection and recovery rates respectively [Karrer and Newman, 2010; Ellison, 2020]. However, these models are still fundamentally based on the mean-field theory or assumptions of propagation dynamics. Therefore, it is necessary to utilize real-world propagation cascades to enhance the performance demonstration and expand the application scenarios of downstream localization methods.

2.2 Source Localization Methods

Due to the convenience and feasibility of snapshot acquisition, many works focus on snapshot based study. Dong et al. introduce a GCN based source identification model to tackle the multiple rumor source detection problem [Dong et al., 2019]. Additionally, some methods construct the dynamic features of propagation before embarking on the source inference process, such as IVGD [Wang et al., 2022a], MCGNN [Shu et al., 2021] and SL_VAE [Ling et al., 2022]. However, these methods do not consider the cross-platform nature of propagation. Although some works focus on source identification in multiplex networks [Paluch et al., 2020b], some parameters of the underlying propagation model required as inputs in the inference are quite difficult to obtain in reality. What’s more, these methods, grounded in simulated data for localization, do not consider the impact of real-world user profiles on propagation. In reality, individuals are the primary drivers of information diffusion. Therefore, distinguishing users while differentiating their characteristics in the source inference process can enhance the applicability in practical scenarios.

3 Method

3.1 Preliminary

Propagation Cascades

Focused on some specific topics or topics, we obtain $K_1$ number of available experienced historical propagation cascades $C^k_{1i}=(X^k_{1i},E^k_{1i},F^k_{1i})$ ($1 \leq k \leq K_1$) from platform Twitter,
and $K_2$ historical cascades $C_k^t=(V_k^t, E_k^t, F_k^t)$ from platform Weibo, where $V_k^t$ is the participant user set with UID in a social media platform, $E_k^t$ is the set of participant’s directed propagation interaction (including comments or retweets from a user to another), and $F_k^t$ is the feature set (i.e., profiles) associated with the users, including verification status (indicating authoritativeness), number of tweets (representing tweet activity), registration date (indicative of user seniority), number of followers (indicating popularity), number of followings (indicating information seeking behavior), ratio of fans to followings (reflecting credibility).

**Historical Relationship Network**

Drawing from $K^*$ historical cascades $C_k^t=(V_k^t, E_k^t, F_k^t)$ in a social media platform $\star$, we construct the historical relationship network $G^t=(V^t, E^t, F^t)$, which is a union graph by combining structural information of different cascades based on the same UIDs. Sincerely, we pick this idea from the field of diffusion inference [Ramezani et al., 2023], where it is widely used as an intuitive yet effective approach when the underlying network is unknown. Specifically, if different cascades are interconnected by the presence of the same UID, it often implies that these cascades are not isolated incidents but indicate potential underlying similarities or connections driven by consistent shared interests or themes. Therefore, a connected historical relationship network represents a group of people (i.e., community) characterized by topic relevance, friendship ties, or similar attributes, even though these may not be explicitly known. Focusing on this distinct identified area, our investigative effort is to locate the sources from a new propagation within $G^t$.

**Problem Definition**

Having constructed the historical relationship network $G^t=(V^t, E^t, F^t)$ in Twitter and $G^t=(V^t, E^t, F^t)$ in Weibo, as for a new propagation cascade of a focal topic or topic spreading in the concerned area $G^t_{P^t}$ and $G^t_{P^t}$, denoted as $C_{P^t}$ and $C_{P^t}$, respectively, we only conveniently and randomly observe two collections of available snapshots $\{V_{t_1}, V_{t_2}, ..., V_{t_n}\}$ and $\{V_{t_1}', V_{t_2}', ..., V_{t_n}'\}$, where $V_{t_i}^t$ is the set of participants in area $G^t$ at a timestamp $t_i$. And we denote the original rumor sources set as $\mathcal{R} \subseteq G^t$. The goal of our method is to predict a source set $\hat{R}$ which can maximize the indicator like $\frac{R^\hat{R}}{R^R}$.

