

## Research Article

# A Hybrid Human Dynamics Model on Analyzing Hotspots in Social Networks

Yunpeng Xiao,<sup>1,2</sup> Bai Wang,<sup>1</sup> Bin Wu,<sup>1</sup> Zhixian Yan,<sup>3</sup>  
Shousheng Jia,<sup>2</sup> and Yanbing Liu<sup>2</sup>

<sup>1</sup> Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia,  
Beijing University of Posts and Telecommunications (BUPT), Beijing, China

<sup>2</sup> Chongqing Engineering Laboratory of Internet and Information Security,  
Chongqing University of Posts and Telecommunications (CQUPT),  
Room 4029, No. 2 Chongwen Road, Nanan District, Chongqing 400065, China

<sup>3</sup> School of Computer and Communication Sciences,  
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

Correspondence should be addressed to Yunpeng Xiao, shineagle2005@hotmail.com

Received 5 July 2012; Accepted 12 September 2012

Academic Editor: Garyfalos Papaschinopoulos

Copyright © 2012 Yunpeng Xiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The increasing development of social networks provides a unique source for analyzing human dynamics in the modern age. In this paper, we analyze the top-one Internet forum in China (“Tianya Club”) and identify the statistical properties of hotspots, which can promptly reflect the crowd events in people’s real-life. Empirical observations indicate that the interhotspot distribution follows a power law. To further understand the mechanism of such dynamic phenomena, we propose a hybrid human dynamic model that combines “memory” of individual and “interaction” among people. To build a rich simulation and evaluate this hybrid model, we apply three different network datasets (i.e., WS network, BA network, and Karate-Club). Our simulation results are consistent with the empirical studies, which indicate that the model can provide a good understanding of the dynamic mechanism of crowd events using such social networking data. We additionally analyze the sensitivity of model parameters and find the optimal model settings.

## 1. Introduction

In China, Internet forum is still one type of the most popular social networking sites for various information propagation and discussion among people. For example, Tianya Club (<http://www.tianya.cn/>, simply noted as “TianYa” in this paper), the 12th most visited website in China, is China’s biggest Internet forum that provides almost of social networking services like BBS, blog, microblog, and photo sharing, and so forth,

([http://en.wikipedia.org/wiki/Tianya\\_Club](http://en.wikipedia.org/wiki/Tianya_Club)). Up to April 2012, Tianya has more than 68 millions registered users and more than one million online users at most of the times. Such forums have tons of information, not only from the perspective of individual behaviors but also in terms of human interactions. Therefore, such social networking sites provide great potential to analyze human behaviors for understanding human dynamics.

In traditional studies, human behavior is usually assumed as random activity and thus be modeled as Poisson processes [1]. This assumption leads to an exponential interevent time distribution of human activities. However, a lot of recent empirical studies have already proved that this assumption is wrong. For example, Barabási first discovered that the time-interval between sending an email and receiving a reply follows a power-law distribution, with heavy tails [2]. Afterwards, a couple of similar statistical properties in human dynamics are empirically discovered by using various datasets, including web browsing [3], short message sending [4], microblogging [5], netizens' behaviors on the forum [6], movie watching [7], and so on.

To understand the intrinsic factor of such heavy-tailed property, Barabási and Vázquez first propose a priority queuing model and successfully explain the phenomenon of human behavior based on task queue [2, 8, 9]. Subsequently, researchers design various human dynamic models for different scenarios, such as aging model [10], optimization mechanism [11], influence of deadline [12], interest-driven model [13, 14], interest and social identity codriven model [5], and relative clock model [15]. These models are largely based on the individual level but not on the crowd level. Recently, there are some emerging crowd-level empirical studies and models, which are largely focusing on network emergencies or terrorist incidents. For example, Johnson et al. propose a self-organizing system that dynamically evolves through the continual coalescence or fragmentation of its constituent groups [16–18]. Galam and Moscovici design group decision-making models by using the percolation theory [19–21]. These researches study the social behavior in the network, considering both “individual behavior” and “interaction between individuals.” However, these works are mainly focused on the social psychology methods and are based on a complete graph (“everyone interacts with everyone”), ignoring the limited structural features of social network. In this paper, we focus on analyzing the real-life social networking datasets. Clauset et al. draw a substitution and competition model for terrorism [22].

Vazquez et al. first propose a memory model to analyze human dynamics [23]. The memory models consider that humans have perceptions of their past activities, and therefore humans accelerate or reduce their activity rates according to their memories. Such memory models provide a good understanding of the possible dynamic mechanism in various scenarios, for example, interevent time statistics of email and letter communications [23], terrorism attacks [24]. In addition to memory, interaction and influence from the neighborhood are used to complement the memory model [24, 25]. For example, Zhu et al. propose a model that combines the role of individual social conformity and self-affirmation psychology for analyzing the possible dynamic mechanism of terrorism attacks [24]. Nevertheless, these human dynamic studies on memory and interaction have several limitations, for example: (1) the interactions are only based on a small group (e.g., 2 agents or 4 neighbors in the 2D lattice network), which are not real-life social network with arbitrary relationships, (2) different nodes have different social identity and social influence in real-life-which is not reflected in these models, (3) in these models, the impact of neighbor nodes is ignored while the status of such neighbor nodes is opposite to the node itself.

In this paper, we study the combined impact of memory and node influence (i.e., interactions) of human dynamics in arbitrary social networks. We analyze the human

behaviors in China’s largest Internet forum (“Tianya Club”), including activities like posting a new topic and adding comments to existing topics. A hotspot in Tianya is the topic with burst comments. We can consider a hotspot as a crowd event in social network. Based on the Tianya datasets, experimental evidence shows that different types of intertime distributions of hotspot topics follow power-law. In addition, we propose a human dynamic model that combines individual habit (i.e., “memory”) and node influence (i.e., “interaction”). While testing with several well-known network datasets, the simulation results of our model are consistent with the empirical observations, which implies that our model offers a suitable explanation of the power-law properties in human dynamics.

