Interaction Models for Detecting Nodal Activities in Temporal Social Media Networks

WINGYAN CHUNG, School of Modeling, Simulation & Training, University of Central Florida, Orlando, Florida, U.S.A. and Department of Computer Science, The University of Hong Kong, Pokfulam, Hong Kong
BINGBING RAO and LIQIANG WANG, Department of Computer Science, University of Central Florida, Orlando, Florida, U.S.A.

Detecting nodal activities in dynamic social networks has strategic importance in many applications, such as online marketing campaigns and homeland security surveillance. How peer-to-peer exchanges in social media can facilitate nodal activity detection is not well explored. Existing models assume network nodes to be static in time and do not adequately consider features from social theories. This research developed and validated two theory-based models, Random Interaction Model (RIM) and Preferential Interaction Model (PIM), to characterize temporal nodal activities in social media networks of human agents. The models capture the network characteristics of randomness and preferential interaction due to community size, human bias, declining connection cost, and rising reachability. The models were compared against three benchmark models (abbreviated as EAM, TAM, and DBMM) using a social media community consisting of 790,462 users who posted over 3,286,473 tweets and formed more than 3,055,797 links during 2013–2015. The experimental results show that both RIM and PIM outperformed EAM and TAM significantly in accuracy across different dates and time windows. Both PIM and RIM scored significantly smaller errors than DBMM did. Structural properties of social networks were found to provide a simple and yet accurate approach to predicting model performances. These results indicate the models' strong capability of accounting for user interactions in real-world social media networks and temporal activity detection. The research should provide new approaches for temporal network activity detection, develop relevant new measures, and report new findings from large social media datasets.

CCS Concepts: • **Information systems → Data analytics**; • **Networks → Social media networks**; • **Human-centered computing → Social media**; • **Applied computing →** *Marketing*;

Additional Key Words and Phrases: Interaction models, dynamic graph modeling, social media analytics, social network analysis, business analytics

ACM Reference format:
Wingyan Chung, Bingbing Rao, and Liqiang Wang. 2019. Interaction Models for Detecting Nodal Activities in Temporal Social Media Networks. *ACM Trans. Manage. Inf. Syst.* 10, 4, Article 14 (December 2019), 30 pages.
https://doi.org/10.1145/3365537

This research was supported in part by funding from Fulbright U.S. Scholar Fellowship (2017–19), Intel Security Curriculum Initiative (grant #23568271), and NSF CAREER program (grant #1622292).
Authors' addresses: W. Chung (corresponding author), School of Modeling, Simulation & Training, University of Central Florida, Orlando, Florida, U.S.A., and Department of Computer Science, The University of Hong Kong, Pokfulam, Hong Kong; email: wchung@ucf.edu; B. Rao and L. Wang, Department of Computer Science, University of Central Florida, Orlando, Florida, U.S.A.; emails: Robin.Rao@knights.ucf.edu, lwang@cs.ucf.edu.
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.
2158-656X/2019/12-ART14 $15.00
https://doi.org/10.1145/3365537

1 INTRODUCTION

Detecting nodal activities in dynamic social networks has strategic importance in many applications. For instance, e-commerce marketers, homeland security experts, and customer relationship managers study social networks of interested actors (e.g., customers, terrorists, and online forum users) to examine activity trends. In recent years, social media has rapidly emerged as a source of intelligence and business value [61]. The economic impact of social media on business is estimated to exceed $1 trillion due to more efficient communication and collaboration [11]. Advanced analytics play an important role in supporting understanding and trend analysis in social media [21]. In particular, social media networks provide a diverse repertoire of theories and frameworks to describe, analyze, and explain emerging behaviors in organizations and in society [31]. These networks capture the information of the agents (modeled as nodes or vertices) and their relationships (modeled as links or edges).

As agent information and relationships change over time, detecting their future activities in social media networks can be helpful for many applications. First, the detection can assist in identifying anomalies that may indicate potential cyber-threats or promising marketing opportunities. Abnormal changes in messaging behavior may be initial signs of new product hype or rumor spreading, providing intelligence to track viral sensation or adversarial behavior [61]. Second, the detection can facilitate projection of future trends of the evolving network being studied. The results can help managers, policy makers, and intelligence experts to foresee the effects of past and current phenomena [14]. Third, temporal components such as changes in product life cycle and stages of innovation adoption can be revealed in the detection. Further analysis can then be performed to model the changes.

Prior research has examined various aspects of dynamic social networks, such as link prediction, user influence and homophily, community structure, activity modeling, and evolution metrics [18, 50, 55, 67, 68]. However, existing models assume that the nodes of social networks are known and static in time [30, 64]. Though relevant, social and economic theories underlying social network link prediction methods are not examined [41]. In addition, current practices for evaluating predictions in social systems need to be better standardized [27]. For example, existing metrics used to measure temporal social media network activities (e.g., mean absolute percentage error [55]) are biased toward selecting low-value predictions [62].

This research developed and validated two models known collectively as "interaction models" for detecting nodal activities in temporal online social networks. The models incorporated information about agents' activities in a large evolving social network and their interaction over time to support future activity detection. Macroscopic information about the network is used to inform simulation and detection of agent behavior. The models exhibit novelty as they (1) capture dynamic nodal activities from their interactions and address the prediction needs in social media networks; (2) cover multiple, consecutive time points of an evolving network; (3) are grounded in social theories and do not assume static nodes of social networks over time.

The experimental results indicate that the models outperformed three benchmark models, two of which are based on traditional exponential smoothing prediction and one of which is based on a static node assumption. The performance difference is consistent across a variety of temporal configurations. A new metrics developed in this research for evaluating social network activity detection was shown to capture temporal activity patterns better than traditional metrics. We shall describe these metrics and how the interaction models use different assumptions of agent interaction in temporal detection.

The contributions of this research are threefold. First, we developed two new models for temporal social network activity detection and have shown their novelty and robustness in detecting

nodal activities in social media networks. Second, we provided new experimental findings to validate the performance of the proposed models, offering insights into understanding agent behaviors in large-scale social networks. Third, we developed a new metrics for evaluating social network nodal activity detection and built a reusable implementation of temporal activity detection for large online social networks. The methods and findings should be useful to researchers and practitioners.

The rest of the article is structured as follows: Section 2 surveys related works on theoretical and technical aspects of social network analysis. Section 3 describes the four models (two proposed and two benchmark models) and their theoretical and technical rationale for temporal network activity detection. Section 4 provides details of the experimental design. Section 5 reports the experimental results and discusses implications. Section 6 concludes the article and suggests future directions.

2 LITERATURE REVIEW

As social media facilitate formation of temporal online networks, the theoretical and technical aspects of social media analytics are important for research development [31]. Social and behavioral theories can help to guide the research into online social network trend detection [64], where traditional theories may not address specific network characteristics, such as size, participants, diffusion, and evolution. Technical aspects of temporal online social networks are founded on social media analytics [13, 21, 70] and on dynamic graph modeling [32, 69]. In particular, temporal network dynamics, a specialty area in social network analysis, can be used to reveal changes of human activities in a network over time [72]. Table 1 summarizes the theoretical and technical aspects of this review.

2.1 Theoretical Aspects of Social Network Analysis

Theories help us understand why and under what circumstances a social network trend detection method works. A number of social, economic, and behavioral theories serve as the foundations of social network analysis [41]. For instance, homophily theory [45], social impact theory [34], social interaction theory [5], and cognitive balance theory [26] have been used to explain formation of social links. This section highlights theories relevant to our research topics. A comprehensive review of these theories and related theoretical questions can be found in Reference [48] and Reference [43], respectively.

2.1.1 Theories Based on Social Position. Homophily theory [45] states that similar entities create social links more likely than dissimilar entities. The famous saying "birds of a feather flock together" reflects the social choice based on similarity. People tend to link to others similar to them, and such linkage strengthens their similarity. In a social network, similarity can be observed between user profiles, geographic proximities, social engagements, and social relationships [44]. Changes in these features over time need to be incorporated for detecting dynamic trends in the networks. Social impact theory states that when other people are the source of impact and the person is the target, impact is a multiplicative power function of the strength, immediacy, and number of other people [34, 60]. Various aspects have been considered to produce social impact, including social influence, social identity [33, 48], and social capital [9, 53]. Social networks provide the conduit of social impact that may be quantified and tracked over time [72]. Social interaction theory states that people make decisions based on the decisions of their social neighbors [5]. Social position plays a key role in impacting the decisions [1]. A social entity's position in a network creates a tendency for itself to conform to its neighborhood and allows information transmission to its neighbors [58]. Social neighborhood is therefore a foundation of the theory to examine activity of social entities over time.

Table 1. Theoretical and Technical Aspects of Social Media Network Trend Detection

Aspect	Basis	Construct/Explanation	Prior Work
Theoretical Aspect	Social position	Homophily theory: similar entities create social links more likely than dissimilar entities	[McPherson et al. 2001; McPherson and Ranger-Moore 1991; Wu et al. 2017]
		Social impact theory: impact is a multiplicative power function of the strength, immediacy, and number of other people who exert the impact	[Latané 1981; Sedikides and Jackson 1990; Kelman 1958; Ngai et al. 2015; Chang and Chuang 2011; Portes 1998]
		Social interaction theory: people make decisions based on their social neighbors' decisions	[Becker 1974; Akerlof 1997; Salancik and Pfeffer 1978]
	Cognition and perception	Cognitive balance theory: entities that link to a third common entity tend to establish a connection	[Heider 1958; Lee et al. 2013; Zheng et al. 2015]
		Primacy and recency models: people are affected by either primary (early) or recent (later) impressions	[Hovland 1957; Deese and Kaufman 1957; Miller and Campbell 1959]
Technical Aspect	Dynamic Graph Modeling	Modeling of dynamic social networks is done through representing entities as nodes and edges in graphs that change with time	[Xu and Hero 2014; Karrer and Newman 2011; Rohe et al. 2011; Leskovec et al. 2010; Jiang and Chen 2016]
	Link prediction	Learning-based approach: derive a model from existing links and use it to classify each non-connected pair of nodes as positive (linked) or negative (not linked) class	[Wu et al. 2017; Gao et al. 2011; Li et al. 2014; Raghavan et al. 2014]
		Similarity-based approach: compute similarities of non-connected pairs of nodes and create links for top-ranked pairs	[Leskovec et al. 2005; Dunlavy et al. 2019; Gao and Liu 2017; Du et al. 2015]
	Node activity detection	Nodal feature characterization: extract node features to identify activities	[Wei and Carley 2015; Zhu et al. 2013]
		Nodal content characterization: examine topics and messages posted by nodes and structural proximities	[Pinto et al. 2015; Qiu et al. 2014; Antoniades and Dovrolis 2015]

2.1.2 Theories Based on Cognition and Perception. Cognitive balance theory states that two social entities that link to a third common entity tend to establish a connection [26]. In a social network, nodes that link to another well-connected node (or "Rockstar" [35]) are likely to associate with each other [71], thereby strengthening the position of the rock-star over time. Human cognition is also limited by the order of information received [28]. The primacy effect states that people are more affected by first impressions than later ones [17], while the recency effect states that people are more influenced by the latest information that they recall than the earlier one [46]. These effects can be amplified in online social networks over time. Detecting dynamic nodal activities in social networks needs to model these effects explicitly.

2.2 Technical Aspects of Social Network Analysis

Technical developments on social network analysis proliferated since the Internet became available in the 1990s [59, 65] and continued to flourish in the 2000s and beyond [15, 29, 47, 49]. Data

about online social network activities and structure, if handled appropriately, could revolutionize our understanding of collective human behavior [66]. The following reviews various technical approaches for dynamic graph modeling, temporal network dynamics focusing on link prediction and node activity detection.

2.2.1 Dynamic Graph Modeling. Dynamic graph models have been developed to represent changing networks in a variety of applications. Xu and Hero III developed a dynamic stochastic blockmodel that performed competitively with a state-of-the-art algorithm and achieved a higher computational efficiency [69]. Karrer and Newman developed a heuristic blockmodel algorithm for community detection [32]. For a network generated from the Stochastic Blockmodel, Rohe et al. bound the number of nodes "misclustered" by spectral clustering [56]. Rossi et al. developed a temporal behavior model that extracts nodal roles in an evolving network using non-negative matrix factorization [57]. The approach is fully automatic but assumes static nodes and does not consider social theories about user interaction and network positions. Jiang and Chen consider nodal attributes for dynamic network analysis and achieved an increased prediction accuracy [30]. Leskovec et al. developed Kronecker graphs to generate realistic networks using a non-standard matrix operation [38]. A comprehensive survey of statistical network models found limited methodological work focusing on evaluating and comparing the predictive ability of various models [2]. Although prior research on link prediction in the relational networks exists, these works focus on combining heterogeneous data to discover new links.

2.2.2 Link-prediction Approaches. Link prediction aims to predict new links or deleted links between nodes for a future time [64]. Leskovec et al. studied a range of real-world graphs and developed a forest-fire model to produce graphs that exhibit properties observed in these graphs [39]. They observed densification of networks over time, as the log-log plot (number of edges e(t) versus number of nodes n(t)) shows increasing trends across different real-world datasets (ArXiv citation, patent citation, autonomous systems, and affiliation). Wu et al. incorporates social theories to model evolution of social network user-item and user-user interactions [68]. They quantified social influence and homophily effect for users' behaviors by using a joint model of user consumption prediction and social link prediction. Dunlavy et al. [19] show that both matrix- and tensor-based techniques are effective for temporal link prediction. Gao et al. developed a link prediction model that integrates the global network structure, the content of nodes, and the local or proximity information of a given node [25]. Li et al. developed a deep learning framework using conditional temporal restricted Boltzmann Machine that predicts links based on individual transition variance and influence introduced by node neighbors [40].

