Measuring Time-Sensitive and Topic-Specific Influence in Social Networks With LSTM and Self-Attention

CHENG ZHENG1, QIN ZHANG2, GUODONG LONG2, CHENGQI ZHANG2, SEAN D. YOUNG3, AND WEI WANG1, (Member, IEEE)

1Department of Computer Science, University of California, Los Angeles, CA 90095, USA
2Center for Artificial Intelligence, Faculty of Engineering and IT (FEIT), University of Technology Sydney, Ultimo, NSW 2007, Australia
3University of California at Irvine, Irvine, CA 92697, USA

Corresponding author: Wei Wang (weiwang@cs.ucla.edu)

The work was supported in part by the National Institutes of Health under Grant U01HG008488, in part by the National Institute of Allergy and Infectious Diseases under Grant R56AI125105 and Grant 7R01AI132030, and in part by the National Institute of Mental Health under Grant 5R01MH106415.

ABSTRACT Influence measurement in social networks is vital to various real-world applications, such as online marketing and political campaigns. In this paper, we investigate the problem of measuring time-sensitive and topic-specific influence based on streaming texts and dynamic social networks. A user’s influence can change rapidly in response to a new event and vary on different topics. For example, the political influence of Douglas Jones increased dramatically after winning the Alabama special election, and then rapidly decreased after the election week. During the same period, however, Douglas Jones’ influence on sports remained low. Most existing approaches can only model the influence based on static social network structures and topic distributions. Furthermore, as popular social networking services embody many features to connect their users, multi-typed interactions make it hard to learn the roles that different interactions play when propagating information. To address these challenges, we propose a Time-sensitive and Topic-specific Influence Measurement (TTIM) method, to jointly model the streaming texts and dynamic social networks. We simulate the influence propagation process with a self-attention mechanism to learn the contributions of different interactions and track the influence dynamics with a matrix-adaptive long short-term memory. To the best of our knowledge, this is the first attempt to measure time-sensitive and topic-specific influence. Furthermore, the TTIM model can be easily adapted to supporting online learning which consumes constant training time on newly arrived data for each timestamp. We comprehensively evaluate the proposed TTIM model on five datasets from Twitter and Reddit. The experimental results demonstrate promising performance compared to the state-of-the-art social influence analysis models and the potential of TTIM in visualizing influence dynamics and topic distribution.

INDEX TERMS Social influence, time-sensitive, topic-specific, LSTM, self-attention.

I. INTRODUCTION Social network influence refers to the ability of a user to change the feelings, attitudes, or behaviors of other users within a network [1]–[3]. Influence measurement has become an essential task in many fields such as online marketing [4], [5] and political campaigns [6]. Due to its practical importance, measuring social influence has drawn growing research interests [7]–[10]. In this paper, we study the problem of influence measurement in temporal social networks: Given streaming posts and interactional activities, the goal is to model the users’ influence dynamics and find the influence distribution on different topics.

The influence of a user varies over time [8]. A person can become influential over a certain period due to a particular event. One example is Douglas Jones,1 the current United States Senator for Alabama. On December 12, 2017, Jones won a special election and became the first Democrat to win a Senate seat in Alabama since 1992. Due to the victory, Douglas Jones’ political influence increased dramatically during the election period and then vanished rapidly after the election. This is reflected by the influence scores

1https://en.wikipedia.org/wiki/Doug_Jones_(politician)
In addition to time, a user’s influence depends heavily on the topics [9], [11]. Users who have high global influence scores may not be influencers on a certain topic and vice versa. An example is shown in Figure 2. We plot the influence scores of two Twitter users on distinct topics. Jeff Dean has a higher overall influence score than Vitalik Buterin, especially on the topics of AI, ML and Big Data. However, Vitalik Buterin is identified as a potential influencer on the topic of Blockchain. The topic-specific influence analysis [9] is of vital importance for many applications.

A vast number of topics in social networks are active and evolve rapidly, which calls for a unified framework to jointly model the influence propagation over time and topics. A question raised from the aforementioned observations is: how can one measure time-sensitive and topic-specific influence? There are three major challenges to approach the problem. First, joint modeling the distribution of influence with respect to time and topics involves combinations of the two types of features, which is impractical to enumerate all possibilities. Recent works [8], [9], [11] focus on using either temporal or topical features but not both. Second, social networks in real-life are composed of multiple types of user interactions. For example, on Twitter [2], [12], [13], we can interact with other users by various features such as follow, retweet and mention, etc. The key question is how to assess the contributions of different interactions when influence propagates. Existing works only consider a single type of interaction or assign equal weights to different types of interactions [14]. Third, the nodes and edges in social networks are countless and evolve rapidly [15]. Therefore, supervised models are unable to take full advantage of the large-scale datasets, because only a small fraction of data is labeled with the ground truth.

