

Will the Winner Take All? Competing Influences in Social Networks Under Information Overload

Chen Feng
Shanghai Jiao Tong University
fengchen@sjtu.edu.cn

Jiahui Sun
Shanghai Jiao Tong University
jhsun1997@sjtu.edu.cn

Luoyi Fu
Shanghai Jiao Tong University
yiluofu@sjtu.edu.cn

ABSTRACT

Influence competition finds its significance in many applications, such as marketing, politics and public events like COVID-19. Existing work tends to believe that the stronger influence will always win and dominate nearly the whole network, i.e., “winner takes all”. However, this finding somewhat contradicts with our common sense that many competing products are actually coexistent, e.g., Android vs. iOS. This contradiction naturally raises the question: *will the winner take all?*

To answer this question, we make a comprehensive study into influence competition by identifying two factors frequently overlooked by prior art: (1) the incomplete observation of real diffusion networks; (2) the existence of information overload and its impact on user behaviors. To this end, we attempt to recover possible diffusion links based on user similarities, which are extracted by embedding users into a latent space. Following this, we further derive the condition under which users will be overloaded, and formulate the competing processes where users’ behaviors differ before and after information overload. By establishing the explicit expressions of competing dynamics, we disclose that information overload acts as the critical “boundary line”, before which the “winner takes all” phenomenon will definitively occur, whereas after information overload the share of influences gradually stabilizes and is jointly affected by their initial spreading conditions, influence powers and the advent of overload. Numerous experiments are conducted to validate our theoretical results where favorable agreement is found. Our work sheds light on the intrinsic driving forces behind real-world dynamics, thus providing useful insights into effective information engineering.

KEYWORDS

information competition, social networks, network embedding

1 INTRODUCTION

Since its outbreak in December 2019, the new coronavirus disease (COVID-19) has been rapidly spreading through 216 countries and areas, causing already 67,210,778 cases and 1,540,777 deaths at the time of writing [23]. People are thrown into a panic for especially effective therapies, providing a hotbed for conflicting ideas, reports and information etc.. A notable example may be the controversy over the effect of sesame oil. At the beginning of 2020, a rumor that applying sesame oil to the body would block the entry of the new coronavirus is widely spreading among people. Although the World Health Organization quickly refuted such misinformation, an increased sale of sesame oil is still reported [13]. Competitions between conflicting ideas (such as whether to prevent COVID-19 by taking a hot bath, drinking alcohol and eating garlic [22]) continue to happen through the year, significantly affecting people’s life.

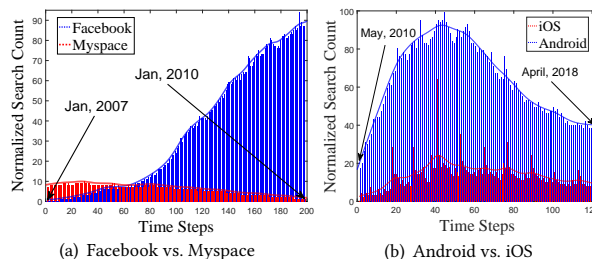


Figure 1: Search interests of competing products over time². (a) In the early years, Myspace is shown to have an edge over Facebook. However, as time goes on, Facebook gradually manifests its advantage and finally dominates nearly the whole market. (b) Across the ten years’ competition between Android and iOS, although Android is shown to be more popular, it does not dominate the whole market, refuting the “winner takes all” phenomenon.

Should we have deeper insights into the competing dynamics of conflicting information, the incidents could have been avoided by accurately engineering the propagation process. Besides pandemic control, similar demands are raised in many other areas such as marketing, politics and social ideology. To fill this gap, a body of work has been dedicated to studying the competition between conflicting information, products, opinions etc.¹. The work of Prakash et al. [14] marks a milestone in investigating the problem, by revealing the well-known “winner takes all” phenomenon in the marketing theory, that the stronger influence will become the winner and finally wipe out the weaker, as shown in Fig. 1(a). Numerous subsequent researches reach the same conclusion that only one influence will survive in the final equilibrium state [10] [24] [19] [20]. However, this finding somewhat contradicts with our daily experience, since many competing products are observed to be coexistent, such as McDonald’s vs. KFC, Boeing vs. Airbus and Android vs. iOS, as illustrated in Fig. 1(b). There is also little literature showing that competing influences will always co-exist. Especially, Koprulu et al. find that the market share of products will gradually stabilize and is basically proportional to their persuasive powers [7], i.e., winner will not take all. The contradictory conclusions and opposite examples naturally lead to the question: *will the winner take all?*

Competition Review. To answer this question, we may need to review the competing process for an overall view. In essence, the competition starts with the injection of influences. Then, the influences propagate along the given network. And the users therein decide which influence to follow. As can be seen, the competition process is an interplay of three parts: the influences themselves

¹For ease of appellation, we refer to them all as “influences”.

²The data is available on Google Trends [4]. In each subgraph, the search count is normalized by the largest count.

(*intrinsic factor*), the users (*subjective factor*) and the underlying networks (*objective factor*). It is worth noting that previous work has delicately considered the intrinsic factor via various characterizations, such as propagation rate [14], persuasive capability [7]. However, the latter two factors are frequently overlooked in existing literature. To elaborate, (1) *subjective factor*: most work ideally assumes a fixed user behavior throughout the whole competing process. However, in reality, as time goes on, more users will become influenced and turn a spreader. Thus, a late decision maker would receive more information which possibly exceeds users' attention limit and leads to information overload [21]. Further studies reveal that overloaded users exhibit a different behavior pattern from before – preferring priority strategies to cope with the overload [15]. (2) *Objective factor*: it is often implicitly assumed in previous work that the influence propagates through a given network. However, due to users' privacy concern, the large scale of social networks etc., the observed network is often only a subgraph of real networks. Many legitimate links are actually not detected.

The New Formulation. In this work, we take all the three factors into account. Specifically, we consider two competing influences I_1 and I_2 propagating through an observed social network $G(V, E)$. The influences are characterized by their influence powers a and b respectively. Each user is featured by their capacity in handling information, i.e., δ_c . Initially, users have enough attention to discriminate the influences and tend to follow influences by the received powers. Beyond threshold δ_c , users are generally overloaded and four typical information processing strategies are considered for users to cope with the overload, with each defined according to some priority, e.g., the arrival time.

To circumvent the incompleteness of the observed network, we look into the generation of social relationships. It is found in social science that network connections generally stem from user similarities [2] and users tend to interact with similar ones [11], known as the “homophily principle”. Thus, we are inspired to recover possible diffusion links based on user similarities, before running the competition. In the attempt to extract user similarities, we find that this idea coincides with an emerging technique in link prediction called *Network Embedding* [5] [9] [18], where users are mapped into a latent space with user similarities encoded as their distances.

Technical Challenges. Proceeding to reveal the competing dynamics, we are confronted with three major challenges.

- User embeddings are typically concerned with practical applications but hardly analytically captured, leaving properties of the latent space unknown (esp. distance distribution) which however are an essential prerequisite for diffusion analyses.
- Four information processing strategies are defined in a broad way for generality, while imposing substantial difficulties in mathematical characterization of user behaviors and accordingly the analysis of competing processes.
- Like existing work, in revealing the competing dynamics, we are faced with the long-standing difficulty of deriving explicit expressions for each influence w.r.t. each time instant.

Our Solution. To meet the above challenges, we first estimate users' distribution in the latent space and further disclose the distribution of user distances. On this basis, the time of overload is derived. Then, the diffusion processes are formalized as differential

equations, based on the characterization of the four information strategies. Finally, we solve the equations and reveal the underlying competing dynamics.

Numerous contributions are made when fulfilling this work, among which the highlights are summarized as follows.

- This work acts as a comprehensive study of influence competition, identifying not only intrinsic factor but also subjective factor and objective factor in the competing process.
- To our knowledge, our work is the first attempt to mathematically delineate the properties of latent space. Moreover, in the case of overload, users' behaviors under four strategies are analytically captured in a unified form.
- We manage to derive the competing dynamics in explicit expressions, showing that the winner will not always “take all”. To explain, before overload, the stronger influence will be the winner and dominate the network, exhibiting the “winner takes all” phenomenon. While under information overload, the result is jointly determined by the initial state, influence power and the time of overload. And the weaker influence could even become the winner and maintain a modest share of market, given enough initial followers and proper overload timing.
- Numerous experiments are carried out on real-world networks and the results are found to be in favorable agreement with our theoretical findings.

Roadmap. We formulate the problem in Sec. 2 and study the properties of the latent space in Sec. 3. Then, we reveal the competing dynamics before and after overload in Sec. 4 and 5. The theoretical findings are further validated in Sec. 7. Finally, we conclude the paper in Sec. 8.

In Sec. 6, we answer the question raised in the title and discuss the implications of our findings in practical events like COVID-19.

2 PROBLEM FORMULATION

The observed social network is denoted as $G(V, E)$, where V is the set of users ($|V| = n$) and E records user connections. Users' degree is assumed to follow the power law distribution with parameter $\lambda > 1$, which is widely observed in the real world [1] [12]. Then, the probability of a user having degree k is C/k^λ , where C is the normalization factor.