Researchers have developed dynamic influence models to study social learning and networking in social media websites [51]. Influence is modeled as conditional dependence between an entity's current state and previous states based on a Markovian assumption and a Hidden Markov Model (HMM) [55]. Community strength was measured to reveal robustness and coherence and was applied to discovering community progression and formation [18]. Simulated experiments were conducted to assess a network-based model that represents both spatial and temporal characteristics of human dynamic behavior [24].

2.2.3 Node Activity Detection Approaches. Network nodes have been characterized to reflect their activities over time. Two classes of dynamic network metrics, emergence and persistence, were developed and applied to temporal network activity prediction [67]. Among four models tested, the exponential aggregation model was found to outperform the average aggregation and linear aggregation models. Temporal detection was formulated as weighted time-series forecasting. Pinto et al. developed a trend detection algorithm to identify topical trends in social networks

using a Hawkes process, a self-exciting point process that considers broadcasting times and user-topic interaction [52]. Zhu et al. developed a socially regularized time-decay model for user activity prediction and cast the problem as a classification of active and inactive users [73].

Temporal stability of network nodes was studied in Reference [54], which developed a method to identify active valuable nodes based on static structural properties and spatio-temporal attributes. Experiments on two online social networks with thousands of nodes showed that the method identified valuable nodes in terms of node stability and influence. In another research [3], a simple probabilistic model was developed to capture the probability of tweet-retweet-follow (TRF) events. The model provides descriptive statistics on TRF events but falls short of predictive power.

2.3 Summary of Research Gaps

Our review shows that most technical approaches predict links based on static snapshots of social networks. For instance, the work described in Reference [40] models node transition and local neighbor influence but assumes the nodes to be static over time. This assumption may not be realistic given that people join and leave social networks frequently. Existing works also do not adequately consider features from social theories; and evaluations of link prediction approaches are inadequate, which may lead to wrong conclusions [64]. Prior research does not examine social and economic theories underlying social network link prediction methods [41].

In the prior works about stochastic blockmodeling (e.g., References [38, 69]), model fitting is the primary focus, thus emphasizing explanation at the expense of prediction. Nodal attributes were shown to increase prediction accuracy of network characteristics [30], but not on an individual nodal activity level. In addition, changes in an agent's (nodal) activities are often modeled as a function of the agent's historical behavior (e.g., Reference [67]) without adequately considering network connectivity, social influence, and social theories. Aggregation functions in linear or exponential forms do not account for the rich interaction among agents, thus adversely affecting detection of nodal activities in temporal networks.

3 INTERACTION MODELS FOR DYNAMIC NODAL ACTIVITY DETECTION

To address the research gaps, we developed two theory-based interaction models to support dynamic social network activity detection and provided their instantiations in Algorithms 1 and 2 and their notation and meaning in Table 2. The two interaction models, namely, "Random Interaction Model" and "Preferential Interaction Model," detect agent activities based on different assumptions of agent behavior. Different from existing models (e.g., References [38, 56]), the proposed models exploit interaction among agents who form the network; cover multiple, consecutive time points of an evolving network; and are grounded in social theories and do not assume static nodes of social networks over time. Existing works such as the Erdös-Rényi [20], Gilert [6], preferential attachment [4], and small-world models aim to represent single time-point snapshots of a network [2]. By contrast, our models aim to represent multiple, consecutive time points of an evolving network.

The theoretical rationale of the interaction models can be explained in four aspects. (1) Homophily theory [44, 45] directs the generation of new dynamic network to resemble existing link structure and network density. This is reflected in line 8 of Algorithm 1 and lines 12, 17, 26, 32 of Algorithm 2; (2) The primacy and recency models of human cognition theory suggests that future human activities depend on their preferences on recent memory or first impression (older memory) [17, 28, 46]. These are modeled in lines 6–7 of Algorithm 1 and in lines 5–6 of Algorithm 2; (3) According to social interaction theory, people make decisions based on their social neighbors' decisions [5]. A social entity's position in a network affects information exchange with neighbors [58] and thus impacts his/her activities. Therefore, social positions of nodes reflect activities in

Table 2. Notation and Its Meaning of Terms Used in Algorithms 1 and 2

Notation	Meaning
i	Index of an agent that represents a node in a network
t	A point in time to represent a snapshot of an evolving network and its nodal activities
V_t	A set of vertices at time t
E_t	A set of edges at time t
$L(w)$	A function that specifies an entity spanning w time steps
$G_t^{L(w)}$	A dynamic graph consisting of vertex set $V_t^{L(w)}$ and edge set $E_t^{L(w)}$, representing an evolving network of nodes and links spanning the period $(t-w, t)$
v_t	Number of vertices contained in $V_t^{L(w)}$
e_t	Number of edges contained in $E_t^{L(w)}$
$\hat{v}_{t+1}$	Predicted number of vertices that will be contained in $V_{t+1}^{L(w)}$
$\hat{e}_{t+1}$	Predicted number of edges that will be contained in $E_{t+1}^{L(w)}$
G'	$G_{t+1}^{L(w)}$, a dynamic graph spanning the period $(t-w+1, t+1)$
d'	Density of G′. A ratio of the predicted number of edges in G′ to the predicted number of possible edges in G'
$C_B(i, G')$	Betweenness centrality of Agent(i) in G'
$P(ij)$	Number of shortest paths connecting Agent(i) and Agent(j)
$P_k(ij)$	Number of shortest paths that connect Agent(i) and Agent(j) and that Agent(k) lies on
p_j	Degree ratio: a ratio of the degree of Agent(j) to the maximum degree of any single agent in the network G'
q_i	Complement of degree ratio of Agent(i). Used in evaluation of link removal
$\mathbf{u}$	A set of agents sorted by ascending order of agents' degree ratios
α	A real-valued number $\in$ [0,1] that controls the primacy and recency effects in prediction

a social network [1, 5, 58]. Accordingly, the interaction models represent nodal activities by using the betweenness centrality algorithm that considers nodal position in a network (line 35 of Algorithm 1; line 38 of Algorithm 2); (4) Social impact theory suggests that impact is a multiplicative, power function of strength, immediacy, and number of other people [33, 34, 60]. This directs the development of the selective procedure of forming preferential interaction in Algorithm 2 (lines 7–37).

3.1 Model Notation

This section provides algorithmic notation and steps of the interaction models. Let $G_t^{L(w)}$ be an undirected network that consists of V_t agents (= nodes or vertices) and E_t links (= edges) at time t. Also, let $L(w)$ be the function that specifies the network to span the most recent w time steps at and before time t. The network is modeled as a dynamic graph:

$$G_t^{L(w)} = \left\{V_t^{L(w)}, E_t^{L(w)}\right\}$$

The proposed interaction models were designed to learn from the network characteristics of randomness and preferential interaction due to community size, human bias, declining connection cost, and rising reachability. Three general steps are used in detecting temporal network activities of agents. (1) Initiation: The models identify existing network's configuration (numbers of human agents and relationships) at time t, and the time frame (w days at and before t) during which agent activities are considered in activity prediction. (2) Simulation of network for time $(t+1)$: Based on existing network, the models predict the configuration (numbers of human agents and

relationships) of the network at $(t+1)$ by using a single exponential smoothing procedure. (a) Derive the density of the network at time $(t+1)$ (Equation (1)) based on the predicted network configuration. (b) Alter agents in the network based on projected interaction of agents:

$$D(g) = \frac{\text{Number of links in Network g}}{\text{Maximum possible number of links in Network g}} \tag{1}$$

(3) Predicting activity levels of agents: The prediction uses the betweenness centrality (BC) formula to represent Agent i's activity level [22, 29], as in Equation (2):

$$C_B(i, G') = \sum_{ijk \in G'; i \neq j, k \notin \{i,j\}} \frac{P_k(ij)/P(ij)}{D(G')}. \tag{2}$$

The aforementioned steps are further explained in Sections 3.3, 3.4, and 3.5. Section 3.2 describes two benchmark models used to compare against the proposed interaction models.

3.2 Benchmark Models

Three models—EAM, TAM, and DBMM—were used as benchmarks to compare against the two interaction models. The Exponential Aggregation Model (EAM) was found in Reference [67] to outperform two other temporal prediction models in empirical studies of four different datasets. The method predicts Agent i's activity level by using a single exponential smoothing method that assigns high weights to activities that are closer to the current time step. The assignment is controlled by a smoothing constant $\alpha \in [0, 1]$ that reflects recency and primacy effects according to cognitive theories [17, 46]. High values of α indicate that recent data are more emphasized in prediction.

EAM has an implicit assumption that an existing trend will continue to be extended in the same intensity as in reference history. However, this may not be always true due to multiplicative effects of human influence [34, 60]. Therefore, we developed another benchmark model named Trend-adjusted-exponential Aggregation Model (TAM), which includes a trend adjustment term $\frac{m_t - m_1}{t-1}$ to incorporate the changing temporal trends present in the community. TAM represents activity of a node as follows:

$$\text{Activity}^{T(\alpha \in [0,1])}(\mathbf{m} \in \mathbb{R}^d) = \left[\alpha^{d-1} m_1 + \sum_{j=2}^{t} (1-\alpha)\alpha^{d-j} m_j\right] + \frac{m_t - m_1}{t-1},$$

where m_t represents actual activity level of the node at time t; α is a parameter controlling primacy and recency effects; and d is the size of time window of reference history.

In addition, the Dynamic Behavioral Mix-membership Model (DBMM) [57] was used to compare against the interaction models to examine the accuracy of role discovery of agents in temporal social networks. Different from EAM and TAM, DBMM does not predict Agent i's activity level but predicts the likelihood of possessing different roles in the network. In DBMM, a network at time t is represented as a node-feature matrix using the approximation $\boldsymbol{G}_t \boldsymbol{F} \approx \boldsymbol{V}_t$, where each row of $\boldsymbol{G}_t$ (an n (= node count) $\times$ r (= role count) matrix) represents an agent (node)'s membership in different roles and each column of $\boldsymbol{F}$ (an $r \times f$ (= feature count) matrix) represents the extent to which a role contributes to feature values. DBMM consists of three steps: network feature extraction, probabilistic role discovery, and behavioral transition modeling. According to Rossi et al. [57], the summary transition model (which defines a summary behavioral snapshot $\boldsymbol{G}_{S(t)}$ shown below) used in DBMM achieved best performance (with $k = 10$, $\alpha = 0.7$) and is thus used in our

implementation:

$$G_{S(t)} = \alpha^{w-1} G_k + \sum_{j=2}^{w} (1-\alpha)\alpha^{w-j} G_j,$$

where $k = t - w$ and w, α are window size and parameter to determine contribution of previous graph snapshots in the model. The prediction for the next network at $(t+1)$ is computed as:

$$\hat{G}_{t+1} = G_t \; G_{S(t)}.$$

3.3 Random Interaction Model

The Random Interaction Model (RIM) detects the activity level of Agent i at time $(t+1)$ by computing his level of interaction with other agents in $G_{t+1}^{L(w)}$, the predicted network at time $(t+1)$. This network is simulated based on the numbers of agents and links predicted at time t using a single exponential smoothing method with $\alpha \in [0, 1]$ (lines 6–7 of Algorithm 1). As the simulated network varies from the observed network at time t, the model adds or deletes agents and links based on a uniform random distribution. The algorithm controls the network composition by using a random variable $r \in [0, 1]$ (lines 10–34 of Algorithm 1). The predicted level of activity of Agent i at time $(t+1)$ is then computed as his betweenness centrality in $G_{t+1}^{L(w)}$ (see Section 3.5 for activity measurement).

3.4 Preferential Interaction Model

The Preferential Interaction Model (PIM) detects the activity of Agent i at time $(t+1)$ by computing his level of interaction with other agents in $G_{t+1}^{L(w)}$, the predicted network at time $(t+1)$. The structure of simulated network at time $(t+1)$ follows that of the observed network at time t, plus adding or deleting agents and links based on an assumption of preferential attachment. PIM differs from RIM mainly by the use of the degree ratio p_j and removal likelihood q_i. Degree ratio measures the connectivity of an agent in the social network (lines 12 and 26 of Algorithm 2). Removal likelihood indicates the likelihood of an agent being removed (lines 17 and 32 of Algorithm 2). Agents with high degree ratios are more preferable to be connected by other agents. This process simulates a community in which users tend to attach to those who already have high connectivity. The predicted level of activity of Agent i at time $(t+1)$ is then computed as his betweenness centrality in $G_{t+1}^{L(w)}$. See Algorithm 2 for the detailed computation. Different from prior work, PIM explicitly models the temporal process of link alteration and allows flexible simulation of future networks by various parameters.