To address the challenges above, we propose an unsupervised model, called Time-sensitive and Topic-specific Influence Measurement (TTIM) model. TTIM consists of influence attention network and matrix-adaptive long short-term memory (LSTM) [16], which can be jointly trained to automate the feature combinations in the first challenge. The proposed influence attention network aggregates node influence representations with attention to different types of interactions [17], [18]. The unsupervised training objective can drive the learning system without supervision from the ground truth. Our proposed framework can also be naturally adapted for online learning. We evaluate the proposed method with five real-world datasets from Twitter and Reddit. To summarize, the primary contributions of this work are:

- To the best of our knowledge, we are the first to simultaneously measure the time-sensitive and topic-specific influence in social networks.
- We propose a unified computational framework, TTIM, to solve the social influence measurement problem. The two sub-networks, influence attention network and matrix-adaptive LSTM, can jointly learn the contributions of different interactions and the influence dynamics in social networks. The framework supports both standard and online learning.
- We use five datasets crawled from Twitter and Reddit to compare TTIM with the state-of-the-art social influence measurement models. The experimental results demonstrate the effectiveness, efficiency, and scalability of the proposed method.

The rest of this paper is organized as follows. Section II presents the problem formulation. The details of the framework are shown in Section III. Section IV presents the datasets and experimental results, comparing our model with state-of-the-art methods. Section V summarizes related work and Section VI concludes this paper.
user nodes and $|\mathcal{V}| = N$. The interactional information is formulated as the adjacency tensor $A_t \in \mathbb{R}^{N \times N \times L}$ for $L$ types of interactions in the $t$-th time interval. The user-topic affinity tensor $X_t \in \mathbb{R}^{N \times M \times D}$ represents the textual information of $N$ users in the $t$-th interval. $M$ is the number of topics in the entire social network and $D$ is the topic embedding dimension.

For example, if the $l$-th type of interactions on Twitter is mention and $A_{t(j)}$ equals to 2, then this represents that user $i$ was mentioned twice by user $j$ in the $t$-th interval. We will detail the generation of tensor $X_t$ in Section III-A. Given the above definition, we introduce the problem formulation.

Problem 1 (Time-Sensitive and Topic-Specific Influence Measurement): Given the temporal attributed graphs $G_t = (\mathcal{V}, A_t, X_t)$, $t = 1, \ldots, T$ that represent the textual and interactional information in social networks, the goal is to output the time-sensitive and topic-specific influence tensor $B \in \mathbb{R}^{N \times T \times M}$ for users $\mathcal{V}$.

Several key questions about Problem 1 need to be answered: 1) How do we extract the user-topic affinity tensor $X_t$? 2) How do we assess the contributions of different types of interactions during the influence propagation? 3) How do we aggregate the textual and interactional information together? 4) How do we measure user influences as a function of topic and time over the graph sequence $G_t(t = 1, \ldots, T)$ in an unsupervised fashion?

III. THE FRAMEWORK OF TTIM

This section introduces the framework of TTIM model. An intuitive illustration is given in Figure 3. At each time interval, there are two types of raw data: textual and interactional. For text data, we utilize the Seeded Latent Dirichlet Allocation (SeededLDA) [19] to perform topic distillation and obtain the user-topic affinity tensor in each time interval. For the temporal graphs of $L$ types of interactions, we design the influence attention network to simulate the influence propagation process and learn the contributions of different interactions. Then the temporal influence is learned by optimizing the unsupervised objective function in a matrix-adaptive LSTM model. We also design an online version of TTIM by slightly altering the pipeline.

A. TOPIC DISTILLATION

In social networks, a user usually has interests on multiple topics. The topic distillation aims to learn the $D$-dimensional vector $X_{t(j)}$ that represents the embedding of user $i$ on topic $j$ at time $t$. Hence, we concatenate the messages posted by the same user in one time interval as one document, resulting in $N \times T$ documents. To obtain the topic focus of users, we utilize the SeededLDA model [19], which can identify latent topics in three fashions,

- **Unsupervised**: Similar to the vanilla LDA [20], the document-topic distribution is learned from the probability distribution with the Dirichlet prior.
• **Supervised**: SeededLDA accepts sets of seed words as the representative of the underlying topics. In this way, we can obtain the document-topic distribution in specific domains.

• **Online**: It is not desirable to retrain the topic model from scratch whenever new data arrive. Instead, with online training, we could progressively update the model by utilizing previous topic-word distribution as seed words to feed to SeededLDA. Combining with the online LSTM model presented in Section III-C, we can train the model incrementally as new data arrive.

In each time interval, we distill $M$ topics and obtain the user-topic affinity tensors $X_t \in \mathbb{R}_N \times M \times D$, $t = 1, \ldots, T$. For user $i$, $X_t(i, j)$ is the term frequencies of top $D$ words belonging to topic $j$. A larger element in the tensor $X_t(i, j)$ indicates the more focus that a user puts on the corresponding topic. The unsupervised SeedLDA is suitable for training TTIM from scratch, where it automatically detects the topics in social posts. The supervised and online fashions are adaptive to the online training of TTIM model.

