

To Live or To Die: Encountering Conflict Information Dissemination over Simple Networks

Jie Wang

Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC 27606
Email: jwang50@ncsu.edu

Wenye Wang

Department of Electrical and Computer Engineering
North Carolina State University, Raleigh, NC 27606
Email: wwang@ncsu.edu

Abstract—In an era of networks in which any individual is connected with one another, such as Internet of Things (IoT) and Online Social Networks (OSNs), the networks are evolving into complex systems, carrying a huge volume of information that may provoke even more. An interesting, yet challenging question is *how such information dissemination evolves, that is, to continue or to stop*. Specifically, we aim to find out the aftermath of epidemic spreading via individuals and conflicting information dissemination. From a holistic, networking view, it is impossible to take every aspect into accounts for complex networks toward these questions. Therefore, we establish a *Susceptible-Infectious-Cured (SIC)* propagation model to examine two simple network topologies, *clique and star*, in terms of *extinction time and half-life time* of information under controllable, epidemic dynamics. For a network of size n , both theoretical and numerical results suggest that extinction time and half-life time are $O(\frac{\log n}{n})$ for clique networks, and $O(\log n)$ for star networks. More interestingly, given an initial network state I_0 , the extinction time is constant ($O(1)$) for cliques, and $O(\log I_0)$ for stars; while the half-life time is $O(\log \frac{1}{I_0})$ for both clique and star networks, respectively. In addition, we developed a method to estimate the conditional infection count distribution, which indicates the scope of information dissemination.

I. INTRODUCTION

After the Boston bombing incident on April 15th, 2013, Reddit users started an online suspect hunt, which led to the wrong person [1], but the rumor information spread rapidly on Twitter, causing huge problems to the wrongly accused. On April 18th, the police released information about the correct suspect in public media, then the rumor soon died out on Reddit and Twitter [2]. A clear observation is that the injection of *positive* or *negative* information seeds and individual spreading over social and communication networks have led to unexpected outcomes. Accordingly, an interesting, yet challenging question is *how such information dissemination evolves, that is, to continue or to stop*. In other words, we have little understanding on the aftermath of epidemic spreading via individuals and conflicting information dissemination. Specifically, we may want to know how much time it takes the negative information to disappear. Such phenomenon can also be seen in the competition of rivalry products, especially when the advantage of one product is dominant. For instance, MySpace was quite popular in 2006, when Facebook just started to open registration to the public. However, due to the

dominant advantage of the latter, the injection of the information (open registration) later overturned the OSN market. Traffic to MySpace shrunk quickly with a diminishing user group, while Facebook eventually took over the market and became the most popular OSN.

In both cases, the injected “positive” information incurred the extinction of an on-going/potential epidemic of the “negative” information. Note that “positive” here does not necessarily mean good, but implying that the information is with a dominant credibility or advantage over the “negative” one. This is an interesting phenomenon of epidemic information dissemination because its broad impact on whether the information continues to propagate or not, that is, *to live or to die*.

The epidemic spreading effect by individuals is not a brand new problem. It was initially developed by epidemiologists to study the spreading behavior of contagions, and emerged as a powerful tool to study information propagation over networks in the context of communication networks, power grids, and biological networks, considering the resemblance between contagions spread and information propagation among individuals. Pastor-Satorras and Vespignani analyzed the lifetime of computer virus on Internet and found the absence of the epidemic threshold on scale-free networks [3]. In the field of power grids security [4], epidemics were used to model cascading failures, *e.g.*, a counter-intuitive finding is that the lowest loads under attack are more harmful than that with higher loads [5]. In studies of biological networks, epidemics are used to predict and prevent the outbreak of transmissible diseases. Kao *et al.* identified the risk of foot-and-mouth disease in 2001 when the catastrophic epidemic occurred [6], with data traces of network of livestock movements in Great Britain. The critical dynamics of a single epidemic in a network include the epidemic threshold [7], [8], the dependency on network topology [9] and epidemic detection [10]. Among them, Ganesh *et al.* derived bounds for the expected extinction time of a Susceptible-Infected-Susceptible (SIS) epidemic in terms of the network size [9], discovered a dichotomy of behaviors with respect to the relation between the effective infection rate and spectral radius of the network, and examined the bound for networks with different topologies, including star, hypercube, complete graph *etc.* Regarding the same SIS model, Krishnasamy *et al.* studied the same bound for bounded susceptibility [11] setting.

Though the epidemic propagation of conflicting information remains an unsolved problem, there have been studies in the problem of competing epidemics, which also involves multiple epidemics. Originated in rivalry products advertising,

This work is partially supported by NSF CNS1423151 and ARO W911NF-15-2-0102.

different propagation models have been developed to describe this problem. Based on whether the network is viewed as a homogeneous population, or a complex ensemble of correlated individuals, different levels of abstractions are adopted. The former is commonly referred to as *population dynamics*, in which the assumption of *well-mixing* population applies, while the latter as *network dynamics* [12], in which the large state space impedes theoretic analysis. Despite the propagation models and levels of abstraction, research questions revolve around the time-invariant (steady) state of the network. For instance, whether the stronger epidemic prevails, if there is a co-existence condition in terms of infection rates, and measures to intentionally suppress one of the two. Among those, N. Pathak *et al.* found the invariant distribution of the size of the infected set of a network under multiple epidemics, with a generalized linear threshold model [13]. Prakash *et al.* [14] proved that the epidemic with stronger propagation properties will eliminate the weaker one in steady state under the SI_1I_2S model (population dynamics) with complete mutual immunity, by analyzing the stability of the system differential equations. Based on the same model, Beutel *et al.* [15] found the co-existence condition for similar SI_1I_2S model with partial mutual immunity. Lin *et al.* [16] considered from both population and network dynamics perspectives, and utilized Mean-Field Approximation (MFA) to enable asymptomatic and numerical analysis for the latter. Newman [17] found the coexistence threshold for two competing epidemics on networks with known degree distributions, under Susceptible-Infected-Recovered (SIR) model. In addition to the aforementioned non-interfering propagation models, interactive models that allow interference during the propagation process, are also developed to study influence maximization [18] strategy, quarantine policy [19] and other counter-measures against viral epidemics. Among those propagation models, both linear threshold model and SIS model allow the belief of an individual to switch back and forth, which is not suitable to study this problem. The reason behind this is perceptions of individuals will change permanently by the more compelling information. The SIR model does not capture the epidemic propagation property of the antidote information in this problem. Interactive models assume interference frequently applied to entities in the system, which is difficult to achieve considering the scale of current OSNs. Therefore, we are motivated to develop a new *information propagation model* in the study of consequences of conflicting information propagation.

On the avenue of OSNs, for almost every post, there is always conflicting, either *positive* or *negative* information. Even breaking news from authoritative agencies may contain misleading information so that corrections have to be released later to control the misleading information from spreading, let alone all kinds of rumor and anti-rumor posts. Ridiculous as it may sound, a rumor #AppleWave#, that iPhones can be charged in a microwave oven after updating to iOS8, emerged on Twitter [20] when iOS8 was introduced, and claimed several victim phones. Later, posts of burnt iPhones went viral and eventually eliminated the rumor. Now recall the prior example of the Reddit suspect hunt. Though the number of individuals involved may differ, similarities reside in the pattern of propagation: First, a piece of negative information spreads in the network, prone to cause an epidemic, which resembles the behavior of an infectious virus. Then another

piece of conflicting positive information with higher credibility is injected, functioning as replicative antidotes. Since the two pieces are conflicting, even contradictory on the same subject, they share the same group of potential recipients (individuals in the network), and the same propagation media (communication links). As a preliminary research, we follow the conventional rough assumption that one contact from the virus/antidote is powerful enough to change the *state* of an individual, *i.e.* the perception of that individual. To better describe the process of an individual adopting an idea, a more delicate multi-state model, like [16], is needed for further in-depth study.

