The Dynamics of Health Sentiments with Competitive Interactions in Social Media

Saike He[1], Xiaolong Zheng[1*], Daniel Zeng[1,2]
[1]The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[2]University of Chinese Academy of Sciences, Beijing, China
{saike.he, xiaolong.zheng, dajun.zeng}@ia.ac.cn
* Corresponding author

Abstract—**Public sentiments affecting health outcomes are increasingly modulated by social media. Existing literature mainly focus on investigating how network structure affects the contagion of health sentiments. However, most of these studies neglect that the interaction topology change in time. In fact, the change of inter-individual connections over time is associated with individual attributes. The mechanism through which individual attributes reshapes the connection topology is mainly governed by the competition between two principles, i.e., homophily (establishing or reinforcing social connections) and homeostasis (preserving the total strength of social connections to each individual). No existing approaches are yet able to accommodate these two competing effects at the same time.**

We thus propose a new statistical model (H2 model, Homophily and Homestasis model) to depict the evolution of temporal network, which is governed by the competition of homophily and homeostasis. In addition, we consider the mediation effect of external shock events, which enables us to separate exogenous confounding factors. Evaluation on Twitter data suggests that H2 model can capture long-range sentiment dynamics and external shock events. In sentiment prediction, H2 consistently outperforms existing methods in terms of error rate. Through the model's shock tensor, we successfully detect several typical events, and reveal that users in negative emotions are more influenced by external shock events than those with positive emotions. Our findings have practical significance for those who supervise and guide health sentiments in online communities.

Keywords—health sentiment; competitive interactions; homophily; homeostasis; social media

I. Introduction

Social media plays an important role in shaping the dynamics of health behaviors and the associated sentiments [1]. For instance, vaccine preventable disease are more likely to outbreak if overall vaccination rates decline [2], or anti-vaccination sentiments domains local communities [3, 4]. The continuously evolving public concern about vaccines demands an increased understanding on how such sentiments spread over time.

Recent studies have enlightened the important role played by the interaction topology on the information diffusion process. One strand of studies utilize explanatory models to infer the underlying spreading cascade [5]. These approaches allows retracing the path taken by a piece of information. Another strand of research aim to predict how a specific diffusion process unfold in a given network [6, 7]. While these work mainly focus on understanding the relationship between online behaviors and the underlying mechanisms governing the transfer and processing of information, most of them neglect that real world systems change their interaction patterns in time. In fact, the change of inter-individual connections over time is associated with individual attributes [8, 9]. The mechanism through which individual attributes reshape the interaction topology is mainly governed by the competition between two principles, i.e., homophily [10] and homeostasis [8, 11].

According to homophily, the ties between individuals are found to be strongly favored by the similarity of their attributes. This principle can affect the diffusion process of information. Considering the social capital available for each individual, we should add a further constraint to describe the mechanism acting on each social tie, i.e., homoestasis. Under the effects of homeostasis, social connections tend to maintain relatively stable. In practice, this second principle considers that the available resources devoted to sustain interactions are finite. The direct result of this effect is that the enhancement of some connection from an individual is counter-balanced by the weakening of other connections of the same individual to the network.

The reinforcing effect caused by homophily and the constraint caused by homestasis, jointly governs individual dynamics and how the dynamics process interact with their social connections. However, no existing approaches are able to accommodate these two competing effects at the same time. We thus propose a new statistical model to depict the evolution of temporal network, which is governed by the competition of homophily and homeostasis. In addition, we consider the mediation effect of external shock events, which enables us to separate exogenous confounding factors from endogenous competition effects. Evaluation on Twitter data suggests that the proposed model can capture long-range sentiment dynamics and lower the error rate in sentiment prediction. Further, it detects several typical shock events and implies how external shocks affect online users holding different sentiments.