Two competing influences I_1 and I_2 are spreading among users, with influence powers a and b respectively. The competing process starts at time t_0 , with initial states $X_1(t_0)$ and $X_2(t_0)$,³ where $X_1(t)$ (resp. $X_2(t)$) denotes the number of users affected by I_1 (resp. I_2) at time t and $X_1(t) + X_2(t) = X(t)$. Then, the share of each influence could be expressed as $X_1(t)/X(t)$ and $X_2(t)/X(t)$.

2.1 Network Embedding

In practice, the subsequent diffusion process is not likely to expand along the observed network exactly, since many legitimate links are not observed due to users' privacy concern, the large scale of real networks etc.. In this regard, we intend to recover possible diffusion links based on user similarities, since social connections

³Without loss of generality (w.l.o.g.), we assume t_0 is equal to $X(t_0)$ in value. Then, according to the diffusion mechanism defined later, at any time step, t is exactly the number of users infected, i.e., $X_1(t) + X_2(t) = t$.

often originate from user similarities [2] and users tend to interact with similar ones [11].

To this end, we adopt the well-known “node2vec” technique [5] to abstract user similarities due to its high accuracy in link prediction (typically over 90%). Specifically, users are mapped into a d dimensional latent space, where each user i is represented by a vector $\mathbf{v}_i \in \mathbb{R}^d$ and the similarity between users i, j is encoded by their distance $l(i, j)$.

The framework of “node2vec” mainly consists of two parts: neighborhood sampling and embedding optimization.

Neighborhood Sampling. For each user $i \in V$, we simulate a biased random walk parameterized by p and q , as defined in Definition 1, to obtain her wide-sense neighborhood $N_S(i)$.

DEFINITION 1. (Sampling Strategy) Consider a random walk which just traversed edge (u, v) and reaches node v . Then, neighbors of v are assigned weight 1 (esp. u itself is assigned $1/p$), and others are of weight $1/q$. The walk next reaches the neighbors of v with probability proportional to their weights.

Embedding Optimization. We next seek to maximize the log-probability of observing each neighbor pair in the latent space

$$\mathcal{O} = \sum_{i \in V} \sum_{j \in N_S(i)} \log P(i, j). \quad (1)$$

Specifically $P(i, j)$ is a softmax function w.r.t. user distance

$$P(i, j) = \frac{\exp(-l(i, j))}{\sum_{w \in V} \exp(-l(i, w))}, \quad (2)$$

where $l(i, j) = \|\mathbf{v}_i - \mathbf{v}_j\|_2^2$, i.e., the squared Euclidean distance, in our work. Accordingly, we obtain the objective function

$$\mathcal{O} = - \sum_{i \in V} \left[\sum_{j \in N_S(i)} l(i, j) + |N_S(i)| \log \sum_{w \in V} \exp(-l(i, w)) \right]. \quad (3)$$

Finally, the objective \mathcal{O} could be optimized by applying the stochastic gradient ascent algorithm.

For an analytical characterization, user embeddings are captured by their distribution in the latent space $h(\mathbf{v}, \cdot, \theta)$, where the parameter θ is estimated by estimation methods like maximum likelihood.

2.2 Competing Mechanism

Based on the given network and user embeddings, we could derive the recovered diffusion network $G(V, E')$, where the edge set E' is the union of the original social relations E and the latent links recovered from similarities. Specifically, we assume a latent link would exist between two users if their distance in the latent space is smaller than $r \in \mathbb{R}^+$, since users are not likely to interact with ones not similar enough. Correspondingly, users' neighborhood consists of both neighbors in $G(V, E)$ and users within r .

On this basis, the competing process unfolds in discrete steps as follows. In each step, a random user i (called *riser*) will arise to make a decision on which influence to follow, based on the influences she receives from her neighbors (if none of her neighbors is influenced, it is natural to assume she will not arise).

In the early stage, since the influenced users are relatively rare, the riser generally will not receive too much information and become overloaded. Therefore, she would be able to discriminate the

power of influences and will follow influences with probability proportional to the cumulative influence power received (just like [20]), as illustrated in Fig. 2. To elaborate, at time t , user i arises to make a decision and finds that three of her neighbors are influenced with one original neighbor (user u) in red, two latent neighbors in blue. Then, user i will become red with probability $\frac{a \times 1}{a \times 1 + b \times 2} = \frac{a}{a + 2b}$, and blue with probability $\frac{2b}{a + 2b}$.

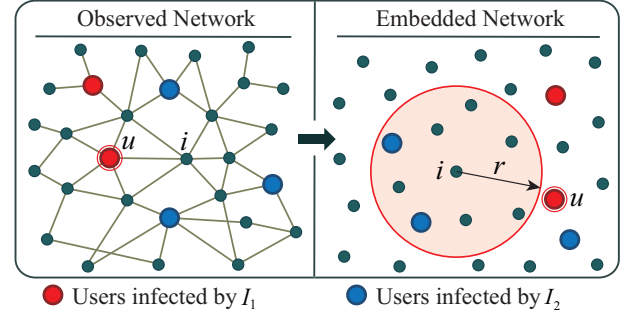


Figure 2: An illustration of the influence propagation model.

As time goes on, more users will become influenced. And a new riser will receive more information when making a decision. The abundance of information will finally exceed user's capacity to consume (denoted as δ_c), since user's attention is limited by [21]. Being overloaded with information, users may not have enough attention to carefully identify the influences and tend to adopt priority strategies to cope with the overload, according to the studies in [6] [15].

To capture this phenomenon, we assume an overloaded user will discriminate the influences as before with probability $p(t) = e^{-\mu(t-t_c)}$ ($t > t_c$), which decreases exponentially over time with a factor of μ , where t_c is the time when users are generally overloaded, i.e., receiving too much information that exceeds users' capacity. On the other hand, with probability $1 - p(t)$, users would turn to priority strategies to help process information. Especially, four typical strategies are considered in Definition 2.

DEFINITION 2. (Overloaded Behaviors) Under information overload, we consider four generic information processing strategies for users to cope with the overload.

- (i) The riser adopts the first influence that reaches her, given arbitrary distribution of information propagation time [19].
- (ii) Contrary to (i), the riser adopts the latest (i.e., the newest) influence, regardless of the distribution of propagation time.
- (iii) The riser follows the influence that her most similar neighbor chooses [11], under any user distributions in the latent space.
- (iv) The riser follows the influence her most popular neighbor (i.e., having the largest degree) adopts, under any degree distribution of the network.

2.3 Problem Statement

We are now ready to state the problem formally. Given an observed social network $G(V, E)$, user distribution $h(\mathbf{v}, \cdot, \theta)$ in \mathbb{R}^d and initial states $X_1(t_0)$ and $X_2(t_0)$, influences I_1 and I_2 with powers a and b are propagating among users according to the competing mechanism characterized by parameters r , δ_c and μ , where in practice the parameters could be derived by experimental estimation.

The goal is to derive the number of users infected by I_1 and I_2 with regard to each time step (i.e., $X_1(t)$ and $X_2(t)$) and thus answer the question “will the winner take all?”.

3 PROPERTIES OF THE DIFFUSION SPACE

In this section, we investigate the properties of the latent space. To obtain explicit results, we apply the d dimensional Gaussian distribution $\mathcal{N}(\mathbf{u}, \sigma^2 \mathbf{I})$ to depict user embeddings like many previous studies [9] [18], where \mathbf{u} is a $d \times 1$ mean vector, σ^2 is the variance and \mathbf{I} is the $d \times d$ identity matrix. The parameters \mathbf{u} and σ^2 could be estimated by the maximum likelihood method.

DEFINITION 3. (Gamma Distribution) A random variable Z follows the Gamma distribution $\Gamma(\alpha, \beta)$, if its probability density function $f(z; \alpha, \beta)$ satisfies

$$f(z; \alpha, \beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\beta z}, & z \geq 0, \\ 0, & z < 0. \end{cases} \quad (4)$$

where $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz$ is the gamma function.

Let random variable (r.v.) L denote the squared Euclidean distance between a random pair of users. We next examine the distribution of L and then derive the probability that the distance between two users is within r by Lemmas 1 and 2 respectively, where the proofs are deferred to the technical report [3] for readability.

LEMMA 1. Given the distribution of user positions $\mathcal{N}(\mathbf{u}, \sigma^2 \mathbf{I})$, the squared Euclidean distance between users L follows the Gamma distribution $\Gamma(\frac{d}{2}, \frac{1}{4\sigma^2})$.

LEMMA 2. In the latent space with user's position distribution $\mathcal{N}(\mathbf{u}, \sigma^2 \mathbf{I})$, the probability that the distance between two random users i, j is smaller than r (i.e., $l(i, j) \leq r$) is

$$p(r, \sigma^2) = \frac{1}{\Gamma(\frac{d}{2})} \int_0^{\frac{r}{4\sigma^2}} z^{\frac{d}{2}-1} e^{-z} dz,$$

which increases with influence range r and decreases with position variances σ^2 .