The open question is how the conflicting information propagates, specifically, *how much time it takes the negative information (or, virus) to disappear and how the perceptions (state) of individuals change during this process*, both about transient states of the network. To answer these questions, we propose a new Susceptible-Infected-Cured (SIC) propagation model to study the epidemic behaviors of conflicting information propagation. For the first question, with both theoretical and numerical analysis, we found that the distinction time of the negative information remains constant in a clique network, while for a star network it is $O(\log I_0)$ with respect to the initial infection count I_0 . The half-life time is $O(\frac{1}{I_0})$ and $O(\log \frac{1}{I_0})$ for clique and star networks respectively. With respect to the size of the network n , distinction time and half-life time in a clique network are both $O(\frac{\log n}{n})$, and both $O(n)$ for that of a star network. For the second question, we provide a method to acquire the conditional infection count distribution.

The rest of the paper is organized as follows. First we introduce the SIC propagation model in Section II. Then we describe the new metrics for the investigation of transient properties of information propagation under the SIC model, followed by the derivation of the upper bounds for simple networks in Section III. Then we estimate the conditional infection count with simulation results in Section V. Finally we conclude this paper in Section VI.

II. SIC INFORMATION PROPAGATION MODEL

In this section, we elaborate in detail the terminologies, assumptions and definitions in the SIC propagation model.

A. Preliminary SIC Propagation Model

The SIC information propagation model determines the dissemination process of an *infectious* virus under the presence of an *infectious* antidote in a network. By *infectious*, we mean both the virus and the antidote can propagate along edges together with data transmitted by individuals.

1) *Network $\mathcal{G}(\mathcal{V}, \mathcal{E})$* : The *network* is described as an undirected, connected and static graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, whose vertex set \mathcal{V} corresponds to nodes in a network, and edge set \mathcal{E} corresponds to links. Since \mathcal{G} is *undirected*, edge $e(i, j) = e(j, i)$, $i, j \in \mathcal{V}$. For any vertex $v \in \mathcal{V}$, its *neighborhood* $\mathcal{N}(v) := \{u \in \mathcal{V} | (u, v) \in \mathcal{E}\}$ is defined as all the vertices that v can exchange data with. Being *static* implies the system is closed, hence size of the network $n(\mathcal{G}) = |\mathcal{V}|$ (noted as n for the ease of notation) remains the same during the whole time.

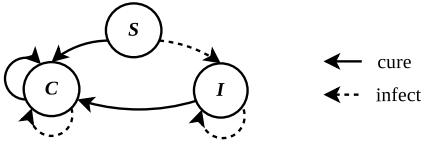


Fig. 1: State Transition Diagram of SIC Epidemic Model.

2) *Propagation Model*: Let x be a virus that is able to infect every vertex $v \in \mathcal{V}$, and a_x be the antidote that is able to cure any vertex infected by x and immune susceptible vertices from x . For such x and a_x , each vertex is associated with a *state*, which can be *susceptible* to x , *infected* by x or *cured* by a_x . Let r.v. $X_v^x(t) : \Omega \rightarrow \Lambda$, where $\Lambda = \{0, 1, -1\} \subset \mathbb{R}$, denote the state of vertex $v \in \mathcal{V}$ regarding virus x at time t . $X_v^x(t)$ takes value in $\{0, 1, -1\}$, with 0, 1, -1 corresponding to susceptible, infected and cured state respectively.

Vertex v is said to be *susceptible* to virus x *iff.* it has not been infected by virus x nor cured by antidote a_x . Susceptible vertices will remain in susceptible state until an *infection* action or a *curing* action is conducted upon it. Vertex v is said to be *infected* by virus x (or a copy of x is passed to v) *at time* t_i *iff.* $\lim_{t \rightarrow t_i^+} X_v^x(t) = 1$ and $\lim_{t \rightarrow t_i^-} X_v^x(t) = 0$. Vertex v is said to be *cured at time* t_c *iff.* $\lim_{t \rightarrow t_c^+} X_v^x(t) = -1$ and $\lim_{t \rightarrow t_c^-} X_v^x(t) \geq 0$.

Due to infections and curings, the state of an individual changes over time. Fig. 1 shows the possible state transitions of a vertex in the network. The transition from susceptible (S) to infected (I) is caused by an infect action, while that from S or I to cured (C) are caused by a curing action. Note that there is no transition from C to I or C to S , indicating cured is the only absorbing state. Since the network is connected, $\lim_{t \rightarrow \infty} X_i^x(t) = -1, \forall i \in \mathcal{V}$.

At any time t , the vertex set \mathcal{V} decomposes into three disjoint subsets according to states: the susceptible set $\mathcal{S}^x(t) = \{v \in \mathcal{V} : X_v^x(t) = 0\}$, the infected set $\mathcal{I}^x(t) = \{v \in \mathcal{V} : X_v^x(t) = 1\}$ and the cured set $\mathcal{C}^x(t) = \{v \in \mathcal{V} : X_v^x(t) = -1\}$, *i.e.* $\mathcal{S}^x(t) \cup \mathcal{I}^x(t) \cup \mathcal{C}^x(t) = \mathcal{V}$. When infected, vertex v will pass copies of virus x to its susceptible neighbors $\mathcal{N}_S^x(t, v) = \mathcal{N}(v) \cap \mathcal{S}^x(t)$. When cured, vertex v will pass copies of a_x to its non-cured neighbors $\mathcal{N}_{NC}^x(t, v) = \mathcal{N}_S^x(t, v) \cup \mathcal{I}^x(t, v) = (\mathcal{N}(v) \cap \mathcal{I}^x(t)) \cup (\mathcal{N}(v) \cap \mathcal{S}^x(t))$. At time t , for an infected vertex $u \in \mathcal{I}^x(t)$, let $t + s_u^x(v)$ denote the time that vertex u infects one of its susceptible neighbors, $v \in \mathcal{N}_S^x(t, v)$. Similarly for a cured vertex $u \in \mathcal{C}^x(t)$ and one of its susceptible or infected neighbor $v \in \mathcal{N}_{NC}^x(t, v)$, let $t + s_u^{a_x}(v)$ denote the time that vertex u pass a copy of antidote to vertex v and cures (or immunes) it.

With these terminologies, we describe two assumptions in the SIC propagation model. (i) Virus x can not infect a vertex that has already been exposed to a_x . Or $\forall \epsilon > 0, \mathbb{P}[X_i^x(t+\epsilon) = 0 | X_i^x(t) = -1] = 0, \mathbb{P}[X_i^x(t+\epsilon) = 0 | X_i^x(t) = 1] = 0$ and $\mathbb{P}[X_i^x(t+\epsilon) = 1 | X_i^x(t) = -1] = 0$. This is later referred to as the *asymmetrical immunity* assumption; (ii) For any vertex u , random intervals $\{s_u^x(v)\}_{v \in \mathcal{N}_S^x(t, v)}$ and $\{s_u^{a_x}(v)\}_{v \in \mathcal{N}_{NC}^x(t, u)}$ are two groups of r.v.'s satisfying (1) pairwise independent; (2) exponentially distributed with a time-invariant parameter.

Remark 1: Assumption (ii) may seem inconsistent with observations. For example, during midnight to dawn, people

exchange information less frequently than during day time, resulting in a slower dissemination of both virus and antidote. However, without the assumption of time homogeneity, explicit theoretical analysis will be difficult, if at all possible. Therefore, we assume the distribution parameter does not change over time, as in most research on epidemics [9], [11], [14].

Those two assumptions allow us to define *virulence* (also known as the *infection rate*) of a virus and the *curing rate* of an antidote, to measure the propagation intensity of a virus or an antidote, or equivalently, how frequently a specific virus or an antidote is exchanged in contacts.