II. MODEL

This section presents the proposed H2 Model (Homophily and Homestasis model). We intend to describe the posting number of user group holding similar sentiments, e.g., positive, negative, or neutral. Suppose that we have a collection of temporal data X of d user groups $X=\{x_1,x_i,\ldots,x_d\}$. Here, x_i is a sequence of user group i (i.e., $x_i=\{x_i^t\}_{t=1}^n$), and n is the duration of all d user groups. Our primary goal is to capture the evolution of X, and uncover how the attributes of each sequence reshape its interactions with others. To differentiate the effects caused by exogenous factors from exogenous group interactions, we also need to quantify external shock events. Overall, we intend to describe the following three properties:

- **(P1)**: Non-linear evolution of user groups.
- **(P2)**: Interaction coefficients between user groups.
- **(P3)**: Effects of external shocks.

Let g_i^t be the estimated posting number of user group i at time tick t, its evolution is governed by the following equations:

$$g_i^{t+1}=g_i^t\left[1+r_i\left(1-\frac{\sum_{j=1}^{d}a_{ij}^t\cdot g_j^t}{K_i}\right)\right],(i=1,\ldots,d) \quad (1)$$

where,

g_i^t: The posting volume of user group i at time tick t.

r_i: Intrinsic growth rate of posting number of user group i, ($r_i\geq 0$).

K_i: Carrying capacity of posting number of user group i when the other user groups are absent ($K_i\geq 0$).

a_{ij}^t: Interaction strength pointing from group j to i at time tick t; a_{ij}^t corresponds to intra-group interactions.

In the proposed model, the percentage of posting number of user group i can be described as:

$$\left(1-\frac{\sum_{j=1}^{d}a_{ij}^t\cdot g_j^t}{K_i}\right) \quad (2)$$

where, a_{ij}^t is the interaction coefficient, which describes the effect rate of user group j on user group i. If there is no inter-group interaction (i.e., $a_{ij}^t=0\,(i\neq j)$), the model deteriorates to traditional time series models that trade each temporal sequence in isolation, such as Auto-regression (AR), Autoregressive integrated moving average (ARIMA) and Kalman filters (KF) [12]. While if $a_{ij}^t\equiv a_{ij}$ independent of the time tick t for user groups i, j, then the model deteriorates to the Lotka-Volterra population model [13], where the interaction coefficients keep constant. In our model setting, the interaction coefficients are set as time-varying and asymmetric. This means that the interaction strength is directed, and varies during the whole observation period. In what follows, we describe the setting of interaction coefficient matrix **A**.

A. Time-varying Competitive Interactions

User interactions involve two aspects, i.e., the interaction frequency, and the effect of each single interaction. Here, we encode the compound effect of interaction frequency and its response into a unique time varying variable a_{ij}^t. In this model, a_{ij}^t depicts three types of interaction, i.e., attractive (negative values), repulsive (positive values), and none (zero values). The interactions a_{ij}^t evolves as:

$$a_{ij}^{t+1}=a_{ij}^t\left[1+\left(s_i^t\cdot p_{ij}^t-\sum_{l=1}^{N}a_{il}^t\cdot p_{il}^t\right)\right] \quad (3)$$

where, s_i^t is the total incoming strength of group i at time tick t, $s_i^t=\sum_{j=1}^{N}a_{ij}^t$, and p_{ij}^t is the degree of local homophily between groups i and j, averaged over time in the interval [t-T, t] [14]:

$$p_{ij}^t=\frac{1}{T}\left|\sum_{l=1}^{T}e^{i\phi_{i,j}^t}\right| \quad (4)$$

where, T is a control parameter that quantifies the amount of memory used by each user group in the updating process. $\phi_{i,j}^t$ is a phase-difference function measuring the attributes distance between user groups i and j at time tick t:

$$\phi_{i,j}^t=\arccos\left(\cos\left(m_i^t,m_j^t\right)\right) \quad (5)$$

where, m_i^t and m_j^t respectively corresponds to the attribute vectors of groups i and j at time tick t; arcso(*) calculate the angle of the Cosine value enclosed. Given the above setting, the quantities p_{ij}^t take values in [0, 1], with $p_{ij}^t=1$ meaning that groups i and j have been perfectly entrained with regard to their attributes along the last T time units [15].

