On Stochastic Confidence of Information Spread in Opportunistic Networks

Yoora Kim[†], *Member, IEEE*, Kyunghan Lee[‡], *Member, IEEE*, and Ness B. Shroff[§], *Fellow, IEEE*
[†]yrkim@ulsan.ac.kr, [‡]khlee@unist.ac.kr, [§]shroff.11@osu.edu

Abstract—Predicting spreading patterns of information or virus has been a popular research topic for which various mathematical tools have been developed. These tools have mainly focused on estimating the average time of spread to a fraction (e.g., α) of the agents, i.e., so-called average α-completion time $E(T_\alpha)$. We claim that understanding stochastic confidence on the time T_α rather than only its average gives more comprehensive knowledge on the spread behavior and wider engineering choices. Obviously, the knowledge also enables us to effectively accelerate or decelerate a spread. To demonstrate the benefits of understanding the distribution of spread time, we introduce a new metric $G_{\alpha,\beta}$ that denotes the time required to guarantee α completion (i.e., penetration) with probability β. Also, we develop a new framework characterizing $G_{\alpha,\beta}$ for various spread parameters such as number of seeders, contact rates between agents, and heterogeneity in contact rates. We apply our technique to a large-scale experimental vehicular trace and show that it is possible to allocate resources for acceleration of spread in a far more elaborated way compared to conventional average-based mathematical tools.

Index Terms—Information spread, CTMC analysis, spread time analysis, spread time distribution

1 Introduction

Spreading patterns of pandemics [1], computer viruses [2], and information [3] have been widely studied in various research disciplines including epidemics, biology, physics, sociology, and computer networks. In these disciplines, most studies have been devoted to characterizing spread behaviors toward a network of mobile agents including humans, vehicles, and mobile devices[1] over time. These studies can be classified into two groups based on their objectives. Interestingly, these objectives lie in opposite directions: delaying or accelerating spread. For the research that deals with biological or electronic viruses, how to slow down the spread has been the most important question to be answered. On the other hand, another set of research work for computer data or information distribution has pursued designing methodologies to speed up the spread.

Whatever the goals are, existing studies have relied on common mathematical techniques such as the branching process, mean-field approximation, and stochastic differential equations [4]. These techniques have mostly been developed to analyze the average behavior of spread under various epidemic models summarized in [5]. The epidemic models are first classified by whether agents are recoverable[2] or not, and then the recoverable cases are further classified by whether the agents become immune after recovery or susceptible again to infection.

Average behavior analysis successfully answers a question on how many nodes are infected (or informed) *on average* under a specific epidemic model after a time duration t from the emergence of a virus (or generation of information). This is often represented by $\mathrm{E}[N_t]$ where N_t denotes the number of infected nodes at time t. Aforementioned mathematical techniques have made various extensions to this analysis. In [6], it is identified how much network topologies affect the speed of virus spreading. The authors in [7] derived a closed-form equation for the critical level of virus infection rate that lets a virus persist in a network when the virus is recoverable with a certain rate. More realistic average spread behaviors of a virus with the heterogeneity inherent in human mobility patterns have been studied through simulations in [8]. In computer networks, [9] analyzed the average propagation behavior of code red worm in the Internet using measurement data from ISPs. In [3], the authors applied understanding on the average behavior of virus spread to information propagation in delay tolerant networks. Similarly, [2] analyzed the average spread behaviors of self-propagating worms in the Internet using branching process.

While there has been a plethora of work on average analysis, the problem of optimal allocation of resources to a network of a set of nodes for slowing down or speeding up spread has been under-explored. Specifically, higher order spread behaviors over time rather than average behaviors have not been well understood. Therefore, the right question to be answered should be what will be the distribution of the number of infected nodes at time t, which is equivalent to what will be the temporal distribution of the event that n nodes are infected. Characterizing the temporal distribution of spread allows us to guarantee the time for spread with desired probabilistic confidence, and it leads to having control knobs for allocating resources to a network with its own purpose of spread. However, understanding the temporal distribution involves non-trivial challenges since it requires to handle a huge dimension of diversity in contact events among nodes in a network.