This paper is organized as follows: after the introduction in Section 1, Section 2 describes the Tianya data; Section 3 shows the empirical results; Section 4 presents our hybrid model on the combination of memory and interaction; Section 5 compares the results of simulation and the empirical ones. Section 6 provides more discussions and Section 7 concludes this paper.

## 2. Data Description

Empirical data are collected from TianYa, which is one of the largest online social networking sites in China. Up to the time of writing, there are 68,360,259 users (with unique IDs) registered in TianYa. The news and topics in Tianya cover all aspects, and therefore it provides a rich dataset to reflect Chinese people’s activities and dynamics. The Tianya data has been studied in [26], analyzing the intercomment time distribution using a simple growing-network based model. In this paper, we study a rich and hybrid model considering both memory and interaction. We analyze interhotspot time distribution between outbreak topics, and evaluate the model with 9-month data from three representative topic sessions, namely, “Social-Life” (Session-A), “Tittle-Tattle” (Session-B), and “Entertainment-Gossip” (Session-C). Tables 1 and 2 present the data summary and the data format, respectively. It is worth noting that *startTime* in Table 2 means the release time of an initial topic, and *endTime* means the release time of the last reply/comment of this topic.

## 3. Empirical Results

This section provides the empirical studies on the Tianya topics. Each topic has an initial post and many following replies (see the topic format in Table 2). We sort topics in a descending order according to the discussion properties (e.g., the number of replies, the number views, or the sum of both) and identify the top  $N$  topics as the hotspots that have maximum discussions, somehow reflecting the crowd events in real-life. Afterwards, we resort these hotspot topics according to their *startTime* or *endTime*, and analyze such interhotspot time distribution of the outbreak topics.

In detail, we have three sessions of 9-month data (see Table 1) and extract hotspots using three ordering choices (i.e., reply number, view number, or the sum of both). We consider five cases of top  $N$  topics, that is,  $N \in \{1000, 2000, 3000, 4000, 5000\}$ . In addition, the interhotspot time can be calculated by either *startTime* or *endTime* of each outbreak topic. Therefore, in total, we have 90 experiments (3 sessions  $\times$  3 orders  $\times$  5 *top-N*  $\times$  2 times). Due to the lack of space, we could not provide all 90 experimental plots but a subset in Figure 1: 3 sessions, 3 orders, using  $N = 1000$  (top 1000 hotspots), using *endTime*, that is,  $3 \times 3 \times 1 \times 1$ . We observe that the intertime distributions of outbreak topic follow power-law and span

**Table 1:** Data summary of 3 sessions (2011.1.1–2011.9.30).

Session	Number of topics	Number of topic replies	Number of total views
A: Social-Life	66,497	984,367	94,709,066
B: Tittle-Tattle	159,806	18,869,664	1,600,825,097
C: Entertainment-Gossip	260,593	30,506,622	3,376,348,257

**Table 2:** Data format.

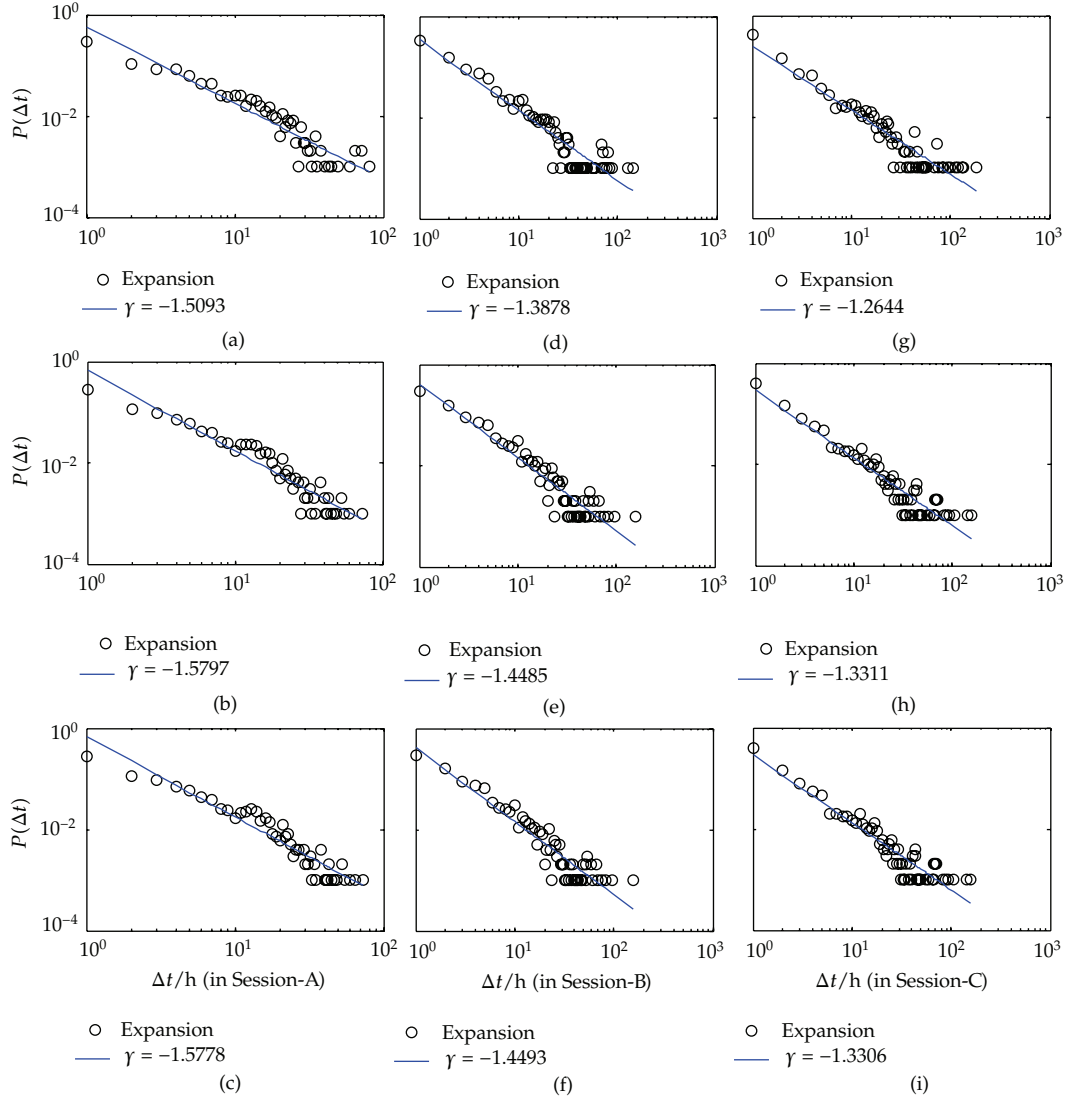
Topic title	Initial author	<i>startTime</i>	<i>endTime</i>	Number of replies	Number of views
...	...	...	...	...	...