DEFINITION 1: For vertex $u \in \mathcal{I}^x(t)$, $v \in \mathcal{N}_S^x(t, v)$, the *virulence* $\beta_{u,v}^x$ of virus x on edge $e(u, v)$, is defined as

$$\beta_{u,v}^x := \lim_{t \rightarrow 0^+} \frac{\mathbb{P}(s_u^x(v) \leq t)}{t}.$$

Remark 2: Consider a network composed of only two connected vertices $\mathcal{V} = \{u, v\}$. Given $X_u^x(t) = 0$ and $X_v^x(t) = 1$, $\mathbb{P}(X_v^x(t+\Delta t) = 1 | X_v^x(t) = 0) = \Delta t \cdot \beta_{u,v}^x \cdot \mathbb{1}_{\{X_u^x(t)=1\}} + o(\Delta t)$, which indicates that virulence $\beta_{u,v}^x$ describes how much the probability of a susceptible vertex v gets infected during interval Δt changes, due to its contact with vertex u , so $\beta_{u,v}^x$ is also known as the *infection rate* of virus x on edge $e(u, v)$. With assumption (ii), it is clear that $s_u^x(v) \sim \text{Exp}(\beta_{u,v}^x)$, and $\mathbb{E}(s_u^x(v)) = \frac{1}{\beta_{u,v}^x}$. For a certain virus x , $\beta_{u,v}^x$ is an innate attribute of edge $e(u, v)$, describing the ability of virus x to infect others (or the propagation speed) via edge $e(u, v)$. Note that when simulating this process in discrete time, quantity $\beta_{u,v}^x$ actually equals to the probability $\mathbb{P}(s_u^x(v) \leq t)$ that vertex v is infected by u during one time step.

DEFINITION 2: For vertex $u \in \mathcal{C}^x(t)$, $v \in \mathcal{N}_{NC}^x(t, v)$, the *curing rate* $\gamma_{u,v}^x$ of antidote a_x on edge $e(u, v)$, is defined as

$$\gamma_{u,v}^x := \lim_{t \rightarrow 0^+} \frac{\mathbb{P}(s_u^{a_x}(v) \leq t)}{t}.$$

DEFINITION 3: *Infection count*, denoted by $I(t)$, is defined as the size of the infected set $\mathcal{I}^x(t)$ at time t .

$$I^x(t) = |\mathcal{I}^x(t)| = \sum_{v \in \mathcal{V}} \mathbb{1}_{\{1\}}(X_v^x(t)).$$

DEFINITION 4: *Cured count*, denoted by $C^x(t)$, is defined as the size of the cured set $\mathcal{C}^x(t)$ at time t .

$$C^x(t) = |\mathcal{C}^x(t)| = \sum_{v \in \mathcal{V}} \mathbb{1}_{\{-1\}}(X_v^x(t)).$$

For a fixed time t , $I^x(t)$ and $C^x(t)$ are r.v.'s on measurable space $(\Omega^n, 2^{\Omega^n}, \mathbb{P})$. Allowing t to take value in domain Γ will result in two random processes $\{I^x(t)\}_{t \in \Gamma}$ and $\{C^x(t)\}_{t \in \Gamma}$, driven by the defined SIC dynamics. From the state transition diagram, it is easy to see that $\lim_{t \rightarrow \infty} X_i^x(t) = -1, \forall i \in \mathcal{V}$, the system will stabilize at all-cured state, that is, $\lim_{t \rightarrow \infty} C^x(t) = n$. We then follow the convention and define $X_v^x(t) = \lim_{t \rightarrow t^+} X_v^x(t)$, then $I^x(t)$ and $C^x(t)$ are both right-continuous. $C^x(t)$ is a counting process, satisfying (i) $C^x(t) \geq 0$; (ii) $C^x(t) \in \mathbb{N}_0$ ($:= \mathbb{N} \cup \{0\}$); (iii) non-decreasing.

From this section on, unless indicated otherwise, we suppress x in $X_v^x(t)$, $\mathcal{S}^x(t)$, $\mathcal{I}^x(t)$, $\mathcal{C}^x(t)$, and write $X_v(t)$, $\mathcal{S}(t)$, $\mathcal{I}(t)$, $\mathcal{C}(t)$ instead, when no confusion is raised.

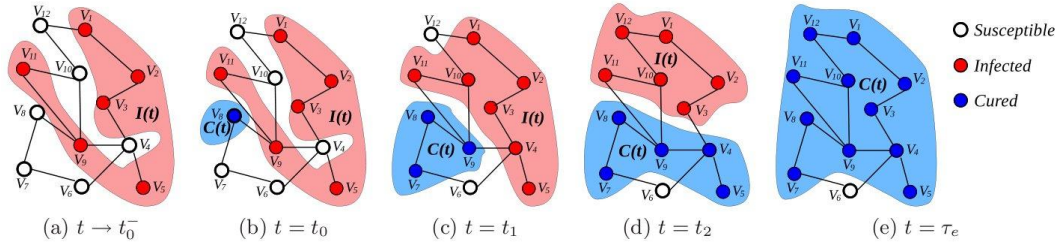


Fig. 2: An Example of a Simple Network under SIC Dynamics.

B. Information Propagation under SIC Dynamics

Fig. 2 shows an example of a simple network with 12 vertices under an SIC dynamics. As shown in Fig. 2(a), initially before antidote is injected in the network, infected set $\{v_1, v_2, v_3, v_5, v_9, v_{11}\}$ (the red shaded area) includes $I_0 = 6$ infected vertices (colored in red). At t_0 , one unit of antidote is given to vertex v_8 , and cures it immediately, so $\mathcal{C}(t_0) = \{v_8\}$, as blue shaded area in Fig. 2(b). The infected set remains unchanged at t_0 , $\lim_{t \rightarrow t_0^-} \mathcal{I}(t) = \mathcal{I}(t_0) = \{v_1, v_2, v_3, v_5, v_9, v_{11}\}$. Fig. 2(c-d) illustrate how states of vertices evolve as virus and antidotes propagate. When the network is observed at $t = t_1$, the cured vertex v_8 has passed a copy of antidote to its susceptible neighbor v_7 and infected neighbor v_9 respectively, so that both of them are in cured state by t_1 . From t_1 to t_2 , the cured set $\mathcal{C}(t_2)$ continues to grow while the infected set $\mathcal{I}(t-2)$ is forced to shrink. Eventually the virus is extinct at τ_e , and $\mathcal{I}(\tau_e)$ diminished to ϕ .

To quickly validate our model, we analyzed the search interests data¹ of OSN mySpace and Facebook during 2007 to 2011, in which the former is modeled as the virus, while the latter as the antidote, considering their precedence and relative popularity. Search interest indicates the number of users. As shown in Fig. 3, normalized search interests of mySpace and Facebook (bars) can be well approximated by an SIC epidemic (dashed lines) on a fully connected network.

Since both OSNs are designed for the same group of users, and people have limited energy to spend on OSNs, it is reasonable to assume that if one is interested in one of the two, he/she will not stay active on the other. Moreover, one's choice is influenced by both one's friends and the size of current user group. In addition, due to a better experience and increasing popularity of Facebook, people who chose Facebook will probably not switch back to mySpace. However, new users, who hasn't tried either, are prone to pick up the first product they encounter. These characteristics coincide with the asymmetrical immunity assumption of our SIC dynamics. On the other hand, influence of these two OSNs are openly presented to the whole network (current users or potential users) by the Internet. Therefore, the network can be considered as fully connected. Let t_0 be the beginning of year 2007, when mySpace had already gained some attention (mySpace started in 2003), while Facebook has just expanded from universities to the general public (Facebook started on September 26,

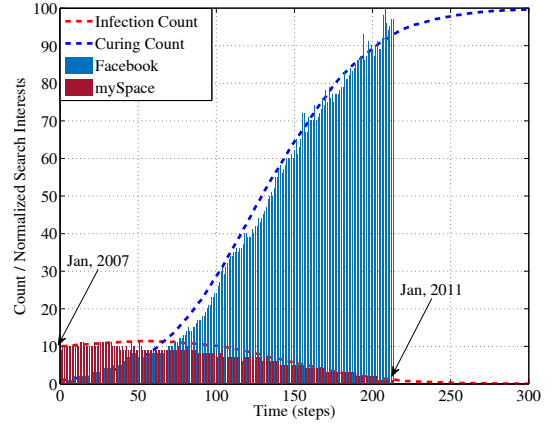


Fig. 3: Search Trends of mySpace and Facebook.