Note that $p(r, \sigma^2)$ is an incomplete gamma function, there is no closed form expression to calculate it. However, we could approximate it with the following formula according to [17]

$$\Gamma(\frac{d}{2}) - \gamma(\frac{d}{2}, \frac{r}{4\sigma^2}),$$

where function $\gamma(a+1, x)$ is defined as $\frac{e^{-x} x^{a+1}}{x-a} [1 - \frac{a}{(x-a)^2} + \frac{2a}{(x-a)^3} + O(\frac{a^2}{(x-a)^4})]$.

With the above results, we are ready to derive the time when new risers will generally feel information overload, i.e., the number of infected neighbors exceeds the threshold δ_c .

THEOREM 1. With probability $1 - o(1)$, a new riser will feel information overload if she arises after time $t_c = \frac{\delta_c}{(1+o(1))p(r, \sigma^2)}$.

PROOF. Consider a random user i with degree d_i in $G(V, E)$. We first examine the number of her actual neighbors, which consists of two parts: the neighbors in the original network and the latent neighbors due to similarity. For the first case, We would like to

show that $d_i = o(n)$ with high probability. Recall that d_i follows the power law distribution parameterized by $\lambda > 1$. When $\lambda \in (1, 2)$, the expectation of user degree is $\mathbb{E}(d_i) = \int_1^n x \frac{C}{x^\lambda} dx = \Theta(n^{2-\lambda})$. By Chernoff bound, we have

$$P(|X - \mathbb{E}(X)| \leq \epsilon \mathbb{E}(X)) \geq 1 - 2e^{-\mathbb{E}(X)\epsilon^2/3}. \quad (5)$$

Let ϵ be $\sqrt{\frac{3 \log n}{n^{2-\lambda+\epsilon'}}$, we have that d_i concentrates around its expectation, which is inferior to n since $\lambda > 1$, with probability $1 - o(1)$. Similarly, when $\lambda \geq 2$, $\mathbb{E}(d_i) = O(\log n)$. By setting $\epsilon = \log n$, we see that $d_i < (\log n)^2 = o(n)$ with probability $1 - o(1)$.

For the second case, the number of her latent neighbors is a random variable (denote it as X) which follows a binomial distribution $B(n - d_i, p(r, \sigma^2))$. Applying the Chernoff bound in Eq. (5) by setting $\epsilon = \sqrt{\frac{3 \log n}{(n-d_i)p(r, \sigma^2)}}$, we see that $X = (1 + o(1))\mathbb{E}(X)$ with probability at least $1 - \frac{1}{(1+o(1))n}$ which approaches 1 as $n \rightarrow \infty$. Thus, we could approximate i 's latent neighborhood with its expectation $(n - d_i)p(r, \sigma^2)$.

To summarize, the size of user's actual neighborhood is $d_i + (n - d_i)p(r, \sigma^2) = (1 + o(1))np(r, \sigma^2)$. Then, at time t , the expected number of infected neighbors is $(1 + o(1))np(r, \sigma^2) \frac{t}{n}$. Setting it to be δ_c , we derive the time when users will be overloaded $t_c = \frac{\delta_c}{(1+o(1))p(r, \sigma^2)}$. \square

Remark. The theoretical results agree well with our common intuitions. To explain, Theorem 1 implies that the time point of overload is directly proportional to users' ability of handling information (i.e., δ_c), and inversely proportional to the probability of having a neighbor, i.e., $p(r, \sigma^2)$. Taking a closer look into $p(r, \sigma^2)$, by Lemma 2 we find that if users are more sociable, i.e., with a larger r , they tend to have more neighbors and are thus more likely to be overloaded. On the other hand, in a sparsely connected network, implying a larger σ^2 , users often have fewer friends, resulting in a smaller probability of being overloaded.

On this basis, we could infer that in $G(V, E')$ users are more likely to be overloaded and have a smaller t_c than in $G(V, E)$, since the recovered network has more edges (thus smaller σ^2) than the observed one (i.e., $|E'| \geq |E|$).

Since user's behaviors before and after overload are different, in subsequent sections, we will analyze the two scenarios respectively.

4 INFLUENCE PROPAGATION BEFORE OVERLOAD

In this section, we consider the influence propagation before time t_c , i.e., users are not overloaded with information.

Recall that the number of users influenced by I_1 and I_2 at time t is $X_1(t)$ and $X_2(t)$ respectively. At the next time step $t + 1$, suppose user i arises to make a decision. In expectation, she will receive $[d_i + (n - d_i)p(r, \sigma)]X_1(t)/n$ pieces of information about I_1 and $[d_i + (n - d_i)p(r, \sigma)]X_2(t)/n$ pieces of information about I_2 . Then, according to the competing mechanism, we have two possible outcomes of i 's decision:

- user i gets influenced by I_1 with probability $\frac{aX_1(t)}{aX_1(t)+bX_2(t)}$;
- user i gets influenced by I_2 with probability $\frac{bX_2(t)}{aX_1(t)+bX_2(t)}$.

Given current states $X_1(t)$ and $X_2(t)$, the mean state of the system at time $t + 1$ could be characterized by the following difference equation set

$$\begin{cases} \mathbb{E}[X_1(t+1) - X_1(t)|X_1(t)] = \frac{aX_1(t)}{aX_1(t) + bX_2(t)}, \\ \mathbb{E}[X_2(t+1) - X_2(t)|X_2(t)] = \frac{bX_2(t)}{aX_1(t) + bX_2(t)}. \end{cases} \quad (6)$$

To solve the equations, we present an equivalent view to the propagation process in a continuous way and convert the difference equations into differential equations. Specifically, we suppose the decision process expands as follows

$$P(\text{A random user makes decision during } (t, t + \tau]) = \tau,$$

$$P(\text{No user makes decision during } (t, t + \tau]) = 1 - \tau,$$

where $\tau \in (0, 1]$. Then, the number of infected users at time t is the same as the original discrete model. For this continuous model, the first equation in Eq. (6) could be written as

$$\mathbb{E}[X_1(t + \tau) - X_1(t)|X_1(t)] = \tau \frac{aX_1(t)}{aX_1(t) + bX_2(t)}. \quad (7)$$

Let $x_1(t)$ (resp. $x_2(t)$) denote the expected number of users infected by I_1 (resp. I_2) in the continuous model, i.e., $x_1(t) = E(X_1(t))$ (resp. $x_2(t) = E(X_2(t))$). Taking the expectation of Eq. (7), we obtain

$$x_1(t + \tau) - x_1(t) = \tau \frac{ax_1(t)}{ax_1(t) + bx_2(t)}. \quad (8)$$

Since Eq. (8) holds for $\tau \in (0, 1]$, we have

$$\lim_{\tau \rightarrow 0} \frac{x_1(t + \tau) - x_1(t)}{\tau} = \frac{ax_1(t)}{ax_1(t) + bx_2(t)}. \quad (9)$$

Then, we conduct similar processes to the second equation in Eq. (6). Finally, by definition of the derivative, we obtain the differential description of the propagation process

$$\begin{cases} \frac{dx_1}{dt} = \frac{ax_1(t)}{ax_1(t) + bx_2(t)}, \\ \frac{dx_2}{dt} = \frac{bx_2(t)}{ax_1(t) + bx_2(t)}. \end{cases} \quad (10)$$

On this basis, we obtain the influence dynamics before information overload by Theorem 2.

THEOREM 2. *At time t , the number of users infected by I_1 and I_2 satisfies the following equations respectively*

$$\begin{cases} C_1 x_1^{\frac{b}{a}}(t) + x_1(t) = t, \\ C_2 x_2^{\frac{a}{b}}(t) + x_2(t) = t, \end{cases} \quad (11)$$

where $C_1 = [x_2(t_0)]/x_1^{\frac{b}{a}}(t_0)$, $C_2 = [x_1(t_0)]/x_2^{\frac{a}{b}}(t_0)$.

PROOF. We obtain the expressions about $x_1(t)$ and $x_2(t)$ by solving the differential equations in Eq. (10). Note that $x_1(t) + x_2(t) = t$ at any time. For the first equation of Eq. (10), by substituting $x_2(t)$ with $t - x_1(t)$, we have

$$\frac{dx_1}{dt} = \frac{ax_1(t)}{(a-b)x_1 + bt}. \quad (12)$$

We next take the reciprocal of each side, and denote $q = \frac{t}{x_1}$. Since q is also a function of x_1 , we derive that $\frac{dt}{dx_1} = q + x_1 \frac{dq}{dx_1}$. Then, Eq.

(12) is equivalent to

$$x_1 \frac{dq}{dx_1} = \frac{a-b}{a} + \left(\frac{b}{a} - 1\right)q.$$

By separating the variables, we have

$$\frac{a}{a-b} \frac{1}{1-q} dq = \frac{1}{x_1} dx_1.$$

Integrating both sides and taking anti-logarithm, after some simplifications, we obtain

$$C_1 x_1^{\frac{b}{a}}(t) + x_1 = t.$$

where C_1 is a constant. Solving C_1 with the initial condition $x_1(t_0)$ and $x_1(t_0) + x_2(t_0) = t_0$, we have $C_1 = [x_2(t_0)]/x_1^{\frac{b}{a}}(t_0)$.