more than two orders of magnitude, with exponent change from  $-1.2644$  to  $-1.5797$ . Similar span of the interval is found in [7]. This range is smaller than the one in [2–5]. A possible reason could be that we use *hour* as the time unit in calculating intertime distribution for our 9 months dataset, while the time unit in these related works is smaller either *minute* or *second*.

Instead of showing the detailed power-law distributions for other 81 (i.e., 90–9 in Figure 1) experiments, for example, different  $N$ , varying ordering strategies. Figure 2 shows the relationship between hotspot number  $N$  and power exponent  $\gamma$  in various experimental settings. In Figure 2(a), we observe that the power exponent  $\gamma$  increases for all sessions when the number of hotspot  $N$  grows. The heavy tail phenomenon tends to disappear when  $N$  becomes larger. As an extreme case, there will be a hotspot in every hour if  $N$  is huge. Of course, such extreme case is meaningless as the topic is not real hotspot if the topic’s ranking order significantly lags behind (e.g.,  $N \gg 5000$ ). Actually, the interhotspot time distributions of outbreak topics in all 3 sessions can lose power-law characteristics gradually when  $N > 5000$ . In addition, we analyze the difference using hotspot topics’ *endTime* or *startTime* in Figure 2(b), and the different ordering strategies (i.e., via reply number, view number, or the sum of both) in Figure 2(c). We observe that using *endTime* can bring larger exponential compared to using *startTime*, this is because we clean those topics from the top  $N$  topics whose release time are before 2011.1.1. And there is no significant difference between different ordering strategies which tell us that when a hotspot topic attracts more replies, more views are attracted also.

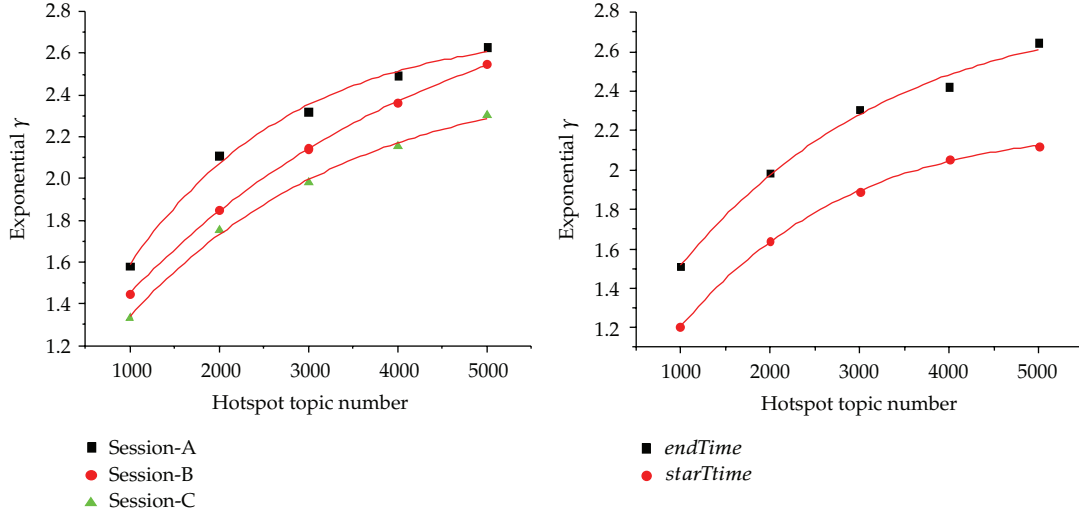
## 4. A Hybrid Model

To understand the intrinsic mechanism of online forum’s outbreak topics (hotspots) that are corresponding to human dynamics in terms of crowd events in social networking, we propose a rich model in this section. This model considers both the inner habit of an individual (called “memory”) and the interaction with social environment (“interaction”); therefore, the model is hybrid. From the memory aspect, a person who was active/inactive in contributing to topics in the past almost keeps the same style in future topics. From the interaction aspect, the behavior of each individual can be affected by the surroundings around us (i.e., neighboring nodes of an individual). Furthermore, people have different social roles in the community, and hereinafter their impacts are distinct from each other. Therefore, we study a hybrid model that combines the impact of memory and interaction in this paper. The ket points of the model are as follows.



