Understanding Popularity Dynamics: Decision-Making With Long-Term Incentives

Xuanyu Cao, Yan Chen, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

***Abstract*—With the explosive growth of big data, human's attention has become a scarce resource to be allocated to the vast amount of data. Numerous items such as online memes, videos are generated everyday, some of which go viral, i.e., attract lots of attention, whereas most diminish quickly without any influence. The recorded people's interactions with these items constitute a rich amount of *popularity dynamics*, e.g., hashtags mention count dynamics, which characterize human behaviors quantitatively. It is crucial to understand the underlying mechanisms of popularity dynamics in order to utilize the valuable attention of people efficiently. In this paper, we propose a game-theoretic model to analyze and understand popularity dynamics. The model takes into account both the instantaneous incentives and long-term incentives during people's decision-making process. We theoretically prove that the proposed game possesses a unique symmetric Nash equilibrium (SNE), which can be computed via a backward induction algorithm. We also demonstrate that, at the SNE, the interaction rate first increases and then decreases, which confirms with the observations from real data. Finally, by using simulations as well as experiments based on real-world popularity dynamics data, we validate the effectiveness of the theory. We find that our theory can fit the real data well and also predict the future dynamics.**

***Index Terms*—Game theory, popularity dynamics, symmetric Nash equilibrium.**

I. Introduction

In the big data era nowadays, people not only read lots of data but also create vast amount of data everyday through interactions. For instance, Twitter users may mention or retweet a hashtag; Youtube users can like or dislike a video; researchers may quote keywords in papers. All of these interactions lead to a notion of *popularity dynamics* such as: Twitter hashtags' mention count dynamics and research keywords' quotation dynamics. The popularity dynamics can describe and track people's interactions with different types of items. In general, people can only pay limited attention to a limited number of items. When the number of items are growing drastically, they can only focus on certain items of their great interest. Meanwhile, in the real world, some items go viral, i.e., appealing to lots of interactions and attentions from people, while most items diminish quickly without any impact. To manage and utilize people's valuable interactions and attention better, it is crucial to understand the underlying mechanisms of the popularity dynamics and thus explain the reason why some items are so successful while others aren't.

Manuscript received January 30, 2016; revised July 9, 2016; accepted September 23, 2016. Date of publication October 28, 2016; date of current version February 7, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61672137. The work of Y. Chen was supported by the Thousand Youth Talents Program of China. The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Vincenzo Matta.

X. Cao and K. J. Ray Liu are with the Department of Electrical and Computer Engineering, University of Maryland, College Park MD 20742 USA (e-mail: apogne@umd.edu; kjrliu@umd.edu).

Y. Chen is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China (e-mail: yan@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSIPN.2016.2623090

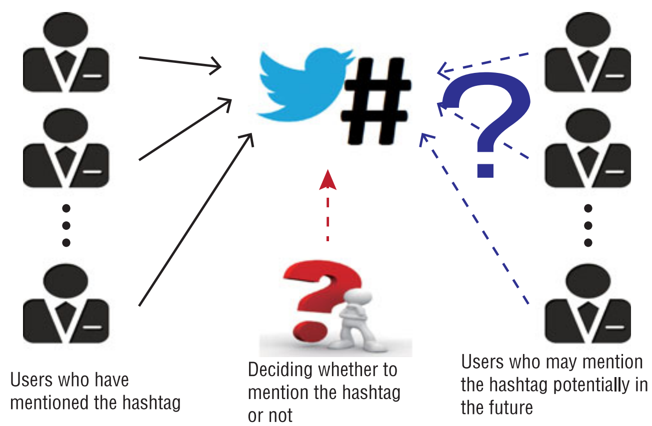


Fig. 1. An illustration of the decision making problem of popularity dynamics. We use the mentioning of a Twitter hashtag as an example here. Consider an arbitrary Twiter user who observes a Twitter hashtag. He needs to decide whether to mention this hashtag or not based on many factors including the intrinsic quality and timeliness of the hashtag, his own interest, current popularity of the hashtag and future actions of other users.

The process of the generation of popularity dynamics is complicated and involves the decision-making of many individuals. Individual's decision is influenced by many factors including the quality and timeliness of the item, the personal preference of the individual and others' decisions. An example of Twitter hashtag is illustrated in Fig. 1. Thus, to model the generation process of the popularity dynamics accurately, we need to take many factors into consideration: the intrinsic attribute of an item, the decaying of the attractiveness of an item as time passes, the heterogeneity of individuals' interests, and the influence of others' decisions, i.e., network externality [1]. Since this involves the interactions between multiple decision makers, game theory [2] can be a very suitable mathematical modeling tool here. By appealing to game theory, we can incorporate all the aforementioned factors into the model of popularity dynamics and the equilibrium of the formulated game can facilitate the understanding or even prediction of users' behaviors in popularity dynamics.

In the literature, game theory has been utilized to model popularity dynamics [3], [4]. However, most of existing game-theoretic models only consider the *instantaneous incentives* of players, i.e., the decision-makers are myopic in the sense that they only decide based on the current state of the system or the current decisions of his neighbors in the network. All the myopic players in the network iteratively update their decisions, leading to a popularity dynamics of an item. On the contrary, in real-world, individuals are usually more farsighted: they may predict the subsequent behaviors of other individuals and then maximize their future benefits based on the predictions. In other words, individuals can have *long-term incentives* which depend on other individuals' future actions. For instance, when a Twitter user is deciding whether to forward a tweet or not, he may take the future influence of the tweet and the future actions of other users into account: will this tweet become popular in the future or will many other users also forward this tweet? This is illustrated in Fig. 1. Different from the previous game-theoretic frameworks for information diffusion dynamics [3], [4], our model incorporates both instantaneous incentives and long-term incentives of individuals. The latter depends on subsequential individuals' actions in the future. Our main contributions in this paper can be epitomized as follows.

1) We propose a game-theoretic framework to model the sequential decision making process of general popularity dynamics. The model incorporates both instantaneous incentives and long-term incentives so that the decision-makers are farsighted enough to take others' future actions into account when making decisions.
2) We theoretically show that the formulated game has a unique symmetric Nash equilibrium (SNE). We observe that the SNE is in pure strategy form and possesses a threshold structure. Furthermore, we design a backward induction algorithm to compute the SNE.
3) From real data, we observe that: (i) most popularity dynamics first increase and then decrease and (ii) for some dynamics, when they are increasing, the increasing speed gradually slows down until they reach the peak and when they are decreasing, the decreasing speed also gradually slows down. We theoretically analyze these properties at the SNE of the proposed game-theoretic model. We find that the equilibrium behavior of the proposed game confirms with real data.
4) The proposed theory is validated by both simulations and experiments based on real data. It is shown that the proposed game-theoretic model can even predict future dynamics of real data.

The roadmap of the rest of the paper is as follows. In Section II, we review the existing literature on popularity dynamics. In Section III, we describe the model in detail and formulate the problem formally. In Section IV, equilibrium analysis is conducted. In Section V, a property of the equilibrium is shown. Simulations and real data experiments are carried out in Section VI. In Section VII, we conclude this work.

II. Related Works

Recently, intensive research efforts have been devoted to network users' behavior dynamics due to its importance [5]. In [6], Ratkiewicz *et al.* studied the popularity dynamics of online webpages and online topics. They proposed a model to combine classic preferential attachment [7] with the random popularity shifts incurred by exogenous factors. Shen *et al.* [8] proposed to use reinforced Poisson processes to model the popularity dynamics and presented a statistical inference approach to predict the future dynamics accordingly.

An important special case of popularity dynamics is the information diffusion dynamics over social networks, which have attracted tremendous research efforts in the recent decade. The abundant literature on information diffusion dynamics can be divided into two categories. In the first category, researchers combine data mining/machine learning techniques with empirical observations from real-world datasets and propose simple models to explain the observed phenomena. Yang and Leskovec [9] studied the temporal shapes of online information dynamics and clustered these temporal dynamics into several patterns, which suggest several types of information dynamics. In [10], the authors empirically studied the temporal dynamics of online memes and discovered interesting phenomena such as an average 2.5 hours lag between the peaks of a phrase in news media and in blogs respectively. The role of social networks, i.e., the influence between network users, in information diffusion is studied in [11] through an experimental approach. In [12]–[14], with machine learning approaches, the underlying implicit diffusion networks are inferred from the observed information cascades to better understand the diffusion processes. Guille and Hacid [15] proposed a predictive model for information diffusion process, which could predict the future information dynamics accurately. In the second category, game-theoretic analyses were conducted to understand the information diffusion processes from the perspective of individual user's decision making. This category has closer relationship with this paper. Jiang *et al.* [3], [4] exploited evolutionary game theory to model and analyze the information diffusion dynamics, where the information diffusion is treated as the consequence of the games played by the network users. In [3], [4], the users were assumed to be myopic and didn't take other individuals' future actions into account when making decisions. Furthermore, in [16], the authors proposed a sequential game model to analyze the voting and answering behaviors in social computing systems such as Stack Overflow, which inspired our model in this work. The differences between the model in [16] and the model here are: (i) we focus on characterizing the temporal dynamics of the interactions while the main goal of [16] was to describe users' behaviors (voting and answering) when facing with certain system states; (ii) we include preferential attachment [7] (a universal phenomenon in network science that items with large popularity are more visible and hence can gain new popularity more easily) into our model while [16] didn't.

There are also many domain specific research literature on popularity dynamics. The citation dynamics were studied in [17] and a universal formula was proposed to characterize the temporal citation dynamics of individual papers. The channel popularity dynamics of Internet Protocol TV were investigated in [18] while the authors in [19] proposed a model predict the dynamics of news. Furthermore, a model for Twitter dynamics was presented in [20] while the dynamics of viral marketing were studied in [21].