Following the similar idea, we could obtain the expression about $x_2(t)$. And this completes the proof. \square

To derive more specific expressions of $x_1(t)$ and $x_2(t)$, we study two special cases in Corollary 1.

COROLLARY 1. (i) *When $\frac{a}{b} = \frac{1}{2}$, at time t , the number of users infected by I_1 (resp. I_2) is*

$$\begin{cases} x_1(t) = \frac{1}{2C_1} \left(\sqrt{4C_1 t + 1} - 1 \right), \\ x_2(t) = t - \frac{1}{2} \left(C_2 \sqrt{4t + C_2^2} - C_2^2 \right). \end{cases}$$

Symmetric results could be derived when $\frac{a}{b} = 2$.

(ii) *When $\frac{a}{b} = 1$, at time t , the number of users infected by I_1 and I_2 is*

$$x_1(t) = \frac{x_1(t_0)}{t_0} t, \quad x_2(t) = \frac{x_2(t_0)}{t_0} t.$$

That is to say, the influence spread of I_1 and I_2 is totally decided by the initial states.

Remark. By Theorem 2, we find that before information overload, the competing dynamics exhibit the ‘‘winner takes all’’ phenomenon. To explain, we assume that I_1 is stronger (i.e., $a > b$) w.l.o.g.. Then, Eq. (11) could be written as

$$x_1(t) = \frac{t}{1 + C_1 x_1^{\frac{b}{a}-1}(t)}, \quad x_2(t) = \left[\frac{t}{C_2 + x_2^{1-\frac{a}{b}}(t)} \right]^{\frac{b}{a}}. \quad (13)$$

Note that $x_1(t)$ and $x_2(t)$ both increase over time. Thus, as $t \rightarrow \infty$, $C_1 x_1^{\frac{b}{a}-1}(t)$ is an infinitesimal $o(1)$ and accordingly $x_1(t)$ scales like t by Eq. (13). Similarly, $x_2^{1-\frac{a}{b}}(t)$ is also in $o(1)$ and $x_2(t)$ scales like $t^{\frac{b}{a}} = o(t)$, i.e., $x_2(t) = o(x_1(t))$. Thus, we see that, as time goes on, if no overload is incurred ($t_c \rightarrow \infty$), I_1 will win the competition and dominate nearly the whole network, i.e., the winner takes all. Especially, if $a = b$ (i.e., no power difference), I_1 and I_2 will share the network according to their ratio in the initial states.

5 INFLUENCE PROPAGATION AFTER OVERLOAD

In this section, we proceed to study the competition at time $t > t_c$ when users have been overloaded. To this end, we first look into users' behavior under information overload. Then we formulate and reveal the competing dynamics of I_1, I_2 .

5.1 User Behavior Under Information Overload

Recall that four information processing strategies are considered in the formulation of our problem in Definition 2. By ordinal, we first study strategies (i) and (i), where users adopt the earliest and latest influence respectively. The corresponding infection probability of a new riser is presented in Lemmata 3 and 4, where the proof of Lemma 4 is deferred to our technical report [3] for readability.

LEMMA 3. *Assume the propagation time of influences follows an arbitrary distribution of p.d.f. $g(x)$. At time step t , the probability that I_1 reaches the riser first is*

$$\frac{X_1(t)}{X_1(t) + X_2(t)}.$$

PROOF. Consider a random riser. At time t , the expected number of her neighbors infected by I_1 (resp. I_2) is $X_1(t)p(r, \sigma)$ (resp. $X_2(t)p(r, \sigma)$), which, for ease of exposition, is written as m (resp. s).

Let Y_1, Y_2, \dots, Y_m (resp. Z_1, Z_2, \dots, Z_s) be the time when the information arrives from users infected by I_1 (resp. I_2). And r.v. Y (resp. Z) denotes the time when the first information about I_1 (resp. I_2) arrives. Then, $Y = \min_{k=1, \dots, m} Y_k$ and $Z = \min_{k=1, \dots, s} Z_k$. User u will adopt I_1 if $Y < Z$ and vice versa. Assume the information from neighbors propagate independently. Then, $\{Y_k\}_{k=1}^m$ and $\{Z_k\}_{k=1}^s$ are *i.i.d.* distributed; Y and Z are independent of each other.

It is easy to see that the p.d.f. of Y is $f_Y(y) = m[1 - G(y)]^{m-1}g(y)$, where $G(\cdot)$ is the cumulative distribution function (c.d.f.) of propagation time. And the p.d.f. of Z is $f_Z(z) = s[1 - G(z)]^{s-1}g(z)$. Since Y and Z are independent, we have

$$\begin{aligned} P(Y < Z) &= \int_0^\infty f_Y(y) \left[\int_y^\infty f_Z(z) dz \right] dy \\ &= \int_0^\infty f_Y(y) \left[\int_y^\infty s[1 - G(z)]^{s-1}g(z) dz \right] dy \\ &= \int_0^\infty m[1 - G(y)]^{m-1}g(y)[1 - G(y)]^s dy \\ &= \frac{m}{m+s} [1 - G(y)]^{m+s} \Big|_0^\infty \\ &= \frac{m}{m+s}. \end{aligned}$$

Recall that $m = X_1(t)p(r, \sigma)$ and $s = X_2(t)p(r, \sigma)$, then the probability that information about I_1 arrives first is $\frac{X_1(t)}{X_1(t) + X_2(t)}$. This completes the proof. \square

LEMMA 4. *Assume the time for influences to reach a user follows an arbitrary distribution of p.d.f. $g(x)$. At time step t , the probability that influence I_1 arrives latest is*

$$\frac{X_1(t)}{X_1(t) + X_2(t)}.$$

Regarding strategies (iii) and (iv), we would like to show that they are mathematically equivalent to cases (i) and (ii) respectively. Note that in strategy (iii), users follow the influence of her most similar neighbor, i.e., the one nearest to her. Assume an arbitrary distribution of user distance $g(x)$ in the latent space. Then, by replacing arrival time by the squared Euclidean distance, the setting

of (iii) is reduced to case (i). Similarly, assume an arbitrary distribution of user degrees $g(x)$. Substituting the arrival time of strategy (ii) by user's degree, we derive the setting of case (iv).

Summarizing the four cases, we find that strategies (i) and (iii) could be captured by Lemma 3, and strategies (i) and (iv) could be delineated by Lemma 4. Particularly, all the four complicated cases implies the same adoption probability.

5.2 Competing Influences Under Information Overload

Recall the diffusion mechanism that under information overload ($t > t_c$), a user will adopt influences based on their influence powers with probability $p(t) = e^{-\mu(t-t_c)}$. And with probability $1-p(t)$, users will resort to the four strategies described in Definition 2. On this basis, similar to the case before overload, we could formulate the behavior of an overloaded riser as follows:

- The new riser will get influenced by I_1 with probability $p(t) \frac{aX_1(t)}{aX_1(t) + bX_2(t)} + (1-p(t)) \frac{X_1(t)}{X_1(t) + X_2(t)}$;
- The new riser will get influenced by I_2 with probability $p(t) \frac{bX_2(t)}{aX_1(t) + bX_2(t)} + (1-p(t)) \frac{X_2(t)}{X_1(t) + X_2(t)}$.

Then, we are ready to formalize the competing process under information overload. Given current states $X_1(t)$ and $X_2(t)$, by applying the continuous analogy technique in Section 4, we obtain the mean state of the system at time $t + 1$

$$\begin{cases} \frac{dx_1}{dt} = p(t) \frac{ax_1(t)}{ax_1(t) + bx_2(t)} + (1-p(t)) \frac{x_1(t)}{x_1(t) + x_2(t)}, \\ \frac{dx_2}{dt} = p(t) \frac{bx_2(t)}{ax_1(t) + bx_2(t)} + (1-p(t)) \frac{x_2(t)}{x_1(t) + x_2(t)}; \end{cases} \quad (14)$$

where $x_1(t)$ (resp. $x_2(t)$) is the expectation of $X_1(t)$ (resp. $X_2(t)$) in the continuous analogy.

In Theorem 3, we characterize the influence spread at any time $t > t_c$, by solving the above Eq. (14).