In order to tackle the challenges involved in this paper, we propose a new analytical framework based on CTMC (continuous

A preliminary version of this paper was published in IEEE INFOCOM'13.
Y. Kim is with the Department of Mathematics, University of Ulsan, Korea.
K. Lee, the corresponding author, is with the School of Electrical and Computer Engineering, UNIST, Korea.
N. B. Shroff is with the Department of Electrical and Computer Engineering and the Department of Computer Science and Engineering, The Ohio State University, USA.

1. We will interchangeably use agents and nodes unless confusion arises.
2. A virus that is not recoverable can be considered to be identical to undeletable or unforgettable information.

time Markov chain), which enables us to fully characterize the temporal aspect of spread behaviors. For simplicity, throughout this paper, we put our emphasis on information distribution among intermittently meeting mobile nodes forming an opportunistic network, i.e., a mobile social network, but our results are easily applicable to general spread of epidemics or information. Our framework is capable of answering many intriguing engineering questions such as "What is the distribution of time for a network to have x% ($0 \leq x \leq 100$) penetration rate?" and "If x% penetration is aimed, when is the time to guarantee that level of penetration with y% ($0 \leq y \leq 100$) of confidence?". It can also answer a more fundamental question involving heterogeneity of nodes in a network, "Does heterogeneity help or hurt spreading?". We show the efficacy of our solution in answering these questions with the use of one of the largest experimental GPS (global positioning system) trace of taxies in Shanghai, China. Our simulation studies on the trace reconfirm that our framework is robust and provides opportunities to engineer the network in a far more elaborated way than existing average-based approaches.

The rest of the paper is organized as follows. In Section 2, we describe our system model and introduce relevant metrics. In Section 3, we develop a new analytical framework. Based on our framework, we characterize the temporal spread behavior and provide its engineering implications in Section 4. We present simulation studies using Shanghai taxi trace in Section 5. In Section 6, we discuss about available techniques that substantially reduce the computational complexity involved in our framework. We conclude our paper in Section 7.

2 Model Description

2.1 Overview of Epidemic Models

In epidemics, an individual is typically classified into either susceptible, infected, or removed (sometimes called recovered and immune) according to its infection status for a contagious disease [4]. A susceptible individual refers to the one who is not infected yet but is prone to be infected. An infected individual refers to the one who already got the disease and is capable of spreading it to susceptible individuals. A removed individual indicates the one who was previously infected but became immune to the disease. These three classifications are conventionally denoted by S, I, and R, respectively and induce SIS, SIR, and SI epidemic models and their variants. For instance, SIR or SIS represents that an individual will become immune or susceptible again after the cure. In this paper, we focus on the SI model in which once a susceptible individual is infected, it stays infected for the remainder of the epidemic process. The SI model fits particularly well with information spread in opportunistic networks, since once a data is delivered to an individual, it is considered that the information included in the data is delivered and recognized by the individual (i.e., permanently infected).

2.2 System Model

We consider a network consisting of N mobile nodes. We assume that mobile nodes in the network can be classified into K different types according to their mobility patterns and infection rates. Hence, all nodes in the same class are assumed to be homogeneous. Note that in our model, K can take any integer value from 1 to N. We denote the collection of the kth type of nodes as class k ($k = 1, \ldots, K$). Let N_k be the number of nodes in class k. Let $\boldsymbol{N} \triangleq (N_1, N_2, \ldots, N_K)$. Then, we have $||\boldsymbol{N}|| \triangleq \sum_k N_k = N$.[3]

The system under our consideration spreads information (or a packet or a virus) as follows. Initially, the information is delivered to a set of selected nodes, which we call *seeders*.[4] Whenever a seeder, say node a, meets a susceptible node not having the information yet, it spreads the information to the susceptible node with probability $\varphi_a \in (0, 1]$. Then, the susceptible node, say node b, successfully receives the information with probability $\psi_b \in (0, 1]$ and becomes infected (or informed). Once the susceptible node becomes infected, it stays infected for the remainder of the spread process and joins disseminating the information in a similar manner as the seeders. The spread process ends when all nodes in the network obtain the information. In our spread model, the probabilities φ_a and ψ_b can be interpreted as the infectivity and the susceptibility of nodes a and b, respectively. For instance, in the case of rumor propagation, φ_a quantifies the tendency of a person a to gossip, while ψ_b quantifies the receptive nature of a listener b to the rumor. For the case of packet forwarding in an opportunistic network, φ_a represents the probability that node a schedules to transmit a packet, and ψ_b represents the probability of successful packet reception at node b, which depends on various factors including contact duration, interference level, number of contending nodes, and wireless channel condition.