## B. INFLUENCE ATTENTION NETWORK

We build the influence attention network to simulate the influence propagation process and learn the contributions of different interactions. Following the formulation in Section II-B, we obtain the adjacency tensor $A_t$, $t = 1, \ldots, T$, corresponding to $L$ types of interactions. Intuitively, different types of interactions play different roles in influence propagation. The majority of existing works only considered a single type of interaction or assigned a weight to interactions [14] according to domain knowledge. Inspired by Graph Attention Networks (GAT) [18], [21] and DeepInf [10], we propose the influence attention network, which can aggregate the node topic distribution with attention on the node’s local neighborhood features and edges in multi-typed social networks.

Specifically, without loss of generality, we sketch the influence attention process focusing on a specific user $i$ in graph snapshot at time $t$. Let $N_{1i} \subseteq \mathbb{N}_N$ be the set of one-hop neighbors of node $i$ at time $t$. Different from GAT or DeepInf, we introduce the attention coefficients for both user-topics affinities and user-user interactions,

$$e_{i, j} = \text{MLP}_\phi(X_{t(i)}, X_{t(j)}, A_{t(i, j)})$$  \hspace{1cm} (1)

where $j \in \mathbb{N}_N$, and the attention coefficient $e_{i, j}$ measures the relative influence that user $i$ has on user $j$. MLP$_\phi$ is a multi-layer neural network with parameters $\phi$. To accommodate users with different neighborhood sizes, we normalize the coefficients with softmax,

$$a_{i, j} = \frac{\exp(e_{i, j})}{\sum_{k \in N_{1i}} \exp(e_{i, k})}$$  \hspace{1cm} (2)

In the influence propagation process, the social network community disseminates messages with multiple rounds of propagation. Therefore, we propose to model the phenomena with multiple influence attention layers by aggregating nodes’ topic distribution vectors in their neighborhood.

The user-topic affinity tensor $X_t(i)$ is utilized as the input node features to the first layer $(F_t^{(0)} = X_t(i))$. The $p$-th influence attention layer performs as follows,

$$F_t^{(p)}(i) = \sigma \left( \sum_{j \in N_{1i}} a_{i, j} F_t^{(p-1)}(j) \right) W_p$$  \hspace{1cm} (3)

where $\sigma(\cdot)$ is a non-linear activation function like ReLU, $F_t^{(p)}(i) \in \mathbb{R}^{N \times M \times D_t}$ is the output node representations, and $W_p \in \mathbb{R}^{d_{t(p-1)} \times d_t}$ is the parameter matrix for this layer. The aggregated feature tensor $F_t$ from the output of the final influence attention layer represents the user topic distribution after influence propagation.

## C. MATRIX-ADAPTIVE LSTM

With the sequence of aggregated feature tensors $F_t$, $t = 1, \ldots, T$, we design a matrix-adaptive LSTM network [22], [23] to learn the time-sensitive and topic-specific influence scores for users. We adopt LSTM [24] motivated by its significant capability for learning long-term dependencies that naturally exist in temporal social network data. Shown in the right part of Figure 4, the matrix-adaptive LSTM accepts a sequence of matrices as input and outputs the state matrices of all time points, working as a many-to-many recurrent model.

The equations from Eq. 4 to Eq. 8 describe the operations in a matrix-adaptive LSTM cell, with the dimension $N$ omitted for simplicity,

$$I_t = \sigma(F_t W_{xi} + H_{t-1} W_{hi} + C_{t-1} W_{ci} + b_i)$$  \hspace{1cm} (4)

$$G_t = \sigma(F_t W_{xf} + H_{t-1} W_{hf} + C_{t-1} W_{cf} + b_f)$$  \hspace{1cm} (5)

$$C_t = G_t \odot C_{t-1} + I_t \odot \tanh(F_t W_{xc} + H_{t-1} W_{hc} + b_c)$$  \hspace{1cm} (6)

$$O_t = \sigma(F_t W_{xo} + H_{t-1} W_{ho} + C_t W_{co} + b_o)$$  \hspace{1cm} (7)

$$H_t = O_t \odot \tanh(C_t)$$  \hspace{1cm} (8)