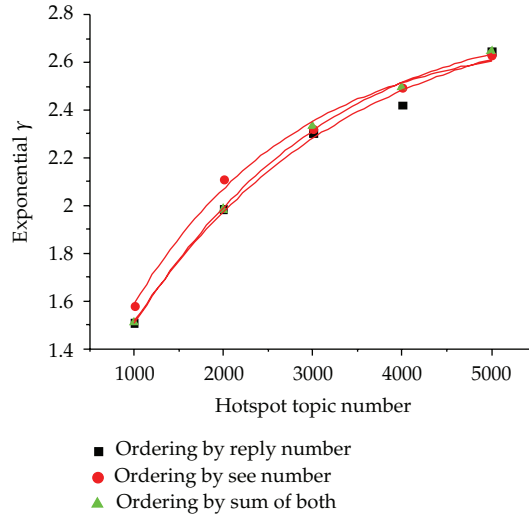
**Figure 1:** Interval distribution of *endTime* of hotspots. (a)–(c) show Session-A, (a) top 1000 posts sorted by reply number (b) top 1000 posts sorted by view number (c) top 1000 posts sorted by sum of reply and view number (d)–(f) show Session-B, similar to (a)–(c). (g)–(i) show Session-C, similar to (a)–(c).

- (1) *Time-discretization*: the time step is discretized in terms of  $\delta_t = 1$  (e.g., one hour in analyzing our Tianya datasets). Therefore, the status of crowd events (e.g., hotspots in our experiment) evolves/changes with timestamp  $t$  (using “hour” as the unit).
- (2) *Social-network*: people (e.g., registered users in Tianya) can be formalized as an undirected graph in terms of a social network.  $G = (V, E)$ ,  $V = \{v_1, v_2, \dots, v_i, \dots, v_n\}$  stands for a node set. Each individual user in the network is expressed as a node  $v_i$  in  $V$ , the number of nodes is  $n = |V|$ . An undirected edge set  $E$  represents social relationships in the crowd, that is,  $N(v_i) = \{v_{i_1}, v_{i_2}, \dots, v_{i_m}\}$  stands for the adjacent node set of node  $v_i$ .  $D(v_i)$  is the degree of node  $v_i$ .  $S(v_i, t)$  is status of node  $v_i$ .



(a) Ordering by *reply number* and choosing topic's *endTime* in three different sessions

(b) Different time scale distribution ordering by reply number and in the same Session-A



(c) Ordering by three different methods and in the same Session-A

**Figure 2:** Relationship between hotspot number  $N$  and power exponent  $\gamma$ .

at timestamp  $t$ . Each node has two possible states, that is,  $S(v_i, t) = \text{ignore} || \text{focus}$ , which represents whether node  $v_i$  concerns the current event in the crowd or not. It is worth noting the order of *ignore* and *focus* is not important in our model, as it does not affect the model's behavioral characteristics in the simulation. We only require  $\text{ignore} \neq \text{focus}$ , and in our simulation we apply  $\text{ignore} < \text{focus}$  as a regular scenario.

- (3) *Crowd-events (hotspots)*: the emergence of a crowd event is as follows: firstly, a user posts a new topic; afterward, more and more people start to participant in this topic (e.g., users reply and view a topic in Tianya); after the participant number

satisfies certain conditions, this topic becomes an outbreak topic (i.e., hotspot). As the time grows, there likely appear new events/topics that incrementally become more important and more interesting. In this way, new crowd events (i.e., hotspots in Tianya) show up and the old ones disappear gradually.

To model the intrinsic human dynamics in the crowd events, we consider both the external factor of interaction mechanism from surroundings in the network and the internal factor of individual's memory mechanism.

### Interaction

Interaction is to model the external factor that stands for the influence from neighboring nodes in the network. Considering different nodes with distinctive impacts, the impact of node  $v_i$  is denoted as  $D(v_i)$ . For a node  $v_j \in N(v_i)$ , the influence of node  $v_j$  to  $v_i$  at timestamp  $t$  is as follows:

$$\text{Affect}(v_j, v_i, t) = \frac{S(v_j, t) * D(v_j)}{\sum_{k=i_1}^{i_m} D(v_k)}, \quad v_j, v_k \in N(v_i), \quad (4.1)$$

where  $S(v_i, t) \in \{\text{ignore}, \text{focus}\}$  is the status of  $v_i$  at timestamp  $t$ . In such case, the total influence to node  $v_i$  from its neighbors is as follows:

$$\text{TotalAffect}(v_i, t) = \sum_{j=i_1}^{i_m} \text{Affect}(v_j, v_i, t), \quad v_j \in N(v_i). \quad (4.2)$$

As  $S(v_i, t) \in \{\text{ignore}, \text{focus}\}$ , we have both  $\text{Affect}(v_j, v_i, t) \in [\text{ignore}, \text{focus}]$  and  $\text{TotalAffect}(v_i, t) \in [\text{ignore}, \text{focus}]$ . Therefore, the distance between node  $v_i$  and the status of its neighboring nodes at time step  $t$  is defined as

$$\text{StatusDistance}(v_i, t) = |\text{TotalAffect}(v_i, t) - S(v_i, t)|. \quad (4.3)$$

As  $\text{TotalAffect}(v_i, t) \in [\text{ignore}, \text{focus}]$  and  $S(v_i, t) \in \{\text{ignore}, \text{focus}\}$ , we have  $\text{StatusDistance} \in [0, |\text{ignore} - \text{focus}|]$ , where  $|\text{ignore} - \text{focus}|$  is the maximum length of status distance. Then, the possibility of status change of node  $v_i$  resulted from external cause is

$$\text{External}(v_i, t) = \frac{\text{StatusDistance}(v_i, t)}{|\text{ignore} - \text{focus}|}. \quad (4.4)$$