2006). The normalized search interest in January 2007 for mySpace is $I_0 = 10$, while that for Facebook is $C_0 = 1$. Note that (i) we are only interested in this time period, because it captures the process of Facebook taking over the market; (ii) we normalized the search counts over the search count of Facebook in January 2011 when the portion of mySpace fell below 1%, *i.e.* can be treated as extinct.

III. TRANSIENT PROPERTIES OF INFORMATION PROPAGATION

Though different metrics are defined in previous research of epidemics, most of them focused on the steady state, that is, the invariant final size of the infected set. The most commonly used technique is to describe the system with differential equations, then find the equilibrium by determining the eigenvalues of the Jacobian matrix. As commented in [12], it is worth noticing that transient behavior of information propagation via epidemics has received much less attention. Therefore, instead of the asymptotic behavior, we aim to study the transient properties of conflicting information propagation in simple networks, with the proposed SIC model. Understanding these properties can clearly demonstrate epidemic behaviors of information propagation in a network.

In an SIC dynamic on the network \mathcal{G} with the initial states as in the previous section, we focus on how size of the infected set $I(t)$ and cured set $C(t)$, evolve over time t .

A. Extinction Time and Half-Life Time

For a network \mathcal{G} under an SIC dynamic, virus x is the information to be propagated, and antidote a_x is the coun-

¹Search interests data are available on Google Trends, <https://www.google.com/trends/>. Search interests are measured by the search counts during a certain period of time (in this case a week), and are further normalized by the search counts of Facebook in January 2011 to mimic a static network.

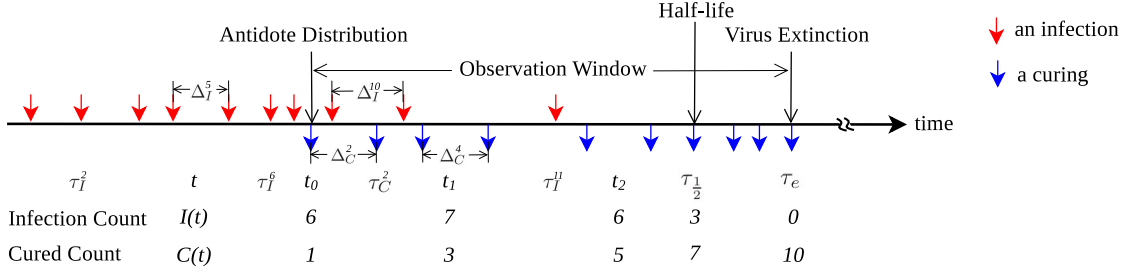


Fig. 4: Extinction Time and Half-life Time under SIC Dynamics.

termeasure, which exhibits an *obstructive* effect on the virus epidemic process. To observe the impacts of positive and negative information propagation, we define two new metrics, *extinction time* and *half-life time* as follows.

DEFINITION 5: In an SIC dynamics of virus x and antidote a_x , the *extinction time* of virus x , denoted as τ_e , is defined as the time that the infected set $\mathcal{I}(t)$ becomes empty, subjecting to the constraint that it is the first time after t_0 .

$$\tau_e := \inf\{t > t_0 : \mathcal{I}(t) = \phi\}.$$

Event $\{\mathcal{I}(t) = \phi\}$ is equivalent to $\{I(t) = 0\}$, so τ_e is a r.v. on measurable space $(\Omega^n, 2^{\Omega^n}, \mathbb{P})$. τ_e takes value in (t_0, ∞) . From the state transition diagram and the fact that \mathcal{G} is connected, the infected set will surely become empty if we observe long enough time, so $\mathbb{P}(\tau_e < \infty) = 1$. Without loss of generality, let $t_0 = 0$, $I(t_0) = I_0$. All the events we discuss afterwards happen in the observation window $[0, \tau_e]$.

DEFINITION 6: For an infection count random process $\{I(t)\}_{t \in [0, \tau_e]}$ in an SIC dynamic, the *half-life time* of the virus epidemic, denoted as $\tau_{\frac{1}{2}}$, is defined as the last time that event $\{I(t) \geq \frac{1}{2}I_0\}$ happens in $[0, \tau_e]$.

$$\tau_{\frac{1}{2}} := \sup\{t \in [0, \tau_e] : I(t) \geq \frac{1}{2}I_0\}.$$

The term *half-life* is from chemical kinetics, which describes the decay of discrete entities. Let t_{max} denote the time that infection count $I(t)$ reaches its peak. In SIC dynamics, though $I(t)$ may not be monotone in $[0, t_{max}]$ or $[t_{max}, \tau_e]$, the trend of infection count in these two intervals are increasing and decreasing in general. Note that unlike in chemical kinetics, where half-life is the mean, we define half-life as the actual time that event $\{I(t) \geq \frac{1}{2}I_0\}$ happens for the last time. Similarly as τ_e , we can show that $\tau_{\frac{1}{2}} : (\Omega^n, 2^{\Omega^n}, \mathbb{P}) \rightarrow [0, \tau_e]$ is a r.v. and we denote its mean as $\mathbb{E}[\tau_{\frac{1}{2}}]$. Clearly, we always have $0 \leq t_{max} < \tau_{\frac{1}{2}}$. Intuitively, $\tau_{\frac{1}{2}}$ is the time required for the infection count function $I(t)$ to drop below half of its initial value ($I_0/2$) and never exceed $I_0/2$ again, so it reflects how fast the infection count decreases under the antidote epidemic, i.e. the efficiency of the antidote dissemination.

Fig. 4 illustrates the implication of *extinction time* and *half-life time* for the same example in Fig. 2, where a red arrow corresponds to an infection, and a blue one represents a curing event. At $\tau_{\frac{1}{2}}$, the infection count of the system drops to 3 ($= \frac{1}{2}I(0)$), and never exceeds 3 again, which means the infection epidemic has been restricted to a limited region, or equivalently, the virus epidemic is under control. At τ_e , the infected set is empty and the virus has been eliminated.

B. How Many Nodes Have Received the Data?

Ideally, we are interested in the distributions of infection count $I(t)$ and cured count $C(t)$, which provide sufficient information to understand the trend of the information propagation. Recall counting process $\{C(t)\}_{t \geq 0}$. The evolution process of information propagation can be described by occurrences of infection and curing events. Hence if all events are recorded in the order of their occurrences, we have a path from (C_0, I_0) to $(n, 0)$ on the $[0, n] \times [0, n] \subset \mathbb{R}^2$ plane. As a result, the temporal evolution of the network state can be described by a discrete time Markov chain $\{(C_k, I_k)\}_{k \in \mathbb{N}_0}$ on state space $B = \{(c, i) \in [0, n] \times [0, n] \mid c + i \leq n, c, n \in \mathbb{N}_0\}$, which captures the trend of an SIC dynamic in a time-compressing manner. This process consists of two correlated components: $\{C_k\}_{k \in \mathbb{N}_0}$ and $\{I_k\}_{k \in \mathbb{N}_0}$. Since the former is a counting process, we assume that the network controller (administrator) can somehow trace the cured count. Then the state of the network is fully described by the distribution of the infection count, so we define the *conditional infection count distribution* to measure the trend of SIC dynamics.

DEFINITION 7: Given Markov chain $\{(C_k, I_k)\}_{k \in \mathbb{N}_0}$ on B . Let $\{B_c\}_{c=0}^n$ be a partition of state space B , where $B_c = \{(c, i) \mid i \in [0, n - c]\}$. Let $\tau_{B_c} = \min\{k \in \mathbb{N} \mid (C_k, I_k) \in B_c\}$ denote the hitting time of set $B_c \in B$. The *conditional infection count distribution*² $\pi_c(i)$, given the curing count at that time equals to c , is defined as

$$\pi_c(i) := \frac{\mathbb{P}_0(I_{\tau_{B_c}} = i)}{\mathbb{P}_0(\tau_{B_c} < \infty)},$$

where $\mathbb{P}_0(\cdot) = \mathbb{P}(\cdot \mid (C_0, I_0) = (c_0, i_0))$ denotes the probability that the Markov chain $\{(C_k, I_k)\}_{k \in \mathbb{N}_0}$ visits a state, on condition that it starts from the initial state $(c_0, i_0) \in B$.