### 3.2 Overall Framework

We have temporal data for each user at each timestamp in each platform. However, while the recurrent module typically processes sequence data, it is also expected to learn network topology information inductively. Hence in a single platform $\star$, before inputting temporal data in a sequential model, it is required to incorporate graph information. Considering the excellent embedding ability of GNNs on graph topology, an intuitive and feasible way is to design a graph-based recurrent layer in the update process. A simplified, vectorized version of the graph-based recurrent layer can be specified as $h_{t_i} = \sigma \left( A^t h_{t_{i-1}} + b_{t_i} + A h_{t_{i-1}} W_{t_i} h + b_{t_i} \right)$, where $\sigma$ is the activation function, $h_{t_i}$ is the hidden layer output for all nodes at a timestamp $t_i$, $A$ is a variant matrix of a certain topological structure. $W_{t_i}$ and $b_{t_i}$ are the learnable input-hidden and hidden-hidden weights, $b_{t_i}$ and $b_{t_i}$ are the corresponding bias. Based on this, we develop an encoder-decoder framework tailored for enhancing temporal data through graph-based feature aggregation, called dual-channel source localization framework (DSL). As illustrated in Fig. 2, after constructing unique characteristics for each user in each snapshot, a self-loop attention based GCN encoder primarily aggregates topology information into temporal data. Further, a dual-structure based KL regularization decoder is designed to sequence and regularize the augmented temporal data from a multi-channel perspective. Finally, a joint loss is designed to optimize the overall performance by enhancing the single-channel specificity and ensuring robust multi-channel alignment.

### 3.3 Features Construction

For a series of propagation snapshots $(V_{t_1}, V_{t_2}, ...)$ of a topic, historical relationships in $G^t$ is used to extract a knowledge based subgraph $G_{t_i}^*$ or adjacency $A_{t_i}^*$ in each timestamp. To ensure the structural consistency of each $A_{t_i}^*$ across different timestamps which is a preprocessing idea for the deep learning module, and to perceive the future potential participants from the experience in $G^t$, first, we unify the largest subgraph $G^t(V^*, E^*)$ or $A^*$ from the focused area $G^t$ by extracting additional one-hop relationships of $V^*_m$.

$$V^*_m = \{ u \mid v \in V^* \} \cup \{ u \mid u \in N^* (v) \} \cup \{ v \mid v \in V^* \},$$

(1)

$$E^* = \{ (v_i, v_j) \mid v_i, v_j \in V^* \} \cup \{ (v_i, v_j) \mid v_i, v_j \in E^* \},$$

(2)

where $N^* (v)$ is the neighbor set of user $v$ in a single historical network $G^t$. Then, a single snapshot $V_{t_i}^*$ can be
mapped onto \( \hat{G}^* \) and is denoted as \( \hat{G}^*_t (\hat{V}_t^*, \hat{E}_t^*, \hat{Y}_t^*) \). Here, \( \hat{Y}_t^*(v_j) = 1 \) indicates that a user \( v_j \) has participated our concerned topic in platform \( * \) at timestamp \( t_i \). Simultaneously, the state and characteristics of users in each timestamp are retained and the topological consistency of different timestamps is also guaranteed (i.e., \(|\hat{V}_t^*| = |\hat{V}_{t_i}^*| = |\hat{V}_{t_i}^*| = |\hat{V}^*|\)).

After obtaining a series of snapshot based subgraphs with topological consistency \( G_{t_1}, \hat{G}_{t_2}, \ldots \), a sequence-to-sequence model is designed to locate the source user of a topic propagation. First in the encoder phase, to better solve the user-level-based source localization task, some unique propagation dynamic features are designed for each user combined with user profiles to differentiate each unique user. There are eight explicit dynamic indicators at a timestamp \( t_i \) (denoted as \( H_{1}^{t_i}, H_{8}^{t_i} \)) are constructed to characterize the time-varying based dynamic features of an individual at a timestamp \( t_i \) in a single platform \( * \). Among them, the ratio of participated neighbors and non-participated neighbors of \( v_j \) at a timestamp \( t_i \) are shown in Eq. (3) and Eq. (4), respectively.

\[
H_1^{t_i}(v_j) = \frac{\sum_{v_k \in N^{G_{t_i}}(v_j)} Y_{t_i}^*(v_k)}{|N^{G_{t_i}}(v_j)|}, \tag{3}
\]
\[
H_2^{t_i}(v_j) = 1 - H_1^{t_i}(v_j). \tag{4}
\]

Additionally, we also pay attention to the degree centrality \( \hat{G}_{t_i}^* \) of each user only has one neighbor and the neighbor participates in the state and characteristics of users in each timestamp are retained and the topological consistency of different timestamps is also guaranteed (i.e., \(|\hat{V}_t^*| = |\hat{V}_{t_i}^*| = |\hat{V}_{t_i}^*| = |\hat{V}^*|\)).