TABLE I
GAME-THEORETIC MODEL FOR POPULARITY DYNAMICS

Game-theoretic model	Popularity dynamics concepts
System state at time t	Cumulative interactions x_t
Players	People arriving at the system
Player type	Relevance of the item to the player $\theta \in [0, 1]$
Action set of each player	$\{\texttt{interacting}, \texttt{not interacting}\}$

III. MODEL

In the generation process of popularity dynamics, multiple intelligent decision makers decide whether to interact with an item or not with the goal of maximizing their own utilities. The system has network externality [1], i.e., the utility of an individual is affected by other individuals' actions, as explained in Section I. Game theory is a mathematical tool to study the decision-making of multiple strategic agents where one's utility is influenced by others' actions, and thus very suitable for modeling the popularity dynamics. Additionally, there are various equilibrium concepts in the game theory literature which can serve as predictions of individuals' behaviors in popularity dynamics and thus promote the understanding of the mechanisms of popularity dynamics. In this section, we will introduce a game-theoretic model of the popularity dynamics in detail.

Suppose an item, item $\mathcal{A}$, is generated. The item can be an online meme, an online video or a keyword in scientific research. People decide whether to interact with item $\mathcal{A}$ or not sequentially. For instance, Twitter users decide whether to mention a hashtag or not; Youtube users decide whether to like a video or not; researchers decide whether to quote a keyword in their papers or not. We view the cumulative interaction dynamics x_t, i.e., the total number of interactions up to time t, as a stochastic dynamical system. We assume people, i.e., players of the game, arrive at the system at discrete time instants $t \in \mathbb{N}$ (one player arrives at each time instant) and decide whether to interact with item $\mathcal{A}$ or not. Players are heterogeneous and have different *types*, which indicate the relevances of the item to the different players. For example, for a Twitter hashtag related to football, football fans have higher types than normal users; for a research keyword related to signal processing, researchers specializing in signal processing have higher types than other researchers. We suppose that each player's type θ is a random variable distributed in $[0, 1]$ with probability density function (PDF) $h(\theta)$. The above concepts are summarized in Table I.

To complete the game-theoretic formulation, we need to define the utilities of the players. As stated in Section I, the utility should encompass both the immediate effect and the future effect of the interactions. Furthermore, due to the preferential attachment property of networks [7], items which already get many interactions should be more visible, i.e., easier to be found by players arriving at the system. Combining all these factors, the proposed model can be illustrated in Fig. 2. When a player arrive at the system with state x, the item will be *visible* to him with some probability related to the current state of the system.

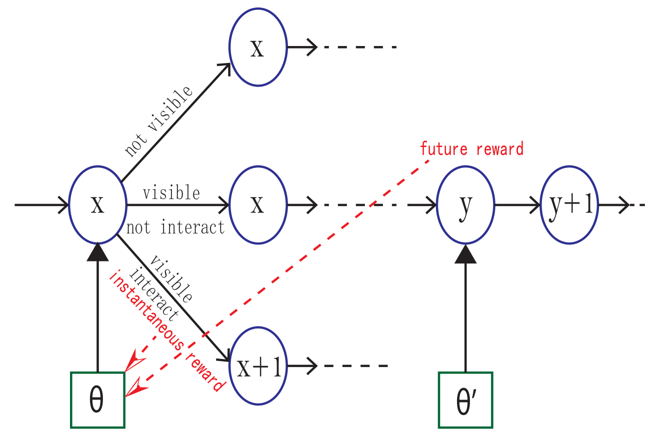


Fig. 2. Illustration of the state transition in the system model. The numbers inside the blue circle are the current states. The numbers inside the green square are the types of the arriving players.

If the item is visible to the player and he chooses to interact with the item, then he will get an *instantaneous reward* which depends on both the type (e.g., hobbies) of the player and the state of the system. Afterwards, whenever there is a new interaction with the item (occurs at say, state y), the aforementioned player will obtain a *future reward* because the current interacting player may pay attention to him. The overall utility of the player will be a discounted sum of the instantaneous reward and all the future rewards. In the following, we specify these components of the model in more detail.

A. Instantaneous Reward

Each player choosing to interact with item $\mathcal{A}$ gets an instantaneous reward $R(x, \theta)$, where x is the state of the system when the interaction occurs and θ is the type of the player. For instance, if a Twitter user is interested in a hashtag, then by mentioning this hashtag, the user will gain some immediate utility. The instantaneous reward depends on the system state since the immediate utilities of an item at different stages (e.g., incipient stage, blooming stage and ending stage) are different. The instantaneous reward also depends on the type of the interacting player because the same item is of different relevances to players of different types: a football fan can get much more utility by mentioning a football related hashtag than a normal Twitter user. Note that $R(x, \theta)$ can also be negative since the interaction implicitly incurs a cost for the player, e.g., by mentioning a hashtag, a Twitter user needs to spend some time and efforts during the manipulation. Now, we impose five assumptions on the function $R(x, \theta)$ as follows.

1) $R(x, \theta)$ decreases with respect to x.
2) $R(x, \theta)$ strictly increases with respect to θ.
3) $R(x, \theta)$ is continuous with respect to θ, for each given x.
4) $R(0, 0) < 0$ and $R(0, 1) > 0$.
5) $\lim_{x\to\infty} R(x, 1) < 0$.

The five assumptions can be justified as follows respectively. (1) Taking timeliness of the item into account, players who interact with the item early (when x is small) should get higher utility than those who interact later (when x is large). For example, a Twitter user mentioning up to date hashtags should gain higher instantaneous reward than a user mentioning outdated hashtags. (2) Those players who find the item more relevant gain higher instantaneous reward by interacting with it. (3) A technical assumption. (4) Initially (i.e., $x = 0$), some players' instantaneous rewards are positive while some are negative. (5) Even for those who find the item very relevant ($\theta = 1$), if the item is very outdated ($x \to \infty$), then it is no longer attractive.

B. Future Reward

For a player $\texttt{B}$ choosing to interact with item $\mathcal{A}$, whenever there is a subsequent interaction with item $\mathcal{A}$, this subsequent interacting player will pay attention to player $\texttt{B}$ with probability $\frac{1}{x}$, where x is the system state when this subsequent interaction occurs. Thus player $\texttt{B}$ will receive an expected reward of $\frac{1}{x}$ due to the possible attention he gets. This reward is called the *future reward* since it is obtained after the interaction occurs. For instance, if a Twitter user $\texttt{B}$ mentions a hashtag $\mathcal{A}$, and later, when hashtag $\mathcal{A}$ has already been mentioned x times, another user $\texttt{C}$ also mentions hashtag $\mathcal{A}$. In such a case, user $\texttt{C}$ may visit those users who have mentioned hashtag $\mathcal{A}$, and with probability $\frac{1}{x}$, user $\texttt{B}$ will be visited by user $\texttt{C}$. We further assume players discount future reward with factor $0 < \lambda < 1$, which is a common assumption in dynamic games and sequential decision making. The instantaneous reward and the future reward together constitute the utility of an interacting player.

C. Visibility Probability

We assume one player arrives at the system at each time instant. Item $\mathcal{A}$ is *visible* to a player with probability $f(x) \in [0, 1]$, where x is the system state when the player arrives. In other words, after a player arrives, he/she will notice item $\mathcal{A}$ with probability $f(x)$. We also impose several assumptions on the visibility probability function $f(x)$ as follows.

1) $f(x)$ increases with x. Justification: Popular items are more visible. This is also refereed as the 'rich gets richer' or preferential attachment phenomenon in network science [7].
2) $f(0) > 0$. Justification: Even the most unpopular item is visible with positive probability.

D. Action Rule and Utility Function

When a player arrives at the system and sees item $\mathcal{A}$, he/she needs to decide whether to interact with the item or not based on the current system state x and his/her type θ. For sake of generality, we allow the players to use mixed action rule $\pi : \mathbb{N} \times [0, 1] :\to [0, 1]$, where $\pi(x, \theta)$ is the probability of choosing the action `interacting` when the state is x and the type of the player is θ. Denote the set of all possible mixed action rules as Π. We denote $g_\pi(x)$ the long-term utility of an interacting player starting from state x while the subsequent players use action rule π.

Denote $p^\pi(x) = \mathbb{E}_\theta[\pi(x, \theta)]$, i.e., the expected probability of a new interaction when the system state is x and users adopt action rule π. Thus, the long term utility $g_\pi(x)$ can be computed recursively as follows. $\forall x \geq 1$:

$$g_\pi(x) = \overbrace{\frac{f(x)p^\pi(x)}{x}}^{\text{instantaneous reward at the current time slot}} + \underbrace{\lambda \left\{ f(x)p^\pi(x)g_\pi(x+1) + [1 - f(x)p^\pi(x)]\, g_\pi(x) \right\}}_{\text{reward in future time slots}}. \quad (1)$$

Denote $u(x, \theta, a, \pi)$ the utility of a type-θ player who enters the system in state x and takes action a while other players adopt action rule π. Thus,

$$u(x, \theta, a, \pi) = \begin{cases} R(x, \theta) + \lambda g_\pi(x+1), & \\ & \text{if } a = \texttt{interacting}, \\ 0 & \text{if } a = \texttt{not interacting}. \end{cases} \quad (2)$$

If a player chooses mixed action, i.e., `interacting` with probability q, then his/her utility is:

$$\mathcal{U}(x, \theta, q, \pi) = q[R(x, \theta) + \lambda g_\pi(x+1)]. \quad (3)$$

E. Solution Concept

In this work, the solution concept is chosen to be the *symmetric Nash equilibrium (SNE)*, which is defined in the following.

Definition 1: An action rule π^* is said to be a symmetric Nash equilibrium (SNE) if:

$$\pi^*(x, \theta) \in \arg \max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*), \forall x \in \mathbb{N}, \theta \in [0, 1]. \quad (4)$$

In an SNE, no player wants to deviate unilaterally, hence the action rule is self enforcing.

IV. Equilibrium Analysis

In this section, we show that there is a unique SNE of the formulated game. A backward induction algorithm for computing this unique SNE is also presented.

The infinite-horizon sequential game is effectively of finite length, given the following lemma, which says that no one will interact with item $\mathcal{A}$ after a certain number of interactions is reached.

Lemma 1: There exists an $\tilde{x} \in \mathbb{N}$ such that $\forall x \geq \tilde{x}, \theta \in [0, 1], \pi \in \Pi$:

$$u(x, \theta, \texttt{interacting}, \pi) < u(x, \theta, \texttt{not interacting}, \pi), \quad (5)$$

i.e., the action `interacting` is strictly dominated by the action `not interacting` regardless of the player type and other players' action rule.