Note that the denominator $\mathbb{P}_0(\tau_{B_c} < \infty)$ is used for normalization, such that $\pi_c(\cdot)$ is a probability measure. Then

$$\begin{aligned} \pi_c(i) &= \frac{\mathbb{P}_0(I_{\tau_{B_c}} = i)}{\sum_{j=0}^{n-c} \mathbb{P}_0(I_{\tau_{B_c}} = j)} \\ &= \frac{\sum_{k=c-C_0}^{2(n-C_0)-I_0} \mathbb{P}_0(I_k = i, C_k = c)}{\sum_{j=0}^{n-c} \sum_{k=c-C_0}^{2(n-C_0)-I_0} \mathbb{P}_0(I_k = j, C_k = c)}. \end{aligned} \quad (1)$$

Taken two simple network topologies, complete graph (clique) and star, into consideration, this paper studies the transient behavior of the SIC dynamics through perspectives of both time and count distribution.

² $\tau_{B_c} = \infty$ when $c < C_0$. In this case, let $\pi_c(i) = 0, \forall i$. Then π_c is a zero measure, but the states involved are only finite, and easy to discuss. Otherwise it is always a probability measure, i.e. a distribution.

IV. BOUNDED INFORMATION DISSEMINATION IN SIMPLE NETWORKS

In this paper, we are interested in the bounded properties of information propagation by controlling epidemic dynamics. For example, to stop or cease dissemination of *negative* information in a network, *positive* information propagation can be carried out by injecting an antidote into the network in the SIC model. Here we study the upper bounds of extinction time and half-life time for two simple network topologies, *clique* and *star* networks³. We follow the definitions of network models that are commonly used in epidemiology and assume that the network is homogeneous in both cases. This means both infection rate β and curing rate γ are constant on every edge of the network.

A. Extinction and Half-Life Time in Clique Networks

Consider first the SIC dynamics on $\mathcal{G} = K_n$, the clique of n vertices. $\forall v_i, v_j \in V(K_n), i \neq j, e(i, j) \in E(K_n)$ and hence $|E(K_n)| = n(n-1)/2$. With full connectivity, clique is used to describe a homo-mixing and well-connected group. A real world example of network structure that resembles a clique can be a household, a neighborhood or a community, within which people are acquainted with each other. Pairwise adjacency in a clique allows us to bound the extinction time τ_e and half-life $\tau_{\frac{1}{2}}$ with the initially distributed antidotes C_0 and initial infection count I_0 , via results from deterministic epidemic model and Continuous Time Markov Chain (CTMC).

THEOREM 1: For an SIC epidemic with curing rate γ , in action on a clique with n nodes, suppose at $t = 0$, the initial infection count is I_0 and antidotes are disseminated to C_0 vertices. The extinction time τ_e and half-life time $\tau_{\frac{1}{2}}$ can be bounded above by the following:

$$\mathbb{E}(\tau_e) \leq \frac{1}{\gamma n} \left(2 + \ln \frac{(n-1)(n-C_0)}{C_0+1} \right) \quad (2)$$

$$\mathbb{E}(\tau_{\frac{1}{2}}) \leq \frac{1}{\gamma n} \left(2 + \ln \frac{(n-1)(n-1-I_0/2)}{I_0/2+2} \right); \quad (3)$$

when $C_0 \geq 2$, we also have

$$\mathbb{E}(\tau_e) \leq \frac{2}{\gamma(n-1+C_0)} \ln(n-C_0) \quad (4)$$

$$\mathbb{E}(\tau_{\frac{1}{2}}) \leq \frac{2}{\gamma(n-1-I_0/2+C_0)} \left(1 + \ln \frac{n-C_0}{I_0/2+1} \right). \quad (5)$$

Proof: $\{C(t)\}$ is a CTMC with transition rate $\gamma c(n-c)$,

$$c \rightarrow c+1 \quad \text{at rate} \quad \gamma c(n-c) \quad \forall C_0 \leq c \leq n-1,$$

and hence $\mathbb{E}(\Delta_C^c) = \frac{1}{\gamma c(n-c)}$, $\forall C_0 \leq c \leq n-1$, since r.v. Δ_C^c is Exponentially distributed with parameter $\gamma c(n-c)$.

Then by Lemma 2 (see Appendix), we have

$$\begin{aligned} \mathbb{E}(\tau_e) &\leq \mathbb{E} \left(\sum_{c=C_0}^{n-1} \Delta_C^c \right) = \sum_{c=C_0}^{n-1} \frac{1}{\gamma c(n-c)} \\ &= \frac{1}{\gamma n} (\mathcal{H}_{n-1} - \mathcal{H}_{C_0} + \mathcal{H}_{n-C_0}) \\ &\leq \frac{1}{\gamma n} (1 + \ln(n-1) - \ln(C_0+1) + 1 + \ln(n-C_0)) \\ &= \frac{1}{\gamma n} \left(2 + \ln \frac{(n-1)(n-C_0)}{C_0+1} \right), \end{aligned}$$

where \mathcal{H}_n is the *Harmonic Number*, and $\ln(n+1) < \mathcal{H}_n \leq \ln(n) + 1$. Similarly, the same method can be used to derive the upper bound of half-life time for clique K_n .

$$\begin{aligned} \mathbb{E}(\tau_{\frac{1}{2}}) &\leq \mathbb{E} \left(\sum_{c=C_0}^{n-1-I_0/2} \Delta_C^c \right) = \sum_{c=C_0}^{n-1-I_0/2} \frac{1}{\gamma c(n-c)} \\ &= \frac{1}{\gamma n} (\mathcal{H}_{n-1-I_0/2} - \mathcal{H}_{I_0/2+1} + \mathcal{H}_{n-1}) \\ &\leq \frac{1}{\gamma n} (1 + \ln(n-1-I_0/2) - \ln(I_0/2+2) + 1 + \ln(n-1)) \\ &= \frac{1}{\gamma n} \left(2 + \ln \frac{(n-1)(n-1-I_0/2)}{I_0/2+2} \right). \end{aligned}$$

Similarly, for $C_0 \geq 2$, we can obtain the presented results. ■

Theorem 1 provides upper bounds for extinction time and half-life time for information dissemination in clique networks. It indicates that as size of the network n grows, both extinction time and half-life time of the same virus with the same initial condition will grow with $O(\frac{\log n}{n})$. More interestingly, as the initial infection count I_0 increases, the extinction time remains the same, while half-life time decreases as $O(\log \frac{1}{I_0})$. As for the impacts of the propagation parameters, both quantities are not dependent on the infection rate β , but are $O(\frac{1}{\gamma})$.

Fig. 5 and 6 show the extinction time and half-life time with simulation results (shown as solid lines) and the upper bounds derived with different initial conditions, respectively. For each group of input parameters, the simulation has been run for 1000 times to get the ensemble mean. As shown in Fig. 5(a) and (b), the extinction time of an SIC epidemics on a clique depends on the number of antidote units (C_0) distributed at time 0, and remains unchanged as initial infection count I_0 increases. However, there is a noticeable gap between the simulation results and the bound when $I_0 \leq 5$. The reason behind this is that when the initial infection count is relatively small, the probability that the virus epidemic never takes off is large, especially when the curing rate is high, for example, the gap is larger in 5(a) when $\gamma = 0.001$ than 5(b) when $\gamma = 0.0005$. As shown in Fig. 6, the tendencies of half life of the virus are correctly described by the bounds as initial infection count I_0 increases. However, due to the granularity issue of the simulation, the upper bound for condition $C_0 = 1$ is slightly less than the actual half life, as can be seen in 6(a) and 6(b), the blue dashed lines are under the blue solid ones. To be more specific, our analysis is on continuous time domain, and thus it is impossible to have two events happen at an exact time point, while in simulation it has to be in discrete time and to better simulate the continuity of time, the number of time steps needed will be huge, resulting in a lengthy simulation.