\[
H_3^{t_i}(v_j) = \frac{\sum_{v_k \in N^{G_{t_i}}(v_j)} Y_{t_i}^*(v_k)}{\max_{u \in \hat{V}_t^*} \left( N^{G_{t_i}}(u) \right)}, \tag{5}
\]
\[
H_4^{t_i}(v_j) = \frac{|N^{G_{t_i}}(v_j)| - \sum_{v_k \in N^{G_{t_i}}(v_j)} Y_{t_i}^*(v_k)}{\max_{u \in \hat{V}_t^*} \left( N^{G_{t_i}}(u) \right)}. \tag{6}
\]

Here, features \( H_1-H_4 \) indicate that we are not solely focused on the dynamic ratio of neighbor users. Both the number of participated and non-participated neighbors emphasize our concern for the precise count of neighbors’ states, not just their proportions. For example, considering that a user only has one neighbor and the neighbor participates in the topic, then \( H_1 \) is a relatively large feature indicator. However, indicator \( H_3 \) of such a user is small. Therefore, both normalized numerical features and proportional features need to be considered. Moreover, the original state \( Y_{t_i}^*(v_j) \) of each user in \( \hat{G}_{t_i}^* \), whether participated \( (H_3) \) or non-participated \( (H_0) \), collectively represents the essential property of the individual. Furthermore, the normalized snapshot sequence index \( (H_7) \) is introduced to indicate the relative timing of each user’s first appearance in the sequence of captured snapshots. Additionally, we also pay attention to the degree centrality \( (H_8) \) which can reflect the celebrity effect in social networks. After these eight propagation dynamic features are obtained, the complete feature embedding \( H_1^{*}(v_j) \in \mathbb{R}^{15} \) at a timestamp \( t_i \) in a single platform \( * \) can be obtained by concatenating the six normalized user profile features in \( \mathcal{F}(v_j) \) (verification status is a binary variable with two dimensions).

### 3.4 Self-loop Attention Based GCN

Since we have the features of users at different timestamps, we can input the temporal information into a recurrent model in a sequence-input-oriented way. However, direct input may result in the loss of prior information about the topological structure. Therefore, we can utilize the topology-based aggregation properties of GNNs. And in each platform, the original GCN at a timestamp \( t_i \) for feature aggregation of \( H_1^{*} \in \mathbb{R}^{|\hat{V}_t^*| \times 15} \) in the spectrum domain convolutions is defined as follows:

\[
H_1^{*} \leftarrow \sigma \left( \hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H_1^{*} W \right), \tag{7}
\]

where \( W \) are the learnable weights in the module. \( \hat{A} = A + I \), where \( A \) is the corresponding adjacency matrix of \( \hat{G}_{t_i}^* \) and \( I \) is an identity matrix. \( \hat{D} \) is the corresponding degree matrix of \( \hat{A} \). However, in the aggregation process of single-layer GCN, we observe that nodes with higher degrees, despite having a larger number of interacting neighbors, often contribute with relatively lower feature weights in the aggregation, impacting both themselves and their neighboring nodes. More details can be seen from Fig. 3. This tendency highlights a nuanced challenge within the single-layer GCN module, where the influence of highly connected nodes might be diminished in the aggregation process. Although as the number of GCN layers increases, the influence of a node is proportional to its degree, the limited extent to which GCNs with an appropriate number of layers enhance the aggregation of features for high-degree nodes may not sufficiently address the required focus on the importance of celebrities. This conclusion is revealed from our comprehensive statistical analysis [Hou et al., 2024b] for propagation cascades datasets in Twitter and Weibo, that is, high-influence users in social networks, such as celebrities with large fans, often engage in topics that attract a vast number of directed participants and have a great impact on these neighbors. In other words, more attention than that of the original GCN with several layers is demanded to enhance the impact of celebrities, and contrarily, the impact of fringe users can be weakened.