As  $\text{StatusDistance} \in [0, |\text{ignore} - \text{focus}|]$ , we have  $\text{External}(v_i, t) \in [0, 1]$ . There are two extreme cases: (1) if all statuses of neighbors of  $v_i$  are consistent with node  $v_i$  at timestamp  $t$ , then at timestamp  $t + 1$ , the probability of node  $v_i$  status changes due to external cause is 0; and on the contrary, (2) if all neighbors of node  $v_i$  have opposite status to node  $v_i$  at timestamp  $t$ , at timestamp  $t + 1$ , the probability of  $v_i$  status changes  $\text{External}(v_i, t) = 1$ .

### Memory

Another factor can result in change of node status is of course the internal cause. Considering the node habit, a person was actively involved in a topic is very likely to participate in future discussion; and a person who was not stick to his opinion in the past also has high chance to change his position in future events/topics. Assume at timestamp  $t$ , our model records the status sequence  $\{S(v_i, t), S(v_i, t-1), \dots, S(v_i, t-\Delta+1)\}$  of node  $v_i$  in the past  $\Delta$  time steps. We calculate the total number of status change in two consecutive steps as  $\Delta_{\text{change}}$ , and therefore the number of status does not change is  $\Delta - \Delta_{\text{change}}$ . In such case, the possibility of status change caused by internal cause at timestamp  $t+1$  can be defined as

$$\text{Internal}(v_i, t) = \frac{\Delta_{\text{change}}}{\Delta}. \quad (4.5)$$

Now, in terms of combing external cause and internal cause together, we can have the probability of status change of node  $v_i$  at timestamp  $t+1$ :

$$P(v_i, t+1) = a * \exp(c * \text{External}(v_i, t)) + b * \exp(d * \text{Internal}(v_i, t)). \quad (4.6)$$

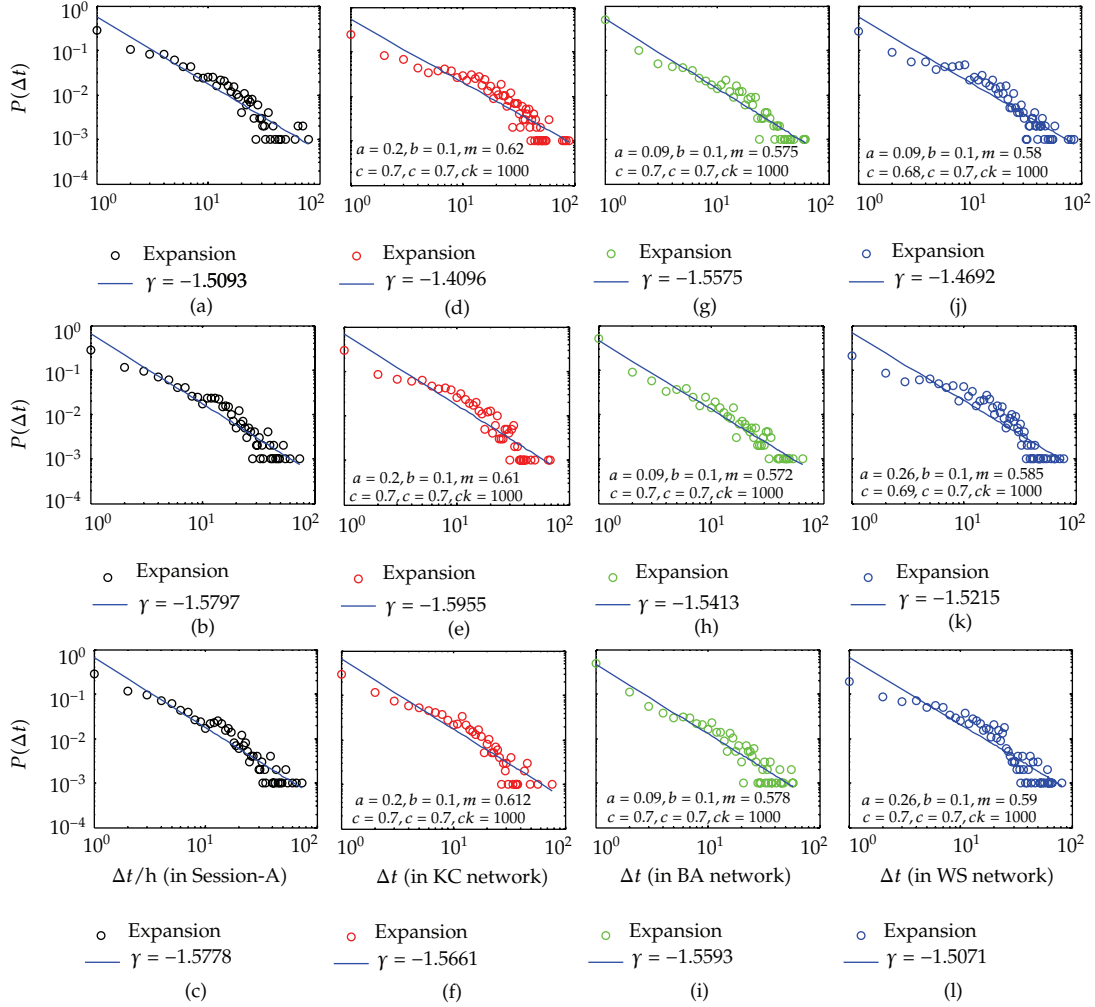
The coefficient  $a$  and  $b$  stand for the crowd acceptance of internal cause and external cause, respectively, and the coefficient  $c$  and  $d$  are the individual acceptance of internal cause and external cause, respectively. In addition, we have two more experimental parameters:  $m$  and  $ck$ .  $m$  represents the ratio of the number of people staying in the status of focus in the whole crowd, and  $m \in [0, 1]$ .  $ck$  is for recording the current number of crowd events, which should be consistent with the number of hotspots in our previous empirical studies in Tianya's datasets.