Proof: The long-term utility of an interacting player starting from state x while subsequent players use action rule π can be upper bounded as follows:

$$\begin{aligned} g_\pi(x) =&\ \mathbb{E}_{\{x_t\}_{t=2}^\infty} \left[\sum_{t=1}^\infty \lambda^{t-1} \frac{f(x_t)p^\pi(x_t)}{x_t} \middle| \pi, x_1 = x \right] \\ \leq&\ \sum_{t=1}^\infty \lambda^{t-1} \frac{1}{x} = \frac{1}{(1-\lambda)x} \to 0, \text{as } x \to \infty. \end{aligned} \quad (6)$$

Algorithm 1: Computation of the unique equilibrium.

Inputs:

Instantaneous reward function $R(x, \theta)$.
Player type PDF $h(\theta)$.
Visibility probability function $f(x)$.
Discount factor λ.

Outputs:

Unique equilibrium action rule $\pi^*(x, \theta)$.

When $x \geq \hat{x} + 1$: $\pi^*(x, \theta) = 0, \theta \in [0, 1]$; $g_{\pi^*}(x) = 0$.
Let $x = \hat{x}$.
while $x \geq 0$ **do**
  if $R(x, 0) + \lambda g_{\pi^*}(x + 1) > 0$ **then**
    $\theta_x^* = 0$,
  else
    θ_x^* is the unique solution of $R(x, \theta) + \lambda g_{\pi^*}(x + 1) = 0$.
  end if
  Compute:

$$\pi^*(x, \theta) = \begin{cases} 1, & \text{if } \theta \geq \theta_x^*, \\ 0, & \text{if } \theta < \theta_x^*, \end{cases} \quad (7)$$

$$p^{\pi^*}(x) = \int_{\theta_x^*}^1 h(\theta) d\theta, \quad (8)$$

$$g_{\pi^*}(x) = \frac{1}{1 - \lambda[1 - f(x)p^{\pi^*}(x)]} \quad (9)$$

$$\left[\frac{f(x)p^{\pi^*}(x)}{x} + \lambda f(x) p^{\pi^*}(x) g_{\pi^*}(x + 1) \right]. \quad (10)$$

  $x \leftarrow x - 1$.
end while

Note that $\lim_{x \to \infty} R(x, 1) < 0$. So, there exists an $\tilde{x} \in \mathbb{N}$, such that $\forall x \geq \tilde{x}$:

$$R(x, 1) + \frac{\lambda}{(1 - \lambda)x} < 0. \quad (11)$$

Hence, $\forall x \geq \tilde{x}, \theta \in [0, 1], \pi \in \Pi$:

$$R(x, \theta) + \lambda g_\pi(x + 1) \leq R(x, 1) + \frac{\lambda}{(1 - \lambda)x} < 0, \quad (12)$$

i.e., $u(x, \theta, \texttt{interacting}, \pi) < u(x, \theta, \texttt{not interacting}, \pi)$ due to the utility expression in (2). ■

Denote $\hat{x} = \max\{x \in \mathbb{N} | R(x, 1) > 0\}$. We design a backward induction algorithm, Algorithm 1, to compute the SNE. We first show that the action rule obtained from Algorithm 1 is indeed an SNE.

Theorem 1: (Existence of the SNE) The action rule π^* computed by Algorithm 1 is an SNE.

Proof: According to Lemma 1, $\forall x \geq \tilde{x}, \theta \in [0, 1]$: $\arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*) = \{0\}$. Thus, $\pi^*(x, \theta) \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*), \forall x \geq \tilde{x}, \theta \in [0, 1]$.

When $\hat{x} + 1 \leq x \leq \tilde{x} - 1$, we have the following. $\forall \theta \in [0, 1]$:

$$\begin{aligned} u(x, \theta, \texttt{interacting}, \pi^*) &= R(x, \theta) + \lambda g_{\pi^*}(x + 1) \\ &= R(x, \theta) \leq R(x, 1) \leq 0. \quad (13) \end{aligned}$$

So, $\pi^*(x, \theta) = 0 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*)$.

When $x = \hat{x}$, we still have $u(\hat{x}, \theta, \texttt{interacting}, \pi^*) = R(x, \theta)$. If $\theta \geq \theta_x^*$, we have $u(\hat{x}, \theta, \texttt{interacting}, \pi^*) \geq 0$ and thus $\pi^*(x, \theta) = 1 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*)$. If $\theta < \theta_x^*$, we have $u(\hat{x}, \theta, \texttt{interacting}, \pi^*) < 0$ and hence $\pi^*(x, \theta) = 0 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*)$.

When $x \leq \hat{x} - 1$, we discuss two cases:

1) *Case 1:* $R(x, 0) + \lambda g_{\pi^*}(x + 1) > 0$. In this case, $\theta_x^* = 0$ and $\pi^*(x, \theta) = 1, \forall \theta \in [0, 1]$. Thus, $u(x, \theta, \texttt{interacting}, \pi^*) \geq R(x, 0) + \lambda g_{\pi^*}(x + 1) > 0, \forall \theta$. Hence, $\pi^*(x, \theta) = 1 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*), \forall \theta$.
2) *Case 2:* $R(x, 0) + \lambda g_{\pi^*}(x + 1) \leq 0$. In such a case, if $\theta \geq \theta_x^*$, then $u(x, \theta, \texttt{interacting}, \pi^*) \geq 0$ and thus $\pi^*(x, \theta) = 1 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*)$. Otherwise, if $\theta < \theta_x^*$, then $u(x, \theta, \texttt{interacting}, \pi^*) < 0$ and $\pi^*(x, \theta) = 0 \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*)$.

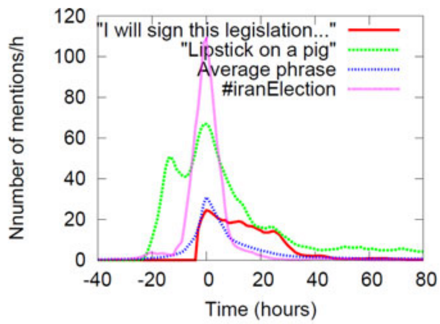
Overall, we always have $\pi^*(x, \theta) \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \pi^*), \forall x, \theta \in [0, 1]$. ■

We further prove that the π^* computed in Algorithm 1 is indeed the unique SNE.

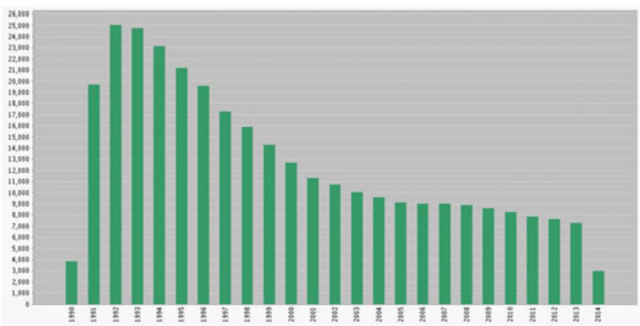
Theorem 2: (Uniqueness of the SNE) Suppose the distribution of player type θ is atomless, i.e., $h(\theta)$ is finite for every $\theta \in [0, 1]$. Then, any SNE $\tilde{\pi}$ differs with π^* for zero mass players, i.e., $\mathbb{P}\{\tilde{\pi}(x, \theta) \neq \pi^*(x, \theta)\} = 0$ for every $x \in \mathbb{N}$.

Proof: Suppose $\tilde{\pi}$ is an SNE, i.e., $\tilde{\pi}(x, \theta) \in \arg\max_{q \in [0,1]} \mathcal{U}(x, \theta, q, \tilde{\pi}), \forall x, \theta \in [0, 1]$. In the following, we show that $\tilde{\pi}$ differs from π^* for zero-mass players. We discuss for different values of x as follows.

1) *Case 1:* $x \geq \tilde{x}$. From Lemma 1, we know that $\forall \theta \in [0, 1]$, $u(x, \theta, \texttt{interacting}, \tilde{\pi}) < 0$. So, $\tilde{\pi}(x, \theta) = 0 = \pi^*(x, \theta)$.
2) *Case 2:* $\hat{x} + 1 \leq x \leq \tilde{x} - 1$. First, we note that $u(\tilde{x} - 1, \theta, \texttt{interacting}, \tilde{\pi}) = R(\tilde{x} - 1, \theta)$. When $\theta < 1$, $u(\tilde{x} - 1, \theta, \texttt{interacting}, \tilde{\pi}) < 0$, and thus $\tilde{\pi}(\tilde{x} - 1, \theta) = 0$. Hence, $p^{\tilde{\pi}}(\tilde{x} - 1) = 0$ and $g_{\tilde{\pi}}(\tilde{x} - 1) = 0$. Suppose we have $\tilde{\pi}(x + 1, \theta) = g_{\tilde{\pi}}(x + 1) = 0$, where $\hat{x} + 1 \leq x \leq \tilde{x} - 2$. Then, for $\theta < 1$, we have $u(x, \theta, \texttt{interacting}, \tilde{\pi}) = R(x, \theta) < 0$. Hence, $\tilde{\pi}(x, \theta) = 0, \forall \theta$. Thus, $p^{\tilde{\pi}}(x) = g_{\tilde{\pi}}(x) = 0$. By induction, we know that $\tilde{\pi}(x, \theta) = g_{\tilde{\pi}}(x) = 0, \forall \hat{x} + 1 \leq x \leq \tilde{x} - 1$ and $\theta < 1$. In particular, we still have $\tilde{\pi}(x, \theta) = \pi^*(x, \theta), \forall \hat{x} + 1 \leq x \leq \tilde{x} - 1$ and $\theta < 1$.
3) *Case 3:* $x = \hat{x}$. $u(\hat{x}, \theta, \texttt{interacting}, \tilde{\pi}) = R(\hat{x}, \theta)$. If $\theta > \theta_{\hat{x}}^*$, we have $u(\hat{x}, \theta, \texttt{interacting}, \tilde{\pi}) > 0$ and thus $\tilde{\pi}(\hat{x}, \theta) = 1$. Similarly, if $\theta < \theta_{\hat{x}}^*$, we have $\tilde{\pi}(\hat{x}, \theta) = 0$. So, $\tilde{\pi}(\hat{x}, \theta) = \pi^*(\hat{x}, \theta), \forall \theta \neq \theta_{\hat{x}}^*$. Hence, $p^{\tilde{\pi}}(\hat{x}) = p^{\pi^*}(\hat{x}), g_{\tilde{\pi}}(\hat{x}) = g_{\pi^*}(\hat{x}), \tilde{\pi}(\hat{x}, \theta) = \pi^*(\hat{x}, \theta), \forall \theta \neq \theta_{\hat{x}}^*$.
4) *Case 4:* $x \leq \hat{x} - 1$. Suppose for some $1 \leq x \leq \hat{x} - 1$, we have $g_{\tilde{\pi}}(x + 1) = g_{\pi^*}(x + 1)$ and $\tilde{\pi}(x + 1, \theta) = \pi^*(x + 1, \theta), \forall \theta \neq \theta_{x+1}^*$. We note that these already hold for $x = \hat{x} - 1$ according to Case 3. Thus, we have $u(x, \theta, \texttt{interacting}, \tilde{\pi}) = R(x, \theta) + \lambda g_{\tilde{\pi}}(x + 1) = R(x, \theta) + \lambda g_{\pi^*}(x + 1)$. If $\theta > \theta_x^*$, we have $u(x, \theta, \texttt{interacting},$



(a)



(b)

Fig. 3. Real-world popularity dynamics [9], [22].