The stochastic characteristic of a pairwise contact process is a critical factor that determines the opportunity of spread and hence the temporal behavior of the spread process. In particular, the time duration between two consecutive contacts of a pair of nodes, called *pairwise inter-contact time*, is an essential factor. In the literature, it has been recently shown that the pairwise inter-contact time can be approximately modeled by an exponential random variable in many scenarios without having too much of discrepancy, e.g., [10]–[12]. In [10], exponential inter-contact patterns are validated experimentally using three different mobility data sets. In [12], the authors considered both the user availability process and the contact process to analyze the distribution of actual inter-transfer time (i.e., time duration between two consecutive available transfer opportunities). They prove that the inter-transfer time distribution becomes close to an exponential distribution, even when the underlying contact dynamics is non-Poisson, provided that the availability process and the contact process operate in a similar time scale. Thus, in this paper we assume that the pairwise inter-contact time between nodes a and b, denoted by $M_{a,b}$, follows an exponential distribution with rate $\lambda_{a,b}$ (> 0):

$$\mathrm{P}(M_{a,b} > t) = \exp(-\lambda_{a,b} t), \quad t \geq 0. \tag{1}$$

Suppose that node a is infected and node b is susceptible. Let $M_{a,b}^{\text{eff}}$ denote the time taken by node a to spread the information to node b. We call $M_{a,b}^{\text{eff}}$ the *infection time* throughout the paper. The infection time $M_{a,b}^{\text{eff}}$ can be obtained from (1) by taking the infectivity φ_a and the susceptibility ψ_b into account as follows:

$$\mathrm{P}(M_{a,b}^{\text{eff}} > t) = \exp(-\lambda_{a,b}^{\text{eff}} t), \quad t \geq 0, \tag{2}$$

where $\lambda_{a,b}^{\text{eff}} = \lambda_{a,b} \varphi_a \psi_b$. Since mobile nodes in the same class are assumed to be stochastically homogeneous, the infection rate $\lambda_{a,b}^{\text{eff}}$

3. Throughout this paper, we use a boldface font for a vector or a matrix notation. For a vector $\boldsymbol{V} = (V_k)$, we interchangeably use the notations $(\boldsymbol{V})_k$ and V_k to denote the kth element of the vector $\boldsymbol{V}$, and define the operation $||\boldsymbol{V}||$ as $||\boldsymbol{V}|| \triangleq \sum_k V_k = \sum_k (\boldsymbol{V})_k$.

4. Note that being selected as seeders can be of willing or unwilling. For instance, a seeder of a virus gets the virus unwillingly.

should be determined by the class index. Thus, we can rewrite the infection rate as $\lambda_{a,b}^{\text{eff}} = \xi_{k(a),k(b)}$, where the subscripts $k(a)$ and $k(b)$ denote the class indices of nodes a and b, respectively. For later use, we define an infection rate matrix $\mathbf{\Lambda}$ as

$$\mathbf{\Lambda} \triangleq \begin{pmatrix} \xi_{1,1} & \xi_{1,2} & \cdots & \xi_{1,K} \\ \xi_{2,1} & \xi_{2,2} & \cdots & \xi_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{K,1} & \xi_{K,2} & \cdots & \xi_{K,K} \end{pmatrix}.$$

Our spread model is general in that it covers a variety of scenarios from homogeneous to completely heterogeneous cases. For instance, when $K = 1$, the spread model reduces to the homogeneous case where any pair of nodes in the network has the same infection rate $\xi_{1,1}$ ($\triangleq \xi$). On the other hand, when $K = N$, it induces the completely heterogeneous case where each node uniquely forms a class. When $K = 2, \ldots, N-1$, the spread model is able to capture heterogeneity arising from multiple communities. In addition to heterogeneity, the spread model is capable of characterizing the impact of various spread parameters (e.g., level of contact rates and group-wide population size) on spread behaviors by varying the values of the rate matrix $\mathbf{\Lambda}$ and the class cardinality vector $\boldsymbol{N}$.