where $\sigma(\cdot)$ denotes the sigmoid function $\sigma(x) = 1/(1 + e^{-x})$, and $I_t, G_t \in [0, 1]^{M \times P}$ are the input and forget gates. $P$ is the size of the hidden states of the LSTM model. $C_t \in \mathbb{R}^{M \times P}$ is the cell state, which is the core of an LSTM cell indicated...
by the longest vertical line in Figure 4. The cell state serves as the information connection between time \( t - 1 \) and time \( t \). The input and forget gates having values normalized to \([0, 1]\) help the cell state control how much information it should take from the input (second term in Eq. 6) and how much is inherited from the previous time interval (first term in Eq. 6). \( O_t \in [0, 1]^{M \times P} \) is the output gate and \( H_t \in \mathbb{R}^{M \times P} \) is the output state. The output gate filters information from the cell state \( C_t \) and passes it to the output state, which serves as the output of the LSTM network. In general, the LSTM network operates in a sequential fashion with \( F_t \) as the initial input. The cell state \( C_t \) at time \( t \) and output state \( H_t \) will be repeatedly fed into the LSTM cell at time \( t + 1 \). The weights \( W_{xc} \in \mathbb{R}^{d \times P}, W_c \in \mathbb{R}^{P \times P} \), and biases \( b_1, b_f, b_i, b_o \in \mathbb{R}^P \) are the model parameters, which are trained by back-propagation with the objective function introduced in Section III-D. The influence tensor \( B \) can be obtained from the concatenation of output states \( H_t \) after a pooling layer. Possible choices for the pooling operation include max, average, and sum.

The matrix-adaptive LSTM network can generalize to support online training. At time \( T' \), we may leverage the model trained at time \( T' - 1 \) to compute the extended user-topic affinity tensor \( X_{T'} \) and the aggregated feature matrix \( F_{T'} \). We can further train the LSTM model starting with the parameters \( (W) \) from the previous LSTM model at time \( T' - 1 \). To capture the temporal dependency, we set a time interval window \( T_W \) as a hyper-parameter: only data arrived during \([T' - T_W, T')\) is used to retrain the LSTM model. This allows the model to converge much faster than retraining from scratch.

### D. Objective Function

In order to measure the time-sensitive and topic-specific influence, we consider three criteria when we build the unsupervised objective function. First, the users with a larger neighborhood and higher affinity should have a higher influence score; Second, active users are more likely to have a high influence score than inactive users; Last, the change in the influence matrix should be smooth. Based on the ideas, the final optimization problem is constructed in Eq. 9 to learn the temporal user-topic influence matrix \( B \in \mathbb{R}^{N \times T \times M} \).

\[
\max L(W, \lambda_t) = \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{t(ij)} (1 + \sum_{k=1}^{N} A_{t(ik)} \cdot \|B_{t(kj)}\|^2)
+ \sum_{t=1}^{T} \sum_{i=1}^{N} \|F_{t(i)}\|^2 \cdot \|B_{t(i)}\|^2
- \sum_{t=2}^{T} \|B_{t(x)} - B_{t(x-1)}\|^2_F
\]

where \( A_{t} \in \mathbb{R}^{N \times N \times L} \) is the adjacency tensor for \( L \) types of interactions in the \( t \)-th time interval; \( F_{t} \) is the aggregated user-topic affinity tensor in the time interval \( t \) introduced in Section III-B. The larger value of \( B_{(i,m)} \) represents user \( i \) has a higher influence on topic \( m \) at time \( t \). \( W \) contains the weight matrices in the LSTM model and influence attention network. \( \xi_1 > 0, \xi_2 = 1 \) are the trade-off parameters to balance the three components. A constraint is added to normalize user influence scores on a topic for each time interval. We use back-propagation through time (BPTT) algorithm to train the model and learn the user influence scores.

With our proposed influence attention network and matrix-adaptive LSTM, we can extend the TTIM model to online fashion. We depict the pseudocode of the online training of the TTIM model in Algorithm 1. With time \( T' \) data arriving, the modified objective function is,

\[
\max L(W, \lambda_t) = \sum_{t=T' - T_W}^{T'} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{t(ij)} (1 + \sum_{k=1}^{N} A_{t(ik)} \cdot \|B_{(t,kj)}\|^2)
+ \xi_1 \sum_{t=T' - T_W}^{T'} \sum_{i=1}^{N} \|F_{t(i)}\|^2 \cdot \|B_{t(i)}\|^2
- \xi_2 \sum_{t=T' - T_W + 1}^{T'} \|B_{(t,x)} - B_{(t-1,x)}\|^2_F
\]

In summary, our TTIM model answers the questions raised in Section II-B with well-designed pipeline: the SeededLDA model learns the user-topic affinity tensor; the adjacency tensors for different interactions are integrated with learnable weights; the influence attention network simulates the influence diffusion; the matrix-adaptive LSTM model captures the long-term dependencies and learns the influence scores following the optimization problem. Streaming texts and dynamic social networks are jointly modeled to measure social influence.