THEOREM 3. *The number of users influenced by I_1 and I_2 satisfies the following equations respectively.*

(i) At time $t_c < t < t_c + 1/\mu$,

$$\begin{cases} C_3 t^{\frac{b-a}{a}} \mu t_c e^{-\frac{a-b}{a} \mu t} x_1^{\frac{b}{a}}(t) + x_1(t) = t, \\ C_4 t^{\frac{a-b}{b}} \mu t_c e^{-\frac{b-a}{b} \mu t} x_2^{\frac{a}{b}}(t) + x_2(t) = t, \end{cases} \quad (15)$$

where

$$\begin{cases} C_3 = x_2(t_c) / t_c^{\frac{b-a}{a}} \mu t_c e^{-\frac{a-b}{a} \mu t_c} x_1^{\frac{b}{a}}(t_c), \\ C_4 = x_1(t_c) / t_c^{\frac{a-b}{b}} \mu t_c e^{-\frac{b-a}{b} \mu t_c} x_2^{\frac{a}{b}}(t_c). \end{cases} \quad (16)$$

(ii) At time $t > t_c + 1/\mu$,

$$x_1(t) = \frac{x_1(t_c + \frac{1}{\mu})}{t_c + \frac{1}{\mu}} t, \quad x_2(t) = \frac{x_2(t_c + \frac{1}{\mu})}{t_c + \frac{1}{\mu}} t. \quad (17)$$

PROOF. Without loss of generality, we focus on the equation about influence I_1 , i.e.,

$$\frac{dx_1}{dt} = p(t) \frac{ax_1(t)}{ax_1(t) + bx_2(t)} + (1-p(t)) \frac{x_1(t)}{x_1(t) + x_2(t)}.$$

Note that $x_1(t) + x_2(t) = t$ at any time, the equation is equivalent to

$$\frac{dx_1}{dt} = p(t) \frac{ax_1(t)}{(a-b)x_1(t) + bt} + (1-p(t)) \frac{x_1(t)}{t}. \quad (18)$$

Let $q = \frac{x_1(t)}{t}$, then $x_1(t) = qt$ and $\frac{dx_1}{dt} = q + t \frac{dq}{dt}$. After some simplification and separating variables, Eq. (18) becomes

$$\frac{b + (a-b)q}{(a-b)q - (a-b)q^2} dq = \frac{p(t)}{t} dt.$$

By splitting the fraction in the left hand side, we have

$$\left(\frac{b}{a-b} + \frac{a}{1-q} \right) dq = \frac{p(t)}{t} dt.$$

Integrating both sides and noting that $q < 1$, we obtain

$$\frac{b}{a-b} \ln(q) - \frac{a}{a-b} \ln(1-q) = \int \frac{p(t)}{t} dt. \quad (19)$$

The r.h.s. is an exponential integral and no closed form expression could be derived. To circumvent this problem, we approximate $p(t)$ with its first-order Maclaurin expansion $\hat{p}(t)$,

$$\hat{p}(t) = \begin{cases} -\mu(t-t_c) + 1, & t_c < t < t_c + 1/\mu; \\ 0, & t \geq t_c + 1/\mu. \end{cases}$$

From Fig. 3, we can see that the gap between $\hat{p}(t)$ and $p(t)$ is small, especially when t is around t_c and large enough. Correspondingly, we discuss Eq. (19) from two parts.

(i) $t_c < t < t_c + 1/\mu$.

Eq. (19) is equivalent to

$$\frac{b}{a-b} \ln(q) - \frac{a}{a-b} \ln(1-q) = \int -\mu + \frac{1+\mu t_c}{t} dt = \ln e^{-\mu t} + \ln(Ct)^{1+\mu t_c},$$

where C is a constant. After some transformation, we have

$$\begin{aligned} \ln \frac{q^b}{(1-q)^a} &= \ln \left[\frac{(Ct)^{1+\mu t_c}}{e^{\mu t}} \right]^{(a-b)} \\ \iff q^b e^{(a-b)\mu t} &= (Ct)^{(1+\mu t_c)(a-b)} (1-q)^a \\ \stackrel{(b)}{\iff} (x_1(t)/t)^b e^{(a-b)\mu t} &= (Ct)^{(1+\mu t_c)(a-b)} (1-x_1(t)/t)^a \\ \iff x_1^b(t) e^{(a-b)\mu t} &= C^{(1+\mu t_c)(a-b)} t^{(a-b)\mu t_c} (t-x_1(t))^a \\ \iff C^{(1+\mu t_c)(b-a)} t^{(b-a)\mu t_c} e^{(a-b)\mu t} x_1^b(t) &= (t-x_1(t))^a \\ \iff C^{\frac{(1+\mu t_c)(b-a)}{a}} t^{\frac{b-a}{a}\mu t_c} e^{\frac{a-b}{a}\mu t} x_1^{\frac{b}{a}}(t) &= t-x_1(t), \end{aligned}$$

where equality (b) is obtained by recovering q with $\frac{x_1(t)}{t}$. Denoting the constant about C as C_3 and solving C_3 with the states at t_c , we derive the solution of the first case.

(ii) $t \geq t_c + 1/\mu$.

In this case, Eq. (19) reduces to

$$\frac{dx_1}{dt} = \frac{x_1(t)}{x_1(t) + x_2(t)},$$

which coincides with a special case of Eq. (10), where $a = b$. By Corollary 1, we obtain that

$$x_1(t) = \frac{x_1(t_c + 1/\mu)}{t_c + 1/\mu} t, \quad x_2(t) = \frac{x_2(t_c + 1/\mu)}{t_c + 1/\mu} t$$

Combining cases (i) and (ii), we complete the proof. \square

Remark. Theorem 3 reveals that under information overload, the competing dynamics are different from before. Assume $a > b$, for $t_c < t < t_c + 1/\mu$, similar to the derivation of Eq. (13), from Eq. (15) we could obtain

$$x_1(t) = t/[1 + C_3 t^{\frac{b-a}{a}\mu t_c} e^{\frac{a-b}{a}\mu t} x_1^{\frac{b}{a}-1}(t)], \quad (20)$$

where the term before $x_1^{\frac{b}{a}-1}(t)$ is significantly larger than that of Eq. (13), implying that if given the same initial states, $x_1(t)$ would be smaller than that before overload, i.e., the stronger influence is impaired after overload. After time t_c , applying similar analyses to Eq. (17), we could see that the impairment is even severer. Specifically, the stronger influence stops increasing and becomes stabilizing, let alone "taking all".

Finally, we make a rigorous comparison between the competing dynamics before and after information overload by Theorem 4. To this end, we treat the two cases as individuals instead of consecutive sub-processes and assume the same initial states, i.e., $x_1(t_0) = x_1'(t_0)$ and $x_2(t_0) = x_2'(t_0)$, where $x_{1,2}(\cdot)$ (resp. $x'_{1,2}(\cdot)$) denotes the influences before (resp. after) information overload (correspondingly referred to as case 1 and case 2 respectively for brevity).

THEOREM 4. *At ant time t , the winner's influence in case 2 is no larger than that in case 1 and the weaker's influence in case 2 is no smaller than that in case 1, i.e.,*

$$\begin{cases} x_1'(t) \leq x_1(t) \text{ and } x_2'(t) \geq x_2(t), & \text{if } a > b; \\ x_1'(t) \geq x_1(t) \text{ and } x_2'(t) \leq x_2(t), & \text{if } a < b. \end{cases} \quad (21a)$$

$$\begin{cases} x_1'(t) \leq x_1(t) \text{ and } x_2'(t) \geq x_2(t), & \text{if } a > b; \\ x_1'(t) \geq x_1(t) \text{ and } x_2'(t) \leq x_2(t), & \text{if } a < b. \end{cases} \quad (21b)$$

PROOF. We focus on proving the results under $a > b$, i.e., I_1 being the winner. And the results under $a < b$ could be proven in a similar way.

To prove the inequality in Eq. (21a), for $x_1(t)$, by Eq. (13) we have

$$x_1(t) = \frac{t}{1 + C_1 x_1^{\frac{b}{a}-1}(t)}, \quad (22)$$

where $C_1 = x_2(t_0)/x_1^{\frac{b}{a}}(t_0)$. For $x_1'(t)$, recall the results in Theorem 3, we could transform it to

$$x_1'(t) = \begin{cases} \frac{t}{1 + C_3' [x_1'(t)]^{\frac{b}{a}-1}}, & t_0 < t < t_0 + 1/\mu \\ t \cdot x_1'(t_0 + \frac{1}{\mu}) / (t_0 + \frac{1}{\mu}), & t \geq t_0 + \frac{1}{\mu} \end{cases} \quad (23a)$$

$$x_1'(t) = \begin{cases} \frac{t}{1 + C_3' [x_1'(t)]^{\frac{b}{a}-1}}, & t_0 < t < t_0 + 1/\mu \\ t \cdot x_1'(t_0 + \frac{1}{\mu}) / (t_0 + \frac{1}{\mu}), & t \geq t_0 + \frac{1}{\mu} \end{cases} \quad (23b)$$

where $C_3' = x_2'(t_0) t^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu t_0} / t_0^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu t_0} [x_1'(t_0)]^{\frac{b}{a}}$. Denoting $\rho(t) = t^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu t} / t_0^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu t_0}$, we could write $C_3' = C_1 \rho(t)$ since $x_1'(t_0) = x_1(t_0)$ and $x_2'(t_0) = x_2(t_0)$.

Seeing that $x_1'(t)$ is a piece-wise function, we compare $x_1(t)$ with the two cases of $x_1'(t)$ respectively.