The adaptive scheme in (3) retains the main characteristics of both homophily and homeostasis. User group j that has similar attributes with group i (e.g. average emotion valence, activity level, and network structure), will enhance their interaction strength, according to homophily; while as a consequence of homeostasis, the interaction strength from the remaining group will be depressed to keep constant the total incoming interaction strength s_i^t, of group i. As a result, all the interactions toward the same group compete for the available social resources.

B. External Shock Events

Apart from endogenous competition mechanisms, it's also notable that online users change their behaviors according to various external shock events [16]. These exogenous activities should be taken into consideration to avoid any potential confounding effects (e.g., attributing merely exogenous effect to endogenous stimuli). Let $\dot{g}_i^t$ be the posting number of user group i at time tick t when considering external shock events, the full model captures exogenous effects with the following equations:

$$\dot{g}_i^t = g_i^t\left[1+e_i^t\right], \quad (i=1,\ldots,d) \tag{6}$$

where, e_i^t corresponds to the additional change rate of posting number of user group i because of exogenous shock events. e_i^t takes values in [-1, +∞).

Given the above definition, the posting number of user group i at time tick t $\dot{g}_i^t$ dependents on both the endogenous competitive interactions and the exogenous activities $\mathbf{E}=\{e_i(t)\}_{i,t=1}^{d,n}$. Each element in $\mathbf{E}$ describes the additional change in posting number, caused by external shock events, such as holidays, etc. For a more compressed representation of $\mathbf{E}$, we decompose $\mathbf{E}$ into two matrices, i.e., effection matrix $\mathbf{B}$ of size ($k\times n_p$) and participation (weight) matrix $\mathbf{W}$ of size ($d\times k$). $\mathbf{B}$ describes a set of k external shocks with duration n_p, while $\mathbf{W}$ describes the participation weight of each sequence for each shock component. The external shock tensor $E=\{e_i^t\}_{i,t=1}^{d,n}$ can be described as the following function:

$$e_i^t = f\left(i,t\,|\,\mathbf{W},\mathrm{B}\right)=\sum_{j=1}^{k} w_{ij} b_j^{\tau}, \quad \left(\tau=\left[t \mod n_p\right]\right) \tag{7}$$

where,

n_p: Period (i.e., 52 weeks in one year).

k: Number of latent shock components.

$\mathbf{W}=\{w_{ij}\}_{i,j=1}^{d,k}$: Participation (weight) matrix, i.e., participation weight of group i for the j-th shock component.

$\mathbf{B}=\{b_j^{\tau}\}_{j,\tau=1}^{k,n_p}$: Effection matrix (consisting k effection dimensions), i.e., temporal effection at time tick τ for the j-th shock component.

The full model is illustrated in Fig. 1. The full parameter set to be learned is $S=\{g^0,r,\mathbf{K},\mathbf{A},\mathbf{W},\mathbf{B}\}$. Table I summarizes the notations used throughout this paper.

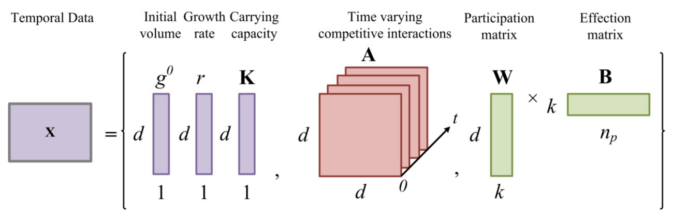

Fig. 1. Model illustration. Given a set of d sequences $\mathbf{X}$ of length n, we depict local properties of each group, i.e., initial volume: g^0, growth rate: r, carrying capacity: $\mathbf{K}$, time varying competitive matrix: $\mathbf{A}$, and a set of k shock components, i.e., participation matrix: $\mathbf{W}$ and Effection matrix: $\mathbf{B}$.

C. Group attributes

To inquiry information diffusion in social media, here we mainly employ three types of group attributes to calculate phase-difference $\phi_{i,j}^t$ defined in (5), namely emotion attributes, behavior attributes and network attributes. These attributes are summarized in Table II. Other more attributes can also be involved flexibly.