³The lower bounds are not studied, partially because the upper bounds are of particular interest to the boundary effects of data, and partially because the proposed SIC model is insufficient to study the lower bounds.

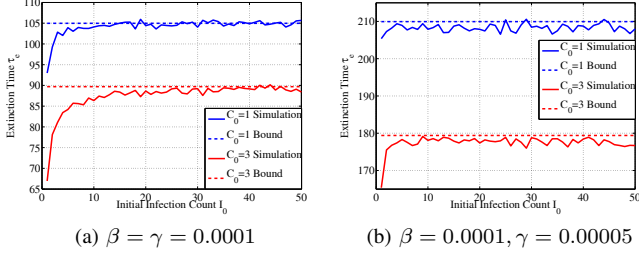


Fig. 5: Extinction Time of Information in a Clique Network.

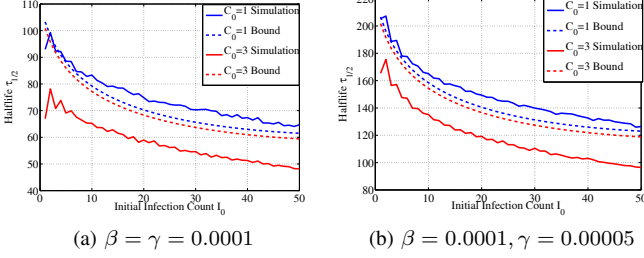


Fig. 6: Half-life Time of Information in a Clique Network.

B. Extinction and Half-Life Time in Star Networks

Now we consider the case of a same SIC dynamic in a *star* network with one hub v_1 and $n - 1$ peripheral vertices.

THEOREM 2: For an SIC epidemic over a *star* network in which the initial infection count satisfy $I_0 \geq 2$, the expected extinction time and half-life follow the following bounds:

$$\mathbb{E}(\tau_e) < \frac{1}{\gamma} \left[\frac{1}{C_0} + 1 + \ln(I_0 - 1) \right] + \frac{\beta}{\gamma} \sum_{k=1}^{\infty} \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k \quad (6)$$

$$\mathbb{E}(\tau_{\frac{1}{2}}) < \frac{1}{\gamma} \left[\frac{1}{C_0} + 1 + \ln\left(\frac{I_0 - 1}{I_0/2}\right) \right] + \frac{\beta}{\gamma} \sum_{k=1}^{\infty} \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k. \quad (7)$$

Proof: $I_0 \geq 2$ indicates that the hub is infected at $t_0 = 0$. Let r.v. τ_0 denote the time when the hub is cured, then $I(t)$ is non-increasing after τ_0 . Let Δ denote the number of new infections until τ_0 , T_k denote the time that the cured hub has disseminated antidotes to k infected peripheral vertices. Equivalently, T_k is the maximum value of k i.i.d exponential r.v.'s with parameter γ . So $\mathbb{E}(T_k) = \frac{1}{\gamma} \mathcal{H}_k$, where $\ln(k+1) < \mathcal{H}_k \leq 1 + \ln k$. Then we write the expected extinction time, $\mathbb{E}(\tau_e) = \mathbb{E}(\tau_0) + \mathbb{E}(T_{\Delta+I_0-1})$, where $\tau_0 \sim \text{Exp}(C_0\gamma)$, with PDF $f_{\tau_0}(t) = C_0\gamma e^{-C_0\gamma t}$, therefore $\mathbb{E}(\tau_0) = \frac{1}{C_0\gamma}$. Therefore

$$\begin{aligned} \mathbb{E}(T_{\Delta+I_0-1}) &= \mathbb{E}(\mathbb{E}(T_{\Delta+I_0-1}|\Delta)) \\ &= \frac{1}{\gamma} \sum_{k=0}^{n-C_0-I_0} \mathcal{H}_{k+I_0-1} \cdot \mathbb{P}(\Delta = k) \\ &< \frac{1}{\gamma} \sum_{k=0}^{n-C_0-I_0} (1 + \ln(k + I_0 - 1)) \cdot \mathbb{P}(\Delta = k). \end{aligned}$$

Let $g(k) = \ln(k + I_0 - 1)$, which means $g(\cdot)$ is concave in $[1, n - C_0 - I_0]$. Then

$$\begin{aligned} \mathbb{E}(T_{\Delta+I_0-1}) &< \frac{1}{\gamma} + \frac{1}{\gamma} \mathbb{E}(g(\Delta)) = \frac{1}{\gamma} + \frac{1}{\gamma} \mathbb{E}(\mathbb{E}(g(\Delta)|\tau_0)) \\ &= \frac{1}{\gamma} + \frac{1}{\gamma} \int_0^{\infty} \mathbb{E}(g(\Delta)|\tau_0 = t) \cdot f_{\tau_0}(t) dt \\ &\stackrel{\text{Jensen}}{\leq} \frac{1}{\gamma} + \frac{1}{\gamma} \int_0^{\infty} g(\mathbb{E}(\Delta|\tau_0 = t)) \cdot f_{\tau_0}(t) dt. \end{aligned}$$

Now consider r.v. $\mathbb{E}(\Delta|\tau_0)$. Since the infections of different susceptible vertices are mutually independent, it is not difficult to find that Δ obeys Binomial distribution $B(n - C_0 - I_0, 1 - e^{-\beta t})$, given fixed τ_0 , that is, $\mathbb{P}(\Delta = k|\tau_0 = t) = \mathbb{P}(k \text{ out of } n - C_0 - I_0 \text{ vertices are infected before } t) = \binom{n-C_0-I_0}{k} (e^{-\beta t})^{n-C_0-I_0-k} (1 - e^{-\beta t})^k$. Therefore $\mathbb{E}(\Delta|\tau_0 = t) = (n - C_0 - I_0) \cdot (1 - e^{-\beta t})$. And we have

$$\mathbb{E}(T_{\Delta+I_0-1}) \leq \frac{1}{\gamma} + \frac{1}{\gamma} \int_0^{\infty} g((n - C_0 - I_0) \cdot (1 - e^{-\beta t})) \cdot f_{\tau_0}(t) dt.$$

By using function $g(k)$ and derivation, we obtain

$$\begin{aligned} \mathbb{E}(T_{\Delta+I_0-1}) &\leq \frac{1}{\gamma} (1 + \ln(I_0 - 1)) \\ &\quad + \frac{\beta}{\gamma} \sum_{k=1}^{\infty} \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{E}(\tau_e) &= \mathbb{E}(\tau_0) + \mathbb{E}(T_{\Delta+I_0-1}) \\ &< \frac{1}{\gamma} \left[\frac{1}{C_0} + 1 + \ln(I_0 - 1) \right] \\ &\quad + \frac{\beta}{\gamma} \sum_{k=1}^{\infty} \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k \end{aligned}$$

Similarly, when considering the half-life time, at τ_0 , there are $\Delta + I_0 - 1$ peripheral vertices that are in the infected state. Thus $T_{\Delta+I_0/2-1} = t$ is when there are $I_0/2$ infected vertices left in the peripheral area, or equivalently, $\Delta + I_0/2 - 1$ out of $\Delta + I_0 - 1$ i.i.d exponential r.v.'s are less than t . Let $g'(k) = \ln \frac{k+I_0-1}{I_0/2}$. Then we have

$$\begin{aligned} \mathbb{E}(\tau_{\frac{1}{2}}) &= \mathbb{E}(\tau_0) + \mathbb{E}(T_{\Delta+I_0/2-1}) \\ &\stackrel{\text{Jensen}}{<} \frac{1}{C_0\gamma} + \frac{1}{\gamma} + \frac{1}{\gamma} \int_0^{\infty} g'((n - C_0 - I_0) \cdot (1 - e^{-\beta t})) \cdot f_{\tau_0}(t) dt \\ &= \frac{1}{\gamma} \left[\frac{1}{C_0} + 1 + \ln\left(\frac{I_0 - 1}{I_0/2}\right) \right] \\ &\quad + \frac{\beta}{\gamma} \sum_{k=1}^{\infty} \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k. \end{aligned}$$

Theorem 2 allows us to estimate the extinction time τ_e and half-life $\tau_{\frac{1}{2}}$ of the virus in a star network. Take τ_e for example.