3.5 Measurement of Agent Activity

There are many metrics to measure the activity level of an agent, such as degree, centrality, and clustering coefficient [29, 59]. In this research, we used Betweenness Centrality (BC) [22] to indicate an agent's activity level, which reflects his influence in a networked community due to its emphasis on connectivity and agent interaction. The non-normalized BC of an Agent k is calculated as the proportion of the number of geodesics between any two distinct agents ($\neq k$; whose connecting paths contain k) to the total number of geodesics between any two distinct agents. The proposed models compute a normalized version of BC to be within [0, 1] by dividing the non-normalized BC by the density of the network (line 35 of Algorithm 1; line 38 of Algorithm 2). Agents with a high BC score play an important role of bridging disparate members and transferring information or goods across different nodes. Different from other measures such as closeness centrality (which measures "time-until-arrival" in network flow), BC reflects the level of agents'

ALGORITHM 1: Predicting the activity of Agent i with Random Interaction model

Input: i, G, α, w, t
Output: Activity of Agent i at time $(t+1)$
Initialize:
$V_{t+1}^{L(w)} \leftarrow \{\}; E_{t+1}^{L(w)} \leftarrow \{\}; G' \leftarrow \{V_{t+1}^{L(w)} E_{t+1}^{L(w)}\}$
Let $G' = G_{t+1}^{L(w)}, v_t = |V_t^{L(w)}|, \hat{v}_{t+1} = |V_{t+1}^{L(w)}|, e_t = |E_t^{L(w)}|, \hat{e}_{t+1} = |E_{t+1}^{L(w)}|$
foreach $i \in V_t^{L(w)}$ **do** $V_{t+1}^{L(w)} \bigcup$ Agent(i);
foreach $i \in E_t^{L(w)}$ **do** $E_{t+1}^{L(w)} \bigcup$ Link(i);
// Predictions for numbers of agents and links and for density
$\hat{v}_{t+1} := \alpha v_t + (1-\alpha)\hat{v}_t, \qquad \hat{v}_2 := v_1$
$\hat{e}_{t+1} := \alpha e_t + (1-\alpha)\hat{e}_t, \qquad \hat{e}_2 := e_1$
Density(G') = $d' = \frac{2\hat{e}_{t+1}}{\hat{v}_{t+1}(\hat{v}_{t+1}-1)}$
// Prediction for G'
if $\hat{v}_{t+1} > v_t$ **then**
  // Add new agents
  for $i = v_t + 1$ **to** $\hat{v}_{t+1}$ **do**
    $V_{t+1}^{L(w)} \bigcup$ Agent(i)
    foreach $j \in V_{t+1}^{L(w)}$ **do**
      Generate a random number $r \in [0 \ldots 1]$
      if $r \leq d'$ **then** $E_{t+1}^{L(w)} \bigcup$ Link(ij);
else if $\hat{v}_{t+1} < v_t$ **then**
  // Delete agents
  foreach $i \in V_{t+1}^{L(w)}$ **do**
    Generate a random number $r \in [0 \ldots 1]$
    if $r \leq \frac{v_t - \hat{v}_{t+1}}{v_t}$ **then**
      $V_{t+1}^{L(w)} \leftarrow V_{t+1}^{L(w)} -$ {Agent(i)}
      $E_{t+1}^{L(w)} \leftarrow E_{t+1}^{L(w)} -$ {Links of Agent(i)}
else
  if $\hat{e}_{t+1} > e_t$ **then**
    foreach $i, j \in V_{t+1}^{L(w)}$ *where* $i \neq j$ **do**
      Generate a random number $r \in [0 \ldots 1]$
      if {Link(ij)} $\notin E_{t+1}^{L(w)}$ *and* $r \leq d'$ **then** Add Link(ij) ;
  else if $\hat{e}_{t+1} < e_t$ **then**
    foreach $j \in E_{t+1}^{L(w)}$ **do**
      Generate a random number $r \in [0 \ldots 1]$
      if $r \leq \frac{e_t - \hat{e}_{t+1}}{e_t}$ **then** Delete Link(j) ;
  else
    do nothing
$C_B(i, G') = \sum_{ijk \in G'; i \neq j, k \notin \{i,j\}} \frac{P_k(ij)/P(ij)}{(\hat{v}_{t+1}-1)(\hat{v}_{t+1}-2)/2}$
return $C_B(i, G')$

activity through interaction and influence in the network (similar to the notion of frequency of arrival in network flow) [7].

4 EXPERIMENTAL DESIGN

This section describes the design of a series of experiments to assess the models' predictive power systematically under three temporal settings: (1) daily activity prediction over consecutive days, (2) varying sizes of reference windows w, and (3) different events that affect agent activities. The

ALGORITHM 2: Predicting the activity of Agent i with Preferential Interaction model

Input: i, G, α, w, t
Output: Activity of Agent i at time $(t+1)$
Initialize:
$V_{t+1}^{L(w)} \leftarrow \{\}; E_{t+1}^{L(w)} \leftarrow \{\}; G' \leftarrow \{V_{t+1}^{L(w)} E_{t+1}^{L(w)}\}$
Let $G' = G_{t+1}^{L(w)}, v_t = |V_t^{L(w)}|, \hat{v}_{t+1} = |V_{t+1}^{L(w)}|, e_t = |E_t^{L(w)}|, \hat{e}_{t+1} = |E_{t+1}^{L(w)}|$
foreach $i \in V_t^{L(w)}$ **do** $V_{t+1}^{L(w)} \bigcup$ Agent(i);
foreach $i \in E_t^{L(w)}$ **do** $E_{t+1}^{L(w)} \bigcup$ Link(i);
$\hat{v}_{t+1} := \alpha v_t + (1-\alpha)\hat{v}_t, \qquad \hat{v}_2 := v_1$
$\hat{e}_{t+1} := \alpha e_t + (1-\alpha)\hat{e}_t, \qquad \hat{e}_2 := e_1$
if $\hat{v}_{t+1} > v_t$ **then**
  $\mathbf{u}$ = SortByAscendingDegreeRatio(Agent $\leftarrow V_t^{L(w)}$)
  for $i = v_t + 1$ **to** $\hat{v}_{t+1}$ **do**
    $r \sim U([0,1])$; $V_{t+1}^{L(w)} \bigcup$ Agent(i) ; $k = \lceil r \times v_t \rceil$
    ; $p_i = p_{u_k}$ // Degree ratio of i
    foreach $j \in V_{t+1}^{L(w)}$ **do**
      $r \sim U([0,1])$; $p_j = \frac{Deg(j)}{\max(Deg(V_{t+1}^{L(w)}))}$
      // Degree ratio of j
      if $r \leq p_j$ **then** $E_{t+1}^{L(w)} \bigcup$ Link(ij);
else if $\hat{v}_{t+1} < v_t$ **then**
  $c = 0$
  foreach $i \in V_{t+1}^{L(w)}$ **do**
    $r \sim U([0,1])$; $q_i = 1 - \frac{Deg(i)}{\max(Deg(V_{t+1}^{L(w)}))}$
    // Removal likelihood of i
    if $r \leq q_i$ *and* $c < (v_t - \hat{v}_{t+1})$ **then**
      $V_{t+1}^{L(w)} \leftarrow V_{t+1}^{L(w)} - \{$Agent$(i)\}$
      $E_{t+1}^{L(w)} \leftarrow E_{t+1}^{L(w)} - \{$Links of Agent$(i)\}$
      $c \leftarrow c + 1$
else
  $c = 0$
  if $\hat{e}_{t+1} > e_t$ **then**
    foreach $i, j \in V_{t+1}^{L(w)}$ *where* $i \neq j$ **do**
      $r \sim U([0,1])$; $p_i = \frac{Deg(i)}{\max(Deg(V_{t+1}^{L(w)}))}$;
      $p_j = \frac{Deg(j)}{\max(Deg(V_{t+1}^{L(w)}))}$
      if $\{$Link$(ij)\} \notin E_{t+1}^{L(w)}$ *and* $r \leq$
      $max(p_i, p_j)$ *and* $c < (\hat{e}_{t+1} - e_t)$
        then Add Link(ij) ;
      $c \leftarrow c + 1$
  else if $\hat{e}_{t+1} < e_t$ **then**
    foreach $i, j \in V_{t+1}^{L(w)}$ *where* $i \neq j$ **do**
      $r \sim U([0,1])$; $q_i = 1 - \frac{Deg(i)}{\max(Deg(V_{t+1}^{L(w)}))}$;
      $q_j = 1 - \frac{Deg(j)}{\max(Deg(V_{t+1}^{L(w)}))}$
      if $\{$Link$(ij)\} \in E_{t+1}^{L(w)}$ *and* $r \leq$
      $max(q_i, q_j)$ *and* $c < (e_t - \hat{e}_{t+1})$
        then Delete Link(ij) ;
      $c \leftarrow c + 1$
  else
    do nothing
$C_B(i, G') = \sum_{ijk \in G'; i \neq j, k \notin \{i,j\}} \frac{P_k(ij)/P(ij)}{(\hat{v}_{t+1}-1)(\hat{v}_{t+1}-2)/2}$
return $C_B(i, G')$

following sections describe the data used in the experiments, the experimental setup, and performance evaluation measures (see Figure 1).

4.1 Data Description

The data used in the experiments were collected from a social media community formed by users who posted messages on the Twitter website. These messages were collected by using a set of carefully selected keywords (see Reference [15] for details of the selection) obtained from reviewing literature on U.S. border security and immigration policy [8, 23, 36, 63]. Examples of the queries include "immigration reform," "US border security," and "US immigration." Between May 2013 and September 2015, our system has collected over 3,286,473 tweets posted by over 790,462 users who formed more than 3,055,797 links. An undirected link is formed between two users (A and B) when User A sends a tweet targeted to User B, or re-tweets another tweet written by User B, or modifies and then sends out a tweet written by User B [15]. These undirected links can represent the transparent and public nature of the links formed in Twitter. The links and the connected users are represented as edges and vertices, respectively, in an undirected graph G. A temporal graph $G_t^{L(w)}$ is formed by considering all the vertices and edges formed within the most recent w days at and prior to time t (see Section 3.1 for notation).

U.S. border security and immigration was chosen as our topic of study because of its timeliness and the dynamic networked community on social media. For example, these sample tweets reveal the diverse opinions expressed in the community:

	RealDonaldTrump: "Amnesty lowers wages and invites more lawlessness. Obama has unilaterally cancelled any chance of immigration reform."
TOP TEXAS ORGANIZING PROJECT	*TXOrganizingProject: "Today is Camino Americano, a rally in Washington D.C. demanding comprehensive immigration reform with a path to ..."*

Following Twitter's guidelines of data collection and data privacy protection, our automated system continually collected on average 4,650 tweets per day using the selected queries. Through searching, retrieving, parsing, filtering, and storing, the system extracts temporal social network data (represented as temporal graphs) from the unstructured social media data. Figure 1 provides a high-level description of the data collection process and the experimental setup that was implemented in a distributed computing environment.

4.2 Datasets and Experimental Setup

Through a series of experiments, we tried to answer these research questions: Which models perform the best in temporal network activity detection in terms of different metrics and window sizes? What is the effect of window size on the performance? How do the structural properties of the temporal networks correlate with the performance? The experiments used two datasets covering two periods of time to represent two different sets of events:

(i) *Dataset 2013*: The dataset contains 316,686 messages posted by 111,212 users during Sept. 1–Oct. 31, 2013. This period was marked by high fluctuation of public sentiment due to strong disappointment on CIR bill impasse in the U.S. House of Representatives and a massive protest occurred on October 6, 2013, in Washington, DC, that led to arrests of hundreds of people including prominent Congress members [16].

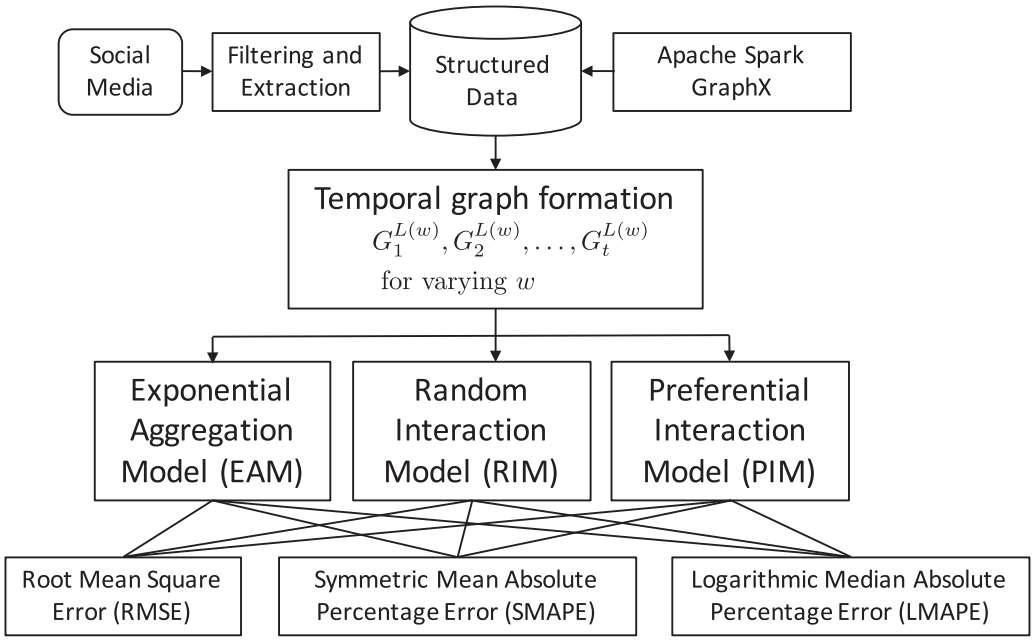


Fig. 1. Data collection and experimental setup.