(i) $t_0 < t < t_0 + 1/\mu$

In this condition, the main difference between $x_1(t)$ and $x_1'(t)$ lies in the time varying coefficient $\rho(t)$. To examine the property of $\rho(t)$, for $t \in [t_0, t_0 + 1/\mu]$, we write t as $t_0 + \Delta$ where Δ is a variable

ranging in $[0, \frac{1}{\mu}]$. Then, we have

$$\begin{aligned} \rho(t) = \rho(\Delta, t_0) &= \frac{(t_0 + \Delta)^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu(t_0+\Delta)}}{t_0^{\frac{b-a}{a}\mu t_0} e^{\frac{a-b}{a}\mu t_0}} \\ &= e^{\frac{a-b}{a}\mu\Delta} \left(1 + \frac{\Delta}{t_0}\right)^{\frac{b-a}{a}\mu t_0} = \left[e/\left(1 + \frac{\Delta}{t_0}\right)\right]^{\frac{t_0}{\Delta} \frac{a-b}{a}\mu\Delta}. \end{aligned} \quad (24)$$

Since $(1+x)^{\frac{1}{x}}$ is monotonically decreasing, we have $\left(1 + \frac{\Delta}{t_0}\right)^{\frac{t_0}{\Delta}} \leq \lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$. Recall that $a > b$, then $\rho_1(t) \geq 1$ and is monotonically increasing over time t .

Combining Eq.s (22) and (23a) and noting that $C_3' = C_1\rho(t)$, we could derive the equation below

$$x_1(t) + C_1 x_1^{\frac{b}{a}}(t) = x_1'(t) + C_1 \rho(t) [x_1'(t)]^{\frac{b}{a}}. \quad (25)$$

The above equation (25) is enough for us to draw our conclusion that $x_1'(t) \leq x_1(t)$. To see this, by contradiction, we assume $x_1'(t) > x_1(t)$. Then, we have

$$x_1'(t) + C_1 \rho(t) [x_1'(t)]^{\frac{b}{a}} > x_1(t) + C_1 \rho(t) [x_1(t)]^{\frac{b}{a}}.$$

Recall Eq. (25), we further derive

$$x_1(t) + C_1 [x_1(t)]^{\frac{b}{a}} > x_1(t) + C_1 \rho(t) [x_1(t)]^{\frac{b}{a}},$$

which implies $\rho(t) < 1$, contradicting with our finding $\rho(t) \geq 1$. Therefore, $x_1'(t) \leq x_1(t)$.

Note that $x_1'(t) + x_2'(t) = x_1(t) + x_2(t) = t$. By $x_1'(t) \leq x_1(t)$, we could derive that $x_2'(t) \geq x_2(t)$.

(ii) $t \geq t_0 + 1/\mu$

For easy of notation, we denote $t_0 + 1/\mu$ as t_0' and consider the influence diffusion thereafter, i.e., treating t_0' as the new start with states consistent with $t \in (t_0, t_0 + 1/\mu)$. Then, the influence spread of I_1 could be expressed as

$$x_1(t) = t/[1 + \widehat{C}_1 x_1^{\frac{b}{a}-1}(t)], \quad t \geq t_0 + 1/\mu, \quad (26)$$

where $\widehat{C}_1 = x_2(t_0')/x_1^{\frac{b}{a}}(t_0')$.

Then, combining Eq.s (26) and (23b) and noting that $t_0' = x_1'(t_0') + x_2'(t_0')$, we have

$$x_1(t) + \frac{x_2(t_0')}{x_1^{\frac{b}{a}}(t_0')} x_1^{\frac{b}{a}}(t) = \frac{x_1'(t_0')}{x_1'(t_0')} = x_1'(t) + \frac{x_2'(t_0')}{x_1'(t_0')} x_1'(t). \quad (27)$$

Again, we tend to prove $x_1'(t) \leq x_1(t)$ by contradiction. To this end, we assume $x_1'(t) > x_1(t)$. Then, Eq. (27) implies that

$$\begin{aligned} x_1(t) + \frac{x_2(t_0')}{x_1^{\frac{b}{a}}(t_0')} x_1^{\frac{b}{a}}(t) &> x_1(t) + \frac{x_2'(t_0')}{x_1'(t_0')} x_1(t) \\ \Leftrightarrow \frac{x_2(t_0')}{x_2'(t_0')} \frac{x_1^{\frac{b}{a}}(t)}{x_1^{\frac{b}{a}}(t_0')} &> [x_1(t)]^{1-\frac{b}{a}} \\ \stackrel{(c)}{\Leftrightarrow} [x_1(t_0')]^{1-\frac{b}{a}} &> [x_1(t)]^{1-\frac{b}{a}} \\ \stackrel{(d)}{\Leftrightarrow} x_1(t_0') &> x_1(t), \end{aligned} \quad (28)$$

where step (c) is due to our results under $t_0 < t < t_0 + 1/\mu$ that $x_1'(t_0') \leq x_1(t_0')$ and $x_2'(t_0') \geq x_2(t_0')$, and step (d) is due to the premise $a > b$. Note that $x_1(t)$ is non-decreasing over time and $t \geq t_0'$. Thus

the conclusion in Eq. (28) contradicts with the monotonicity of $x_1(t)$. And we have $x_1'(t) \leq x_1(t)$ for $t \geq t_0 + 1/\mu$.

Following similar technique in case (i), we derive $x_2'(t) \geq x_2(t)$, since $x_1'(t) + x_2'(t) = x_1(t) + x_2(t) = t$.

Combining cases (i) and (ii), we finish the proof of Eq. (21a). To obtain the symmetric result Eq. (21b), we only need to swap a and b and thus the detailed derivations are omitted. \square

Remark. Theorem 4 indicates that in the case of overload, the advantage of the stronger influence is impaired and the weaker influence could gain more followers, resulting in a smaller difference between I_1 and I_2 . That is, the competition after overload seems to be milder than that before overload.

Synthesizing Theorems 2, 3 and 4, we see that in the beginning, the share of the stronger influence constantly increases and the weaker one decreases. However, with the advent of overload, the competition becomes milder and after a short time ($1/\mu$), both influences tend to stabilize, as if the competition is over.

6 DISCUSSION

To answer the question posed in the title and illuminate the implications of our theoretical findings, we present a detailed discussion on the following issues.

Question: *Will the winner take all?*

Answer: The winner could take all under certain condition but not always. The result is a joint effect of the intrinsic, subjective and objective factors. Generally speaking, if users are not to be overloaded, the stronger influence will win out and dominate the network, exhibiting the ‘‘winner takes all’’ phenomenon. In the case of information overload, the result depends on the stable state, which roughly begins at the time of overload and keeps steady till the final state.

Question: *What determines the final states?*

Answer: Before overload, the final state is solely determined by the influence power. The influence with a larger power will dominate nearly the whole network. In the case of overload, the final state is jointly determined by the initial state, influence power and the time of overload. Specifically, to win the competition, an influence is expected to have more initial followers, larger influence power and a later advent of overload, to establish and enhance its advantage before the advent of overload.

Further, the overload is likely to arrive late if the network is sparsely connected. And users are robust to superfluous information if they are less active in social interactions.

Question: *Are the theoretical results of practical significance in events like COVID-19?*

Answer: Absolutely. Take the ‘‘sesame treatment’’ as an example. Suppose the rumor has more initial followers and larger influence power, due to its broadcast on media and users’ panic for new treatments. In turn, the truth is situated in a negative condition. By our findings, to win the competition, we should first publish the truth timely to earn more initial followers. Secondly, we could attach scientific evidences, advice of the authority etc. to the propagation of the truth, as a means of elevating its influence power.

Thirdly, the use of information overload is a little tricky. Specifically, when the share of the truth is smaller, the overload is desired

to restrain the false, and vice versa. To harness the overload, we could adjust the network connectivity by blocking or recommending links; control the influence range by promoting or restricting user's activeness; affect user's sense of overload by flooding them with information or filtering information for them.

7 EXPERIMENTS

In this section, we validate our theoretical findings by simulating influence competitions on real-world social networks.

7.1 Dataset Description

Two datasets are included for experiments, with detailed information as follows. As can be seen, the two networks are of nearly the same size of users, while BlogCatalog is shown to have 12 times larger amount of edges than Ca-HepTH, implying much denser user connections, i.e., smaller σ^2 .

- **BlogCatalog** is a social blog directory website, managing the bloggers and their blogs [16]. This dataset contains 10,312 users and 333,983 edges.

- **Ca-HepTh** is a collaboration network with 9,877 users and 25,998 edges, covering collaborations between authors who submitted papers to the HepTh category in arXiv [8].

In preparation for running competing processes, we first convert the networks into a latent space by the “node2vec” technique, where the user distance is instantiated as the squared Euclidean distance and the parameters are set by default (esp. $d = 128$). For an intuitive vision, we visualize user embeddings of two random dimensions in Fig. 4. It is shown that the embeddings of BlogCatalog are evidently denser than that of Ca-HepTh, due to the gap in the density of edges. This observation is also verified by the estimation of the variance σ^2 , which is found to be 0.0147 and 0.0916 respectively.