Though the sum of the series $\left\{ \frac{1}{C_0\gamma + k\beta} \cdot \left(\frac{n - C_0 - I_0}{n - C_0 - 1} \right)^k \right\}_{k \in \mathbb{N}}$ is hard to calculate, as $k \rightarrow \infty$, the term of this series goes to zero faster than an exponential decay. Therefore, the sum can be estimated by summing up first few (finite many) terms. In addition, it reveals an interesting property of the star topology, that the hub is of great importance because all the copies of antidote or virus are passed through the hub except for the initial distribution of antidote at $t_0 = 0$. As the size of

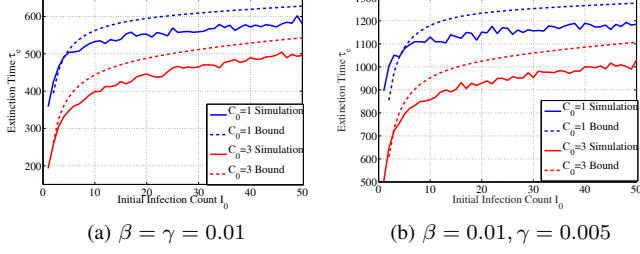


Fig. 7: Extinction Time of Information in a Star Network.

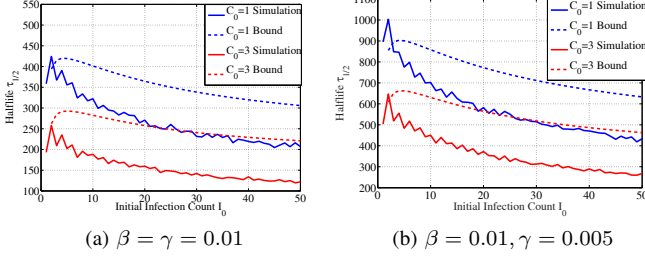


Fig. 8: Half-life Time of Information in a Star Network.

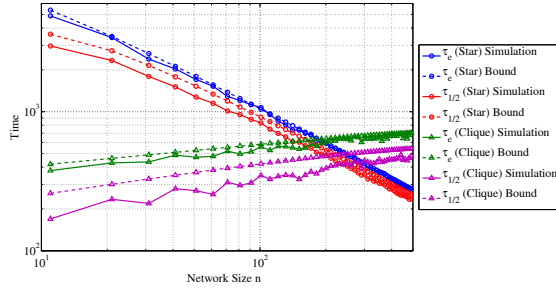


Fig. 9: Impact of Network Size n on Extinction Time and Half-life Time in both Clique and Star Topology.

network n increases, the extinction time and half-life time grow as $O(\log n)$, due to the second term in Eq. (6) and (7). And as the initial infection count I_0 increases, the extinction time grows as $O(\log I_0)$, while half-life time decreases, as $O(\log \frac{1}{I_0})$. Theorem 2 also indicates that both quantities are not dependent on the infection rate β , but are $O(\frac{1}{\gamma})$.

Fig. 7 and 8 show the extinction time and half-life of an SIC dynamics on a star with different initial conditions. As shown in Fig. 7, the extinction time of SIC dynamics on star topology grows slower than that of clique topology as initial infection count I_0 increases. It is worth noticing that even the rates (both infection and curing) are much higher than of clique topology, it takes much longer to eliminate the virus or force the infected count to be cut in half in star networks.

Fig. 9 shows the impact of network size n on τ_e and $\tau_{\frac{1}{2}}$ for both topologies. We fix $C_0 = 1$, $I_0 = 10$, and examined $\beta = \gamma = 0.01$ for star topology and $\beta = \gamma = 0.0001$ for clique topology, because otherwise the two temporal quantities will be dramatically apart and comparison will be difficult. Both the simulation results and the bounds indicate that τ_e and $\tau_{\frac{1}{2}}$ are $O(\frac{\log n}{n})$ for clique networks, while $O(\log n)$ for star networks.

V. INFECTION AND CURED COUNTS IN SIMPLE NETWORKS

Recall that we are interested in conditional infection count distribution $\{\pi_c\}_{c \in [0, n]}$, because they provide information to illustrate the transient behavior of the SIC epidemic in a network.

The transient behavior of chain $\{(C_k, I_k)\}_{k \in \mathbb{N}_0}$ depends on two factors: (i) one-step transition probability, which includes information of network topology, virus and antidote; (ii) initial state (C_0, I_0) . Let $p_{c,i}$, $r_{c,i}$, $q_{c,i}$ denote the probability that the next event is an infection, a curing of a susceptible vertex or a curing of an infected vertex respectively, when the current state of the chain is (c, i) . $p(\cdot)$, $r(\cdot)$ and $q(\cdot)$ are functions of infection rate matrix $\beta = \{\beta_{i,j}\}_{i,j \in \mathcal{V}}$, infection rate matrix $\gamma = \{\gamma_{i,j}\}_{i,j \in \mathcal{V}^4}$, infection count i , and cured count c .

We refer to $P = (p_{c,i})_{c,i \in B}$, $R = (r_{c,i})_{c,i \in B}$ and $Q = (q_{c,i})_{c,i \in B}$ as *partial transition (probability) matrices*⁵, since all together they uniquely determine the transition probability of the system. Based on those matrices, we give the following lemma to derive the conditional infection count distribution for a given initial state (C_0, I_0) .

LEMMA 1: For an SIC dynamics on a n -vertex network with partial transition matrix P , R and Q , and initial state (C_0, I_0) . Let $\mathbf{p}_c = \{p_{c,0}, p_{c,1}, \dots, p_{c,n}\}$, \mathbf{r}_c and \mathbf{q}_c denote the c -th column vector of matrices P , R and Q respectively. Let $\theta_{c,i} = \mathbb{P}_0(I_{\tau_{B_c}} = i)$ denote the probability that chain $(C_k, I_k)_{k \in \mathbb{N}_0}$ ever hits (c, i) , and vector $\boldsymbol{\theta}_c = \{\theta_{c,i} | i \in [0, n]\}$. Then

$$\boldsymbol{\theta}_{C_0} = \mathbf{p}_{C_0}^{Prod} \quad (8)$$

$$\boldsymbol{\theta}_{c+1} = [\boldsymbol{\theta}_c \circ \mathbf{r}_c + DSh(\boldsymbol{\theta}_c \circ \mathbf{q}_c)] \circ (\mathbf{1} + \mathbf{p}_c^{Prod}), \quad (9)$$

where $(\cdot \circ \cdot)$ denotes the Schur product of two vectors, $\mathbf{1}$ denotes an all-one vector of length $n + 1$, $\mathbf{p}_c^{Prod} = \{\prod_{i=0}^n p_{c,i}, \prod_{i=1}^n p_{c,i}, \dots, p_{c,n}\}$ and $DSh(\mathbf{q}_c) = \{q_{c,1}, q_{c,2}, \dots, q_{c,n}, 0\}$ denote the acyclic down shifting operation of vector \mathbf{q}_c .

Proof: There is only one way to reach state $(C_0, I_0 + k)$ from initial state (C_0, I_0) in k steps, that is, through k steps moving upward. Therefore, $\mathbb{P}((C_k, I_k) = (C_0, I_0 + k) | (C_0, I_0)) = p_{C_0, I_0} p_{C_0, I_0+1} \dots p_{C_0, I_0+k-1}$. Similarly, there is only one way to get to state $(C_0 + k, I_0 - k)$ from initial state (C_0, I_0) in k steps, that is, through k steps moving downward. So $\mathbb{P}((C_k, I_k) = (C_0 + k, I_0 - k) | (C_0, I_0)) = q_{C_0, I_0} q_{C_0+1, I_0-1} \dots q_{C_0+k-1, I_0-k-1}$. Then iteratively applying equation $\mathbb{P}((C_{k+1}, I_{k+1}) = (c, i) | (C_0, I_0)) = \mathbb{P}((C_k, I_k) = (c, i-1) | (C_0, I_0)) * p_{c, i-1} + \mathbb{P}((C_k, I_k) = (c-1, i) | (C_0, I_0)) * r_{c-1, i} + \mathbb{P}((C_k, I_k) = (c-1, i+1) | (C_0, I_0)) * q_{c, i-1}$ will yield the results. ■

By plugging $\mathbb{P}_0(I_{\tau_{B_c}} = i) = \theta_{c,i}$ in Eq. (1), the conditional infection count distribution $\pi_c(\cdot)$ can be determined, allowing us to see the evolving trend provided information of the cured count.