2.3 Performance Metrics

In this section, we describe our performance metrics in detail. Let $S_k(t)$ be the number of susceptible nodes in class k at time t and $I_k(t)$ be the number of infected nodes in class k at time t. Then, we have $S_k(t)+I_k(t) = N_k$ for all k and t. The first performance metric of our interest is α-completion time as defined below.

Definition 1 (α-completion time)**.** For $\alpha \in [0,1]$, let T_α denote the minimum time required to infect (i.e., penetrate) α fraction of the total population, i.e.,

$$T_\alpha \triangleq \inf\left\{t \geq 0 : \sum_{k=1}^{K} I_k(t) \geq \alpha N\right\}. \tag{3}$$

We call T_α the *α-completion time*.

The α-completion time T_α is closely connected with existing studies that have characterized the average number of infected nodes at time t (i.e., $\mathrm{E}\big[\sum_k I_k(t)\big]$) using various mathematical tools, because $\mathrm{E}[T_\alpha]$ is a dual of $\mathrm{E}\big[\sum_k I_k(t)\big]$. However, to better understand the spread behavior and to better design spread prevention or acceleration methods, it is essential to characterize the distribution of T_α beyond simply the mean. To this end, we introduce a new metric, called (α,β)-guaranteed time, as defined next.

Definition 2 ((α,β)-guaranteed time)**.** For $\alpha \in [0,1]$ and $\beta \in [0,1]$, let $G_{\alpha,\beta}$ denote the minimum time required to guarantee spread to α fraction of the total population with probability at least β. It is then given by

$$G_{\alpha,\beta} \triangleq \inf\big\{t \geq 0 : \mathrm{P}(T_\alpha \leq t) \geq \beta\big\}. \tag{4}$$

We call $G_{\alpha,\beta}$ the *(α,β)-guaranteed time*.

Note that the quantity β in (4) can be interpreted as the probability that the actual spread time T_α does not exceed the time $G_{\alpha,\beta}$. In that sense, $G_{\alpha,\beta}$ can be used to predict not only the range of spread time but also the confidence of the prediction: the higher we set the value of β, the greater the prediction gets confident. Thus, $G_{\alpha,\beta}$ facilitates avoiding underestimating or overestimating the required resources for spreading information to a network. The ratio $R_{\alpha,\beta}$ defined below describes just how much $\mathrm{E}[T_\alpha]$ underestimates or overestimates the spread time compared to the guaranteed time.

Definition 3 ((α,β)-guaranteed to average time ratio)**.** For $\alpha \in [0,1]$ and $\beta \in [0,1]$, let $R_{\alpha,\beta}$ denote the ratio of the guaranteed time to the average α-completion time, i.e.,

$$R_{\alpha,\beta} \triangleq \frac{G_{\alpha,\beta}}{\mathrm{E}[T_\alpha]}. \tag{5}$$

We call $R_{\alpha,\beta}$ the *(α,β)-guaranteed to average time ratio*.

Finally, we define the set of seeders in each class. Let $s_k \triangleq I_k(0)$ denote the number of seeders in class k, and let $\boldsymbol{s} \triangleq (s_k)$ be the seeder vector. If $||\boldsymbol{s}|| \geq \alpha N$, then we have a trivial result that $T_\alpha = 0$, $G_{\alpha,\beta} = 0$, and $R_{\alpha,\beta} = 1$ for any $\beta \in [0,1]$. Therefore, in the rest of the paper, we only consider the regime of $||\boldsymbol{s}|| < \alpha N$. For a given $\boldsymbol{s} = (s_k)$, the s_k number of seeders are chosen randomly in each class k, since mobile nodes in the same class are stochastically homogeneous.

3 Temporal Analysis Framework

In this section, we develop a framework for analyzing the performance metrics T_α, $G_{\alpha,\beta}$, and $R_{\alpha,\beta}$. We first explain the main ideas and technical approaches that lead us to present our temporal analysis framework. We then provide a step-by-step procedure for computing the performance metrics.