---

**Algorithm 1** TTIM-Online With New Data Arriving at Time \( T' \)

**Require:** \( A_{t(ij)}, F_{t} \), where \( t = T' - T_W, \ldots, T' \), documents \( d_{T'} \), previous SeededLDA(\( T' - 1 \)), previous LSTM(\( T' - 1 \)), training epoch \( n_{epoch} \), hyperparameters \( \alpha, \xi_1, \xi_2, T_W \)

1. Load the topic-word distribution from SeededLDA(\( T' - 1 \)) as seed distribution for SeededLDA(\( T' \))
2. Train SeededLDA(\( T' \)) with \( d_{T'} \)
3. Compute \( X_{T'(ij)} \)
4. Compute \( F_{T'} \) using Eq. 1, 2, 3
5. Load \( W \) from LSTM(\( T' - 1 \)) to LSTM(\( T' \))
6. for epoch = 1; epoch \( \leq n_{epoch} \) do
5. Compute \( I_t, G_t, C_t, O_t, H_t \) using Eq. 4 - Eq. 8
9. end for
10. Compute \( L(W, \lambda_t) \) using Eq. 10
11. Backpropagate and update \( W \)
12. end for
IV. EXPERIMENTS
In this section, We evaluate our proposed method with extensive experiments. First, we introduce the labeled datasets to quantitatively evaluate our model with the influencer detection task, shown in Section IV-A. Second, we qualitatively evaluate the time-sensitive and topic-specific property of TTIM model with large unlabeled datasets in Section IV-B. Third, the proposed TTIM-Online method is shown to be efficient in training and achieve competitive results in Section IV-C. Finally, we conduct the parameter sensitivity and scalability analysis in Section IV-D.

A. EXPERIMENTS WITH LABELED DATASETS
In this section, we detail the experimental results on the influencer detection task with three labeled datasets. The task aims to identify the top influential individuals from all users in social networks. First, we introduce the datasets and baselines, followed by the influencer detection results.

1) DATASETS
We created three manually labeled datasets from Twitter (Politics set and Technology set) and Reddit (Reddit set). Dataset statistics are shown in Table 2. More preprocessing details can be found in the supplementary materials.

| Dataset   | observation window | # time intervals | # users | # posts |
|-----------|--------------------|------------------|---------|---------|
| Politics  | 2017.10.22 - 2017.12.30 | 10                | 1,031/64 | 1,840,552 |
| Technology| 2018.01.07 - 2018.01.13 | 7                 | 1,122/80  | 141,835  |
| Reddit    | 2015.05.01 - 2015.05.31 | 31                | 35,267/100 | 126,125 |
| LV-shooting| 2017.10.01 - 2017.10.11 | 11               | 2,859,809 | 17,635,937 |
| General   | 2016.08.01 - 2019.07.31 | 36               | 1,893,174 | 15,953,165 |

Note: * refers to the number of influencers we manually labeled.

In the two datasets from Twitter, we labeled the influencers by selecting users with a large group of followers, active involvement in the politics/technology topics and top global influence on other users' actions. The labels were selected from the majority votes of three human labelers. The Politics set contains 1,031 users who send politics-related tweets, and 64 of them are labeled as influencers. There are 10 one-week intervals and 1,840,552 tweets in total. The Technology set contains 1,122 users who send technology-related tweets, and 80 of them are labeled as influencers. There are 7 one-day intervals and 141,835 tweets in total.

Reddit is an online discussion forum where users post and comment on contents in different topical communities. In the Reddit platform, users can upvote posts that they are in favor of, so the number of upvotes can indicate the influence of posts and their senders. We labeled users whose posts received the most upvotes as influencers. The Reddit set is from May 2015 Reddit comment dump and it contains 35,267 users, with 100 labeled influencers. We build a user-to-user interaction graph, connecting users if one user comments under another user’s post. There are 31 one-day intervals and 126,125 posts/comments in Reddit set.

2) BASELINES
We compare the proposed TTIM model with the following seven representative baselines:

- **Followers.** The feature used by this baseline is the number of the user’s followers. Note that we only have access to this feature on Twitter, not with Reddit.
- **TwitterRank** [11] is an extension of the PageRank algorithm, which uses LDA to find some topics, and then calculates the rank of users with respect to topics based on their influence on followers and their interests in these topics.
- **Topical Affinity Propagation (TAP)** [25] is a topical affinity propagation model built on a factor graph to identify the topic-specific social influence.
- **ReFluence** [26] is a statistical and analytical model based on Edelman’s topology of influence to determine the user’s role and influence on each other. Here we treat the “Idea Starter” and “Amplifier” defined in this baseline as influencers and the others as normal users.
- **RR-LT** model [8] uses a function of edge weights and the self-weight of nodes to represent influence probabilities under the Linear Threshold model. Polling-based methods and a sample of random reversely reachable sets are used to approximate the influence of nodes.
- **RR-IC** model [8] uses propagation probability, polling, and random reversely reachable sets to track influencers under the Independent Cascade model.
- **CoupledGNN** [27] applies two coupled graph neural networks to iteratively model and predicts the network-aware popularity.