$\tilde{\pi}) > 0$ and thus $\tilde{\pi}(x, \theta) = 1$. Otherwise, if $\theta < \theta_x^*$, we have $u(x, \theta, \texttt{interacting}, \tilde{\pi}) < 0$ and $\tilde{\pi}(x, \theta) = 0$. Consequently, we have $\tilde{\pi}(x, \theta) = \pi^*(x, \theta), \forall \theta \neq \theta_x^*$ and hence $g_{\tilde{\pi}}(x) = g_{\pi^*}(x)$. So, by induction, we have $\tilde{\pi}(x, \theta) = \pi^*(x, \theta), \forall \theta \neq \theta_x^*, \forall x \leq \hat{x} - 1, \theta \neq \theta_x^*$.

In all, we have $\tilde{\pi} = \pi^*$ except for a zero mass amount of players. ■

Remark 1: The unique SNE of the game is in pure strategy form and possesses a threshold structure. For every state x, there exists a threshold θ_x^* such that a player of type θ will interact with item $\mathcal{A}$ if and only if $\theta \geq \theta_x^*$.

V. Popularity Dynamics at the Equilibrium

In this section, we first observe some properties of popularity dynamics from real data. Then, we analyze the corresponding properties at the equilibrium of the proposed game. We find that the equilibrium behavior of the proposed game-theoretic model confirms with the real data.

A. Observations from Real Data

In Fig. 3, we plot mention dynamics of popular memes and sum citation dynamics of all the papers published in *Nature* in 1990. Here, we use the dynamics of memes and the citation dynamics of papers as examples of popularity dynamics.

We observe that, typically, the popularity dynamics of an item will first increase and then decrease, leading to a peak in the dynamics. This is a general *first order* property of popularity dynamics. Thus, a natural question is: does the equilibrium behavior of the proposed game-theoretic model possess this property? Intuitively, it should. The reason is as follows. At

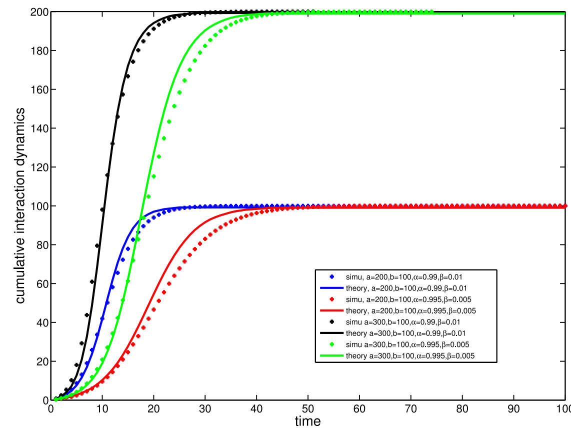


Fig. 4. Simulation results under different parameter setups.

first, according to our model, the visibility probability is low since the item has few interactions. As time goes, the item accumulates more interactions so that the visibility probability increases and the interaction rate, i.e., the dynamics observed in Fig. 3, also increases. When time is sufficiently large, the visibility probability basically saturates. With the augment of the cumulative interactions, the instantaneous reward and long-term reward decreases so that few players will further interact with the item, leading to a decrease in interaction rate. In next subsection, we will formally state and prove this first order property.

Furthermore, we observe that some popularity dynamics, especially the citation dynamics of papers as in Fig. 3-(b), Fig. 6 and Fig. 7-(c)(d), have the following *second order* property: when it is increasing, its increasing speed gradually slows down and when it is decreasing, its decreasing speed also gradually slows down. This behavior is reasonable. Many items' interaction rates increase drastically when they first come out and keep increasing (but at a lower speed) until they reach the peak. Later, after the items are no longer that popular, their interaction rates decrease quickly and will keep decreasing for some time (but at a lower speed). In next subsection, we will formally state and prove this second order property under certain assumptions.

B. Properties at the Equilibrium

Generally, the unique SNE should be computed using the backward induction as specified in Algorithm 1, which is hard to analyze. To facilitate analysis, we further restrict attention to models satisfying the following three assumptions:

1) *(Linear reward)* $R(x, \theta) = -x + a\theta - b$, where $a > b > 0$.
2) *(Uniform player type distribution)* $h(\theta) = 1, \forall \theta \in [0, 1]$.
3) *(Saturated visibility probability)* There exists a $\check{x} \in \mathbb{N}$ less than $\hat{x} = \lfloor a - b \rfloor$ such that: $\forall x \geq \check{x} : f(x) = 1$, and $\forall 1 \leq x \leq \check{x}$:

$$\frac{f(x)}{f(x-1)} \geq 1 + \frac{1 + \frac{\lambda}{x}}{a - b - \check{x}}. \tag{14}$$

These three assumptions can be justified as follows. (1) Linear reward is used to simplify the analysis and it indeed increases

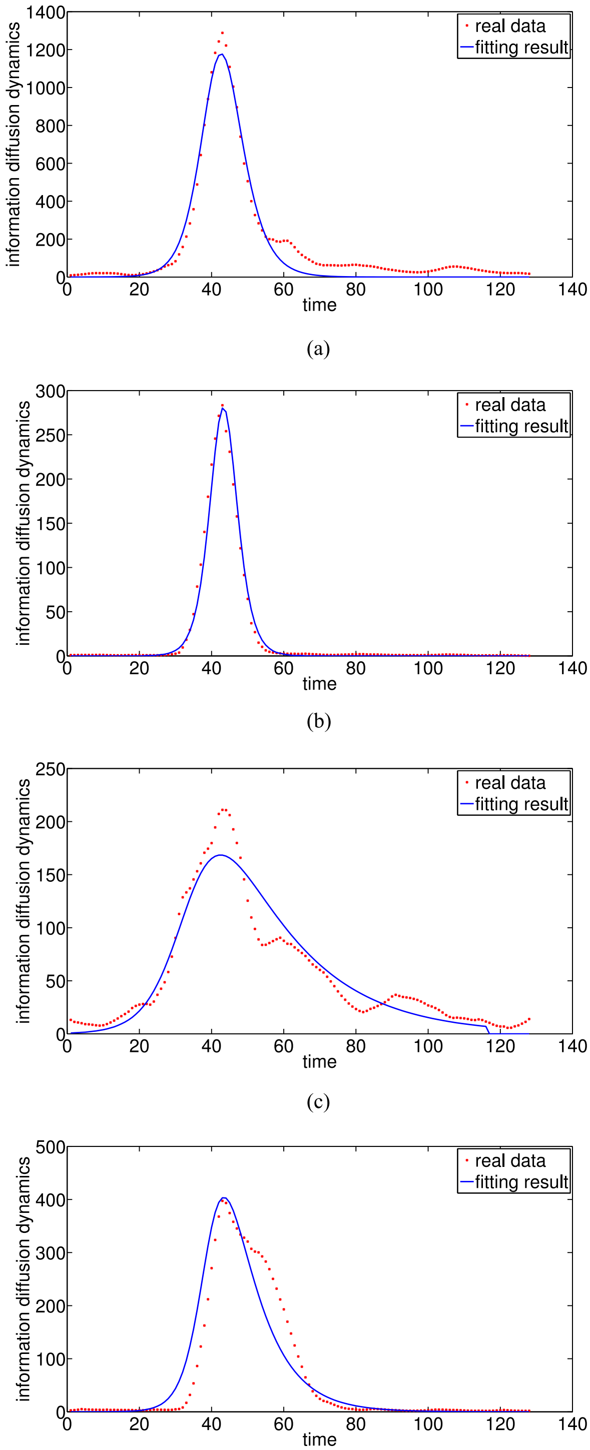


Fig. 5. Fitting Twitter hashtag dynamics.

with θ and decreases with x, which coincides with the assumptions in Section III. (2) Uniform player type distribution is also to simplify calculations though our analysis is applicable to more complicated distributions in principle. (3) When the number of interactions is large enough, the item becomes ‘famous’ enough so that it is visible to everyone arriving at the system. Before this saturation occurs, however, it increases at a speed not too

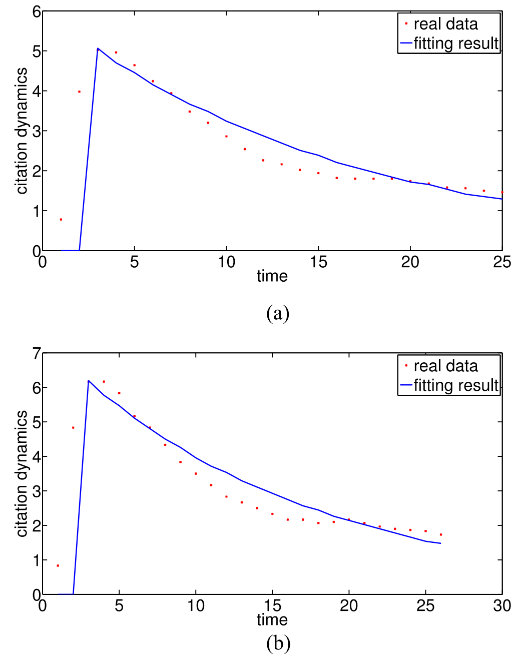


Fig. 6. Fitting paper citation dynamics.

slow. Note that the R.H.S. of (14) is very close to 1 since the numerator of the ratio is close to 1 while the denominator is some integer much larger than 1 generally. So, the assumption is indeed very weak.