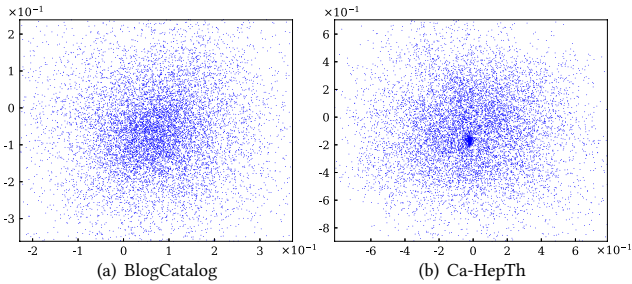


Figure 4: Visualization of the embedded networks.

7.2 Experimental Settings

Basically, we set influence range r to be 4 and μ to be 10. Before information overload, for effective validation, we assume t_c is large enough by setting $\delta_c = |V|$. Two pairs of initial states are considered: 40% I_1 versus 60% I_2 (denoted as S1) and the reverse case 60% I_1 versus 40% I_2 (denoted as S2), with total initial followers $X_1(t_0) + X_2(t_0) = 40$. Further, I_1 is selected to be the weaker one with influence power $a = 1$. And I_2 is tested under power $b = 2$ and 3 for comparison.

In the case of overload, the influence powers are set to be $a = 1$ and $b = 2$. To test the effect of overload timing, the competition is run under both $\delta_c = 30$ and $\delta_c = 500$. Moreover, to demonstrate that the weaker influence is possible to be the winner, besides S1

we consider a new case S3 (80% I_1 versus 20% I_2) where I_1 is given a larger initial advantage.

7.3 Experimental Results

The competing dynamics, featured by the share of each influence, are reported in Figs. 5 and 6, corresponding to cases before and after information overload respectively.

7.3.1 Before Overload. For both initial states S1 and S2 in Fig. 5, the stronger influence (i.e., I_2) is shown to be the winner and gradually dominate the network, verifying the “winner takes all” phenomenon revealed by Theorem 2.

Specifically, under initial state S1, as one may imagine, with a greater influence power and more initial followers, I_2 undoubtedly becomes the winner and quickly dominates nearly the whole network, as shown in Fig. 5(a). More precisely, it is observed that the line marked by “ Δ ” ($I_2, b = 3$) grows faster than that marked “ \square ” ($I_2, b = 2$), since I_2 gains more advantage over I_1 when $b = 3$, as predicted by Theorem 2.

Under initial state S2, although I_2 is given less initial followers than I_1 , as can be seen in Fig. 5(b), I_2 still managed to win the competition with a share approaching 100%. This observation serves as a firm validation of Theorem 2 which implies that the stronger influence will win nearly the whole network anyway, regardless of the initial condition.

7.3.2 After Overload. The competing process after information overload evidently differs from that before overload, exhibiting more complex competing dynamics.

The results under initial state S1 are reported in Fig. 6(a). It is shown that even with a greater influence power and more initial followers, I_2 is not assured to dominate the whole network as it could have done without overload (the dashed grey line), and only maintains a stable share of market, as predicted by Theorem 3. Comparing the lines marked “ Δ ” ($I_2, \delta_c = 500$) and “ \square ” ($I_2, \delta_c = 30$), we see that the final share of I_2 is increased when $\delta_c = 500$, since by Theorem 1 the advent of overload is postponed with a larger δ_c and therefore I_2 is given more time to increase. Moreover, we find that the share of the stronger influence without overload (the dashed grey line) is larger than that under overload (lines marked “ Δ ” and “ \square ”), thus verifying the results in Theorem 4.

We further simulate the competition under initial state S3 to show that the stronger influence I_2 is not assured to be the winner, let alone “taking all”. Instead, the weaker one I_1 is possible to win the competition. As can be seen in Fig. 6(b), starting from a larger initial share of 80%, I_1 is constantly decreasing in the beginning. However, with a small $\delta_c = 30$ (i.e., early advent of overload), the tendency of decline is stopped timely before losing its advantage (Theorem 3). Resultantly, I_1 becomes the final winner though with a weaker power. Conversely, under $\delta_c = 500$, the time of overload is postponed (Theorem 1). Thus, the share of I_1 continues to decrease and becomes inferior to I_2 finally.

Comparing the results under an early and late advent of overload, we could infer that given the same initial states, the weaker influence in $G(V, E')$ would have a larger share than that in $G(V, E)$, since $|E'| \geq |E|$ due to the link recovery which implies a smaller σ^2 and an early advent of overload.

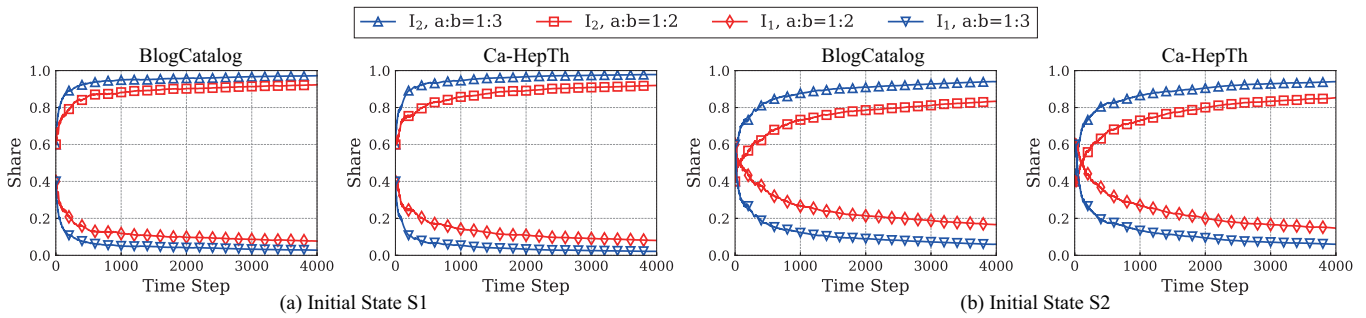


Figure 5: Competing dynamics BEFORE information overload.

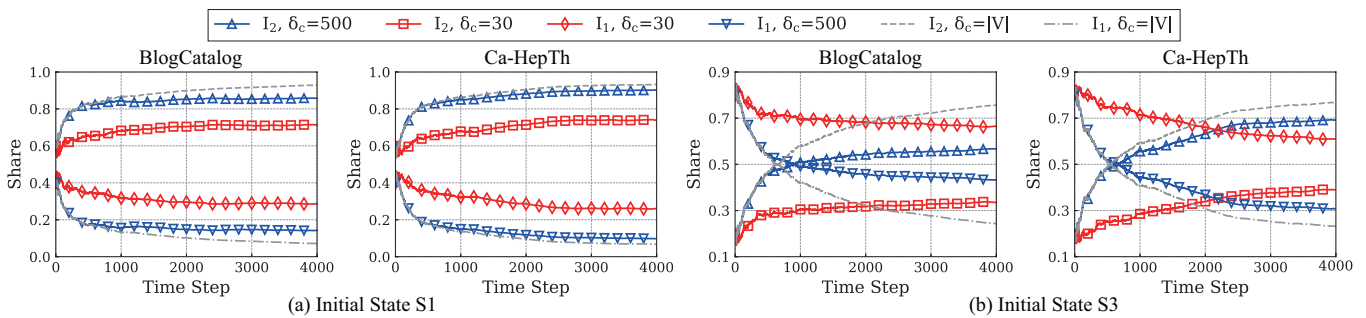


Figure 6: Competing dynamics AFTER information overload.

By Figures 5 and 6, we validate the theoretical findings which are found to be in good agreement with the experimental results.

8 CONCLUSION

In this paper, we present a comprehensive study into the competing dynamics in social networks. Different from existing work, we considered not only the intrinsic factor but also the subjective and objective factors. Specifically, the network is mapped into the latent space to recover possible diffusion links. And users are allowed to adopt different information strategies when overloaded. In this context, we formulate the cases before and after overload. Finally, explicit expressions of competing dynamics are derived. It is found that before information overload, the competing process exhibits the “winner takes all” phenomenon. While in the case of overload, the final state is determined by the stable state which is a joint result of the considered factors.