TABLE I. NOTATIONS.

Symbol	Definition
d	Number of user groups
n	Duration of sequences
X	d co-evolving time sequences (i.e., $X=\{x_1,\ldots,x_d\}$)
x_i	Temporal sequence i (i.e., $x_i=\{x_i^1,\ldots,x_i^n\}$)
x_i^t	Value of sequence i at time tick t.
g_i^t	Estimated posting number of user group i at time tick t
$\dot{g}_i^t$	Estimated posting number of user group i at time tick t when considering external shock events
g^0	Initial group size i.e., $\{g_i^0\}_{i=1}^d$
r	Growth rate i.e., $[5]_{i=1}^d$
$\boldsymbol{K}$	Carrying capacity i.e., $[17]_{i=1}^d$
$\mathbf{A}$	Interaction matrix ($d\times d$) i.e., $\mathbf{A}=\{a_{ij}\}_{i,j=1}^{d,d}$
n_p	Period (i.e., 52 weeks)
k	Number of hidden shock events
$\mathbf{E}$	External shock tensor ($d\times$ n) i.e., $\mathrm{E}=\{e_i(t)\}_{i,t=1}^{d,n}$
$\mathbf{W}$	Participation (weight) matrix ($d\times k$) i.e.,$\mathbf{W}=\{w_{ij}\}_{i,j=1}^{d,k}$
$\mathbf{B}$	Effection matrix ($k\times n_p$) i.e., $\mathbf{B}=\{b_j(\tau)\}_{j,\tau=1}^{k,np}$

TABLE II. GROUP ATTRIBUTES.

Type	Attribute	Description
Emotion	aveEmoVal$_i^t$	Average emotion valence of individuals in group i at time tick t.
	aveEmoStren$_i^t$	Average emotion strength of individuals in group i at time tick t.
Behavior	avgActLevel$_i^t$	Average activity level (posting number) of individuals in group i at time tick t.
	diffuCoef$_i^t$	Ratio of diffusion (reply or reposting) in group i at time tick t.
Network	avgFollowNum$_i^t$	Average follower number in group i at time tick t.
	networkEmbed$_i^t$	Average ratio of common friends between individuals in group i at time tick t.

Emotion attributes are designed to capture the relationship between emotion status of online users and their posting number; behavior attributes are expected to clarify whether user activity level and posting number are correlated; network attributes may uncover the effect of structure characteristics on posting number. Based on these attributes, we are able to generate an attribute vector m_i^t for each user group i at time tick t. Then, $\phi_{i,j}^t$ can be readily calculated according to (5).

D. Optimization algorithm

We now turn to the optimization of model parameters. Specifically, we need to solve two issues: (1) to find an optimal set of shock components (i.e., **W** and **B**), and (2) to efficiently estimate full parameter set *S* that best capture the important patterns in the posting behavior of different user groups.

To determine the number of component *k*, we elaborate an efficient coding scheme, which permits automatic configuration of appropriate sizes for **W** and **B**. This coding scheme is based on the minimum description length (MDL) principle. The basic idea behind this scheme is that the more we can compress the data, the more we can learn about its underlying patterns. The description complexity of model parameter set *S* are summarized in Table III.

TABLE III. DESCRIPTION COMPLEXITY OF THE PROPOSED MODEL.

Item	Description complexity (bits)
$d + n$	$\log^*(d) + \log^*(n)$
$\{g^0, r, \mathbf{K}\}$	$c_F \cdot (d \times 3)$
$\mathbf{A}$	$c_F \cdot (d \times d - d)$
$\{k, \mathbf{W}, \mathbf{B}\}$	$\log^*(k) + \log^*(n_p) + c_F \cdot (dk + kn_p)$

Given the full parameter set *S*, the original data *X* can be encoded with Huffman coding [18]. Huffman coding assign a number of bits (i.e., the negative log-likelihood) to each value in *X*. The encoding cost of *X* given parameter *S* is:

$$\mathrm{Cos}\,t_C\left(X|S\right)=\sum_{i,t=1}^{d,n}\log_2 p^{-1}_{Gauss\left(\mu,\sigma^2\right)}\left(x_i^t-\dot{g}_i^t\right) \quad (8)$$

where, x_i^t and $\dot{g}_i^t$ are the original and estimated posting number of user group *i* at time tick *t*; μ and σ^2 are the mean and variance of the distance between the x_i^t and $\dot{g}_i^t$.