(ii) *Dataset 2015*: The dataset contains 137,363 messages posted by 64,583 users during May 1–June 30, 2015. This period marked a turbulent time during which Donald Trump announced (on June 16, 2015) his candidacy for the U.S. presidency and made controversial remarks about border security in the U.S. Other candidates including Hillary Clinton and 16 Republican candidates announced their candidacies mostly in the same year. The analysis of this dataset can help to understand the effect of the U.S. presidential election on social media responses.

Two parameters in the models require tuning and selecting values: the window size (w) and the smoothing parameter (α). The value of w controls how many days of previous data were used to predict the value of the period's next day. The value of α (ranging from 0 to 1) determines the weight assigned to primacy or recency effects in the prediction. Given our research focus on the predictive and simulation capabilities of the interaction models, a comprehensive study of all possible combinations of the parameters and their sensitivities is beyond the scope of this research.

4.2.1 Model Performance for Various Dates with a Specified w. First, the four models (PIM, RIM, EAM, and TAM) were evaluated over a series of dates on which predictions were made based on a pre-defined w days of data. In the 61-day coverage of each dataset, we ran the models to predict once per three days. We also selected w to be 18 so the models had the same reference window to produce predictions. This size was set empirically to produce high efficiency and significance of predictions. Hence, the predictions were made for the activity levels on the 13th day of the first month, and then the 16th, 19th, 22nd, and so on. This procedure generated 17 sets of predicted values for the 2013 dataset and 16 sets for the 2015 dataset. Each set contains 12 ($= 4 \times 3$) values that are the combinations of the four models and three metrics (RMSE, SMAPE, and LMAPE). For each model, we empirically determined the optimal smoothing constant α^* that enabled the model to achieve its best predictive performance within the period.

4.2.2 Effect of Temporal Window Size. Second, the four models were evaluated over a set of different temporal window sizes w (i.e., time spans in days), where $w = 4, 5, 6, \ldots, 20$. Using these window sizes, the model performances for predicting each period's last activity levels (recorded on 10/30/2013 for Dataset 2013 and on 6/28/2015 for Dataset 2015) were aggregated and compared empirically to identify statistical differences.

4.2.3 Accuracy in Behavioral Role Discovery. Third, the interaction models were compared against the Dynamic Behavioral Mix-membership Model (DBMM) [57] to evaluate the accuracy in discovering agents' behavioral roles. Each of the three models (DBMM, PIM, and RIM) was run, respectively, using the 2013 and 2015 datasets, each with two different window sizes (w = 12 and w = 18). To enable PIM and RIM to predict agent roles, the first two steps of DBMM (network feature extraction and probabilistic role discovery) were applied to predict node membership $\hat{G}_{t+1}$ with the parameter values (k = 10, $\alpha = 0.7$) recommended in Reference [57].

4.2.4 Structural Analysis of Model Performances. Fourth, we analyzed the relationship between performance of the four models (PIM, RIM, EAM, and TAM) and structural properties of the temporal social networks involved. Five structural properties were selected to provide information of the social networks: total number of vertices (V), total number of edges (E), network density (D), global clustering coefficient (C), and number of connected components (M). Their formulas are available in Reference [29].

4.3 Performance Evaluation

A measure that accounts for agent position in a network, the betweenness centrality (BC) score, represents the ability of an agent in bridging other users in the social media network. Due to its capability to model user interaction in social media networks, it is used as the ground truth to evaluate our models' performances (see Section 3.5 for further explanation). Based on this measure, four metrics were used to compare between the interaction models and benchmarks.

4.3.1 Root Mean Square Error (RMSE). Root Mean Square Error (RMSE) is one of the most frequently used measures of predictive performance. RMSE is calculated as the square root of the average of the squared differences between the actual values ($A_{t,i}$) and predicted values ($F_{t,i}$). n_t is the number of users in the network at time t. The RMSE formula is shown in Equation (3):

$$RMSE(t) = \sqrt{\frac{\sum_{i=1}^{n_t} (A_{t,i} - F_{t,i})^2}{n_t}} \tag{3}$$

4.3.2 Symmetric Mean Absolute Percentage Error (SMAPE). Symmetric Mean Absolute Percentage Error (SMAPE) is a variant of Mean Absolute Percentage Error (MAPE), which expresses accuracy as a percentage value instead of absolute value to adjust for extreme values in the calculation. Different from MAPE, SMAPE uses the average of the absolute actual and forecast values as the denominator to further smooth the deviation between the two. Equation (4) shows the formula:

$$SMAPE(t) = \frac{100\%}{n_t} \sum_{i=1}^{n_t} \frac{|A_{t,i} - F_{t,i}|}{(|A_{t,i}| + |F_{t,i}|)/2} \tag{4}$$

4.3.3 Logarithmic Mean Absolute Percentage Error (LMAPE). Research shows that MAPE and its variants (such as SMAPE) are biased toward predictions that are too low [62]. As the activity levels of temporal network agents can change dramatically, the predictive models may be required to produce extreme values that may create biases if MAPE is used. To address the problem, we

developed a new measure called Logarithmic Mean Absolute Percentage Error (LMAPE; see Equation (5)). LMAPE is designed to provide unbiased evaluation of predictive models that may make extreme forecasts. The measure considers the absolute differences of the logarithm of actual and forecast values and divides the differences by the geometric means of the actual values. The use of geometric mean helps to alleviate the reliance of arithmetic mean that is found to cause bias in evaluation [62]:

$$LMAPE(t) = \frac{100\%}{n_t} \sum_{i=1}^{n_t} \frac{|\ln(A_{t,i}) - \ln(F_{t,i})|}{GM(A_{t,i})}, \tag{5}$$

$$\text{where } GM(A_{t,i}) = \left(\prod_{i=1}^{n_t} A_{t,i} \right)^{1/n_t}$$

4.3.4 Frobenius Loss (FL). Frobenius Loss (FL) can be used to reflect the approximation error between the estimated agent membership $\hat{\boldsymbol{G}}_{t+1}$ and the true agent membership $\hat{\boldsymbol{G}}_{t+1}$ when the models are used to predict agent roles. Equation (6) shows the computation of FL [57]:

$$FL(t+1) = \|\boldsymbol{G}_{t+1} - \hat{\boldsymbol{G}}_{t+1}\|_F \tag{6}$$

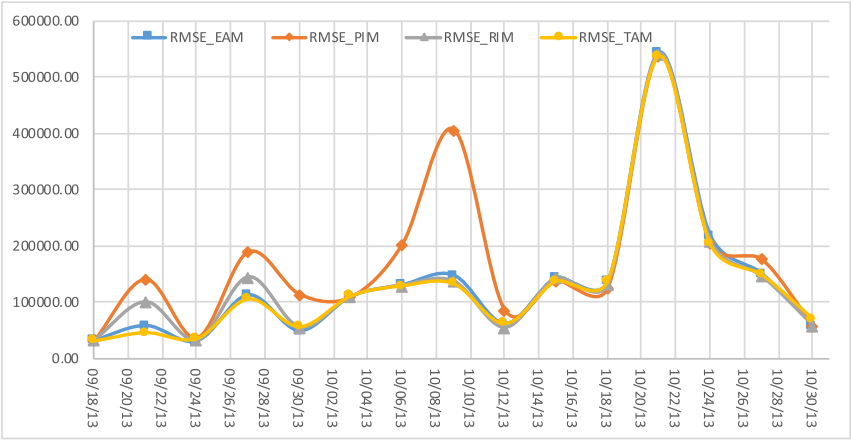
5 EXPERIMENTAL FINDING

This section reports results of a series of experiments comparing among the models using different time periods, window sizes, and performance measures. We provide statistical findings and charts to explain the performance and to support the observations. We analyze relationship between temporal properties of the networks and the results, with a view to explain the performance differences. To provide further contextual explanation, we conducted case studies to show agents' interaction patterns within the temporal social media networks.

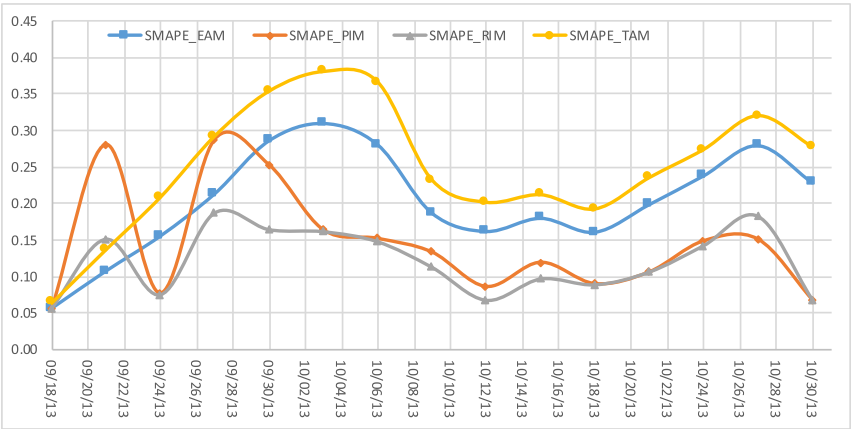
5.1 Model Comparison across Different Dates

Figures 2 and 3 show comparisons of model performance by date across the two datasets (2013 and 2015) using three metrics (RMSE, SMAPE, and LMAPE; smaller values indicate better performance; $n = 15$ (number of data points) for Dataset 2013 (from 9/18/13 to 10/30/13 with a three-day increment), $n = 14$ for Dataset 2015 (from 5/18/15 to 6/27/15 with a three-day increment), and $w = 18$ for both datasets (this value of w was selected to provide the best performance within a dataset)). The performance comparison for both datasets indicates that PIM and RIM generally achieved better performance than EAM and TAM in terms of SMAPE and LMAPE (see Figure 2(b–c) and Figure 3(b–c)), whereas performance of TAM and EAM was generally worse than the other models. Based on RMSE, the four models produced similar performance. Therefore, as shown in Figure 2(a) and Figure 3(a), these differences are not clearly distinguished.

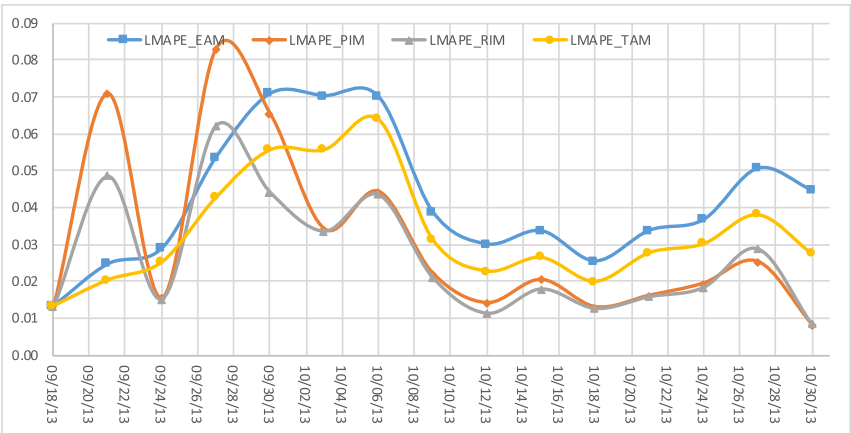
To identify the significance of the performance differences, we conducted statistical comparisons of model performances using the two datasets. The Welch two-sample pairwise t test was used, assuming unequal variances of the sample data. The results are shown in Table 3. Each pair of values embraced within parentheses refers to performance values of two models being compared (e.g., the LMAPEs of EAM and RIM are 0.042 and 0.026, respectively, for Dataset 2013). Each value is the average of a model's performance over all dates studied in a dataset. A lower value indicates better performance of the respective model (e.g., RIM in the aforementioned example) and is underlined if the performances of the models are significantly different (as shown in a significant p-value). Significant results are marked with one or more asterisks next to the p-values of the respective comparison (e.g., p = 0.0024 for EAM vs. RIM based on LMAPE).