## 5. Simulation

To validate our hybrid model, we build simulations using three well-known social network datasets, that is, WS network [27] (100 nodes, 4 initial neighbors, 0.1 rewiring probability), BA network [28] (100 nodes, 4 links by new node), and Zachary's Karate Club (KC) network [29]. For the value of  $S(v_i, t)$ , we set ignore as  $-1$  and focus as  $1$ . The initial status of each node is randomly assigned according to the uniform distribution. The human internal memory is not unlimited, and we set  $\Delta = 7$  like the literature [30, 31]. As mentioned in Section 4, our model has six main parameters, that is,  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $m$ , and  $ck$ . For each simulation, we fix the setting of the 6 parameters, run the model 50 times with different initial assignment, and calculate the average results from 50 independent runs.

In Section 3, we discussed that there are 90 empirical experiments in total. By using the three different social networks (i.e., WS, BA, KC), we have  $90 \times 3 = 270$  simulations in total to verify our model. Due to the lack of paper space, we first only pick the empirical results of Session-A ("People-Life") in the Tianya dataset, namely, Figures 1(a) and 1(c) as the target of the simulations. The simulation results using these three social networks are shown in Figure 3. Here crowd event counter  $ck$  is set to 1000, which is consistent with the hotspot number ( $N$ ) of the empirical results in Figure 1. Figure 3 verifies that our model simulations are consistent with the empirical results. More detailed discussions will be provided in Section 6, and now we focus on analyse sensitivity of important parameters like  $m$  and  $ck$ .



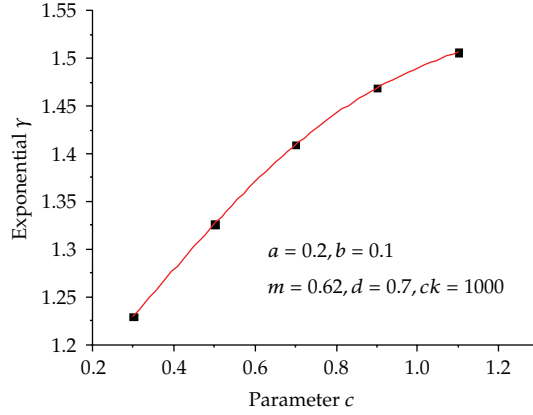


**Figure 3:** Comparing empirical results with simulation results. Figure 3(a)–3(c) with black color are the same in Figure 1(a)–1(c). Figure 3(d)–3(f) with red color are the simulation results on the KC network. Figure 3(g)–3(i) with green color are the simulation results on the BA network. Figure 3(j)–3(l) with blue color are the simulation results on the WS network.

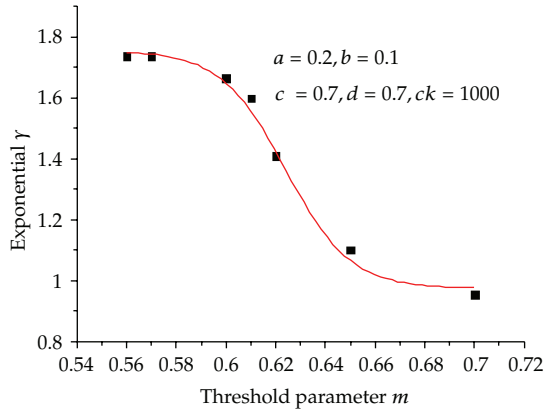
**Table 3:** Parameters  $a$  and  $b$ , running on 3 social networks with 3 TianYa sessions.

	Session-A	Session-B	Session-C
WS network	$a = 0.2, b = 0.1$	$a = 0.07, b = 0.1$	$a = 0.06, b = 0.08$
BA network	$a = 0.09, b = 0.1$	$a = 0.09, b = 0.1$	$a = 0.08, b = 0.09$
Karate Club network	$a = 0.2, b = 0.1$	$a = 0.12, b = 0.1$	$a = 0.08, b = 0.1$

From our comprehensive simulation results (including Figure 3), we observe that the value of  $a, b$  remains stable when the simulation reaches a good performance with regards to the empirical results. Table 3 shows the most suitable parametric settings of  $a, b$  for the three session data in TianYa, corresponding to the three social networks.



**Figure 4:** Effect of parameter  $c$  on power exponent  $\gamma$ .



**Figure 5:** Effect of parameter  $m$  on power exponent  $\gamma$ .

From our simulation, we also observe that  $c$  and  $m$  are the main factors that influence the value of power exponent ( $\gamma$ ), and the effect of parameter  $d$  to  $\gamma$  is not significant. Figures 4 and 5 show the sensitivity of  $c$  and  $m$ , respectively. We fix the values of other parameters, for example,  $a = 0.2$ ,  $b = 0.1$ ,  $d = 0.7$ , and  $ck = 1000$ . From Figure 4, we observe that while  $c$  changed from 0.3 to 1.1,  $\gamma$  varies from 1.2295 to 1.5065; and from Figure 5, we observe that while  $m$  changed from 0.56 to 0.70,  $\gamma$  changes from 0.956 to 1.7382. This  $\gamma$  scope covers the range of  $\gamma$  in the empirical experiments. When  $m > 0.72$ , the interhotspot time distribution starts to lose the power-law characteristics. With the increase of  $m$ , the frequency of hotspots decreases. This is consistent with the ground-truth cases: in a large social network (many registered IDs in TianYa), there might be many people that are interested in a hotspot in a given session (e.g., Session-A on people life), but it is still impossible to attract all users to be interested in this hotspot topic.