3) EXPERIMENTAL SETTINGS
The proposed TTIM is implemented in the Tensorflow framework [28]. The training optimizer is Adam [29] with a learning rate as 0.0005, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The experiments are conducted on a Linux server with a 16G memory Tesla V100 GPU, 20 Intel Xeon E5-2698 CPUs, and 512 GB memory. We use the grid search to tune hyper-parameters. The topic embedding dimension $D$ is chosen from $\{5, 10, 20, 30\}$. The trade-off parameters in equation 9 ($\xi_1, \xi_2$) are searched from $10^{-4}$ to $10^{4}$ with a step of $10^1$. We initialize the weight matrices in the proposed TTIM model with Xavier initialization [30].

To identify influencers in the labeled datasets, we sum the learned influence score of all topics and max-pool over time to obtain each user’s influence score. The output predictions will be the top-$k$ users with the highest influence scores. We evaluate the influencer detection performance by precision.
**TABLE 3.** Experimental results on the labeled datasets.

| Methods           | Politics set | Technology set | Reddit set |
|-------------------|--------------|----------------|------------|
|                   | Prec, F1, AUC| Prec, F1, AUC  | Prec, F1, AUC |
| Followers         | 0.484        | 0.582          | -          |
| TAP               | 0.374        | 0.613          | 0.445      |
| TwitterRank       | 0.363        | 0.620          | 0.392      |
| ReFluence         | 0.394        | 0.330          | 0.471      |
| RR-LT             | 0.734        | 0.307          | 0.390      |
| RR-JC             | 0.641        | 0.241          | 0.527      |
| CoupledGNN        | 0.673        | 0.638          | 0.537      |
| TTIM w/ Attention | 0.612        | 0.683          | 0.632      |
| TTIM w/ LSTM      | 0.654        | 0.715          | 0.658      |
| TTIM-Online       | 0.779        | 0.786          | 0.691      |
| TTIM              | **0.789**    | **0.805**      | **0.761**  |

**FIGURE 5.** The precision of top-\(k\) influencers detection on the labeled datasets.

4) EXPERIMENTAL RESULTS

Table 3 shows the detailed results of detecting the top-\(k\) influencers, where \(k\) is the number of positive samples in the ground truth (\(k = 64\) for Politics set, \(k = 80\) for Technology set, and \(k = 100\) for Reddit set).\(^5\) The best results are highlighted in bold. The results show that TTIM is effective and outperforms other baselines in precision, F1, and AUC on these datasets. Figure 5 further illustrates the precision at \(k\), where \(k\) is the number of top influencers identified by each method. We can observe that TTIM always maintains a higher precision level than the baselines and has 100\% precision over the top 20 on Politics and Technology sets. These observations verify the outstanding ability of TTIM to detect the top influencers. We attribute the significant improvement to the following two reasons. First, TTIM considers diverse sources including the text contents and multiple interactions. Compared against baselines like CoupledGNN which treated interactions equally, TTIM automatically learns the different weights of interactions via attention mechanism. Second, TTIM well models the temporal data by using LSTM to learn the influence score with the streaming text and dynamic social networks integrated seamlessly.

We conduct the ablation study by removing the attention mechanism (TTIM w/ Attention) and LSTM (TTIM w/ LSTM) one by one at a time. As the results in the Table 3 show, each module contributes to the performance improvement and the proposed TTIM benefits from the influence propagation process learned by the influence attention network and the time-sensitive pattern learned by the matrix-adaptive LSTM.

We also highlight the capability of TTIM on retrieving the time-sensitive and topic-specific influence score of users with labeled Twitter datasets, shown in Figure 6(a),(c) for the Politics set and in Figure 6(b),(d) for the Technology set. We can observe some interesting phenomena. For example, in Figure 6 (a), there is not only a peak for Twitter user Douglas Jones but also a similar trend for user TheDailyEdge and TeaPainUSA. A probable reason for this could be they are in the same political party as Douglas Jones and share the influential benefits from the election event. Another finding is that user Vitalik Buterin’s influence score is mainly limited to the topic blockchain. This could be the reason why his influence trend is similar to the bitcoin price during that period.

\(^4\) F1 = \(\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}\)

\(^5\) Here, precision and F1 values are always equal since \(k\) equals the number of true labels, making the number of false positives and false negatives equivalent.
C. Zheng et al.: Measuring Time-Sensitive and Topic-Specific Influence in Social Networks

B. EXPERIMENTS WITH UNLABELED DATASETS

In this section, we discuss the experimental results on influence measurement on the large unlabeled datasets.