Manually setting the coefficient of each user through numerous experiments is a time-consuming process. Therefore, we propose a self-loop attention based GCN to revise the coefficient weight of propagation dynamic features and profiles \( H_1^{*} (v_j) \) of each user during the process of information aggregation. In this way, celebrities can weaken the average of feature influence by its neighbors, so as to better reflect the real-world level of influence of their characteristics during the aggregation process. As shown in Eq. (8), we add a learnable diagonal matrix to personalize the element value on the diagonal of the matrix \( \Lambda \in \mathbb{R}^{|\hat{V}_t^*| \times |\hat{V}_t^*|} \).

\[
H_1^{*} \leftarrow \sigma \left[ \left( \hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} + \Lambda \right) H_1^{*} W \right]. \tag{8}
\]

The diagonal elements of this matrix are controlled by a multi-head attention mechanism. And a single-layer BP neural network \( \tilde{a} \in \mathbb{R}^{15} \) is applied for each head of the attention
of the hidden layer at the timestamp $t_i$ as $h_{\tilde{t}_i} \in \mathbb{R}^{|V_i| \times 2}$. Similarly, the backward output is represented as $\tilde{h}^{t_i}_{\tilde{t}_i} \in \mathbb{R}^{|V_i| \times 4}$.

Finally, the two hidden layer states of the timestamp $t_i$ are concatenated, i.e., $\tilde{O}^*_i = \left[ h^{t_i}_{\tilde{t}_i}, \tilde{h}^{t_i}_{\tilde{t}_i} \right] \in \mathbb{R}^{|V_i| \times 4}$.

### 3.6 E-VAE for User’s Latent Distribution Space

Having identified the source features on a single platform from $O^*_i$ in a coarse-grained way, it is important to note that even when appearing in different platforms, certain characteristics of a similar topic, such as the behaviors of participants involved and their attitudes towards the topic, tend to remain potentially consistent. These invariant features provide valuable insights into improving the detection performance and robustness of source localization. Based on the above considerations, we start from the perspective of Weibo and Twitter data and propose a dual-structure based KL regularization to capture the consistent implicit features of the cascades with similar topics in different platforms, which not only identifies the unique factors of each platform but also bridges all channels to reveal a comprehensive alignment of user behavior and decision-making outcomes.

In practice, users with equivalent behavioral characteristics, influence, or profiles, are likely to exhibit decision-making and actions stemming from approximate distribution spaces when dealing with the same topics. Therefore, the source probability of similar users across different platforms, as a reflection of their decision outcomes, is also highly likely to originate from similar distribution spaces. Drawing insights from the phenomenon analysis of the propagation cascades [Hou et al., 2024b; Hou et al., 2024a], we have developed an exponential variational autoencoder (E-VAE) instead of normal distribution sampling.

$$\log q_{\phi}(\mathbf{z} | \tilde{O}^*_i(v_j)) = \log \text{Exp} \left( \mathbf{z}; \lambda^{(v_j)} \right),$$  

where $q_{\phi}(\mathbf{z} | \tilde{O}^*_i(v_j))$ represents the approximate posterior distribution of the latent variable $\mathbf{z}$ parameterized by $\phi$, conditioned on the source probability evaluation of a user $v_j$ at a timestamp $t_i$ in a single platform $*$. Eq. (10) implies that the distribution of the latent variable $\mathbf{z}$ is inferred from the data output by the sequence model. The sequence LSTM model encapsulates temporal dependencies, and then the coarse-grained source features predicted by LSTM are leveraged to guide the VAE framework in learning the latent distribution spaces of user’s decision-making outcomes. In this case, samples are drawn from an exponential distribution. Our goal is to model the latent variables $\mathbf{z}$ by using an exponential distribution when an observed data point user $v_j$ is given. Therefore, $\log \text{Exp}(\mathbf{z}; \lambda^{(v_j)})$ represents the natural logarithm of the probability density function of an exponential distribution with a rate parameter $\lambda^{(v_j)}$. Moreover, to ensure efficient backpropagation of gradients and avoid the potential issue of exploding exponentials within E-VAE, we use a combination of specific transformations and the reparameterization trick\(^1\).

\(^1\)The interpretability of E-VAE and the corresponding KL divergence are proven in the Appendix.
3.7 Loss Function

Specifically, given users from both platforms, their latent distribution space of decision-making outcomes from the perspective of source probability can be obtained in Sec. 3.6. We now introduce the KL divergence constraints for approximating the distribution space of similar users across two platforms. Utilizing the chi-squared test from statistical analysis [Hou et al., 2024b], we evaluate the influence of each user based on user profiles of $F^{k_1}$ and $F^{k_2}$. Then, as for the cascades with similar topics in different platforms, we read out the influence of users, rank the users on each platform in descending order of their influence, and perform cross-platform matching and alignment of users. Further, following an interpretable derivation, we design a decision-making regularization loss function to minimize the KL divergence between the two distributions for users in different platforms but with similar influences. Without loss of generality, we define $P_f(\hat{Y}) = q_\theta(z|O^F(v_j))$ as the distribution space of decision-making outcomes (e.g., source probability) for a user $v_j$ in the Twitter platform, and $P_g(\hat{Y}) = q_\phi(z|O^G(v'_j))$ as that for a user $v'_j$ in the Weibo platform.