Furthermore, the hotspots counter  $ck$  increases from 1000 to 2000, 3000, 4000, and 5000, corresponding to the empirical studies in Section 3. As shown in Figure 6, we observe that the effect of  $ck$  on  $\gamma$  is consistent with the empirical experiment in Figure 2(c). When  $ck > 5000$ , the interhotspot time distribution that was generated by our model will lose power-law characteristics gradually. By fixing  $ck$  and adjusting other parameters, we can achieve all reasonable  $\gamma$  that covers the range of power exponent in the empirical experiment.

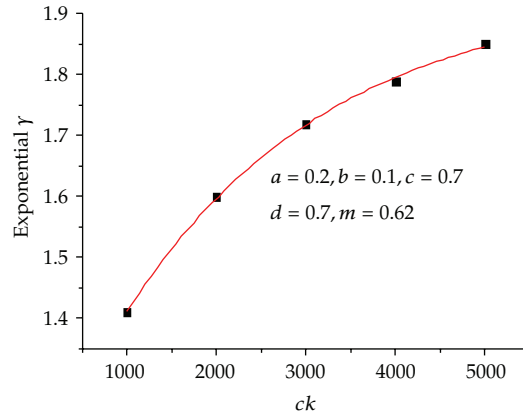


Figure 6: Effect of parameter  $ck$  on power exponent  $\gamma$ .

## 6. Discussion

From the simulation in Section 5, we identify the stability of the model coefficients  $a, b$  values for a specific network (i.e., WS, BA, KC) in a specific TianYa topic session (i.e., A-Social-Life, B-Tittle-Tattle C-Entertainment-Gossip). For example, in the simulation of A-social-life using the BA network, we observe the stable value of parameter  $a, b$  from Figure 3(g)–3(i). Such stability means that a social network has stable internal acceptances and external acceptances to individuals in general. Furthermore, the values of  $a, b$  can vary with regards to different topic sessions and different social networks, which indicates that different crowd/networks present different internal and external acceptance.

Regarding parameter  $d$  of the internal effects, it has little influence on the hotspots. This means that the influence of some individual variation to the hotspots/crowd events is quite low and even can be ignored. On the contrary, the external impact  $c$  has significant influence on the power exponent  $\gamma$ . It indicates that individual people are highly affected by external context (e.g., the behaviors of soundings in the network). Therefore, in a social network, the stronger influence of surroundings and social propaganda, the higher chance of crowd event can outbreak (e.g., appear of a new hotspot in an Internet forum like TianYa). Nevertheless, Figure 4 shows that when parameter  $c$  grows to a certain level,  $\gamma$  tends to be stable, which indicates that the influence of environment and interaction between nodes is not infinite. The individual internal factor also takes effects, for example, some users never join the discussion of a hotspot topic.

Based on the comparison between our model simulations and the empirical results, the outbreak of hotspots in social networks and the interhotspot time distribution highly depend on the two aspects, that is, the internal memory mechanisms and the external interaction in the networks. In particular, the mutual influence between nodes is the main factor to the final power exponent ( $\gamma$ ). Suitable parameters are required in the model simulation, and too high or too low parameters can result in impractical simulations that are far away from real-life crowd events in social networks. For a large topic session in online social forum like TianYa, the value of  $m$  indicates how many users are interested in a hotspot. If  $m$  is too large, almost every user focuses on the same hotspot, and this situation is quite irregular; and on the other hand, if  $m$  is too small, such hotspots are not real interesting crowd events.

Therefore, for simulations on a selected topic session in TianYa, we find suitable settings of four model parameters (i.e.,  $a$ ,  $b$ ,  $c$ , and  $d$ ) and then fix them and study the sensitivity of  $m$ . This is because for a specific session with a chosen network, the network internal/external acceptance and the individual internal/external acceptance are stable.

## 7. Conclusions

Social networking sites like Internet forum (e.g., TianYa in China) provide a unique way for rapid information propagation and discussion. Research on the laws underlying user behaviors on such social networking sites means a lot in understanding human dynamics and in turn can provide better services. Traditional studies on such human dynamics are largely limited to a simple model, either trivial memory mechanism or simple interactions with only a small set of neighbors (say 2–4). In this paper, we first provide a hybrid and rich model, that is, able to combine the impact of individual memory and interactions among users in a large social network. We try not to simply plug the two parts together, but build a stronger model with a sound mathematical integration of various useful parameters during our modeling and simulation. We designed a hybrid model that can fully integrate both sides. Moreover, when we discuss “interactions”, a set of structural-level network features and node influences upon the social network are deeply considered. The reason is that nodes of social network can have different habit and social influence. We simulated our hybrid model with three well-known networking datasets and evaluated it with large-scale top-one Internet forum in China. We focused on analyzing hotspots (i.e., outbreak topics) in different topic sessions. Based on the comparison between our simulation and empirical studies, we observe similar power-law interhotspot time distribution using different networks. Therefore, our model can offer an understanding of the dynamic mechanism of crowd events in social networks.

In this paper, the node influence is measured by node degree. To further improve our hybrid model, we will apply advanced metrics in quantifying node influence. For example, we will consider link analysis algorithms like PageRank to model node diffusion. In addition, we will model the evolution of social networks and study its effects on hotspots, to better understand human dynamics in an evolving social networking context.