The total code length for *X* with respect to a given parameter set *S* given by:

$$\mathrm{Cos}\,t_T\left(X;S\right)=\log^*\left(d\right)+\log^*\left(n\right)+\mathrm{Cos}\,t_M\left(g^0,r,\mathbf{K}\right)$$
$$+\mathrm{Cos}\,t_M\left(\mathbf{A}\right)+\mathrm{Cos}\,t_M\left(k,\mathbf{W},\mathbf{B}\right)+\mathrm{Cos}\,t_C\left(X|S\right) \quad (9)$$

The optimal number of shock components (k_{opt}) could be automatically determined by minimize (9) with regard to *k*:

$$k_{opt}=\arg\min \mathrm{Cos}\,t_T\left(X;S\right) \quad (10)$$

III. DATA

In 2015, the United States experienced a large, multi-state measles outbreak. Totally, 189 people from 24 states and the District of Columbia were reported to have measles. This outbreak has incurred fierce discussion on Twitter, with large portion of users against measles vaccination. To study how public sentiments spread in Twitter, we use Twitter4J API to collect all tweets that contain any one of the keywords listed in Table IV[1]. Keywords listed in the second line of Table IV guarantees the recall rate of the data collected. The relevance of the collected data was checked manually. The time span of the collected data is from Jan. 1st 2015 to Aug. 1st 2015. Except for the original tweets, their corresponding retweets and replies are also collected. Totally, we have collected 216,783 tweets written by 100,987 Twitter users. For a reliable analysis, users with less than five posts are pruned out. This manipulation finally generates a datasets with 5,633 users and 90,807 posts (i.e. 16 tweets per user). After data collection, SentiStrength [19] is employed to label tweets as negative, positive or neutral with respect to the intent of getting vaccinated against measles. Of all the used tweets, 13,644 were classified as positive (POS, pro-vaccination), 27,957 as negative (anti-vaccination), and 49,206 as neutral (undetermined). Statistics of the dataset is show in Table V.

TABLE IV. KEYWORDS USED FOR DATA COLLECTION.

Category	Keywords
measles	measle, morbilli, rubeola, rubeolla
vaccination	vacinate, vaccinate, vaccination

TABLE V. DATA STATISTICS.

User Group	Tweet Nr.	Retweet Nr.	Reply Nr.	Total
POS	11158	1582	904	13644
NEU	39033	7403	2770	49206
NEG	22039	3883	2035	27957
Total	72230	12868	5709	90807

IV. RESULTS

We analyze user sentiments in Tweet during measles outbreak in the USA from three aspects. First, we evaluate how our model fit the empirical data of user group size. Secondly, we check the model's performance in prediction. Thirdly, we check whether the model's shock tensor could detect typical external shocks, and how these shocks impact online user sentiments.

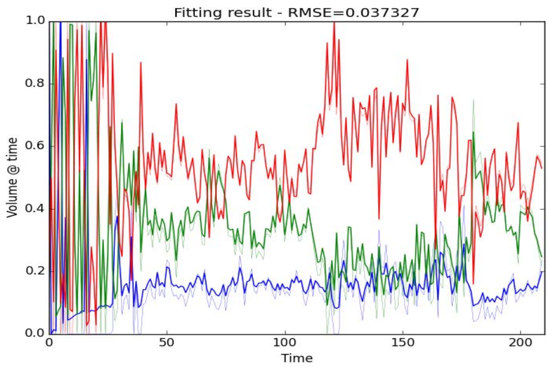

Fig. 2. Fitting results. Our model (deep color lines) fits the original data (light color lines) very well. Red, green and bluess repestively corresponding to user groups holding neutral, negative and positive sentiment.