\[
\mathcal{L}_c(P_f(\hat{Y}), P_g(\hat{Y})) = \frac{1}{2} [KL(P_f(\hat{Y}) \parallel P_g(\hat{Y})) + KL(P_g(\hat{Y}) \parallel P_f(\hat{Y}))].
\]

Further, a unique loss function for DSLF is designed to better adapt to the task of source localization:

\[
\mathcal{L}_{loss} = \mathcal{L}_{E-VAE}(\hat{R}^{p_1}, R^{p_1}) + \mathcal{L}_{E-VAE}(\hat{R}^{p_2}, R^{p_2}) + \alpha * (\mathcal{L}_{w-Entropy}(\hat{O}^{p_1}) + \mathcal{L}_{w-Entropy}(\hat{O}^{p_2})) + \beta * \mathcal{L}_c(P_f(\hat{Y}), P_g(\hat{Y})).
\]

The Loss encompasses the weighted (focusing on a small number of source nodes) binary cross entropy loss of source prediction task for each platform, E-VAE loss (including reconstruction error and KL divergence between the latent representation and the standard exponential distribution) for each platform, and regularization loss across both platforms. Here, $\alpha$ and $\beta$ are served as the coefficients to balance each item.

4 Experiments

4.1 Experimental Setup

We use three datasets collected from two real-world social media platforms, Weibo and Twitter, for locating the sources of cascades with similar topics in different platforms [Liu et al., 2015; Ma et al., 2016; Ma et al., 2017].

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Twitter</th>
<th>Weibo</th>
</tr>
</thead>
<tbody>
<tr>
<td>#users</td>
<td>677,640</td>
<td>2,856,741</td>
</tr>
<tr>
<td>#users in $G^*$</td>
<td>677,058</td>
<td>2,856,519</td>
</tr>
<tr>
<td>#relations in $G^*$</td>
<td>828,546</td>
<td>3,508,596</td>
</tr>
<tr>
<td>#cascades</td>
<td>2,308</td>
<td>4,664</td>
</tr>
<tr>
<td>#topic-aligned cascades</td>
<td>1,831</td>
<td>1,831</td>
</tr>
</tbody>
</table>

Table 1: Statistics and relevant information of the datasets.

And we consider TGASI [Hou et al., 2023], IVGD [Wang et al., 2022a], SL_VAE [Ling et al., 2022], GCSSI [Dong et al., 2022], and MCGNN [Shu et al., 2021] for comparison. And to demonstrate the source prediction performance of all methods rigorously, the widely used $F_1$-score = $2 \times \frac{Precision \times Recall}{Precision + Recall}$ is chosen as the evaluation metric [Wang et al., 2023].

In our experiments, we employ a 10-fold cross-validation strategy to divide the training and test datasets. Further, DSLF utilizes the training dataset for learning, and then the final result is output by averaging the prediction across each fold in the test dataset. Moreover, an early stopping mechanism is designed to avoid over-fitting in the training process. For optimization, the Adam optimizer is used, configured with a learning rate of 0.0005 for all model parameters.

It’s noteworthy that to demonstrate the generalizability of DSLF, we employ widely recognized simulated synthetic datasets for source localization, generated based on the IC and HSI propagation models, as commonly used in traditional localization methods [Ling et al., 2022; Wang et al., 2022a]. Similar to the settings of those studies, we select 10% of the nodes as ground-truth sources and simulated the IC and HSI propagation processes on a network with 4,039 nodes and 88,234 edges. The infection rate of each node follows the uniform distribution of $U(0.05, 0.15)$ [Hou et al., 2023]. And we independently generate 1,000 sets of propagation data, each including several snapshots with distinct timestamps. Since the experimental datasets from the SOTA methods we follow are single-layer networks, only one channel of DSLF is used and the cross-platform based regularization loss is removed.