Denote $r(x) = f(x)p^{\pi^*}(x)$, i.e., the probability that there is a new interaction at state x in the SNE. We first show the first order property of the SNE.

Theorem 3: (First order characterization of the SNE) Suppose the assumptions (1)(2)(3) hold, the SNE π^* satisfies the following:

1) For $1 \le x \le \check{x}$: $r(x) \ge r(x-1)$;
2) For $x \ge \check{x}$: $r(x) \ge r(x+1)$.

In other words, the interaction rate $r(x)$ first increases and then decreases.

Proof: According to the assumptions of linear reward and uniform player type distribution, we can obtain closed form expression of the iterative update of θ_x^* and $p^{\pi^*}(x)$ in Algorithm 1 as follows: $\forall x \le \hat{x} = \lfloor a - b \rfloor$:

$$p^{\pi^*}(x) = 1 - \frac{1}{a}(x + b - \lambda g_{\pi^*}(x+1))^+, \tag{15}$$

$$g_{\pi^*}(x) = \frac{1}{1 - \lambda(1 - f(x)p^{\pi^*}(x))} \left[\frac{f(x)p^{\pi^*}(x)}{x} + \lambda f(x)p^{\pi^*}(x)g_{\pi^*}(x+1)\right], \tag{16}$$

where $x^+ \triangleq \max\{x, 0\}$.

We first consider the case $x \ge \check{x}$. In the following, we show that g_{π^*} is decreasing for $\check{x} \le x \le \hat{x}+1$. When $\check{x} \le x \le \hat{x}$,

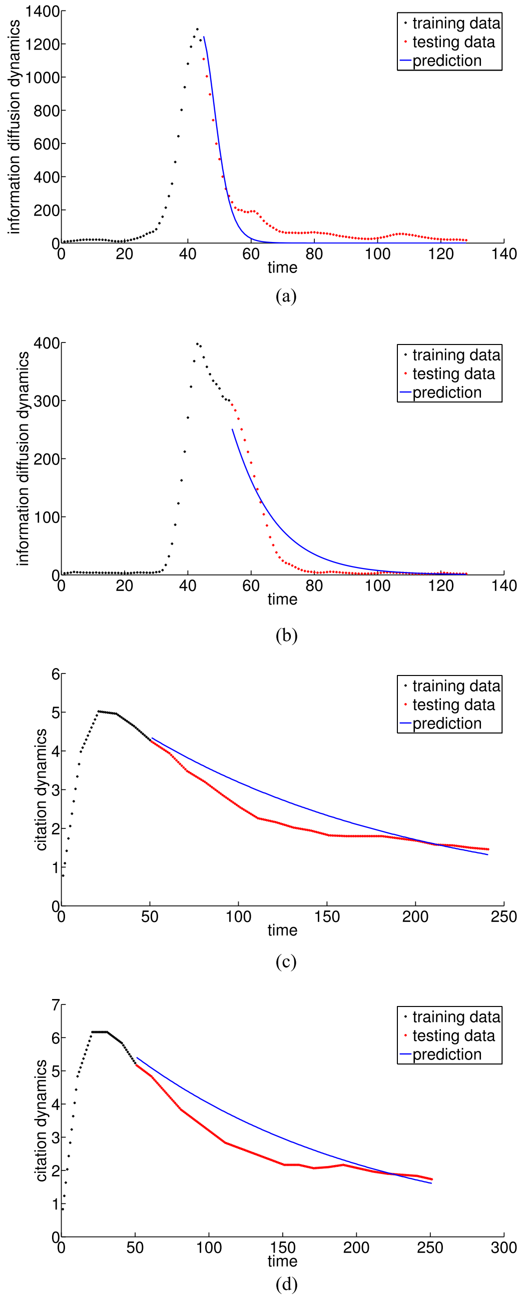


Fig. 7. Predicting future dynamics.

noticing that $f(x) = 1$, we rewrite (16) as:

$$\begin{aligned} &g_{\pi^*}(x) - g_{\pi^*}(x+1) \\ &= \frac{p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))}\left[\frac{1}{x} - \frac{1-\lambda}{p^{\pi^*}(x)} g_{\pi^*}(x+1)\right]. \quad (17) \end{aligned}$$

Since $g_{\pi^*}(\hat{x}+1) = 0$, we have:

$$\begin{aligned} 0 \;&=\; g_{\pi^*}(\hat{x}+1) \le g_{\pi^*}(\hat{x}) \\ &=\; \frac{p^{\pi^*}(\hat{x})}{[1-\lambda(1-p^{\pi^*}(\hat{x}))]\hat{x}} \le \frac{p^{\pi^*}(x)}{(1-\lambda)\hat{x}}. \quad (18) \end{aligned}$$

Suppose $g_{\pi^*}(x) \ge g_{\pi^*}(x+1)$ and $g_{\pi^*}(x) \le \frac{p^{\pi^*}(x)}{(1-\lambda)x}, \forall m \le x \le \hat{x}$ for some $\check{x}+1 \le m \le \hat{x}$ (note that we already show that these hold for $m = \hat{x}$). We next show $g_{\pi^*}(m-1) \ge g_{\pi^*}(m)$ and $g_{\pi^*}(m-1) \le \frac{p^{\pi^*}(m-1)}{(1-\lambda)(m-1)}$.

According to (15), for $m \le x \le \hat{x}$:

$$p^{\pi^*}(x-1) \;=\; 1 - \frac{1}{a}(x - 1 + b - \lambda g_{\pi^*}(x))^+, \quad (19)$$

$$p^{\pi^*}(x) \;=\; 1 - \frac{1}{a}(x + b - \lambda g_{\pi^*}(x+1))^+. \quad (20)$$

Since $g_{\pi^*}(x) \ge g_{\pi^*}(x+1), \forall m \le x \le \hat{x}$, comparing the above two expressions, we have $p^{\pi^*}(x-1) \ge p^{\pi^*}(x), \forall m \le x \le \hat{x}$. In particular, $p^{\pi^*}(m-1) \ge p^{\pi^*}(m)$, thus,

$$g_{\pi^*}(m) \le \frac{p^{\pi^*}(m)}{(1-\lambda)m} \le \frac{p^{\pi^*}(m-1)}{(1-\lambda)(m-1)}. \quad (21)$$

Hence, by (17) and (21), we obtain:

$$\begin{aligned} &g_{\pi^*}(m-1) - g_{\pi^*}(m) \\ &= \frac{p^{\pi^*}(m-1)}{1-\lambda(1-p^{\pi^*}(m-1))}\left[\frac{1}{m-1} - \frac{1-\lambda}{p^{\pi^*}(m-1)} g_{\pi^*}(m)\right] \\ &\ge 0. \quad (22) \end{aligned}$$

So, $g_{\pi^*}(m-1) \ge g_{\pi^*}(m)$. In addition:

$$\begin{aligned} g_{\pi^*}(m-1) \;&=\; \mathbb{E}\left[\sum_{t=1}^{\infty} \lambda^{t-1} \frac{p^{\pi^*}(x_t)}{x_t} \middle| \pi^*, m-1\right] \\ &\le \sum_{t=1}^{\infty} \lambda^{t-1} \frac{p^{\pi^*}(m-1)}{m-1} = \frac{p^{\pi^*}(m-1)}{(1-\lambda)(m-1)}. \quad (23) \end{aligned}$$

Hence, by induction, we have $g_{\pi^*}(x) \ge g_{\pi^*}(x+1)$ and $g_{\pi^*}(x) \le \frac{p^{\pi^*}(x)}{(1-\lambda)x}, \forall \check{x} \le x \le \hat{x}$. Thus, by (15), we have: $p^{\pi^*}(x-1) \ge p^{\pi^*}(x), \forall \check{x} \le x \le \hat{x}$. Note that $p^{\pi^*}(x) = 0, \forall x \ge \hat{x}+1$ since $\pi^*(x,\theta) = 0, \forall x \ge \hat{x}+1, \theta \in [0,1]$. Thus, for $x \ge \check{x}$: $p^{\pi^*}(x) \ge p^{\pi^*}(x+1)$. So, for $x \ge \check{x}$: $r(x) \ge r(x+1)$.