1) DATASETS

We utilize two unlabeled datasets. The LV-shooting set contains 1% of all tweets during the period from October 1, 2017 to October 11, 2017. On the night of October 1, 2017, a gunman fired more than 1,100 rounds to a crowd of concertgoers at the Route 91 Harvest music festival in Las Vegas, leaving 58 people dead and 851 injured. This event aroused a huge response on social media platforms, so we crawled the tweets over the following 11 days. After data preprocessing, the dataset contained 2,859,809 users, 17,635,937 tweets and 11 one-day time intervals. Another dataset General set contains 1% of tweets in three years (Aug 2016 - Jul 2019). The dataset contained 1,893,174 users and 15,953,165 tweets, with 36 one-month time intervals.

2) RESULTS

Table 4 shows the top-5 topics with their top keywords and top influencers. We name these topics to simplify the presentation. Intuitively, it is clear that the influencers are very relevant to the corresponding topics. For example, one would expect Donald Trump, Mike Pence, and Hillary Clinton to be influential on politics-related topics, just as one would expect the public figures such as Rihanna (singer) and Jake Tapper (journalist), and online video-sharing platform (Youtube) to be influential in the praying activities after the Las Vegas shooting tragedy.

We explore the time-sensitive and topic-specific property of the influence score respectively in Figure 7(a), (b) and Figure 7(c), (d). We show the influence scores of the top-5 influencers detected by TTIM in these two datasets. In Figure 7(a), four users had an influence peak on October 1, 2017, just after the Las Vegas Shooting happened, except BleacherReport (which is a sports platform). From Figure 7(b), we can observe that Donald Trump and two news platforms Fox News and The New York Times have obvious peaks during the period of the presidential election (October to November 2016) and the presidential inauguration (December 2016 to January 2017), which is reasonable. We can see the account for Donald J. Trump has relatively high influence over the period in both datasets, because of his activeness on Twitter. Figure 8 shows the 3D plots of the influence score of Donald Trump in the LV-shooting and General datasets, respectively. We can see that the influence score varies significantly along the dimensions of time and topic. In summary, our proposed TTIM model captures the time-sensitive and topic-specific influence on a large scale and can identify influencers with various time granularity.

C. ONLINE TRAINING

We furnish a comparison between the standard TTIM model and TTIM-Online. When new data arrive at time $T'$:

- TTIM is retrained with all data arrive so far, i.e., during the time $[1, T']$
- TTIM-Online starts from the previous model trained based on data of $[1, T' - 1]$, and updates the model only using the recent data, as detailed in Algorithm 1.

---

6https://en.wikipedia.org/wiki/2017_Las_Vegas_shooting
Both algorithms run until the models converge. The time interval window $T_W$ in the LSTM model in TTIM-Online is set as 3. We show the TTIM-Online performance in Table 3 and Figure 5. TTIM-Online still outperforms baseline models measured by precision, F1, and AUC on labeled datasets. Figure 9 shows the detailed precision and training time of TTIM and TTIM-Online in Politics and Technology sets. In Figure 9(a), we observe that both algorithms achieve similar precision in detecting influencers. In the beginning, the arrival of new data will enhance the precision, until the model saturates and its performance reaches a plateau. However, Figure 9(b) shows the online version of the TTIM model is much more efficient and scalable than the standard TTIM model. The training time of the online TTIM model remains at a constant level for each timestamp, whereas the training time taken by the standard TTIM model at each time stamp grows linearly as the number of timestamps increases.
has attracted many interests in influence dynamics analysis. We ignore the influence of specific topics on the mensuration, method that jointly modeled text and followship. And if we changes as a stream of edge weight updates.

Compares with Coupled-GNNs and its analogs, our TTIM aims to measure the macro-level influence and model its dynamics, which is vital to global influencer identification. Compared with Deepinf and NNMLinf, our proposed TTIM method considers the specific topics during the exploring of influence, not only the cascading effects (i.e. time-sensitive effects).

Besides the vanilla problem, if we ignore the influence of time on the mensuration, topic-specific influencer detection has been studied in several previous works [25], [36]–[38]. TwitterRank [11] used both network structure and topic similarity in calculating user influence on Twitter. Bi et al. [9] proposed a Bernoulli-multinomial mixture method that jointly modeled text and followship. And if we ignore the influence of specific topics on the mensuration, influence dynamics analysis has attracted many interests considering the evolving nature of social networks [39], [40]. Aggarwal et al. [41] proposed the influential node discovery in dynamic networks with the forward and backward trace approach. Yang et al. [8] studied influential node tracking and influence maximization [42], [43] by modeling dynamic changes as a stream of edge weight updates.

In summary, there is no existing work measuring time-sensitive and topic-specific influence in social networks. That motivates us to propose the LSTM and self-attention based TTIM, which integrates streaming texts and multiplex interactions to measure the temporal social influence on various topics.