3.1 Technical Approach

According to Definition 1, we need the distribution of the total number of infected nodes $\sum_k I_k(t)$ as a function of time t. Directly solving it appears to be intractable, unless we know how the overall infected nodes are distributed to each class.[5] To this end, our approach is to rewrite the spread process as a joint *level-phase* process, where *level* tracks the total number of infected nodes, and *phase* supplements the level by specifying which sample instance realizes the level. Then, the α-completion time T_α becomes equivalent to the time taken by the level-phase process to reach the level $\lceil \alpha N \rceil$, where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. In our analysis, we characterize the joint level-phase process in terms of its temporal distribution. From the characterization, we can identify the distribution and moments of T_α, which in turn yield the (α,β)-guaranteed time $G_{\alpha,\beta}$ and the ratio $R_{\alpha,\beta}$.

3.2 Temporal Analysis Framework

We describe step-by-step procedures to obtain $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$ for a given set of system parameters. In this section, we only brief key results. The necessity of each step and technical derivations involved in each step are explained in detail in the following section.

5. In [13], it considers a similar problem in the context of content delivery time. The authors analyze a bound on the content delivery time to a certain portion of a network, while its exact distribution is left unsolved (See Lemma 5). The approach in Lemma 5 of [13] is mainly based on the edge expansion of a graph, but it is not applicable for capturing how the nodes having the content are distributed to each class.

Step 1 (Level-phase representation). In this step, we describe mathematical notions of level and phase. First, we define a set $\mathcal{F}$ in a K-dimensional space $(\mathbb{Z}_{\geq 0})^K$ by

$$\mathcal{F} \triangleq \big\{\boldsymbol{f}=(f_k)_{1\leq k\leq K} : 0 \leq f_k \leq N_k, 1 \leq ||\boldsymbol{f}|| \leq N\big\}. \quad (7)$$

The set $\mathcal{F}$ represents the collection of all feasible combinations of the infected node distribution to each class. For $\boldsymbol{f} \in \mathcal{F}$, let $L(\boldsymbol{f}) \triangleq ||\boldsymbol{f}||$. We call $L(\boldsymbol{f})$ the *level* of $\boldsymbol{f}$. According to the levels, we can partition the set $\mathcal{F}$ into N disjoint subsets $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_N$ to denote the set of possible node distributions at each level, as defined below:

$$\mathcal{F}_i \triangleq \big\{\boldsymbol{f} \in \mathcal{F} : L(\boldsymbol{f}) = i\big\}, \qquad 1 \leq i \leq N.$$

Without loss of generality, we assume that the elements in $\mathcal{F}_i$ are ordered in a lexicographical manner, and then denote the position of $\boldsymbol{f}$ in the subset $\mathcal{F}_{L(\boldsymbol{f})}$ by $J(\boldsymbol{f})$. We call $J(\boldsymbol{f})$ the *phase* of $\boldsymbol{f}$ and distinguish $\boldsymbol{f}$ from others in the same level.

Now, we define a map $\mathcal{X} : \mathcal{F} \to \mathbb{N} \times \mathbb{N}$ for $\boldsymbol{f} \in \mathcal{F}$ by

$$\mathcal{X}(\boldsymbol{f}) \triangleq \big(L(\boldsymbol{f}), J(\boldsymbol{f})\big).$$

That is, the map $\mathcal{X}$ transforms each $\boldsymbol{f}$ in $\mathcal{F}$ into a pair of level and phase. The resulting level-phase space is given by

$$\mathcal{U} = \big\{(i,j) : 1 \leq i \leq N, 1 \leq j \leq |\mathcal{F}_i|\big\}, \quad (8)$$

where $|\mathcal{F}_i|$ denotes the cardinality of the set $\mathcal{F}_i$. See Fig. 1 for an example. There is a one-to-one correspondence between $\boldsymbol{f}$ and $\mathcal{X}(\boldsymbol{f})$. Thus, the inverse function $\mathcal{X}^{-1} : \mathcal{U} \to \mathcal{F}$ exists, and we henceforth use the relations $\mathcal{X}(\boldsymbol{f}) = (i,j)$ and $\boldsymbol{f} = \mathcal{X}^{-1}(i,j)$ interchangeably.