4.2 Overall Experimental Results

The source detection performance based on the real-world dataset and the paradigm of the traditional localization methods is illustrated in Tab. 2. When benchmarked against the optimal baseline TGASI, DSLF exhibits an average improvement of 82.1% in real-world datasets and a 15.2% enhancement in simulated datasets from the propagation models. In conclusion, DSLF outperforms all the SOTA methods based on rigorous metrics in all datasets. There are three key reasons for the significant improvement in real-world datasets: (1) The dynamic user features and profiles are constructed to enhance individual uniqueness in user-driven-based real-world propagation scenarios. (2) The self-loop attention-based GCN module refines neighborhood aggregation, improving graph and profile representation quality. (3) The KL regularization module bolsters cross-platform robustness by aligning behavioral distributions of similar user groups.

Furthermore, in simulated datasets, the detection performance of DSLF, although superior to SOTA methods, shows a less pronounced improvement compared to real-world datasets, for two reasons: (1) The simulated data lacks user profile information. (2) The KL regularization module can not be used in single-layer scenarios. It’s noteworthy that DSLF transfers effectively in single-layer networks, indicating better generalizability. In summary, its continued higher
4.3 Ablation Study for Modules in DSLF

We further study the influence of designed components of DSLF in the source detection performance to prove each part of the contribution. As shown in Tab. 3, it will lead to a performance decrease or a delay in the convergence speed of model training no matter removing or exchanging any critical modules. Due to the limited space, more discussion is demonstrated in the Appendix.

Table 4: The performance evaluation of variant models by masking one of the user profiles.

<table>
<thead>
<tr>
<th>Variants</th>
<th>Twitter</th>
<th>Weibo</th>
</tr>
</thead>
<tbody>
<tr>
<td>-Ratio</td>
<td>0.871 (↑4.59%)</td>
<td>0.855 (↓5.50%)</td>
</tr>
<tr>
<td>-Status</td>
<td>0.881 (↓3.41%)</td>
<td>0.859 (↓5.01%)</td>
</tr>
<tr>
<td>-Date</td>
<td>0.886 (↓2.82%)</td>
<td>0.867 (↓4.04%)</td>
</tr>
<tr>
<td>-Fans</td>
<td>0.890 (↑2.36%)</td>
<td>0.874 (↑3.20%)</td>
</tr>
<tr>
<td>-Followings</td>
<td>0.898 (↓1.45%)</td>
<td>0.886 (↓1.81%)</td>
</tr>
<tr>
<td>-Tweets</td>
<td>0.905 (↑0.66%)</td>
<td>0.891 (↓1.23%)</td>
</tr>
<tr>
<td>DSLF</td>
<td>0.911</td>
<td>0.902</td>
</tr>
</tbody>
</table>

Table 5: The performance evaluation in different layers scenarios.

- Ratio of fans to followings suggests greater influence, while a lower ratio indicates a stronger capacity for information-seeking. Abundant numerical meaning ensures its highest importance. Furthermore, another notable conclusion is that the sum of each decline in performance resulting from individually masking each of the six features (15.29% summed decrease) exceeds that of all are masked simultaneously (a 6.30% decrease from the -Profile variant in Tab. 3). This suggests the presence of interaction effects among these features.

4.5 Scalability

The developed dataset is based on dual-channel propagation and further DSLF is verified based on two layers. In fact, DSLF can be extended to any arbitrary layers of networks. This scalability is demonstrated by feature-driven learning processes adaptable to different scales of networks, universal applicability of self-loop attention for diagonal enhancement, flexible VAE module sampled from different distributions (we also provide another VAE sampling from the normal distribution in our source code), and flexible dual-channel KL constraints for any two layers in multiplex scenarios. More descriptions can be found in the Appendix. And to prove the feasibility, we randomly divided all 6,900 cascades from Weibo or Twitter datasets into more platforms to test the time complexity of DSLF.

5 Conclusion

By developing a topic-aligned propagation cascade dataset enriched with user profiles, we bridge the gap between simulation and reality in source localization. Further, the DSLF method is developed which can infer the source in networks with any number of layers. Ablation experiments demonstrate the effectiveness of crawled user profiles, constructed propagation dynamic features, self-loop attention-based GCN, and dual-structure KL regularization modules of DSLF. In the future, we will focus on collecting multi-layered data to expand and refine our models, broadening the scope of our research and its applicability.
Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. U22B2036, 62271411, U22A2098, 11931015, 61976181), Fok Ying-Tong Education Foundation, China (No. 171105), the Fundamental Research Funds for the Central Universities (No. D5000230366), and the Tencent Foundation and XPLORER PRIZE. Dr. Xianghua Li is the corresponding author.

References


[Simmie et al., 2013] D. Simmie, M.G. Vigliotti, and C. Hankin. Ranking twitter influence by combining net-