A. Model Accuracy

Fig. 2 shows the fit of H2 model to the posting number of user groups with different sentiments (in ratio). The good agreement supports for the proposed model. To test the

[1] Keywords are selected according to CHV Wiki: http://consumerhealthvocab.chpc.utah.edu/CHVwiki/

rationale of the proposed H2 model, we also compare its fitting ability with three alternative models, i.e., the model only considers homophily (H model), the Lotka-Volterra population model (LV model) [20], and AR model. Fig. 3 shows the root mean square error (RMSE) between the original and estimated posting number from different models. A lower value indicates a better fitting accuracy. As shown in Fig. 3, our approach achieved highest fitting accuracy. Since the LV model cannot capture external shocks, it was strongly affected by multiple spikes and failed to capture co-evolving dynamics. H Model has the ability to capture the homophily principle, but it was not completely successful in capturing the competition between homophily and homoestasis.

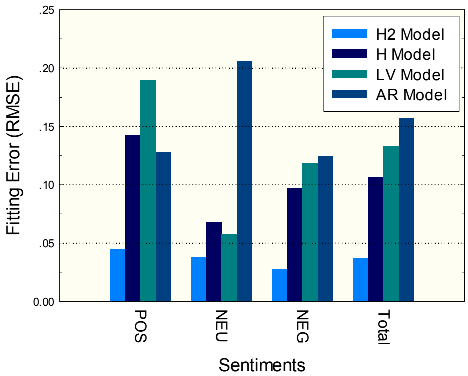

Fig. 3. Fitting Error of Different Models.

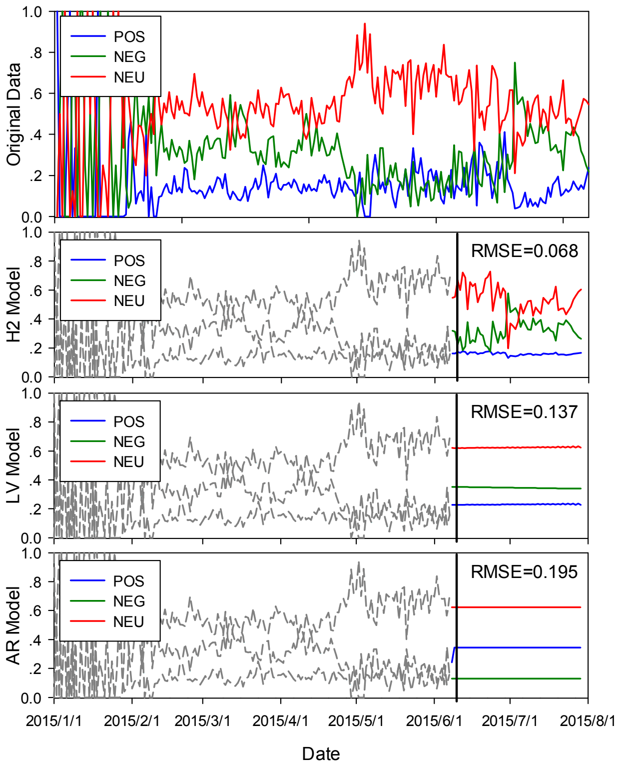

Fig. 4. Prediction Performance of Different Models.

B. Prediction Performance

This subsection present the prediction ability of the proposed model compared with two alternatives, i.e. LV model and AR model. Fig. 4 shows the prediction performance of different models. We train the model parameters by using the 3/4 values for each user group (black dashed lines in Fig. 4), and then predict the rest (colored lines). In Fig. 4, from the top to the bottom rows show the original data, and the prediction results of H2, LV, and AR, respectively. As shown in Fig. 4, our model successfully predicts the long-range evolution of user sentiments, while LV fails to capture external shocks, and AR fails to capture the non-linear evolutions.