REFERENCES

- [1] Albert-László Barabási and Réka Albert. 1999. Emergence of scaling in random networks. *science* 286, 5439 (1999), 509–512.
- [2] Munmun De Choudhury, Hari Sundaram, Ajita John, Doree Duncan Seligmann, and Aisling Kelliher. 2010. ‘Birds of a Feather’: Does User Homophily Impact Information Diffusion in Social Media? *arXiv preprint arXiv:1006.1702* (2010).
- [3] Chen Feng, Jiahui Sun, and Luoyi Fu. 2020. Will the Winner Take All? Competing Influences in Social Networks Under Information Overload. <https://www.dropbox.com/s/c1gg51xntlxedtm/content%20transmission-twc%20supp-camera%20ready.pdf?dl=0>.
- [4] Google Trends. (accessed December 9, 2020). Search count of Myspace and Facebook, iOS and Android. <https://trends.google.com/trends/>.
- [5] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *Proc. Int. Conf. Knowl. Discov. Data Mining, KDD*. ACM, 855–864.
- [6] Jacob Jacoby. 1984. Perspectives on information overload. *Journal of consumer research* 10, 4 (1984), 432–435.
- [7] Irem Koprulu, Yoora Kim, and Ness B Shroff. 2019. Battle of opinions over evolving social networks. *IEEE/ACM Trans. Netw.* 27, 2 (2019), 532–545.
- [8] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. 2007. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discovery from Data*

- 1, 1 (2007), 2–es.
- [9] Lizi Liao, Xiangnan He, Hanwang Zhang, and Tat-Seng Chua. 2018. Attributed social network embedding. *IEEE Trans. Knowl. Data Eng.* 30, 12 (2018), 2257–2270.
- [10] Ji Liu, Philip E Paré, Angelia Nedich, Choon Yik Tang, Carolyn L Beck, and Tamer Basar. 2019. Analysis and control of a continuous-time bi-virus model. *IEEE Trans. Automat. Contr.* (2019).
- [11] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Ann. Rev. Sociol.* 27, 1 (2001), 415–444.
- [12] Lev Muchnik, Sen Pei, Lucas Parra, Saulo Reis, José S Andrade Jr, Shlomo Havlin, and Hernán A Makse. 2013. Origins of power-law degree distribution in the heterogeneity of human activity in social networks. *Sci. Rep.* 3, 1 (2013), 1–8.
- [13] Tomohiro Osaki. (accessed Dec. 9, 2020). Sesame oil? Granite baths? Sea lettuce? COVID-19 rumors spread like a virus in Japan. <https://www.japantimes.co.jp/news/2020/03/08/national/coronavirus-cure-rumors/>.
- [14] B Aditya Prakash, Alex Beutel, Roni Rosenfeld, and Christos Faloutsos. 2012. Winner takes all: competing viruses or ideas on fair-play networks. In *Proc. Int. Conf. World Wide Web*. ACM, 1037–1046.
- [15] Manuel Gomez Rodriguez, Krishna Gummadi, and Bernhard Schoelkopf. 2014. Quantifying information overload in social media and its impact on social contagions. In *Int. Conf. Weblogs and Social Media, ICWSM*. AAAI, 170–179.
- [16] Lei Tang and Huan Liu. 2009. Relational learning via latent social dimensions. In *Proc. Int. Conf. Knowl. Discov. Data Mining, KDD*. 817–826.
- [17] FG Tricomi. 1950. Asymptotische eigenschaften der unvollständigen gammafunktion. *Mathematische Zeitschrift* 53, 2 (1950), 136–148.
- [18] Daixin Wang, Peng Cui, and Wenwu Zhu. 2016. Structural deep network embedding. In *Proc. Int. Conf. Knowl. Discov. Data Mining, KDD*. 1225–1234.
- [19] Jie Wang and Wenye Wang. 2016. To live or to die: Encountering conflict information dissemination over simple networks. In *Proc. Conf. Comput. Commun., INFOCOM*. IEEE, 1–9.
- [20] Xuetao Wei, Nicholas C Valler, B Aditya Prakash, Iulian Neamtii, Michalis Faloutsos, and Christos Faloutsos. 2013. Competing memes propagation on networks: A network science perspective. *IEEE J. Sel. Areas Commun.* 31, 6 (2013), 1049–1060.
- [21] Lilian Weng, Alessandro Flammini, Alessandro Vespignani, and Filippo Menczer. 2012. Competition among memes in a world with limited attention. *Scientific Rep.* 2 (2012), 335.
- [22] World Health Organization. (accessed Dec. 9, 2020). Coronavirus disease (COVID-19) advice for the public: Mythbusters. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters>.
- [23] World Health Organization. (accessed Dec. 9, 2020). WHO Coronavirus Disease (COVID-19) Dashboard. <https://covid19.who.int/table>.
- [24] Lu-Xing Yang, Xiaofan Yang, and Yuan Yan Tang. 2017. A bi-virus competing spreading model with generic infection rates. *IEEE Trans. Netw. Sci. Eng.* 5, 1 (2017), 2–13.

A PROOF OF LEMMA 1

Consider two random users i and j . Let random vectors $\mathbf{V}_i = (V_{i,1}, V_{i,2}, \dots, V_{i,d})$ and $\mathbf{V}_j = (V_{j,1}, V_{j,2}, \dots, V_{j,d})$ denote their positions in the diffusion space respectively.

Then, the distance L could be written as $L = \sum_{k=1}^d (V_{i,k} - V_{j,k})^2$, i.e., the sum of their squared distance in each dimension. Recall that the covariance matrix of the Gaussian distribution is an identity matrix. Then, the position of users in each dimension is also Gaussianly distributed and independent of each other. Thus, we only need to derive the distribution in each dimension and find out the distribution of the summation over d dimensions.

As we know, $V_{i,k}$ and $V_{j,k}$ follow the Gaussian distribution $N(\mu_k, \sigma^2)$. Then, their difference $V_{i,k} - V_{j,k}$ follows the Gaussian distribution of $N(0, 2\sigma^2)$. Note that $(V_{i,k} - V_{j,k})/(\sqrt{2}\sigma)$ is a standard normal distribution. Thus, by the definition of χ^2 distribution, we obtain that $\frac{L}{2\sigma^2} = \sum_{k=1}^d \left(\frac{V_{i,k} - V_{j,k}}{\sqrt{2}\sigma}\right)^2$ follows the χ_d^2 distribution, which coincides with Gamma distribution $\Gamma(\frac{d}{2}, \frac{1}{2})$. By changing the variable, it is easy to see that L follows $\Gamma(\frac{d}{2}, \frac{1}{4\sigma^2})$.

B PROOF OF LEMMA 2

By Lemma 1, we see that the p.d.f. of the squared distance L between two random users could be written as

$$f_L(z) = \begin{cases} z^{\frac{d}{2}-1} e^{-\frac{z}{4\sigma^2}}, & z \geq 0, \\ 0, & z < 0. \end{cases} \quad (29)$$

Then, the probability that $L \leq r$ could be expressed as

$$p(r, \sigma^2) = \int_0^r \frac{z^{\frac{d}{2}-1} e^{-\frac{z}{4\sigma^2}}}{(4\sigma^2)^{\frac{d}{2}} \Gamma(\frac{d}{2})} dz \stackrel{(a)}{=} \frac{1}{\Gamma(\frac{d}{2})} \int_0^{\frac{r}{4\sigma^2}} z^{\frac{d}{2}-1} e^{-z} dz,$$

where equality (a) is derived by changing variable z with $\frac{z}{4\sigma^2}$. To see the monotonicity of $p(r, \sigma^2)$, we next calculate its partial derivatives. Note that $p(r, \sigma^2)$ is a definite integral with variable upper limit, we have

$$\begin{aligned} \frac{\partial}{\partial r} p(r, \sigma^2) &= \frac{1}{4\sigma^2 \Gamma(\frac{d}{2})} \left(\frac{r}{4\sigma^2}\right)^{\frac{d}{2}-1} e^{-\frac{r}{4\sigma^2}}, \\ \frac{\partial}{\partial \sigma^2} p(r, \sigma^2) &= \frac{-r}{4\sigma^4 \Gamma(\frac{d}{2})} \left(\frac{r}{4\sigma^2}\right)^{\frac{d}{2}-1} e^{-\frac{r}{4\sigma^2}}. \end{aligned}$$

Since $r > 0$, it is easy to see that $\frac{\partial}{\partial r} p(r, \sigma^2) > 0$ and $\frac{\partial}{\partial \sigma^2} p(r, \sigma^2) < 0$, which implies that $p(r, \sigma^2)$ increases with r and decreases with σ^2 .

C PROOF OF LEMMA 4

We adopt the same notations as the proof of Lemma 3 (such as $\{Y_k\}_{k=1}^m, \{Z_k\}_{k=1}^s$). And we define $Y = \max_{k=1, \dots, m} Y_k$ and $Z = \max_{k=1, \dots, s} Z_k$.

Then, the p.d.f. of Y (resp. Z) is $f_Y(y) = mG^{m-1}(y)g(y)$ (resp. $f_Z(z) = sG^{s-1}(z)g(z)$). Note that Y and Z are independent since $\{Y_k\}_{k=1}^m, \{Z_k\}_{k=1}^s$ are independent of each other. The probability

that a user will be infected by I_1 is

$$\begin{aligned} P(Y < Z) &= \int_0^\infty f_Y(y) \left[\int_0^y f_Z(z) dz \right] dy \\ &= \int_0^\infty f_Y(y) \left[\int_0^y n[1 - G(z)]^{s-1} g(z) dz \right] dy \\ &= \int_0^\infty m[G(y)]^{m-1} g(y)[G(y)]^s dy \\ &= \frac{m}{m+s} [G(y)]^{m+s} \Big|_0^\infty \\ &= \frac{m}{m+s}. \end{aligned}$$

Recall the definitions of m and s , we obtain the probability of information about I_1 being the latest $\frac{X_1(t)}{X_1(t) + X_2(t)}$.