Next, we consider the case $x \le \check{x} (\le \hat{x})$. In such a case, we rewrite the update equations (15) and (16) in terms of g_{π^*} and r as follows:

$$\begin{aligned} g_{\pi^*}(x) \;&=\; \frac{r(x)}{1-\lambda(1-r(x))}\left[\frac{1}{x} + \lambda g_{\pi^*}(x+1)\right], \\ &\qquad \forall 1 \le x \le \check{x}, \quad (24) \\ r(x) \;&=\; \left[1 - \frac{1}{a}(x + b - \lambda g_{\pi^*}(x+1))^+\right] f(x), \\ &\qquad \forall 0 \le x \le \check{x}. \quad (25) \end{aligned}$$

Rewriting (24) yields: $\forall 1 \leq x \leq \check{x}$:

$$\begin{aligned} & g_{\pi^*}(x) - g_{\pi^*}(x+1) \\ & = \frac{r(x)}{1-\lambda(1-r(x))}\left[\frac{1}{x} - \frac{1-\lambda}{r(x)} g_{\pi^*}(x+1)\right] \\ & \leq \frac{r(x)}{1-\lambda(1-r(x))}\frac{1}{x} \\ & \leq \frac{1}{x}. \end{aligned} \tag{26}$$

For $1 \leq x \leq \check{x}$:

$$\frac{r(x)}{r(x-1)} = \frac{1-\frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+}{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+}\frac{f(x)}{f(x-1)}. \tag{27}$$

For $1 \leq x \leq \check{x}$, from (26) we have:

$$\begin{aligned} & (x+b-\lambda g_{\pi^*}(x+1)) - (x-1+b-\lambda g_{\pi^*}(x)) \\ & = 1 + \lambda[g_{\pi^*}(x) - g_{\pi^*}(x+1)] \leq 1 + \frac{\lambda}{x}. \end{aligned} \tag{28}$$

So,

$$(x+b-\lambda g_{\pi^*}(x+1))^+ - (x-1+b-\lambda g_{\pi^*}(x))^+ \leq 1 + \frac{\lambda}{x}. \tag{29}$$

Hence,

$$\begin{aligned} & \left[1 - \frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+\right] \\ & - \left[1 - \frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+\right] \\ & \leq \frac{1}{a}\left(1 + \frac{\lambda}{x}\right). \end{aligned} \tag{30}$$

We further know that:

$$\begin{aligned} 1 - \frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+ & \geq 1 - \frac{1}{a}(x+b) \\ & \geq 1 - \frac{1}{a}(\check{x}+b). \end{aligned} \tag{31}$$

Thus,

$$\frac{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+}{1-\frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+} - 1 \tag{32}$$

$$\leq \frac{\frac{1}{a}\left(1+\frac{\lambda}{x}\right)}{1-(x+b-\lambda g_{\pi^*}(x+1))^+} \tag{33}$$

$$\leq \frac{\frac{1}{a}\left(1+\frac{\lambda}{x}\right)}{1-\frac{1}{a}(\check{x}+b)} \tag{34}$$

$$= \frac{1+\frac{\lambda}{x}}{a-b-\check{x}}. \tag{35}$$

Note that for $1 \leq x \leq \check{x}$:

$$1 + \frac{1+\frac{\lambda}{x}}{a-b-\check{x}} \leq \frac{f(x)}{f(x-1)}. \tag{36}$$

Combining (35) and (36) yields:

$$\frac{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+}{1-\frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+} \leq \frac{f(x)}{f(x-1)}, \tag{37}$$

which, according to (27), is equivalent to:

$$r(x) \geq r(x-1), \tag{38}$$

where $1 \leq x \leq \check{x}$. ■

Next we turn to the second order property of the SNE.

Theorem 4: (Second order characterization of the SNE) Suppose that the assumptions (1)(2)(3) hold. Further assume that (i) $\lambda \leq \frac{1}{a-b}$ and (ii) $\forall 2 \leq x \leq \check{x}$: $f(x) + \left(1 + \frac{1+\frac{\lambda}{x}}{a-b-\check{x}}\right) f(x-2) \leq 2f(x-1)$. Then the SNE π^* satisfies the following:

1) For $2 \leq x \leq \check{x}$: $0 \leq r(x) - r(x-1) \leq r(x-1) - r(x-2)$;
2) For $x \geq \check{x}+2$: $0 \leq r(x-1) - r(x) \leq r(x-2) - r(x-1)$.

In other words, we have: (a) when $r(x)$ is increasing, its increasing speed gradually slows down; (b) when $r(x)$ is decreasing, its decreasing speed also gradually slows down.

Proof: We first consider the case $\check{x}+2 \leq x \leq \hat{x}-1$. From $\lambda \leq \frac{1}{a-b}$, we get:

$$\lambda\left(a-b-\check{x}+\frac{\lambda}{(1-\lambda)(\check{x}+1)}\right) \leq 1. \tag{39}$$

Since $p^{\pi^*}(x)$ is decreasing for $x \geq \check{x}$, we have:

$$\begin{aligned} p^{\pi^*}(x) \leq\ & p^{\pi^*}(\check{x}) \\ =\ & 1 - \frac{1}{a}(\check{x}+b-\lambda g_{\pi^*}(\check{x}+1)) \\ \leq\ & 1 - \frac{1}{a}(\check{x}+b) + \frac{\lambda}{a(1-\lambda)(\check{x}+1)} \\ \leq\ & \frac{1}{\lambda a}, \end{aligned} \tag{40}$$

where in the second last inequality and last inequality we make use of (6) and (39) respectively. Furthermore,

$$\frac{p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))} = \frac{1}{\lambda + \frac{1-\lambda}{p^{\pi^*}(x)}} \leq 1. \tag{41}$$

So, together with (40), we have:

$$\frac{\lambda p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))} \leq \frac{1}{ap^{\pi^*}(x)}. \tag{42}$$

For any $\check{x} \leq x \leq \hat{x}$:

$$\begin{aligned} & x+b-\lambda g_{\pi^*}(x+1) \\ & \geq \check{x}+b-\lambda g_{\pi^*}(\check{x}+1) \\ & \geq \check{x}+b-\frac{\lambda}{(1-\lambda)(\check{x}+1)} \\ & \geq 0. \end{aligned} \tag{43}$$

So, from (15), for any $\check{x}+2 \leq x \leq \hat{x}-1$, we obtain:

$$p^{\pi^*}(x-1) - p^{\pi^*}(x) = \frac{1}{a} + \frac{\lambda}{a} g_{\pi^*}(x) - \frac{\lambda}{a} g_{\pi^*}(x+1) \geq \frac{1}{a}, \tag{44}$$

where the last inequality is due to the monotonicity of $g_{\pi^*}(x)$ for $x \geq \check{x}$. Combining (42) and (44) yields $\forall \check{x}+2 \leq x \leq \hat{x}-1$:

$$\frac{\lambda p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))} \leq \frac{p^{\pi^*}(x-1)-p^{\pi^*}(x)}{p^{\pi^*}(x)}. \tag{45}$$

From (16), noticing that $f(x)$=1, we have $\forall \check{x}+2 \leq x \leq \hat{x}-1$:

$$g_{\pi^*}(x+1) \geq \frac{1}{1-\lambda(1-p^{\pi^*}(x))}\frac{p^{\pi^*}(x)}{x}, \tag{46}$$

and thus

$$\begin{aligned}
&\frac{g_{\pi^*}(x)}{g_{\pi^*}(x+1)} \\
&\leq \frac{p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*(x)})}\left\{\frac{1}{x}\frac{x[1-\lambda(1-p^{\pi^*}(x))]}{p^{\pi^*}(x)}+\lambda\right\} \\
&= 1+\frac{\lambda p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))} \\
&\leq \frac{p^{\pi^*}(x-1)}{p^{\pi^*}(x)},
\end{aligned} \tag{47}$$

where the last inequality is due to (45). Hence, $\forall \check{x}+2 \leq x \leq \hat{x}-1$:

$$\frac{g_{\pi^*}(x+1)}{p^{\pi^*}(x)} \geq \frac{g_{\pi^*}(x)}{p^{\pi^*}(x-1)}. \tag{48}$$

From (16), we have:

$$\begin{aligned}
&g_{\pi^*}(x)-g_{\pi^*}(x+1) \\
&= \frac{p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))}\left[\frac{1}{x}-\frac{1-\lambda}{p^{\pi^*}(x)}g_{\pi^*}(x+1)\right],
\end{aligned} \tag{49}$$

$$\begin{aligned}
&g_{\pi^*}(x-1)-g_{\pi^*}(x) \\
&= \frac{p^{\pi^*}(x-1)}{1-\lambda(1-p^{\pi^*}(x-1))}\left[\frac{1}{x-1}-\frac{1-\lambda}{p^{\pi^*}(x-1)}g_{\pi^*}(x)\right].
\end{aligned} \tag{50}$$

From monotonicity of $p^{\pi^*}(x)$ on $x \geq \check{x}$, we have:

$$\frac{p^{\pi^*}(x)}{1-\lambda(1-p^{\pi^*}(x))} \leq \frac{p^{\pi^*}(x-1)}{1-\lambda(1-p^{\pi^*}(x-1))}. \tag{51}$$

Combining (48) and (51) yields:

$$g_{\pi^*}(x)-g_{\pi^*}(x+1) \leq g_{\pi^*}(x-1)-g_{\pi^*}(x). \tag{52}$$

From (15) and (43), we get:

$$\begin{aligned}
&2p^{\pi^*}(x-1)-p^{\pi^*}(x)-p^{\pi^*}(x-2) \\
&= \frac{\lambda}{a}\left[2g_{\pi^*}(x)-g_{\pi^*}(x-1)-g_{\pi^*}(x+1)\right] \\
&\leq 0.
\end{aligned} \tag{53}$$

So, $r(x-1)-r(x) \leq r(x-2)-r(x-1), \forall \check{x}+2 \leq x \leq \hat{x}-1$. Furthermore, since $0 = 2g_{\pi^*}(\hat{x}+1) \leq g_{\pi^*}(\hat{x})+g_{\pi^*}(\hat{x}+2)$, we have:

$$\begin{aligned}
&2\left[1-\frac{1}{a}(\hat{x}+b-\lambda g_{\pi^*}(\hat{x}+1))\right] \\
&\leq 1-\frac{1}{a}(\hat{x}-1+b-\lambda g_{\pi^*}(\hat{x})) \\
&\quad +1-\frac{1}{a}(x+1+b-\lambda g_{\pi^*}(\hat{x}+2)) \\
&< 1-\frac{1}{a}(\hat{x}-1+b-\lambda g_{\pi^*}(\hat{x})).
\end{aligned} \tag{54}$$

Hence,

$$p^{\pi^*}(\hat{x}) \leq \frac{1}{2}p^{\pi^*}(\hat{x}-1) = \frac{1}{2}(p^{\pi^*}(\hat{x}-1)+p^{\pi^*}(\hat{x}+1)) \tag{55}$$

Thus, we have $r(\hat{x})-r(\hat{x}+1) \leq r(\hat{x}-1)-r(\hat{x})$. Since $\lambda \leq \frac{1}{a-b}$, we have $\hat{x}-1 \leq (1-\lambda(1-p^{\pi^*}(\hat{x})))\hat{x}$. Together with $p^{\pi^*}(\hat{x}) \leq \frac{1}{2}p^{\pi^*}(\hat{x}-1)$, we have:

$$\begin{aligned}
&\frac{p^{\pi^*}(\hat{x}-1)}{\hat{x}-1} \\
&\geq \frac{2}{1-\lambda(1-p^{\pi^*}(\hat{x}))}\frac{p^{\pi^*}(\hat{x})}{\hat{x}} \\
&= 2g_{\pi^*}(\hat{x}) \\
&\geq (2-2\lambda+\lambda p^{\pi^*}(\hat{x}-1))g_{\pi^*}(\hat{x}).
\end{aligned} \tag{56}$$

So,

$$\begin{aligned}
&g_{\pi^*}(\hat{x}-1) \\
&= \frac{1}{1-\lambda(1-p^{\pi^*}(\hat{x}-1))} \\
&\quad \times\left[\frac{p^{\pi^*}(\hat{x}-1)}{\hat{x}-1}+\lambda p^{\pi^*}(\hat{x}-1)g_{\pi^*}(\hat{x})\right] \\
&\geq 2g_{\pi^*}(\hat{x}).
\end{aligned} \tag{57}$$

Thereby,

$$\begin{aligned}
&2p^{\pi^*}(\hat{x}-1)-p^{\pi^*}(\hat{x})-p^{\pi^*}(\hat{x}-2) \\
&= \frac{\lambda}{a}[2g_{\pi^*}(\hat{x})-g_{\pi^*}(\hat{x}-1)] \leq 0.
\end{aligned} \tag{58}$$

So, $r(\hat{x}-1)-r(\hat{x}) \leq r(\hat{x}-2)-r(\hat{x}-1)$. Hence, overall, $r(x-1)-r(x) \leq r(x-2)-r(x-1), \forall x \geq \check{x}+2$.