C. External shocks events

The effects of external shocks are captured by the participation matrix **W** and the Effection matrix **B**. The values of the learned **W** are -1.157, 5.394 and 3.826 respectively for POS, NEU, and NEG user group. This results suggest that external shock events affect POS group and NEG group differently, i.e., users in NEG group are more influenced by external shocks (absolute value is larger). Fig. 5 summarizes external shock events detected. There are two drops, one occurs when CDC (Centers for Disease Control and Prevention) reported 235 measles cases; and another one occurs during the World Immunization Week. These two events depressed anti-vaccination sentiments in Twitter (W • B<0). There is also a peak when one death due to measles was reported on Jul 2 2015. After this peak, anti-vaccination sentiment weakens gradually (as shown by the decrease of curve b1).

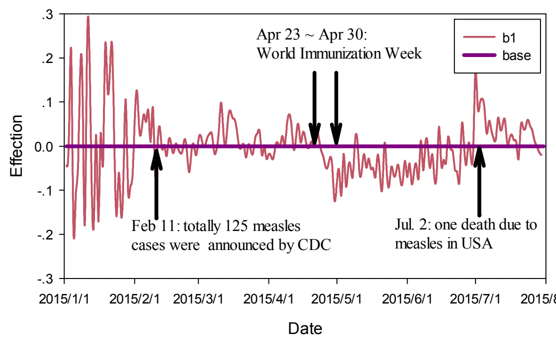

Fig. 5. External Shock Events.

V. LITERATURE REVIEW

The related work falls into the following large subgroups:

Homophily: Homophily principle accounts for the tendency that ties between individuals are strongly favored by the similarity of their attributes. Under this effect, the interactions between similar individuals are enhanced. Such enhancement can affect the extent of sentiments or behaviors' adoption in a population. At the individual level, homophilous ties can promote the spread of information between individuals [21]. While at the community level, homophily will reduce overall information coverage, thereby increasing intergroup inequality across diverse populations [22].

Homestasis: Homeostasis is the tendency for a system to maintain a relatively stable, constant state of balance. When any deviation from homeostasis occurs, the system responds by enacting negative feedback to bring the system back to a state of equilibrium. When this principle applies to society, the enhancement of some interactions from an individuals is counter-balanced by the weakening of other interactions of the same individual to his friends.

Dynamic Models: Traditional approaches applied to analyze temporal data include auto-regression (AR), linear dynamical systems (LDS), Kalman filters (KF). These approaches could not capture nonlinear dynamics. To deal with this problem, various non-linear models are proposed, such as the Lotka-Volterra (LV) model, the logistic function (LF) [23], and the susceptible-infected (SI) model. These models incorporate domain knowledge. However, they are not intended to capture co-evolving online activities. Further, all of them examine temporal data on static complex networks. Thus, they are not inadequate for describing many real-world networks, which are intrinsically time-varying [24].

In short, none of the existing methods examine non-linear dynamics in a time-varying manner, and the underlying mechanism governing the dynamic process demands further investigation.

VI. Conclusions and Future Work

We have proposed a new statistical model (H2 model) to depict the evolution of temporal network, which is governed by the competition of homophily and homeostasis. Evaluation on Twitter data suggests that H2 model can capture long-range sentiment dynamics and lower the error rate in sentiment prediction. Through the model's shock tensor, we successfully detect several typical events, and reveal that users in negative emotions are more influenced by external shock events.

In our future work, we intend to clarify whether user attributes correlate with interaction strength. As network structure affects social interactions occurring on it, we are aslo interested in uncover the relationship between network diversity and sentiment diffusion in online communities.

Acknowledgment

This work was supported in part by the following grants: the National Natural Science Foundation of China under Grant Nos. 71402177, 71472175, and 71103180; and the Ministry of Health under Grant No. 2013ZX10004218.

References

[1] N. A. Christakis, and J. H. Fowler, "The collective dynamics of smoking in a large social network," *New England journal of medicine,* vol. 358, no. 21, pp. 2249-2258, 2008.

[2] V. A. Jansen, N. Stollenwerk, H. J. Jensen, M. Ramsay, W. Edmunds, and C. Rhodes, "Measles outbreaks in a population with declining vaccine uptake," *Science,* vol. 301, no. 5634, pp. 804-804, 2003.