Step 2 (Initial condition). In Step 2, we formalize the initial condition concerning the seeder vector $\boldsymbol{s} = (s_k)$. First, for a given $\alpha \in [0,1]$, we define

$$\mathcal{T}_\alpha \triangleq \big\{(i,j) \in \mathcal{U} : i \leq \lceil \alpha N \rceil - 1\big\},$$

which denotes the subspace of $\mathcal{U}$ with level lower than αN. Next, we define a vector $\boldsymbol{h}_\alpha(\boldsymbol{s})$ by

$$\boldsymbol{h}_\alpha(\boldsymbol{s}) \triangleq \big(\mathbf{1}_{\{\mathcal{X}(\boldsymbol{s})=(i,j)\}}\big)_{(i,j)\in\mathcal{T}_\alpha}, \quad (9)$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. Note that by definition, $\mathcal{X}(\boldsymbol{s})$ returns the level and phase of the seeder vector $\boldsymbol{s}$. Accordingly, $\boldsymbol{h}_\alpha(\boldsymbol{s})$ pinpoints the initial state in the space $\mathcal{T}_\alpha$ by reflecting the value of $\boldsymbol{s}$.

Step 3 (Transition rate matrix). In this step, we obtain a matrix $\boldsymbol{F}_\alpha$ that represents transition rate from the space $\mathcal{T}_\alpha$ to itself. The matrix $\boldsymbol{F}_\alpha$ is given in (6) at the bottom of this page. In (6), $\boldsymbol{R}_i$ is a $|\mathcal{F}_i| \times |\mathcal{F}_{i+1}|$ matrix representing transition rate from level i to level $i+1$; $\boldsymbol{e}$ is a column vector of ones; $diag(\boldsymbol{R}_i\boldsymbol{e})$ is a $|\mathcal{F}_i| \times |\mathcal{F}_i|$ diagonal matrix with the main diagonal $\boldsymbol{R}_i\boldsymbol{e}$; and $\mathbf{0}$ is a zero matrix.

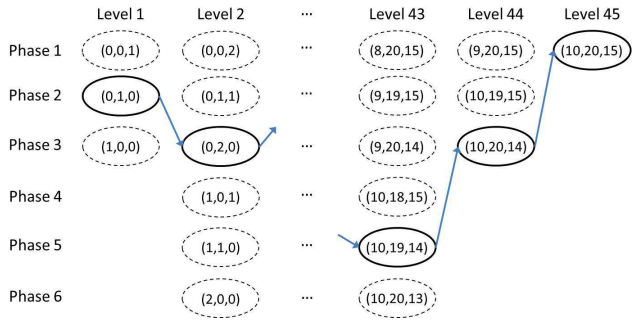


Fig. 1. Example of a sample path for a level-phase process: we set $K = 3$, $\boldsymbol{s} = (0,1,0)$, and $\boldsymbol{N} = (10,20,15)$. The level-phase representation of $\boldsymbol{s}$ and $\boldsymbol{N}$ is given by $\mathcal{X}(\boldsymbol{s}) = (1,2)$ and $\mathcal{X}(\boldsymbol{N}) = (45,1)$, respectively. Thus, the process starts from $(1,2)$ and ends up with $(45,1)$.

To compute the matrix $\boldsymbol{R}_i$, we define a function $\lambda : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ for $\boldsymbol{f}, \boldsymbol{g} \in \mathcal{F}$ by

$$\lambda(\boldsymbol{f},\boldsymbol{g}) \triangleq \begin{cases} (\boldsymbol{N}-\boldsymbol{f})_k(\boldsymbol{f}\boldsymbol{\Lambda})_k & \text{if } \boldsymbol{g} = \boldsymbol{f} + \boldsymbol{e}_k \text{ for some } k, \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

where $\boldsymbol{e}_k$ denotes the kth unit vector. The value $\lambda(\boldsymbol{f},\boldsymbol{g})$ indeed is the transition rate from $\boldsymbol{f}$ to $\boldsymbol{g}$ (which will be shown in Section 3.3). In particular, the case $\boldsymbol{g} = \boldsymbol{f} + \boldsymbol{e}_k$ accounts for the rate to infect one additional node in class k, provided that the current infected node distribution is given by $\boldsymbol{f}$. The other cases become null events, yielding zero transition rate. Then, the matrix $\boldsymbol{R}_i$ is obtained by using $\lambda(\cdot,\cdot)$ as

$$\boldsymbol{R}_i = \Big[\lambda\Big(\mathcal{X}^{-1}(i,j), \mathcal{X}^{-1}(i+1,j')\Big)\Big]_{1\leq j\leq|\mathcal{F}_i|, 1\leq j'\leq|\mathcal{F}_{i+1}|}. \quad (11)$$

Step 4 (Formulas for $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$). Once we have the vector $\boldsymbol{h}_\alpha(\boldsymbol{s})$ in (9) and the matrix $\boldsymbol{F}_\alpha$ in (6) at Steps 2 and 3, we can obtain the following Theorem on $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$.