Now, consider $2 \leq x \leq \check{x}$. Because $g_{\pi^*}(x+1) \leq \frac{1}{(1-\lambda)(x+1)} \leq \frac{1}{(1-\lambda)x}$, we have $\frac{1}{x}-\frac{1-\lambda}{r(x)}g_{\pi^*}(x+1) \geq \frac{1}{x}\left(1-\frac{1}{r(x)}\right)$. Hence, from (26), we get:

$$\begin{aligned}
&g_{\pi^*}(x)-g_{\pi^*}(x+1) \\
&\geq \frac{r(x)}{1-\lambda(1-r(x))}\frac{1}{x}\left(1-\frac{1}{r(x)}\right) \\
&= -\frac{1-r(x)}{1-\lambda(1-r(x))}\frac{1}{x} \\
&\geq -\frac{1}{\lambda}.
\end{aligned} \tag{59}$$

Thus,

$$(x+b-\lambda g_{\pi^*}(x+1))^+ \geq (x-1+b-\lambda g_{\pi^*}(x))^+. \quad (60)$$

So,

$$\frac{1-\frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+}{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+} \leq 1. \quad (61)$$

Together with (35), we know that:

$$\begin{aligned} &\frac{1-\frac{1}{a}(x+b-\lambda g_{\pi^*}(x+1))^+}{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+} f(x) \\ &\quad + \frac{1-\frac{1}{a}(x-2+b-\lambda g_{\pi^*}(x-1))^+}{1-\frac{1}{a}(x-1+b-\lambda g_{\pi^*}(x))^+} f(x-2) \\ &\leq f(x) + \left(1+\frac{1+\frac{\lambda}{x}}{a-b-\breve{x}}\right) f(x-2) \\ &\leq 2f(x-1). \end{aligned} \quad (62)$$

Thus, $r(x)+r(x-2) \leq 2r(x-1)$, i.e., $r(x)-r(x-1) \leq r(x-1)-r(x-2), \forall 2 \leq x \leq \breve{x}$. ■

Remark 2: Assumption (i) of Theorem 4 requires the discount factor λ to be sufficiently small, or in other words, players of the popularity dynamics game are myopic and don't care about future rewards very much. Assumption (ii) is basically equivalent to $f(x)-f(x-1) \leq f(x-1)-f(x-2), \forall 2 \leq x \leq \breve{x}$ since the ratio in the parenthesis of (ii) is usually very small. This requires $f(x)$'s increasing speed is slowing down as x approaches $\breve{x}$, which is a reasonable assumption. Moreover, we notice that Theorem 4 does not cover all the situations of popularity dynamics. There are real-world popularity dynamics, such as those in Fig. 3-(a), which have more complicated second order patterns. For example, during increasing phase of the dynamics, the increasing speed can first increase and then decrease. Due to the intricacy of these second order patterns, we don't give theoretical discussions about them here.

VI. Simulations and Real Data Experiments

In this section, we conduct simulations and real data experiments to validate the theoretical results obtained. We choose the form of instantaneous reward function to be linear, i.e., $R(x,\theta) = -x + a\theta - b$, where $a > b > 0$.

A. Simulations

In our simulation, we define the visibility probability function $f(x)$ in the following form:

$$f(x) = \alpha\left(1-e^{-\beta x}\right) + 1 - \alpha, \quad (63)$$

where α, β are parameters controlling the initial visibility probability and the increasing speed of the visibility probability. We set the discount factor to be $\lambda = 0.5$. For different parameter setups of a, b, α, β, we stochastically simulate the equilibrium popularity dynamics calculated by Algorithm 1 many times and then take average of them. Here, the equilibrium behaviors are stochastic because (i) the user types are random variables; (ii) whether the item is visible to the arriving player is random. We also theoretically compute the expected equilibrium popularity dynamics by Algorithm 1, which serve as the theoretical dynamics. Specifically, for theoretical dynamics, at each time instant, we replace the actual stochastic equilibrium behavior with the expected equilibrium behavior. This deviation may affect the system state at the next time instant, which in turn influence the equilibrium behaviors at the next time instant since players' strategies depend on the system state. In other words, the deviation caused by using the expected equilibrium behaviors to approximate the actual stochastic equilibrium behaviors may propagate and accumulate. The simulations are aimed at verifying that this approximation does not hurt much,

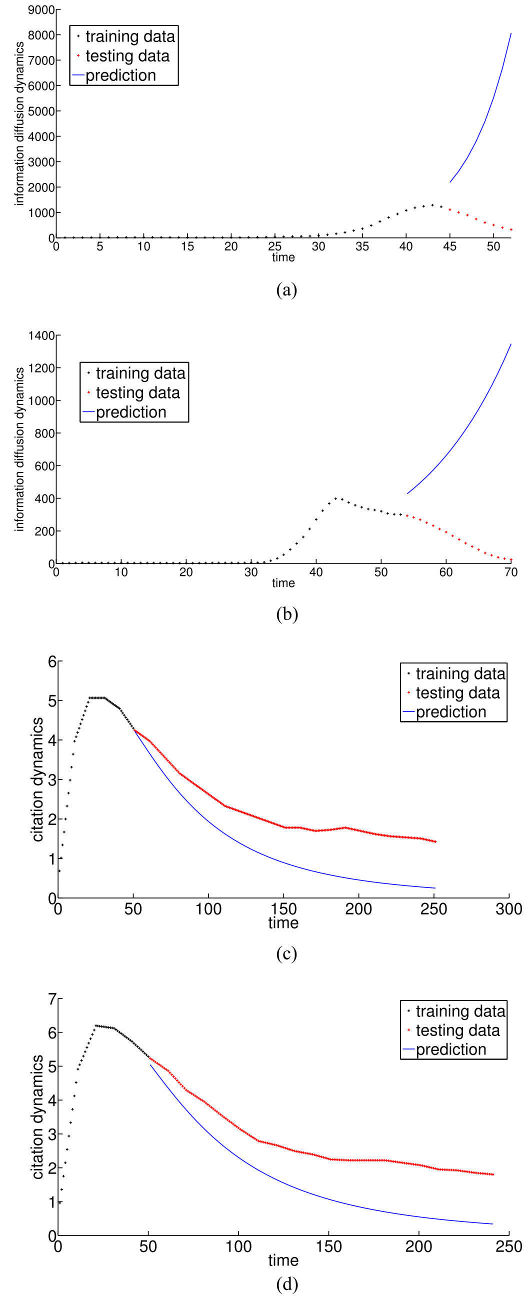


Fig. 8. Prediction results of the method in [17]

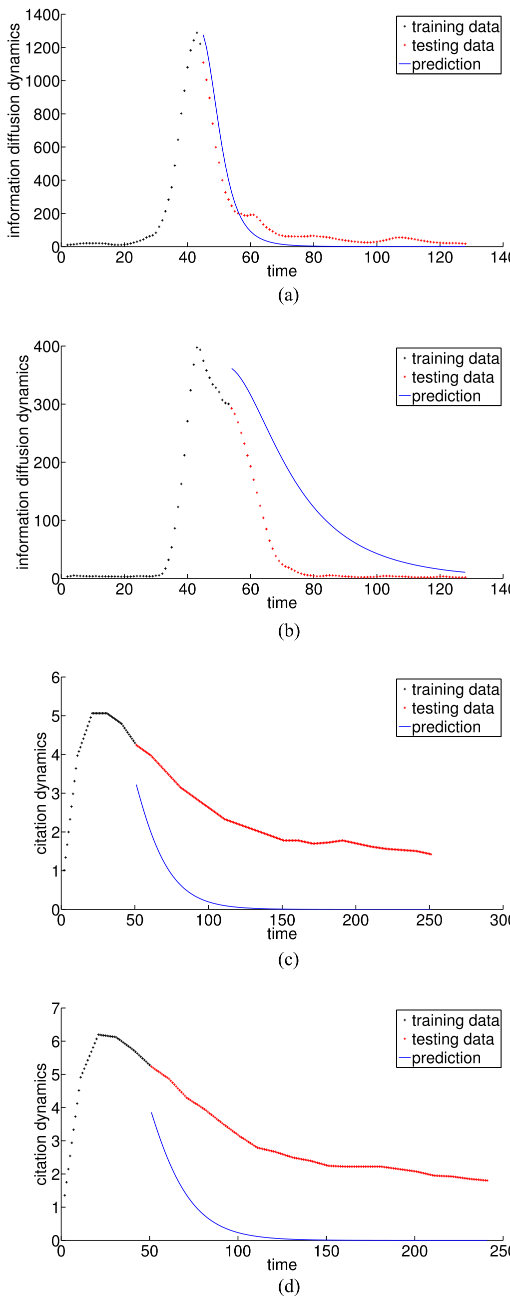


Fig. 9. Prediction results of the method in [4].