[3] M. Salathé, and S. Bonhoeffer, "The effect of opinion clustering on disease outbreaks," *Journal of The Royal Society Interface,* vol. 5, no. 29, pp. 1505-1508, 2008.

[4] S. B. Omer, K. S. Enger, L. H. Moulton, N. A. Halsey, S. Stokley, and D. A. Salmon, "Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis," *American Journal of Epidemiology,* vol. 168, no. 12, pp. 1389-1396, 2008.

[5] M. De Choudhury, Y.-R. Lin, H. Sundaram, K. S. Candan, L. Xie, and A. Kelliher, "How does the data sampling strategy impact the discovery of information diffusion in social media?," *ICWSM,* vol. 10, pp. 34-41, 2010.

[6] J. Goldenberg, B. Libai, and E. Muller, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing letters,* vol. 12, no. 3, pp. 211-223, 2001.

[7] J. Yang, and J. Leskovec, "Modeling information diffusion in implicit networks," in Data Mining (ICDM), 2010 IEEE 10th International Conference on, 2010, pp. 599-608.

[8] R. Gutiérrez, A. Amann, S. Assenza, J. Gómez-Gardenes, V. Latora, and S. Boccaletti, "Emerging meso-and macroscales from synchronization of adaptive networks," *Physical review letters,* vol. 107, no. 23, pp. 234103, 2011.

[9] G. Kossinets, and D. J. Watts, "Empirical analysis of an evolving social network," *science,* vol. 311, no. 5757, pp. 88-90, 2006.

[10] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, pp. 415-444, 2001.

[11] G. G. Turrigiano, and S. B. Nelson, "Homeostatic plasticity in the developing nervous system," *Nature Reviews Neuroscience,* vol. 5, no. 2, pp. 97-107, 2004.

[12] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multidimensional trajectories," in Data Engineering, 2002. Proceedings. 18th International Conference on, 2002, pp. 673-684.

[13] R. M. May, "Qualitative stability in model ecosystems," *Ecology*, pp. 638-641, 1973.

[14] J. Gómez-Gardenes, Y. Moreno, and A. Arenas, "Paths to synchronization on complex networks," *Physical review letters,* vol. 98, no. 3, pp. 034101, 2007.

[15] S. He, X. Zheng, D. Zeng, C. Luo, and Z. Zhang, "Exploring Entrainment Patterns of Human Emotion in Social Media," *PloS one,* vol. 11, no. 3, pp. e0150630, 2016.

[16] Y. Matsubara, Y. Sakurai, W. G. Van Panhuis, and C. Faloutsos, "FUNNEL: automatic mining of spatially coevolving epidemics," in Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, 2014, pp. 105-114.

[17] K. Kirst-Ashman, *Human Behavior in the Macro Social Environment*: Cengage Learning, 2010.

[18] C. Böhm, C. Faloutsos, J.-Y. Pan, and C. Plant, "RIC: Parameter-free noise-robust clustering," *ACM Transactions on Knowledge Discovery from Data (TKDD),* vol. 1, no. 3, pp. 10, 2007.

[19] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment strength detection in short informal text," *Journal of the American Society for Information Science and Technology,* vol. 61, no. 12, pp. 2544-2558, 2010.

[20] J. D. Murray, *Mathematical Biology. II Spatial Models and Biomedical Applications {Interdisciplinary Applied Mathematics V. 18}*: Springer-Verlag New York Incorporated, 2001.

[21] E. M. Rogers, *Diffusion of innovations*: Simon and Schuster, 2010.

[22] P. DiMaggio, and F. Garip, "How network externalities can exacerbate intergroup inequality," *American Journal of Sociology,* vol. 116, no. 6, pp. 1887-1933, 2011.

[23] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez, *Mathematical models in population biology and epidemiology*: Springer, 2001.

[24] P. Holme, "Network reachability of real-world contact sequences," *Physical Review E,* vol. 71, no. 4, pp. 046119, 2005.