## Acknowledgments

This work is supported by NSFC(60905025, 90924029, 61074128). Joint Construction Science and Technology Research Program of the Chongqing Municipal Education Committee under Grants of KJ110529.

## References

- [1] F. A. Haight, *Handbook of the Poisson distribution*, John Wiley & Sons, New York, 1967.
- [2] A. L. Barabási, “The origin of bursts and heavy tails in human dynamics,” *Nature*, vol. 435, no. 7039, pp. 207–211, 2005.
- [3] Z. Dezso, E. Almaas, A. Lukacs, B. Racz, I. Szakadat, and A.-L. Barabasi, “Dynamics of information access on the web,” *Phys Rev E*, vol. E73, Article ID 066132, 2006.
- [4] W. Hong, X. P. Han, T. Zhou, and B. H. Wang, “Heavy-tailed statistics in short-Message communication,” *Chinese Physics Letters*, vol. 26, no. 2, Article ID 028902, 2009.
- [5] Q. Yan, L. Yi, and L. Wu, “Human dynamic model co-driven by interest and social identity in the microblog community,” *Physica A*, vol. 391, pp. 1540–1545, 2012.

- [6] J. Yu, Y. Hu, M. Yu, and Z. Di, "Analyzing netizens' view and reply behaviors on the forum," *Physica A*, vol. 389, no. 16, pp. 3267–3273, 2010.
- [7] T. Zhou, H. Kiet, B. Kim et al., "Role of activity in human dynamics," *Europhysics Letters*, vol. 82, no. 2, pp. 28002–28006, 2008.
- [8] A. Vázquez, J. G. Oliveira, Z. Dezsö, K. I. Goh, I. Kondor, and A. L. Barabási, "Modeling bursts and heavy tails in human dynamics," *Physical Review E*, vol. 73, no. 3, Article ID 036127, pp. 1–19, 2006.
- [9] A. Vázquez, "Exact results for the Barabási model of human dynamics," *Physical Review Letters*, vol. 95, no. 24, Article ID 248701, pp. 1–4, 2005.
- [10] P. Blanchard and M. O. Hongler, "Modeling human activity in the spirit of Barabási's queueing systems," *Physical Review E*, vol. 75, no. 2, Article ID 026102, 2007.
- [11] L. Dall'Asta, M. Marsili, and P. Pin, "Optimization in task-completion networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 2, Article ID P02003, 2008.
- [12] Z. Deng, N. Zhang, and J. Li, "Inuence of deadline on human dynamic model," in *Dynamic Model of Human Behavior*, J. L. Guo, T. Zhou, N. Zhang, and J. M. Li, Eds., p. 2934, Shanghai System Science Publishing House, Hong Kong, 2008.
- [13] X. P. Han, T. Zhou, and B. H. Wang, "Modeling human dynamics with adaptive interest," *New Journal of Physics*, vol. 10, Article ID 073010, 2008.
- [14] M. Shang, G. Chen, S. Dai et al., "Interest-driven model for human dynamics," *Chinese Physics Letters*, vol. 27, no. 4, Article ID 048701, 2010.
- [15] T. Zhou, Z. Zhao, Z. Yang et al., "Relative clock verifies endogenous bursts of human dynamics," *Europhysics Letters*, vol. 97, p. 18006, 2012.
- [16] N. Johnson, M. Spagat, J. Restrepo et al., "From old wars to new wars and global terrorism," arXiv:physics/0506213, 2005.
- [17] N. Johnson, M. Spagat, J. A. Restrepo et al., "Universal patterns underlying ongoing wars and terrorism," arXiv physics: 0605035v1, 2006.
- [18] N. Johnson, *Complexity in Humuan Conflict*, Springer, New York, NY, USA, 2008.
- [19] S. Galam and S. Moscovici, "Towards a theory of collective phenomena: consensus and attitude changes in groups," *European Journal of Social Psychology*, vol. 21, no. 1, pp. 49–74, 1991.
- [20] S. Galam, "Rational group decision making: a random field Ising model at  $t = 0$ ," *Physica A*, vol. 238, no. 1–4, pp. 66–80, 1997.
- [21] S. Galam, "Sociophysics: a review of galam models," *International Journal of Modern Physics C*, vol. 19, no. 3, pp. 409–440, 2008.
- [22] A. Clauset, L. Heger, and M. Young, "Substitution and competition in the israelpalestine conflict," *Chinese Physics Letters*, vol. 27, p. 068902, 2010.
- [23] A. Vazquez, "Impact of memory on human dynamics," *Physica A*, vol. 373, pp. 747–752, 2007.
- [24] J. F. Zhu, X. P. Han, and B. H. Wang, "Statistical property and model for the inter-event time of terrorism attacks," *Chinese Physics Letters*, vol. 27, no. 6, Article ID 068902, 2010.
- [25] J. G. Oliveira and A. Vazquez, "Impact of interactions on human dynamics," *Physica A*, vol. 388, no. 2-3, pp. 187–192, 2009.
- [26] Y. Wu, C. Zhou, M. Chen, J. Xiao, and J. Kurths, "Human comment dynamics in on-line social systems," *Physica A*, vol. 389, no. 24, pp. 5832–5837, 2010.
- [27] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [28] A. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [29] W. Zachary, "An information ow model for conict and fission in small groups," *Journal of Anthropological Research*, vol. R 33, no. 4, pp. 452–473, 1977.
- [30] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information," *Psychological Review*, vol. 63, p. 8197, 1956.
- [31] A. Baddeley, "Developments in the concept of working memory," *Psychological Bulletin*, vol. 101, p. 353, 1994.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