(a): y-axis = RMSE, x-axis = date

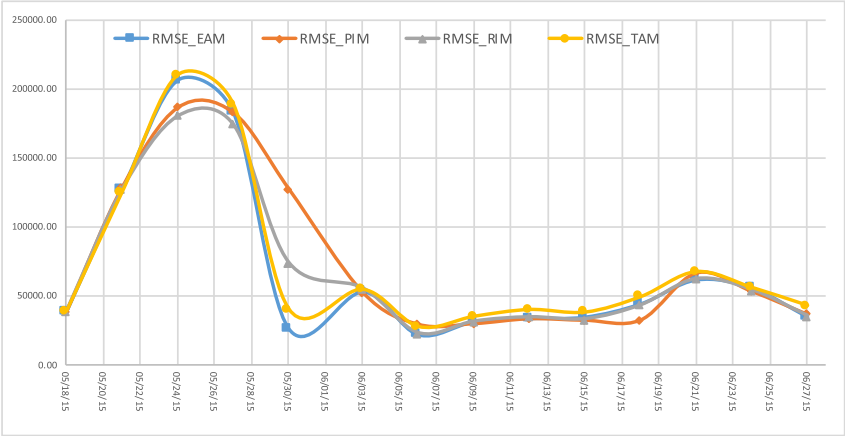


(b): y-axis = SMAPE, x-axis = date

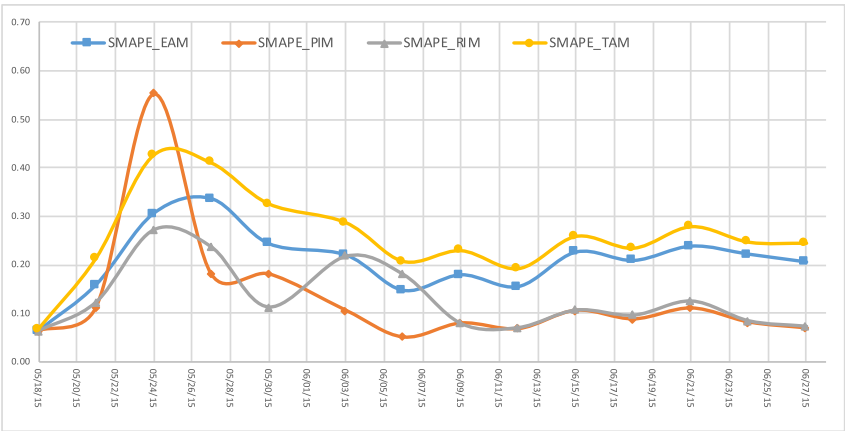


(c): y-axis = LMAPE, x-axis = date

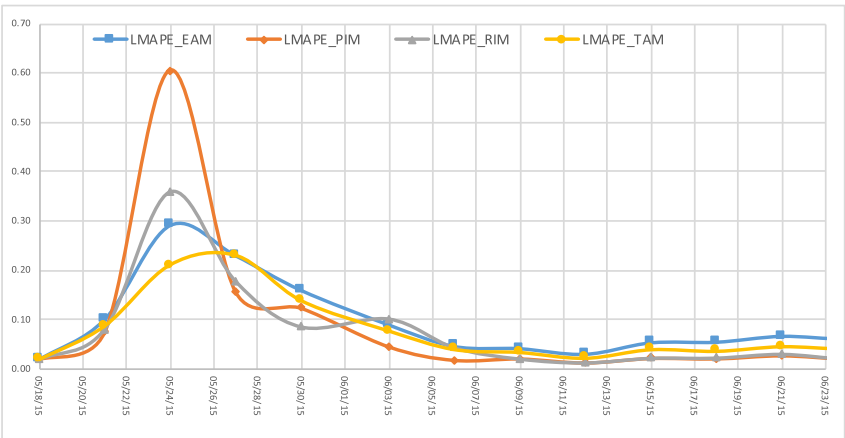
Fig. 2. Model comparison across different dates for Dataset 2013.



(a): y-axis = RMSE, x-axis = date



(b): y-axis = SMAPE, x-axis = date



(c): y-axis = LMAPE, x-axis = date

Fig. 3. Model comparison across different dates for Dataset 2015.

Table 3. Performance and P-values of Model Comparison by Date

Models	RMSE	SMAPE	LMAPE
Dataset 2013	*Performance*		
EAM vs. RIM	(132744.58, 135077.55)	(0.2024, 0.1201)	(0.0416, 0.0264)
EAM vs. PIM	(132744.58, 170775.41)	(0.2024, 0.1445)	(0.0416, 0.0310)
PIM vs. RIM	(170775.41, 135077.55)	(0.1445, 0.1201)	(0.0310, 0.0246)
TAM vs. RIM	(130429.73, 135077.55)	(0.2495, 0.1201)	(0.0416, 0.0264)
TAM vs. PIM	(130429.73, 170775.41)	(0.2495, 0.1445)	(0.0416, 0.0310)
Dataset 2013	*p-value*		
EAM vs. RIM	0.5477	4.02028E-05***	0.0024**
EAM vs. PIM	0.0543$^{\prime}$	0.0219*	0.0848$^{\prime}$
PIM vs. RIM	0.0681$^{\prime}$	0.0531$^{\prime}$	0.0618$^{\prime}$
TAM vs. RIM	0.3233	2.95072E-06***	0.0706$^{\prime}$
TAM vs. PIM	0.0526$^{\prime}$	0.0007***	0.6694
Dataset 2015	*Performance*		
EAM vs. RIM	(68492.15, 69267.52)	(0.2075, 0.1320)	(0.0953, 0.0734)
EAM vs. PIM	(68492.15, 74151.85)	(0.2075, 0.1323)	(0.0953, 0.0853)
PIM vs. RIM	(74151.85, 69267.52)	(0.1323, 0.1320)	(0.0853, 0.0734)
TAM vs. RIM	(72699.44, 69267.52)	(0.2587, 0.1320)	(0.0771, 0.0734)
TAM vs. PIM	(72699.44, 74151.85)	(0.2587, 0.1323)	(0.0771, 0.0853)
Dataset 2015	*p-value*		
EAM vs. RIM	0.8515	0.0003***	0.0345*
EAM vs. PIM	0.4680	0.0161*	0.6986
PIM vs. RIM	0.2417	0.9906	0.5355
TAM vs. RIM	0.3486	9.82525E-06***	0.7787
TAM vs. PIM	0.8360	0.0069**	0.7887

Note: An underlined value indicates significantly better performance of a model in the pair. Asterisks indicate statistical significance: $^{\prime}$p < 0.1; *p < 0.05; **p < 0.01; ***p < 0.001.

The results show that PIM and RIM outperformed both TAM and EAM significantly based on SMAPE (for both 2013 and 2015 datasets) and on LMAPE (for five of eight cases in the two datasets). By contrast, the comparisons between RIM and PIM yielded diverse results. Using the 2013 dataset, RIM outperformed PIM significantly only at the $p < 0.1$ level (but not at $p < 0.05$ or finer levels). Using the 2015 dataset, the performances of the RIM and PIM are not significantly different based SMAPE and LMAPE. Because RMSE is suitable mainly for comparing among linear models, performance differences among models that describe both exponential and network effects (e.g., RIM and PIM) may not be captured precisely by RMSE. Thus, many of the differences in RMSE are not significant.

We observe a small number of spikes in errors on certain dates (e.g., 9/27/2013, 10/21/2013 in Figure 2(a)). These spikes may be caused by computational operations in specific metrics and by the interaction models' adjustment in nodes and links. *First*, the squaring operation in RMSE may contribute to the spikes shown in Figure 2(a). For instance, on 10/21/2013 shown in the figure, the social media network consists of 49,725 vertices and 120,295 edges, which make it one of the largest networks in the 2013 dataset. Together with randomness of user behavior, the size and complexity of networks contribute to amplify errors that are computed through the squaring operation in the RMSE metrics. By contrast, such amplification is not as prominent in Figure 2(b) and in Figure 2(c), which do not use squaring in SMAPE and LMAPE, respectively. *Second*, the

node and link adjustment processes used in PIM and RIM may contribute to the spikes shown in Figure 2 and Figure 3. For instance, for the spikes shown on 9/21/2013, 9/27/2013, and 5/24/2015 in Figure 2(b–c) and Figure 3(a–c), all these dates have bursty upward trends of users and links, for which RIM and PIM are required to create new agents and links in the simulated networks at $t + 1$ (see Algorithms 1 and 2). These data might have contributed to larger-than-required predictions (and thus the spikes) of the interaction models. Table 10 provides additional information about the structural information of the networks on these dates. By contrast, EAM assumes nodes to be static and does not have the similar dynamic adjustment process that PIM and RIM have. As a result, EAM tends to make relatively smooth prediction, producing fewer spikes in errors. Future studies may explore the mechanism to adjust nodes and links to increase prediction accuracy.

5.2 Model Comparison across Different Window Sizes

Figures 4 and 5 show comparisons of model performances by window size across the two datasets using three metrics (RMSE, SMAPE, and LMAPE). For each window size (ranging from 4 to 20 with an increment of 1), the best alpha (that resulted in highest performance values) was used for each model in its running. The optimal alpha value is identified for each experimental configuration (with one model, one window size, one prediction date, and one metrics) by running a given model five times using five alpha values: [0.1, 0.3, 0.5, 0.7, 0.9], and then choosing the alpha that yielded the best performance (lowest error) in that configuration. The prediction periods are same as those listed in Section 5.1.

The performance comparison for both datasets (Figure 4(b–c) and Figure 5(b–c)) indicates that, in terms of SMAPE and LMAPE, PIM and RIM outperformed EAM and TAM in almost all window sizes. Based on RMSE, the four models produced similar performances (Figure 4(a) and Figure 5(a)). A possible reason is that RMSE fails to explain the network and exponential effects of PIM and RIM and treated all four models similarly.

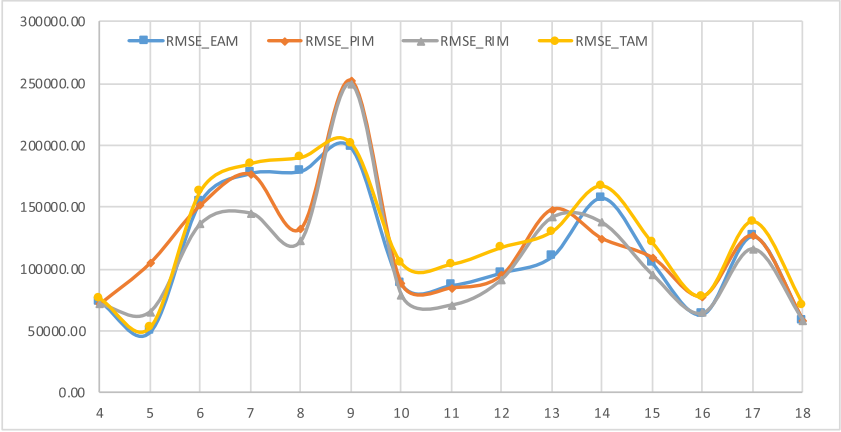
Table 4 shows the results of statistical pairwise comparisons of the model performances. As in Table 3, each pair of values embraced within parentheses refers to performance values of two models being compared. Each value is the average of a model's performance for predicting the last activity levels in each dataset (activity levels on 10/30/2013 for Dataset 2013 and on 6/28/2015 for Dataset 2015) using different window sizes ($w = 4, 5, 6, \ldots, 20$). A lower value indicates better performance of the respective model.

The results show that both RIM and PIM outperformed EAM and TAM significantly based on both SMAPE and LMAPE, while no significant difference in model performances was found based on RMSE. Thus, H3 was rejected in eight of the twelve cases.

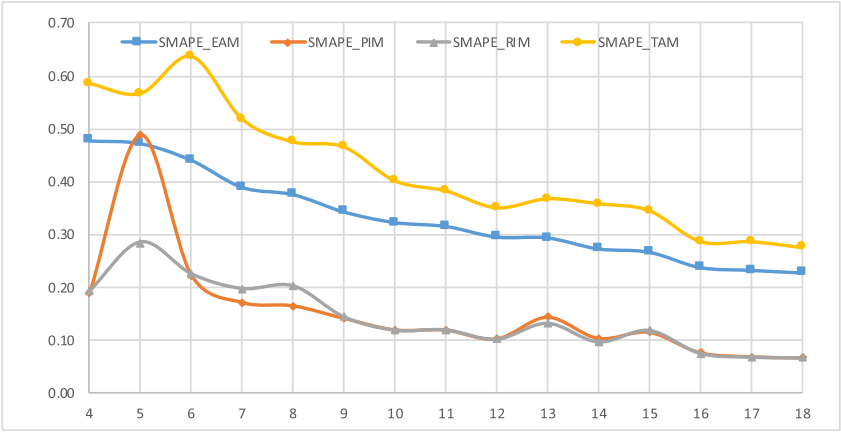
Using the 2013 dataset, RIM outperformed PIM significantly based on RMSE. Using the 2015 dataset, PIM outperformed RIM significantly based on all three metrics (RMSE, SMAPE, and LMAPE). Due to inconsistent results found in two datasets, additional analyses of the environments and structural properties are needed to explain these findings.