VI. CONCLUSION
This paper explores the problem of measuring time-sensitive and topic-specific influence in social networks. A computational framework, Time-sensitive and Topic-specific Influence Measurement, is proposed based on influence attention network and matrix-adaptive LSTM. With multiple types of interactions and streaming texts, the influence attention network simulates the influence diffusion with self-attention. The matrix-adaptive LSTM captures the long-term dependencies and learns the influence scores following the optimization problem. Comprehensive evaluations of the proposed method are conducted with five datasets from Twitter and Reddit. The experimental results show superior performance of TTIM over the state-of-the-art social influence analysis models. By applying the proposed TTIM model to Twitter data of a large scale, we can visualize the influence dynamics and topic distributions in social networks.

ACKNOWLEDGMENT
The work was partially supported by National Institutes of Health (U01HG008488), National Institute of Allergy and Infectious Diseases (R56AI125105 and 7RO1AI132030), and the National Institute of Mental Health (5R01MH106415). The authors thank the anonymous reviewers for their careful reading and insightful comments on our manuscript.

REFERENCES

[1] H. C. Kelman, “Compliance, identification, and internalization processes of attitude change,” J. Conflict Resolution, vol. 2, no. 1, pp. 51–60, Mar. 1958.
[2] A. S. Badashian and E. Stroulia, “Measuring user influence in GitHub: The million follower fallacy,” in Proc. 3rd Int. Workshop CrowdSourcing Softw. Eng. CSf-SE, May 2016, pp. 15–21.
[3] F. Riquelme and P. González-Cantergiani, “Measuring user influence on Twitter: A survey,” Inf. Process. Manage., vol. 52, no. 5, pp. 949–975, Sep. 2016.
[4] M. Richardson and P. Domingos, “Mining knowledge-sharing sites for viral marketing,” in Proc. 8th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining KDD, 2002, pp. 61–70.
[5] P. Domingos and M. Richardson, “Mining the network value of customers,” in Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining KDD, 2001, pp. 57–66.
[6] S. Ranganath, X. Hu, J. Tang, and H. Liu, “Understanding and identifying advocates for political campaigns on social media,” in Proc. 9th ACM Int. Conf. Web Search Data Mining WSDM, 2016, pp. 43–52.
[7] J. Tang, S. Wu, and J. Sun, “Confluence: Conformity influence in large social networks,” in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining KDD, 2013, pp. 347–355.
[8] Y. Yang, Z. Wang, J. Pei, and E. Chen, “Tracking influential individuals in dynamic networks,” IEEE Trans. Knowl. Data Eng., vol. 29, no. 11, pp. 2615–2628, Nov. 2017.
[9] B. Bi, Y. Tian, Y. Sismanis, A. Balmin, and J. Cho, “Scalable topic-specific influence analysis on microblogs,” in Proc. 7th ACM Int. Conf. Web Search Data Mining WSDM, 2014, pp. 513–522.
[10] J. Qiu, J. Tang, H. Ma, Y. Dong, K. Wang, and J. Tang, “Deepinf: Social influence prediction with deep learning,” in Proc. KDD, 2018, pp. 2110–2119.
CHENGQI ZHANG received the bachelor’s degree in computer science from Fudan University, in March 1982, the master’s degree in computer science from Jilin University, in March 1985, the Ph.D. degree in computer science from The University of Queensland, in October 1991, and the Doctor of Science degree from Deakin University, in October 2002. He has been appointed as a Distinguished Professor with the University of Technology Sydney, from February 2017 to February 2022, the Executive Director of UTS Data Science, from January 2017 to January 2021, an Honorary Professor with The University of Queensland, from January 2015 to December 2017, an Adjunct Professor with the University of New South Wales, from March 2017 to March 2020, and has been a Research Professor of information technology with UTS, since December 2001. In addition, he has been selected as the Chairman of the Australian Computer Society National Committee for Artificial Intelligence, since November 2005, and the IEEE Computer Society Technical Committee of Intelligent Informatics (TCII), since June 2014.

SEAN D. YOUNG received the Ph.D. degree in psychology and the master’s degree in health services research from Stanford University. He is currently the Executive Director of the University of California Institute for Prediction Technology, the UCLA Center for Digital Behavior, and a Medical School and Informatics Professor with the UCI Departments of Emergency Medicine and Informatics. Before joining UCI, he was a Medical School Professor with the UCLA Department of Family Medicine, where he continues to hold a joint appointment.

WEI WANG (Member, IEEE) received the Ph.D. degree in computer science from the University of California at Los Angeles, Los Angeles, in 1999. She is currently the Leonard Kleinrock Chair Professor in computer science with the University of California at Los Angeles, and the Director of the Scalable Analytics Institute (ScAi). Her research interests include big data analytics, data mining, bioinformatics and computational biology, and databases. She was a Professor in computer science and a member of the Carolina Center for Genomic Sciences and Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, from 2002 to 2012, and a Research Staff Member with the IBM Thomas J. Watson Research Center, from 1999 to 2002.