Theorem 1. The CDF of T_α is given by

$$H_\alpha(t) \triangleq \mathrm{P}(T_\alpha \leq t) = 1 - \boldsymbol{h}_\alpha(\boldsymbol{s}) \exp(\boldsymbol{F}_\alpha t)\, \boldsymbol{e}. \quad (12)$$

The inverse function of $H_\alpha(\cdot)$ exists and gives $G_{\alpha,\beta}$ as

$$G_{\alpha,\beta} = H_\alpha^{-1}(\beta). \quad (13)$$

Also, the inverse matrix of $\boldsymbol{F}_\alpha$ exists and gives $R_{\alpha,\beta}$ as

$$R_{\alpha,\beta} = \frac{H_\alpha^{-1}(\beta)}{-\boldsymbol{h}_\alpha(\boldsymbol{s})(\boldsymbol{F}_\alpha)^{-1}\boldsymbol{e}}. \quad (14)$$

Proof: See Section 3.3. ■

We summarize a step-by-step guide for computing the performance metrics $G_{\alpha,\beta}$, and $R_{\alpha,\beta}$ in the following table.

$$\boldsymbol{F}_\alpha \triangleq \begin{bmatrix} -diag(\boldsymbol{R}_1\boldsymbol{e}) & \boldsymbol{R}_1 & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -diag(\boldsymbol{R}_2\boldsymbol{e}) & \boldsymbol{R}_2 & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -diag(\boldsymbol{R}_3\boldsymbol{e}) & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & -diag(\boldsymbol{R}_{\lceil\alpha N\rceil-2}\boldsymbol{e}) & \boldsymbol{R}_{\lceil\alpha N\rceil-2} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & -diag(\boldsymbol{R}_{\lceil\alpha N\rceil-1}\boldsymbol{e}) \end{bmatrix}. \quad (6)$$

Algorithm 1: Step-by-step guide for computing $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$

Input: set α, penetration rate
Input: set β, guaranteeness
Input: set $\boldsymbol{\Lambda}$, infection rate matrix
Input: set $\boldsymbol{s}$, seeder
Input: set K, number of classes
Input: set $\boldsymbol{N}$, number of nodes in each class

Output: (α,β)-guaranteed time $G_{\alpha,\beta}$
Output: (α,β)-guaranteed to average time ratio $R_{\alpha,\beta}$

Step 1. Generate level-phase space $\mathcal{U} \leftarrow (K, \boldsymbol{N})$ by Eq. (8)
Step 2. Compute initial condition $\boldsymbol{h}_\alpha \leftarrow (\mathcal{U},\alpha,\boldsymbol{s})$ by Eq. (9)
Step 3-1. Compute matrix $\boldsymbol{R}_i \leftarrow (\mathcal{U}, \boldsymbol{\Lambda}, \boldsymbol{N})$ by Eq. (11)
Step 3-2. Compute rate matrix $\boldsymbol{F}_\alpha \leftarrow \boldsymbol{R}_i$ by Eq. (6)
Step 4-1. Compute CDF $H_\alpha \leftarrow (\boldsymbol{h}_\alpha, \boldsymbol{F}_\alpha)$ by Eq. (12)
Step 4-2. Compute time $G_{\alpha,\beta} \leftarrow (\beta, H_\alpha)$ by Eq. (13)
Step 4-3. Compute ratio $R_{\alpha,\beta} \leftarrow (\boldsymbol{h}_\alpha,\boldsymbol{F}_\alpha,H_\alpha,\beta)$ by Eq. (14)