5.3 Model Comparison Using Frobenius Loss

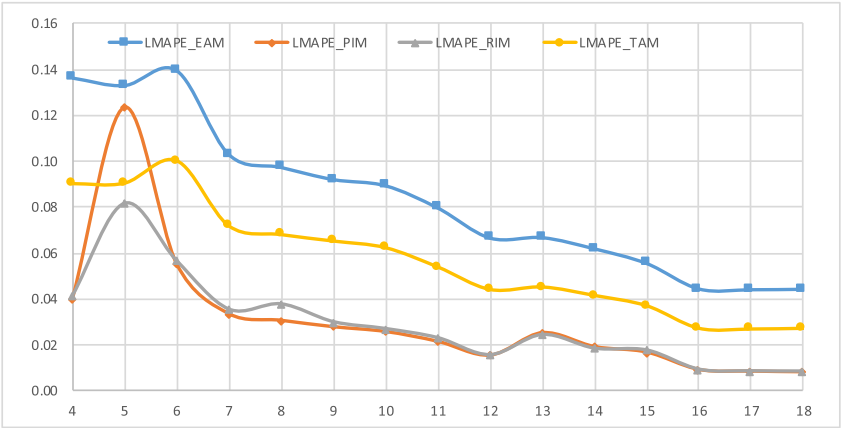
As shown in Table 5, the interaction models achieved on average a lower Frobenius Loss (FL) than DBMM did in all the eight comparisons in both the 2013 and 2015 datasets, indicating a generally better capability in agents' behavioral role discovery. The two interaction models significantly outperformed DBMM in all four comparisons ($p < 0.001$) in the 2015 dataset. PIM outperformed DBMM significantly ($p = 0.01576$) in the 2013 dataset ($w = 12$, $n = 40$), and RIM obtained a better FL than DBMM did with moderate p values (0.1005 and 0.1551 for $w = 12$ and $w = 18$, respectively). One reason for the superior performance of the interaction models is their realistic adjustment of node composition in evolving networks. This is especially true in online social media in which the



(a): y-axis = RMSE, x-axis = *w*

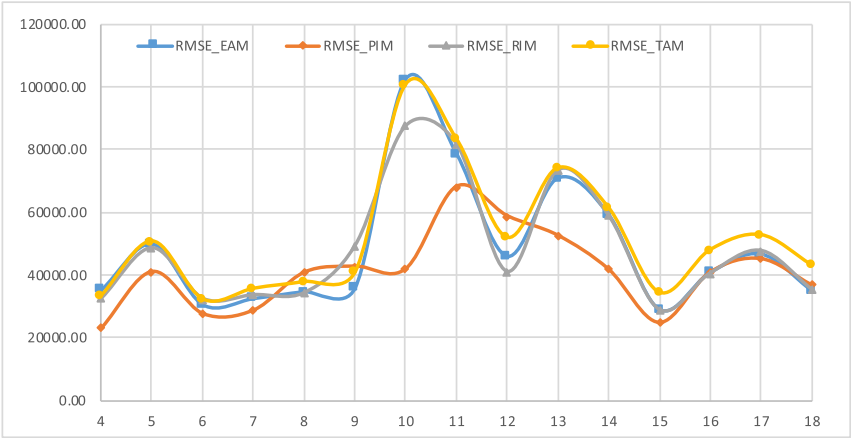


(b): y-axis = SMAPE, x-axis = *w*

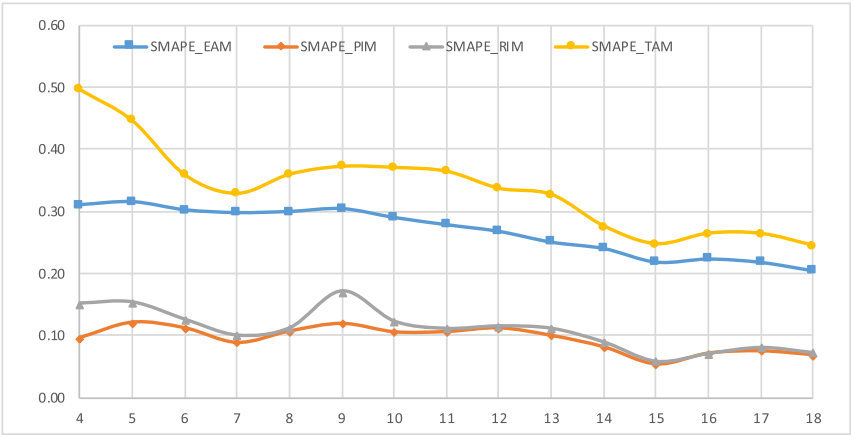


(c): y-axis = LMAPE, x-axis = *w*

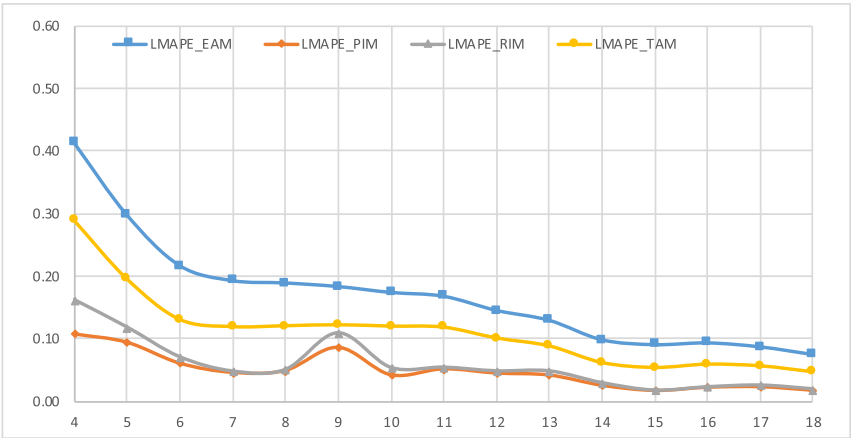
Fig. 4. Model comparison across different window sizes for Dataset **2013**



(a): y-axis = RMSE, x-axis = w



(b): y-axis = SMAPE, x-axis = w



(c): y-axis = LMAPE, x-axis = w

Fig. 5. Model comparison across different window sizes for Dataset 2015.

Table 4. Performance and P-values of Model Comparison by w

Models	RMSE	SMAPE	LMAPE
Dataset 2013	*Performance*		
EAM vs. RIM	(115269.73, 110059.43)	(0.3318, 0.1439)	(0.0839, 0.0290)
EAM vs. PIM	(115269.73, 120501.04)	(0.3318, 0.1538)	(0.0839, 0.0307)
PIM vs. RIM	(120501.04, 110059.43)	(0.1538, 0.1439)	(0.0307, 0.0290)
TAM vs. RIM	(126938.53, 110059.43)	(0.4208, 0.1439)	(0.0569, 0.0290)
TAM vs. PIM	(126938.53, 120501.04)	(0.4208, 0.1538)	(0.0569, 0.0307)
Dataset 2013	*p-value*		
EAM vs. RIM	0.4397	2.89204E-12***	9.71997E-09***
EAM vs. PIM	0.4735	3.80808E-08***	2.27542E-07***
PIM vs. RIM	0.0062**	0.4936	0.5619
TAM vs. RIM	0.0280*	6.1488E-11***	1.05328E-07***
TAM vs. PIM	0.4082	5.88899E-09***	0.0001***
Dataset 2015	*Performance*		
EAM vs. RIM	(47690.12, 48065.58)	(0.2600, 0.1062)	(0.1589, 0.0542)
EAM vs. PIM	(47690.12, 39917.28)	(0.2600, 0.0913)	(0.1589, 0.0454)
PIM vs. RIM	(39917.28, 48065.58)	(0.0913, 0.1062)	(0.0454, 0.0542)
TAM vs. RIM	(51531.55, 48065.58)	(0.3242, 0.1062)	(0.1049, 0.0542)
TAM vs. PIM	(51531.55, 39917.28)	(0.3242, 0.0913)	(0.1049, 0.0454)
Dataset 2015	*p-value*		
EAM vs. RIM	0.7799	5.82636E-15***	6.98436E-07***
EAM vs. PIM	0.0622$^{\wedge}$	2.51315E-14***	2.43422E-06***
PIM vs. RIM	0.0238*	0.0017**	0.0178*
TAM vs. RIM	0.0090**	5.79323E-12***	1.57455E-06***
TAM vs. PIM	0.0045**	3.85823E-11***	9.45915E-06***

Note: An underlined value indicates significantly better performance of a model in the pair. Asterisks indicate statistical significance: $^{\wedge}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 5. Performance and P-values of Model Comparison by Frobenius Loss

Parameter	DBMM vs. PIM	DBMM vs. RIM	RIM vs. PIM
Dataset 2013	*Performance*		
w=12, n=40	(79.42, 72.17)	(79.42, 74.30)	(74.30, 72.17)
w=18, n=34	(100.69, 99.92)	(100.69, 95.62)	(95.62, 99.92)
Dataset 2013	*p-value*		
w=12, n=40	0.01576 *	0.1005$^{\wedge}$	0.2673
w=18, n=34	0.795	0.1551$^{\wedge}$	0.3461
Dataset 2015	*Performance*		
w=12, n=40	(49.25, 42.34)	(49.25, 42.97)	(42.97, 42.34)
w=18, n=34	(56.07, 47.36)	(56.07, 48.78)	(48.78, 47.36)
Dataset 2015	*p-value*		
w=12, n=40	9.993e-06 ***	3.262e-05 ***	0.668
w=18, n=34	1.906e-04 ***	4.796e-04 ***	0.4458

Note: An underlined value indicates significantly better performance of a model in the pair. Asterisks indicate statistical significance: $^{\wedge}p < 0.2$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.
w is the window size used in the models and n is the sample size (number of days' data).

Table 6. Pearson Correlations (and P-values) between Structural Properties and Model Performances for Datasets 2013 and 2015

Correlation for Dataset 2013

	RMSE_EAM	RMSE_TAM	RMSE_PIM	RMSE_RIM	SMAPE_EAM	SMAPE_TAM	SMAPE_PIM	SMAPE_RIM	LMAPE_EAM	LMAPE_TAM	LMAPE_PIM	LMAPE_RIM
NumVertices_13	0.251983	0.246338	0.232552	0.210932	−0.227823	−0.365386	−0.481636	−0.440020	−0.600474	−0.718916	−0.677971	−0.735199
NumEdges_13	0.171389	0.167132	0.163567	0.130969	−0.340960	−0.474337	−0.500433	−0.466678	−0.694097	−0.804960	−0.693455	−0.743010
Density_13	−0.378742	−0.373043	−0.354124	−0.352905	0.282294	0.396002	0.435608	0.397663	0.588011	0.696339	0.604860	0.645586
Clustering_13	−0.262539	−0.252270	−0.219916	−0.252480	0.245832	0.390168	0.604349	0.471853	0.606045	0.714359	0.769093	0.746683
ConnectedCompo	0.338293	0.328962	0.323361	0.314009	−0.503724	−0.625061	−0.459404	−0.441456	−0.758702	−0.804753	−0.626318	−0.641852

p-value of cor.test for Dataset 2013

	RMSE_EAM	RMSE_TAM	RMSE_PIM	RMSE_RIM	SMAPE_EAM	SMAPE_TAM	SMAPE_PIM	SMAPE_RIM	LMAPE_EAM	LMAPE_TAM	LMAPE_PIM	LMAPE_RIM
NumVertices_13	0.329222	0.340533	0.369076	0.416416	0.379165	0.149245	0.050278	0.077141	0.010811	0.001146	0.002781	0.000772
NumEdges_13	0.510717	0.521427	0.530474	0.616348	0.180478	0.054388	0.040766	0.058965	0.001994	0.000097	0.002022	0.000632
Density_13	0.133833	0.140268	0.163148	0.164704	0.272296	0.115598	0.080504	0.113941	0.013044	0.001901	0.010101	0.005125
Clustering_13	0.308665	0.328654	0.396368	0.328238	0.341557	0.121553	0.010182	0.055843	0.009916	0.001274	0.000308	0.000574
ConnectedCompo	0.184133	0.197302	0.205497	0.219666	0.039253	0.007297	0.063571	0.076069	0.000414	0.000098	0.007146	0.005475

Correlation for Dataset 2015

	RMSE_EAM	RMSE_TAM	RMSE_PIM	RMSE_RIM	SMAPE_EAM	SMAPE_TAM	SMAPE_PIM	SMAPE_RIM	LMAPE_EAM	LMAPE_TAM	LMAPE_PIM	LMAPE_RIM
NumVertices_15	0.3061	0.2013	0.3585	0.2887	0.6148	−0.6512	−0.0520	−0.0679	−0.7613	−0.7454	−0.7570	−0.8160
NumEdges_15	0.2561	0.1536	0.3082	0.2456	0.6053	−0.6655	−0.0990	−0.1153	−0.7724	−0.7455	−0.7877	−0.8307
Density_15	−0.4803	−0.3522	−0.3292	−0.3278	−0.6297	0.5187	0.0280	−0.0819	0.8269	0.9582	0.7272	0.8687
Clustering_15	0.0751	0.0201	−0.0957	0.0144	0.6037	−0.0017	0.0603	−0.0788	−0.5306	−0.5370	−0.4478	−0.5978
ConnectedCompo	0.4658	0.3028	0.5191	0.4678	0.4939	−0.6499	0.1505	−0.0033	−0.7516	−0.7515	−0.6237	−0.7848

p-value of cor.test for Dataset 2015

	RMSE_EAM	RMSE_TAM	RMSE_PIM	RMSE_RIM	SMAPE_EAM	SMAPE_TAM	SMAPE_PIM	SMAPE_RIM	LMAPE_EAM	LMAPE_TAM	LMAPE_PIM	LMAPE_RIM
NumVertices_15	0.2489	0.4548	0.1727	0.2782	0.0113	0.0063	0.8483	0.8028	0.0006	0.0009	0.0007	0.0001
NumEdges_15	0.3383	0.5701	0.2456	0.3591	0.0130	0.0049	0.7154	0.6706	0.0005	0.0009	0.0003	0.0001
Density_15	0.0597	0.1809	0.2132	0.2152	0.0089	0.0395	0.9179	0.7631	0.0001	0.0000	0.0014	0.0000
Clustering_15	0.7823	0.9411	0.7246	0.9579	0.0133	0.9949	0.8243	0.7716	0.0345	0.0320	0.0820	0.0145
ConnectedCompo	0.0690	0.2544	0.0393	0.0677	0.0518	0.0064	0.5780	0.9903	0.0008	0.0008	0.0098	0.0003

composition of social networks can change dramatically over a short time. By contrast, DBMM assumes nodes to be static and their composition unchanged over time. Although DBMM estimates dynamic behavior of nodes, the absence of change in nodal composition in evolving networks makes the role estimation less effective and unrealistic. In addition, the rank-r approximation used in DBMM makes another static assumption of roles and nodal features. This is in contrast with the more flexible prediction of agents and links in PIM and RIM based on relevant social theories that can be used to explain dynamic human behavior. However, the only comparison showing no significant difference between DBMM and PIM is when w = 18 in the 2013 dataset (p = 0.795). It should be noted that the parameter value $\alpha = 0.7$ used in PIM and RIM (as recommended in Reference [57]) is not empirically tested to be optimal for the two interaction models. Future studies may explore the effect of parameter values (w and α) on agent role discovery.