As an illustration of the conditional infection count distribution, we examined the aforementioned Facebook v.s. mySpace search interests data again with the estimated propagation

⁴Note that $\beta_{i,j} = \gamma_{i,j} = 0$, $\forall (i, j) \notin \mathcal{E}$

⁵ P , R and Q for a clique or star network are quite easy and hence omitted.

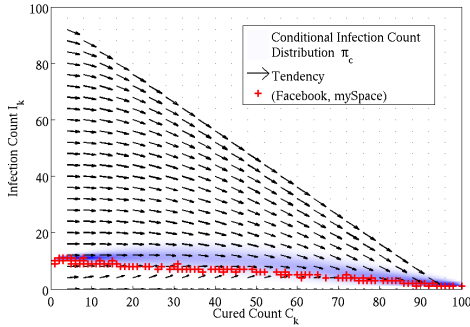


Fig. 10: Search Trends of mySpace v.s. Facebook in Field Plot.

parameters. We first calculate the partial transition matrices P , R and Q , and illustrate their influence on the trend with vectors (black arrows in Fig. 10) indicating expected state transition tendency. Then the achieved conditional infection count distribution is shown by the background color, where the darker the blue indicates the higher the probability. Data points with red '+' markers represent the real search data arranged in the format of (Facebook, mySpace). All the data points form a "trail" through the field, describing the evolution trend of the dynamics, where the background blue shade (i.e. the conditional infection count distribution data) indicates the area where this trial is most probable to occur. As can be seen, the estimated conditional infection count distribution coincide with the collected data points, proving the validity of the proposed SIC information propagation model.

VI. CONCLUSION

In this paper, we proposed an SIC information propagation model to study the impact of conflicting information dissemination. For simple networks with clique and star topology, we derived theoretical upper bounds for extinction time and half-life time of data, and validated the results with simulations. In addition, the conditional infection count distribution of an SIC dynamic is analyzed with given initial state and propagation parameters. We hope these will be useful in studying the conflicting information propagation problem.

VII. APPENDIX

LEMMA 2: Consider a n -vertex network $\mathcal{G}(\mathcal{V}, \mathcal{E})$ under an SIC dynamic, with initial infection count I_0 and cured count C_0 . We have the following upper bounds for extinction time and half-life time of the virus.

$$\tau_e \leq \sum_{c=C_0}^n \Delta_C^c, \quad (10)$$

$$\text{and } \tau_{\frac{1}{2}} \leq \sum_{c=C_0}^{n-1-I_0/2} \Delta_C^c. \quad (11)$$

Proof: Lemma 2 comes from the asymmetric immunity assumption. Since a cured vertex will never be infected again, the cured count $C(t)$ is non-decreasing, then the half-life time $\tau_{\frac{1}{2}}$ is bounded above by the spreading time of the antidote. Recall our definition of time interval between curings $\Delta_C^c = \tau_C^c - \tau_C^{c-1}$. At $t_1 = \inf\{t > 0 : C(t) = n - I_0/2\}$, $I(t_1) + S(t_1) = n - C(t_1) = I_0/2$, therefore $I(t) \leq I_0/2, \forall t \geq t_1$. ■

REFERENCES

- [1] H. Alexander, "Boston bomb: Reddit apologises for identifying wrong suspects." <http://www.telegraph.co.uk/news/worldnews/northamerica/usa/10012382/>, 2013.
- [2] D. Lee, "Boston bombing: How internet detectives got it very wrong." <http://www.bbc.com/news/technology-22214511>, 2013.
- [3] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical review letters*, vol. 86, no. 14, p. 3200, 2001.
- [4] M. Wei and W. Wang, "Greenbench: A benchmark for observing power grid vulnerability under data-centric threats," in *INFOCOM, 2014 Proceedings IEEE*, pp. 2625–2633, IEEE, 2014.
- [5] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the us power grid," *Safety Science*, vol. 47, no. 10, pp. 1332–1336, 2009.
- [6] R. Kao, L. Danon, D. Green, and I. Kiss, "Demographic structure and pathogen dynamics on the network of livestock movements in great britain," *Proc. of the Royal Society of London B: Biological Sciences*, vol. 273, no. 1597, pp. 1999–2007, 2006.
- [7] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, "Epidemic thresholds in real networks," *ACM Trans. Inf. Syst. Secur.*, vol. 10, pp. 1:1–1:26, Jan. 2008.
- [8] B. Prakash, D. Chakrabarti, N. Valler, M. Faloutsos, and C. Faloutsos, "Threshold conditions for arbitrary cascade models on arbitrary networks," *Knowledge and Information Systems*, vol. 33, no. 3, pp. 549–575, 2012.
- [9] A. Ganesh, L. Massoulié, and D. Towsley, "The effect of network topology on the spread of epidemics," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2, pp. 1455–1466 vol. 2, March 2005.
- [10] C. Milling, C. Caramanis, S. Mannor, and S. Shakkottai, "Detecting epidemics using highly noisy data," in *Proceedings of the Fourteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing, MobiHoc '13*, (New York, USA), pp. 177–186, ACM, 2013.
- [11] S. Krishnasamy, S. Banerjee, and S. Shakkottai, "The behavior of epidemics under bounded susceptibility," in *The 2014 ACM International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS '14*, (New York, NY, USA), pp. 263–275, ACM, 2014.
- [12] C. Nowzari, V. Preciado, and G. Pappas, "Analysis and Control of Epidemics: A survey of spreading processes on complex networks,"
- [13] N. Pathak, A. Banerjee, and J. Srivastava, "A generalized linear threshold model for multiple cascades," in *Proceedings of the 2010 IEEE International Conference on Data Mining, ICDM '10*, (Washington, DC, USA), pp. 965–970, IEEE Computer Society, 2010.
- [14] B. A. Prakash, A. Beutel, R. Rosenfeld, and C. Faloutsos, "Winner takes all: Competing viruses or ideas on fair-play networks," in *Proceedings of the 21st International Conference on World Wide Web, WWW '12*, (New York, NY, USA), pp. 1037–1046, ACM, 2012.
- [15] A. Beutel, B. A. Prakash, R. Rosenfeld, and C. Faloutsos, "Interacting viruses in networks: Can both survive?," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12*, (New York, USA), pp. 426–434, ACM, 2012.
- [16] Y. Lin, J. Lui, K. Jung, and S. Lim, "Modeling multi-state diffusion process in complex networks: Theory and applications," in *Signal-Image Technology Internet-Based Systems (SITIS), 2013 International Conference on*, pp. 506–513, Dec 2013.
- [17] M. E. J. Newman, "Threshold Effects for Two Pathogens Spreading on a Network," *Physical Review Letters*, vol. 95, p. 108701, Sept. 2005.
- [18] S. Bharathi, D. Kempe, and M. Salek, "Competitive influence maximization in social networks," in *Proceedings of the 3rd International Conference on Internet and Network Economics, WINE'07*, (Berlin, Heidelberg), pp. 306–311, Springer-Verlag, 2007.
- [19] C. Borgs, J. Chayes, A. Ganesh, and A. Saberi, "How to distribute antidote to control epidemics," *Random Structures & Algorithms*, vol. 37, no. 2, pp. 204–222, 2010.
- [20] Z. D. Boren, "Fake Advert for Apple Wave Convinces People to Charge Their iPhones in Microwave." <http://www.ibtimes.co.uk/fake-advert-apple-wave-convinces-people-charge-their-iphones-microwave-1467253>, 2014.