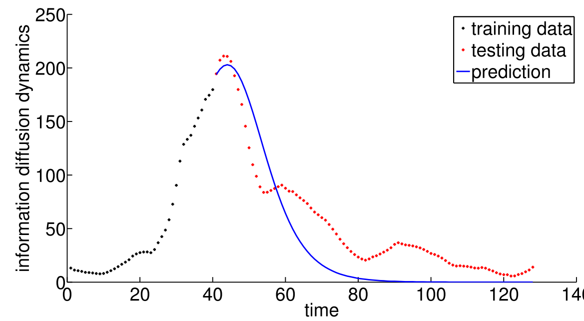


Fig. 10. Prediction before reaching the peak of the dynamics: Twitter hashtag #Tehran.

i.e., the theoretical dynamics can still match well with the simulated ones. The theoretical cumulative dynamics as well as the corresponding simulated cumulative dynamics are shown in Fig. 4, from which we observe that (i) the theoretical dynamics indeed match well with the simulated dynamics; (ii) the proposed game-theoretic model can flexibly generate popularity dynamics of different shapes by tuning the parameters.

B. Real Data Experiments

In this subsection, real data experiments are carried out to verify that the proposed theory matches well with the real-world popularity dynamics. The datasets we use here are Twitter hashtag dataset [9] and the citation data of papers from the Web of Science [22]. The Twitter hashtag data are the mentioning counts of popular hashtags from June to December 2009. We first use the equilibrium computed by Algorithm 1 to fit the mention dynamics of four popular Twitter hashtags in Fig. 5. To fit a popularity dynamics, we use the dynamics data to estimate the parameters of the proposed model and then use the estimated parameters to generate a theoretical dynamics, which is the fitting result. We observe that the theoretical fitting dynamics match well with the real-world dynamics. We further fit the average citation dynamics of the papers published in Nature 1990 and Science 1990, respectively, in Fig. 6. We remark that the fitting is still very accurate, though the temporal shape of the citation dynamics are very different from that of the Twitter hashtag dynamics, confirming the universality of our theory for popularity dynamics.

Additionally, we can even exploit the equilibrium of the proposed game to predict future dynamics for real data. To this end, we use part of the dynamics data to train the proposed game-theoretic model, i.e., estimate the parameters in the model, and then predict future dynamics by using the trained model. The prediction results are reported in Fig. 7, from which we see that the prediction is quite accurate. To highlight the advantage of the proposed approach, we compare with the prediction results of two existing methods, namely the methods in [17] and [4], which are reported in Figures 8 and 9, respectively. The four dynamics to be predicted are the same as those in Fig. 7. First, we note that the approach in [17] is proposed for citation dynamics. From Fig. 8, we observe that the method of [17] fails in predicting the dynamics of the two hashtags (subfigures (a) and (b)). Even for prediction of citation dynamics (subfigures (c) and (d)), our approach outperforms the method in [17]. Second, noting that the method in [4] is designed for information diffusion dynamics, we observe that our approach still outperforms it when predicting the dynamics of two hashtags (Fig. 9-(a) and Fig. 9-(b)). When it comes to the prediction of citation dynamics, our approach is much better than the method in [4] (Fig. 9-(c) and Fig. 9-(d)). These comparisons demonstrate that our proposed approach is universally good for general popularity dynamics. Even compared with methods specifically designed for a certain kind of popularity dynamics (e.g., [4] for information diffusion and [17] for citations), our method is still better. In addition, the performance enhancement over the method in [4] can be ascribed to the fact that the model in [4] is merely based on instantaneous incentives while our model incorporates long-term incentives as well, which suggests the importance and necessity of taking long-term incentives of individuals into account.

Generally, the prediction is accurate when the training period includes the peak of the dynamics. However, sometimes, we may even predict future dynamics accurately without knowing the peak, which is illustrated by a Twitter hashtag #Tehran in Fig. 10.

VII. Conclusion

In this paper, a sequential game is proposed to characterize the mechanisms of popularity dynamics. We prove that the proposed game has a unique SNE, which is a pure strategy action rule with a threshold structure and can be computed using a backward induction algorithm. Moreover, at the equilibrium of the proposed game, we analyze some properties observed from the real data, demonstrating that the equilibrium behavior of the proposed game confirms with real-world popularity dynamics. The theory is validated by both simulations and experiments based on real data. The proposed model can even predict future dynamics.

References

[1] C.-Y. Wang, Y. Chen, and K. J. R. Liu, "Sequential chinese restaurant game," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 571–584, Feb. 2013.

[2] S. Tadelis, *Game Theory: An Introduction*. Princeton, NJ, USA: Princeton Univ. Press, 2013.

[3] C. Jiang, Y. Chen, and K. J. R. Liu, "Graphical evolutionary game for information diffusion over social networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 4, pp. 524–536, Aug. 2014.

[4] C. Jiang, Y. Chen, and K. J. R. Liu, "Evolutionary dynamics of information diffusion over social networks," *IEEE Trans. Signal Process.*, vol. 62, no. 17, pp. 4573–4586, Sep. 2014.

[5] J. Leskovec, *Dynamics of Large Networks*. Ann Arbor, MI, USA: ProQuest, 2008.

[6] J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani, "Characterizing and modeling the dynamics of online popularity," *Phys. Rev. Lett.*, vol. 105, no. 15, 2010, Art. no. 158701.

[7] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[8] H.-W. Shen, D. Wang, C. Song, and A.-L. Barabási, "Modeling and predicting popularity dynamics via reinforced poisson processes," in *Proc. 28th AAAI Conf. Artif. Intell.*, 2014.

[9] J. Yang and J. Leskovec, "Patterns of temporal variation in online media," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, 2011, pp. 177–186.

[10] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2009, pp. 497–506.

[11] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 519–528.

[12] M. Gomez-rodriguez, D. Balduzzi, and B. Schlkopf, "Uncovering the temporal dynamics of diffusion networks," in *Proc. 28th Int. Conf. Mach. Learn.*, 2011, pp. 561–568.

[13] J. Yang and J. Leskovec, "Modeling information diffusion in implicit networks," in *Proc. 2010 IEEE 10th Int. Conf. Data Mining*, 2010, pp. 599–608.

[14] M. G. Rodriguez, J. Leskovec, D. Balduzzi, and B. Schölkopf, "Uncovering the structure and temporal dynamics of information propagation," *Netw. Sci.*, vol. 2, no. 01, pp. 26–65, 2014.

[15] A. Guille and H. Hacid, "A predictive model for the temporal dynamics of information diffusion in online social networks," in *Proc. 21st Int. Conf. Companion World Wide Web*, 2012, pp. 1145–1152.

[16] Y. Gao, Y. Chen, and K. J. R. Liu, "Understanding sequential user behavior in social computing: To answer or to vote?," *IEEE Trans. Netw. Sci. Eng.*, vol. 2, no. 3, pp. 112–126, Jul./Sep. 2015.

[17] D. Wang, C. Song, and A.-L. Barabási, "Quantifying long-term scientific impact," *Science*, vol. 342, no. 6154, pp. 127–132, 2013.

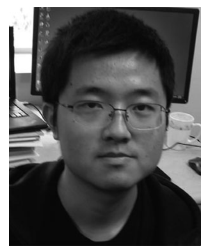
[18] T. Qiu, Z. Ge, S. Lee, J. Wang, Q. Zhao, and J. Xu, "Modeling channel popularity dynamics in a large IPTV system," in *Proc. ACM SIGMETRICS Perform. Eval. Rev.*, vol. 37, 2009, pp. 275–286.

[19] K. Lerman and T. Hogg, "Using a model of social dynamics to predict popularity of news," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 621–630.

[20] J. Ko, H. Kwon, H. Kim, K. Lee, and M. Choi, "Model for twitter dynamics: Public attention and time series of tweeting," *Phys. A, Stat. Mech. Appl.*, vol. 404, pp. 142–149, 2014.

[21] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Trans. Web*, vol. 1, no. 1 2007, Art. no. 5.

[22] 2016. [Online]. Available: http://wokinfo.com/



Xuanyu Cao received the bachelor's degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2013. He received the Jimmy Lin scholarship from the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA, where he is now working toward the Ph.D. degree. His general research interest is network science with current focus on game-theoretic analysis of user behavior dynamics as well as mechanism design and optimization for resource trading. He received the first prizes in Chinese National Mathematics Contest in 2007 and 2008.

Yan Chen (SM'14) received the Bachelor's degree from the University of Science and Technology of China, Hefei, China, in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, in 2007, and the Ph.D. degree from the University of Maryland, College Park, MD, USA, in 2011. Being a founding member, he joined Origin Wireless, Inc., as a Principal Technologist in 2013. He is currently a Professor with the University of Electronic Science and Technology of China, Chengdu, China. His research interests include multimedia, signal processing, game theory, and wireless communications.

Dr. Chen received multiple honors and awards, including the Best Student Paper Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing, in 2016, the best paper award at the IEEE Global Telecommunications Conference, in 2013, the Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, the University of Maryland in 2011, and the Chinese Government Award for outstanding students abroad in 2010.

K. J. Ray Liu (F'03) was named a Distinguished Scholar-Teacher of the University of Maryland, College Park, MD, USA, in 2007, where he is Christine Kim Eminent Professor of information technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on smart radios for smart life.

He is a member of the IEEE Board of Director. He was the President of the IEEE Signal Processing Society, where he has served as a Vice President of Publications and Board of Governor. He has also served as the Editor-in-Chief of the *IEEE Signal Processing Magazine*. He received the 2016 IEEE Leon K. Kirchmayer Technical Field Award on graduate teaching and mentoring, the IEEE Signal Processing Society 2014 Society Award, and the IEEE Signal Processing Society 2009 Technical Achievement Award. He also received teaching and research recognitions from the University of Maryland including University-Level Invention of the Year Award, the College-Level Poole and Kent Senior Faculty Teaching Award, the Outstanding Faculty Research Award, and Outstanding the Faculty Service Award, all from A. James Clark School of Engineering.

Recognized by Thomson Reuters as a Highly Cited Researcher, he is a fellow of AAAS.