5.4 Structural Analysis of Social Networks

To understand the relationship between the model performance and the structure of the social networks in each dataset, we conducted a correlation analysis to identify the degree to which structural properties relate to performances. Table 6 shows the correlations (and p-values of testing their significance) between each pair of structural properties and model performances for Datasets 2013 and 2015. Shaded cells in the first and third tables highlight correlations with absolute values above 0.4 (light shade) or 0.7 (dark shade). Shaded cells in the second and fourth tables highlight p-values below 0.05 (light shade) or 0.01 (dark shade).

The results show significant correlations between LMAPE (of all models) and the five structural properties. For Dataset 2013, all 20 cases reported significant p-values, and all the absolute values of correlations are above 0.5, with the highest being 0.805 (LMAPE of TAM and clustering coefficient). For Dataset 2015, 14 out of 15 cases reported significant p-values, and all the absolute values of the significant correlations are above 0.5, with the highest being 0.958 (LMAPE of RIM and density). The p-values in 10 of 15 cases in Dataset 2015 are less than 0.001, indicating extremely significant

Table 7. Percentage of Overlap between ACTUAL Users and Users Predicted by Model

Model	EAM		PIM		RIM		TAM	
Date	Top20	All	Top20	All	Top20	All	Top20	All
10/3/13	55%	93.60%	50%	93.641%*	70%*	93.60%	30%	93.60%
10/6/13	65%*	97.27%	55%	97.281%*	65%*	97.27%	60%	97.27%
10/9/13	80%	99.59%	70%	99.678%*	90%*	99.58%	80%	99.59%
6/15/15	15%	96.078%*	80%*	95.98%	75%	95.99%	80%*	96.078%*
6/18/15	30%	96.79%	75%	97.05%*	80%*	96.80%	70%	96.79%
6/21/15	50%	95.792%*	75%*	95.792%*	75%*	95.792%*	65%	95.792%*

*indicates highest percentage overlap across all models in either "Top20" or "All" category.

correlations between LMAPE of the four models and structural properties. Therefore, the findings reveal strong predictability of structural properties on LMAPE of all models in both datasets.

The results also show high to moderate correlations between SMAPE (of EAM, TAM, and RIM) and the structural properties. For Dataset 2013, 4 out of 20 cases reported significant p-values, and all these cases have absolute correlations above 0.5. For Dataset 2015, 8 cases (all involving EAM or TAM) reported significant p-values while 10 cases (involving PIM and RIM) have insignificant p-values. The findings indicate that structural properties have moderately strong predictability on SMAPE of EAM in Dataset 2015. But for RIM, the predictability is moderately strong in Dataset 2013 but is weak in Dataset 2015. The predictability for SMAPE of PIM is weak in Dataset 2015. The structural properties have weak and insignificant correlations (<0.4 in absolute value) with RMSE of the four models in most cases. All p-values are not significant in Dataset 2013 and in Dataset 2015, with the only exception being the correlation between the number of connected components and RMSE of PIM in Dataset 2015 ($p = 0.03934$, correlation = 0.5191). Therefore, the results show that structural properties are poor predictors of RMSE of all models.

5.5 Case Studies of User Interaction Patterns

To enable contextual reasoning and explanation of the models' predictions, we conducted case studies of top 20 users (who have the highest activity scores) and all users, and provided their interaction patterns to show different results of RIM and PIM compared with the two benchmark models. The case studies examine three dates of user interaction patterns in each of the two datasets (2013 and 2015). The three dates chosen from the 2013 dataset were 10/3/2013, 10/9/2013, and 10/9/2013, during which a massive protest took place in Washington, D.C. in support of comprehensive immigration reform in the United States. This event resulted in over 200 people arrested by police, among them 8 Democratic members of the U.S. Congress. The three dates chosen from the 2015 dataset were 6/15/2015, 6/18/2015, and 6/21/2015, during (and before) which prominent members of the Republican Party, Donald Trump and Jeb Bush (among others), both announced their candidacies for the U.S. presidency (on 6/15/2015), triggering much reaction in social media.

First, for each model, we computed the percentage of overlap between the actual users (in the Top20 or All lists) and the model's predicted users in the same two lists. The results, shown in Table 7, reveal that among all four models, RIM achieved (on five out of six days) the best matching (marked with *) between the actual top-20 users and the predicted top-20 users; PIM is ranked second (two out of six days). When considering all users, PIM achieved (on five out of six days) the best matching between actual users and predicted users. These results confirm the superior performance of PIM and RIM in predicting nodal activities in the social media networks of the selected dates.

Second, we selected examples of top-20 users to examine cases where all models predicted them successfully (i.e., matching actual top-20 users) in contrast with cases where only RIM and PIM

Table 8. Interaction Patterns of Users Predicted by All Four Models to be among Top20

Date	Public User Name	Message Type					Follower Count	Friend Count
		Target Tweet	Retweet	Modified Tweet	Self retweet	Duplicate Tweet		
10/3/13	Immigraxion	2	569	34	0	2	323	180
10/6/13	SpeakerBoehner	285	698	19	4	112	803,118	7,704
10/9/13	SteveWorks4You	61	1,765	15	0	1	19,320	4,857
6/15/15	AnnCoulter	136	706	2	4	1	611,756	387
6/18/15	JebBush	157	225	2	5	16	130,736	111
6/21/15	HillaryClinton	79	577	2	1	9	3,499,414	16

Table 9. Interaction Patterns of Users Predicted by Only PIM and RIM to be among Top20

Date	Public User Name	Message Type					Follower Count	Friend Count
		Target Tweet	Retweet	Modified Tweet	Self retweet	Duplicate Tweet		
10/3/13	politico	93	190	0	0	1	1,185,623	999
10/6/13	nytimes	67	353	4	0	2	17,324,215	991
10/9/13	FAIRImmigration	52	487	160	0	12	7,821	2,869
6/15/15	HillaryClinton	77	474	2	1	6	3,499,414	16
6/18/15	realDonaldTrump	129	71	0	1	0	3,071,548	46
6/21/15	ianbremmer	3	185	0	0	0	110,509	988

Table 10. Structural Information of Social Media Networks on Selected Dates

Date	Max(degree)	Min(degree)	Mean(degre	NumVertice	NumEdges	Density	ClusteringCoeff	#ConnectedComponents
9/21/13	4,164	1	5	28,823	84,611	2.04E-04	1.20E-02	2,328
9/27/13	3,706	1	5	24,126	64,100	2.20E-04	1.20E-02	2,004
5/24/15	1,316	1	2	11,581	17,298	2.58E-04	5.00E-03	1,307

predicted them successfully. Interaction patterns of these two groups of users can be used to distinguish predictive capabilities of the models. As shown in Table 8, the users predicted by all four models to be among Top20 generally have significantly more retweets, more self retweets, and more duplicate tweets. By contrast, users predicted by only PIM and RIM (see Table 9) generally have fewer targeted tweets and fewer modified tweets. These message types may be appropriately captured by PIM and RIM due to their unique nature in identifying interactive users. Although the user "HilliaryClinton" appears in both tables, the dates that the user appears are different with different interaction patterns: 6/21/2015 in Table 8 and 6/15/2015 in Table 9. It is possible that PIM and RIM capture interaction patterns of the user better on 6/15/2015 due to their capabilities of detecting nodal activities better in an environment with highly opposing views (two Republican candidates announced candidacies on that date, whereas "HillaryClinton" belongs to Democratic Party).

5.6 Implications

The experimental findings provide insights in understanding the model performances, temporal features, structural properties of networks, user interaction patterns, and attributes of datasets. Their implications for predicting dynamic social network activities are discussed below. *First*, the proposed interaction models (RIM and PIM) performed significantly better than the benchmark models (EAM, TAM, and DBMM) across datasets of different time frames and content, indicating the models' strong capability of accounting for user interactions in activity prediction across dif-

ferent contexts. This result confirms that the proposed models can be useful for detecting activity levels in large dynamic online social networks. *Second*, the time window size (w) (the length of history that the models used to predict activity) was found to demonstrate the performance differences between the interaction models and the benchmark models. Both RIM and PIM outperformed TAM and EAM across different w. The result implies that user interaction is captured sufficiently by the two proposed models consistently, despite changes in w. *Third*, PIM outperformed RIM across different w in Dataset 2015, indicating that the performance of PIM is less affected by changing w. It follows that, when the length of prior history is relatively short ($4 \leq w \leq 20$), PIM would be a preferred model for detecting activity levels of social network users. For longer prior histories ($w > 20$), additional data and studies will be needed to understand performance differences. *Fourth*, both PIM and RIM outperformed DBMM significantly in terms of agents' behavioral role discovery, indicating their superior performance due to a realistic treatment of dynamic network composition. *Fifth*, structural properties of the social networks correlated significantly with LMAPE of all four models and with SMAPE of EAM. The result shows that macroscopic, structural factors of networks have strong to moderately strong predictability of model performances (as measured by LMAPE and SMAPE respectively). This allows assessment of temporal network activity detection models with relative ease of computation and yet with high accuracy. Furthermore, interaction patterns of selected top users reveal that PIM and RIM provide superior predictive capabilities especially in social media networks composed of highly opposing views.

6 CONCLUSION

As social media networks are increasingly used to represent human activities in online environments, effective methods for detecting temporal nodal activities are needed. This research has developed and validated a new class of models that examine user interactions to support activity detection in dynamic social networks. The models incorporated information about agents' activities in the network and their interaction over time to detection at the next time step. The innovation of the models consists of novel simulation of online social networks, accurate detection based not only on historical activities but also on agent interaction, and adaptive link addition or removal based on ongoing activity patterns.

6.1 Summary of Findings

The experimental findings show that the two interaction models outperformed the benchmark models significantly across different dates and across different temporal window sizes. PIM was also found to significantly outperform RIM across different window sizes in Dataset 2015, demonstrating the best accuracy among all four models when w ranges from 4 to 20. For a fixed window size, performances of RIM and PIM are not significantly different across dates. This finding shows that both PIM and RIM capture unique activity patterns that are complementary to each other. Behaviors exhibited in random interaction and preferential interaction both contribute to future activity levels detected by these models. Therefore, both patterns demonstrate significant presence in user activities in the temporal networks. In addition, the two interaction patterns (preferential interaction and random interaction) do not have a significant difference in their contribution to user activities. Such lack of clear distinction is not because their individual effects are not significant, but possibly because users adopt these behavior patterns selectively, contingent on the circumstances that they face. In other words, although a majority of users favors preferential interaction, most activities are characterized by random interaction as users make contingent choices to adjust to reality. Examining these contingencies can possibly reveal patterns of strategic behavioral adoption and its relationship with external factors.

Regarding evaluation metrics, the experimental results demonstrate significant advantages of SMAPE and LMAPE over RMSE. The results show that the linear assumption of RMSE often fails to characterize the differences in model performance, making it less capable to measure activities that exhibit primarily non-linear patterns. The findings show that SMAPE and LMAPE captured the random interaction and exponential activity patterns very well, and hence should be more suitable for measuring online social network activities than traditional approaches such as RMSE (e.g., Reference [27]).

Results of the correlation analysis show that structural properties of social networks provide a simple and yet accurate approach to predicting model performance (measured by LMAPE and SMAPE). These properties allow linear projection of the performance values that are based on non-linear model prediction. However, the directions and intensity of correlation vary according to datasets and performance measures. Future research may examine in more detail these variations. As online social network activities tend to change dramatically over time, there is a strong potential to use these properties to support advanced analysis and understanding of temporal network activities.

6.2 Contributions and Limitations

This research should provide several contributions. The interaction models for temporal network prediction are shown to advance traditional models (e.g., References [57, 67]) by explicitly modeling agent interaction and by simulating network behavior. New experimental findings obtained in this research can be used to inform model development and performance. Moreover, the research developed a reusable implementation of temporal prediction for large networks in a social media community, thus contributing to standardizing the evaluation of temporal social networks [27]. The computational framework developed in this research provides a reusable graph construction and transformation pipeline that is not available in prior research [37]. The methods and findings should be useful to computer and information scientists, intelligence experts (e.g., for cyber-surveillance [14]), social researchers (e.g., for public- and health-policy decision making [42]), and business practitioners (e.g., business analytics [10, 12]), among others.

There are several limitations in this research. The datasets used in this research came from one specific source (Twitter). Although different time spans of datasets and different analysis aspects were considered, the choice of data source may limit the generalizability of the findings. There is a lack of prior theoretical study of temporal nodal activity detection in online social networks. The identification of relevant theoretical and technical aspects was non-trivial in this research, before the technical infrastructure was built. The resources available to collect and handle the data were limited, restricting additional analysis that could be performed.

6.3 Future Directions

In the future, this research may be extended in several ways. Extending the temporal window w to study its impact on predictive performance using different models may help to characterize the domains being predicted for and the models being used. Multiple domains could be studied to examine their effect on model performance. Additional types of interaction and datasets may be modeled to capture their unique activity patterns. Social links other than geodesics used in BC calculation may be explored (e.g., direct links). Comparison of which behavioral assumptions work well in which predictive circumstances should advance understanding of model design and deployment.

ACKNOWLEDGMENTS

We thank the editors and reviewers for their valuable suggestions.

REFERENCES

[1] G. A. Akerlof. 1997. Social distance and social decisions. *Econometrica* 65, 5 (1997), 1005–1027.

[2] Anna Goldenberg, Alice X. Zheng, Stephen E. Fienberg, and Edoardo M. Airoldi 2010. A survey of statistical network models. *Found. Trends Mach. Learn.* 2, 2 (Feb. 2010), 129–233

[3] D. Antoniades and C. Dovrolis. 2015. Co-evolutionary dynamics in social networks: A case study of Twitter. *Comput. Soc. Netw.* 2, 14 (2015).

[4] A.-L. Barabási and A. Réka. 1999. Emergence of scaling in random networks. *Science*. 286, 5439 (1999), 509–512.

[5] G. S. Becker. 1974. A theory of social interactions. 1974. *J. Polit. Econ.* 82, 6 (1974) 1063–1093.

[6] B. Bollobás. 2001. *Random Graphs*. 2001. Cambridge University Press, New York.

[7] S. P. Borgatti. 2005. Centrality and network flow. *Soc. Netw.* 27, 1 (2005), 55–71.

[8] J. Bush, T. F. McLarty III, and E. H. Alden. 2009. *U.S. Immigration Policy*. Council on Foreign Relations, New York.

[9] H. H. Chang and S.-S. Chuang. 2011. Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Inform. Manag.* 48, 1 (2011), 9–18.

[10] H. Chen, R. Chiang, and V. Storey. 2012. Business intelligence and analytics. *MIS Quart.* 36, 4 (2012) 1165—1188.

[11] M. Chui, J. Manyika, J. Bughin, R. Dobbs, C. Roxburgh, H. Sarrazin, G. Sands, and M. Westergren. 2012. The Social Economy: Unlocking Value and Productivity through Social Technologies. McKinsey & Company.

[12] W. Chung. 2014. BizPro: Extracting and categorizing business intelligence factors from textual news articles. *Int. J. Inform. Manag.* 34, 2, (2014), 272–284.

[13] W. Chung. 2016. Social media analytics: Security and privacy issues. *J. Inform. Priv. Secur.* 12, 3 (2016) 105–106.

[14] W. Chung and D. Zeng. 2018. Social-media-based policy informatics: Cases on cybersecurity and public health informatics. In *Policy Analytics, Modelling, and Informatics*, J. Ramon Gil-Garcia, Theresa A. Pardo, and Luis F. Luna-Reyes (Eds.). Springer, Cham.

[15] W. Chung and D. Zeng. 2016. Social-media-based public policy informatics: Sentiment and network analyses of U.S. immigration and border security. *J. Assoc. Inform. Sci. Technol.* 67, 7 (2016), 1588–1606.

[16] P. Constable and Morello. 2013. C. Marchers urge Congress to pass immigration reform; Several congressmen arrested. *The Washington Post*. Retrieved from http://www.washingtonpost.com/local/marchers-urge-congress-to-pass-immigration-reform/2013/10/08/9714e5c0-2f8f-11e3-9ccc-2252bdb14df5_story.html.

[17] J. Deese and R. A. Kaufman. 1957. Serial effects in recall of unorganized and sequentially organized verbal material. *J. Exper. Psychol.* 54, 3 (1957), 180–187.

[18] N. Du, X. W. Jia, J. Gao, V. Gopalakrishnan, and A. D. Zhang. 2015. Tracking temporal community strength in dynamic networks. *IEEE Trans. Knowl. Data Eng.* 27, 11 (2015), 3125–3137.

[19] D. M. Dunlavy, T. G. Kolda, and E. Acar. 2011. Temporal link prediction using matrix and tensor factorizations. *ACM Trans. Knowl. Disc. Data* 5, 2 (2011).

[20] P. Erdős and A. Rényi. 1959. On random graphs. *Pub. Mathe.* 6 (1959), 290–297.

[21] W. Fan and M. D. Gordon. 2014. The power of social media analytics. *Commun. ACM* 57, 6 (2014), 74–81.

[22] L. C. Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry* 40, 1 (1977), 35–41.

[23] J. Gans, E. M. Replogle. and D. J. Tichenor (Eds.). 2012. *Debates on U.S. Immigration*. Sage Publications, Inc.

[24] C. Gao and J. Liu. 2017. Network-based modeling for characterizing human collective behaviors during extreme events. *IEEE Trans. Syst., Man, Cyber.: Syst.* 47, 1 (2017), 171–183.

[25] S. Gao, L. Denoyer, and P. Gallinari. 2011. Temporal link prediction by integrating content and structure information. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. ACM Press, 1169–1174.

[26] F. Heider. 1958. *The Psychology of Interpersonal Relations*. Erlbaum, Hillside, NJ.

[27] J. M. Hofman, A. Sharma, and D. J. Watts. 2017. Prediction and explanation in social systems. *Science*, 355, 6324 (2017), 486–488.

[28] C. I. Hovland. 1957. *The Order of Presentation in Persuasion*. Yale University Press, Inc.

[29] M. O. Jackson. 2008. *Social and Economic Networks*. Princeton University Press, Princeton, NJ.

[30] S. Jiang and H. C. Chen. 2016. NATERGM: A model for examining the role of nodal attributes in dynamic social media networks. *IEEE Trans. Knowl. Data Eng.* 28, 3 (2016), 729–740.

[31] G. C. Kane, M. Alavi, G. J. Labianca, and S. P. Borgatti. 2014. What's different about social media networks? A framework and research agenda. *MIS Quart.* 38, 1 (2014), 275–304.

[32] B. Karrer and M. E. J. Newman. 2011. Stochastic blockmodels and community structure in networks. *Phys. Rev. E* 83, 1 (2011).

[33] H. C. Kelman. 1958. Compliance, identification, and internalization three processes of attitude change. *J. Confl. Resol.* 2, 1 (1958), 51–60.

[34] B. Latané. 1981. The psychology of social impact. *Amer. Psychol.* 36, 4 (1981), 343–356.

[35] M. J. Lee, B. Ferwerda, J. Choi, J. Hahn, J. Y. Moon, and J. Kim. 2013. GitHub developers use rockstars to overcome overflow of news. *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM. 133–138.

[36] M. C. LeMay. 2004. *U.S. Immigration: A Reference Handbook.* ABC-CLIO, Santa Barbara, CA.
[37] A. Lenharth, D. Nguyen, and K. Pingali. 2016. Parallel graph analytics. *Commun. ACM* 59, 5 (2016), 78–87.
[38] J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani. 2010. Kronecker graphs: An approach to modeling networks. *J. Mach. Learn. Res.* 11 (2010), 985–1042.
[39] J. Leskovec, J. Kleinberg, and C. Faloutsos. 2005. Graphs over time: Densification laws, shrinking diameters, and possible explanations. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining.* ACM Press.
[40] X. Li, N. Du, H. Li, K. Li, J. Gao, and A. Zhang. 2014. A deep learning approach to link prediction in dynamic networks. In *Proceedings of the SIAM International Conference on Data Mining.* Society of Industrial and Applied Mathematics.
[41] Z. Li, X. Fang, and O. R. L. Sheng. 2017. A survey of link recommendation for social networks: Methods, theoretical foundations, and future research directions. *ACM Trans. Manag. Inform. Syst.* 9, 1 (2017), Article 1: 1–26.
[42] Y. Liang, X. Zheng, D. D. Zeng, X. Zhou, S. J. Leischow, and W. Chung. 2015. Characterizing social interaction in tobacco-oriented social networks: An empirical analysis. *Nat. Sci. Rep.* 5 (2015).
[43] L. A. McFarland and R. E. Ployhart. 2015. Social media: A contextual framework to guide research and practice. *J. Appl. Psychol.* 100, 6 (2015), 1653–1677.
[44] J. M. McPherson and J. R. Ranger-Moore. 1991. Evolution on a dancing landscape: Organizations and networks in dynamic blau space. *Soc. Forces* 70, 1 (1991), 19–42.
[45] M. McPherson, L. Smith-Lovin, and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Ann. Rev. Sociol.* 27 (2001), 415–444.
[46] N. Miller and D. T. Campbell. 1959. Recency and primacy in persuasion as a function of the timing of speeches and measurements. *J. Abnorm. Soc. Psychol.* 59, 1 (1959), 1–9.
[47] M. E. J. Newman. 2006. Modularity and community structure in networks. *Proc. Nat. Acad. Sci. USA* 103, 23 (2006), 8577–8582.
[48] E. W. T. Ngai, S. S. C. Tao, and K. K. L. Moon. 2015. Social media research: Theories, constructs, and conceptual frameworks. *Int. J. Inform. Manag.* 35, 1 (2015), 33–44.
[49] E. Otte and R. Rousseau. 2002. Social network analysis: A powerful strategy, also for the information sciences. *J. Inform. Sci.* 28, 6 (2002), 441–453.
[50] R. Palovics and A. A. Benczur. 2015. Temporal influence over the Last.fm social network. *Soc. Netw. Anal. Mining* 5, 4 (2015).
[51] W. Pan, W. Dong, M. Cebrian, T. Kim, J. H. Fowler, and A. S. Pentland. 2012. Modeling dynamical influence in human interaction. *IEEE Sig. Proc. Mag.* 29, 2 (2012), 77–86.
[52] J. C. L. Pinto, T. Chahed, and E. Altman. 2015. Trend detection in social networks using Hawkes processes. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.* 1441–1448.
[53] A. Portes. 1998. Social capital: Its origins and applications in modern sociology. *Ann. Rev. Sociol.* 24, 1 (1998), 1–24.
[54] D. H. Qiu, H. Li, and Y. Li. 2014. Identification of active valuable nodes in temporal online social network with attributes. *Int. J. Inform. Technol. Decis. Making* 13, 4 (2014), 839–864.
[55] V. Raghavan, G. V. Steeg, A. Galstyan, and A. G. Tartakovsky. 2014. Modeling temporal activity patterns in dynamic social networks. *IEEE Trans. Comput. Soc. Syst.* 1, 1 (2014), 89–107.
[56] K. Rohe, S. Chatterjee, and B. Yu. 2011. Spectral clustering and the high-dimensional stochastic blockmodel. *Ann. Stat.* 39, 4 (2011), 1878–1915.
[57] R. A. Rossi, J. Neville, B. Gallagher, and K. Henderson. 2013. Modeling dynamic behavior in large evolving graph. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining.* ACM Press, 667–676.
[58] G. R. Salancik and J. A. Pfeffer. 1978. Social information processing approach to job attitudes and task design. *Admin. Sci. Quart.* 23, 2 (1978), 224–253.
[59] J. Scott. 2000. *Social Network Analysis: A Handbook.* Sage Publications.
[60] C. Sedikides and J. M. Jackson. 1990. Social impact theory: A field test of source strength, source immediacy and number of targets. *Basic Appl. Soc. Psychol.* 11, 3 (1990), 273–281.
[61] D. Tayouri. 2016. Social media as an intelligence goldmine. *Cyber Secur. Rev.* (Spring) (2016), 27–30.
[62] C. Tofallis. 2014. A better measure of relative prediction accuracy for model selection and model estimation. *J. Operat. Res. Soc.* 66, 8 (2014), 1352–1362.
[63] U.S. Commission on Immigration Reform. 1994. *U.S. Immigration Policy—Restoring credibility: A Report to Congress.* U.S. Commission on Immigration Reform, Washington, DC.
[64] P. Wang, B. W. Xu, Y. R. Wu, and X. Y. Zhou. 2015. Link prediction in social networks: The state-of-the-art. *Sci. China-Inform. Sci.* 58, 1 (2015).
[65] S. Wasserman and K. Faust. 1994. *Social Network Analysis: Methods and Applications.* Cambridge University Press.
[66] D. J. Watts. 2007. A twenty-first century science. *Nature* 445, 7127 (2007), 489.

[67] W. Wei and K. M. Carley. 2015. Measuring temporal patterns in dynamic social networks. *ACM Trans. Knowl. Disc. Data* 10, 1 (2015), 1–27.
[68] L. Wu, Y. Ge, Q. Liu, E. H. Chen, R. C. Hong, J. P. Du, and M. Wang. 2017. Modeling the evolution of users' preferences and social links in social networking services. *IEEE Trans. Knowl. Data Eng.* 29, 6 (2017), 1240–1253.
[69] K. S. Xu and A. O. Hero. 2014. Dynamic stochastic blockmodels for time-evolving social networks. *IEEE J. Select. Topics Sig. Proc.* 8, 4 (2014), 552–562.
[70] D. Zeng. 2015. Policy informatics for smart policy-making. *IEEE Intell. Syst.* 30, 6 (2015), 2–3.
[71] X. Zheng, D. Zeng, and F.-Y. Wang. 2015. Social balance in signed networks. *Inform. Syst. Front.* 17, 5 (2015), 1077–1095.
[72] X. L. Zheng, Y. G. Zhong, D. Zeng, and F. Y. Wang. 2012. Social influence and spread dynamics in social networks. *Front. Comput. Sci.* 6, 5 (2012), 611–620.
[73] Y. Zhu, E. Zhong, S. J. Pan, X. Wang, M. Zhou, and Q. Yang. 2013. Predicting user activity level in social networks. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management.* ACM Press, 159–168.

Received March 2018; revised August 2019; accepted September 2019