Remark 1. In Theorem 1, $\exp(\boldsymbol{F}_\alpha t)$ is a matrix exponential defined by $\exp(\boldsymbol{F}_\alpha t) \triangleq \sum_{n=0}^{\infty} \frac{1}{n!}(\boldsymbol{F}_\alpha t)^n$. When the diagonal entries of $\boldsymbol{F}_\alpha$ are all distinct, $\exp(\boldsymbol{F}_\alpha t)$ can be written rather simply in terms of its diagonal entries as follows:

$$\exp(\boldsymbol{F}_\alpha t) = \boldsymbol{V} \begin{bmatrix} e^{-\zeta_1 t} & 0 & 0 & \dots & 0 \\ 0 & e^{-\zeta_2 t} & 0 & \dots & 0 \\ 0 & 0 & e^{-\zeta_3 t} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & e^{-\zeta_M t} \end{bmatrix} \boldsymbol{V}^{-1},$$

where $M \triangleq |\mathcal{F}_1| + \ldots + |\mathcal{F}_{\lceil \alpha N \rceil - 1}|$ is the size of the matrix $\boldsymbol{F}_\alpha$; $\zeta_m\,(1 \le m \le M)$ is the absolute value of the mth diagonal entry of $\boldsymbol{F}_\alpha$; and $\boldsymbol{V}$ is a $M \times M$ matrix whose mth column vector is the eigenvector of $\boldsymbol{F}_\alpha$ corresponding to $-\zeta_m$ (which in fact is the eigenvalue of $\boldsymbol{F}_\alpha$). Hence, the tail distribution $\mathrm{P}(T_\alpha > t) = 1 - H_\alpha(t)$ in Theorem 1 becomes a weighted linear sum of exponential functions whose decay rates come from the main diagonal of $\boldsymbol{F}_\alpha$. The proof is given in Appendix C

Remark 2. When $K = 1$, our result in Theorem 1 is simplified as follows. (i) Phase j is always 1 for any level i, and the level-phase space is given by $\mathcal{U} = \{(i,1); i = 1,\ldots,N\}$. Hence, for any $(i,1) \in \mathcal{U}$, we have $\mathcal{X}^{-1}(i,1) = i$. (ii) The rate matrix $\boldsymbol{\Lambda}$ reduces to the constant $\boldsymbol{\Lambda} = \xi$. (iii) The submatrix $\boldsymbol{R}_i$ inside $\boldsymbol{F}_\alpha$ becomes a scalar representing transition rate from $(i,1)$ to $(i+1,1)$. By applying (i)-(iii) to the function (10) in Step 3, we obtain $\boldsymbol{R}_i = \lambda(i,i+1) = i(N-i)\xi$. As a result, T_α follows a phase-type distribution with simplified representation $(\boldsymbol{h}_\alpha(\boldsymbol{s}), \boldsymbol{F}_\alpha)$, which gives

$$T_\alpha = \sum_{i=||\boldsymbol{s}||}^{\lceil \alpha N \rceil - 1} Z_i, \tag{15}$$

where Z_i are independent exponential random variables with rates $\boldsymbol{R}_i = i(N-i)\xi$ and are the duration of time that the system stays in level i.

Remark 3. Our framework is designed to accommodate various spread factors systematically. First, Step 1 reflects network configuration parameters including the number of classes K, the size of each class N_k, and the network size N. Next, Step 2 reflects the selection of the seeder vector $\boldsymbol{s}$. Lastly, Step 3 reflects parameters for the infection rates such as infectivity, susceptibility, and contact rates using the rate matrix $\boldsymbol{\Lambda}$. Compared to our prior work in [14] where K-dimensional Markov chain is used for analysis, the idea of adopting level and phase in this paper enables us to handle the diversity in contact events and group formations without increasing the dimension of state space. Even for a completely heterogeneous network (i.e., the case $K = N$), we can analyze the system using a two-dimensional process. Moreover, the solution in Theorem 1 provides a unified structural form for any $K \in \mathbb{N}$, while the one in [14] has difficulty in expressing the case $K \ge 3$.

3.3 Proof of Theorem 1

To prove Theorem 1, we start with the easiest case $\alpha = 1$. We then extend our analysis to general cases $\alpha \in [0,1)$.

3.3.1 *Proof for the case* $\alpha = 1$

As outlined in Section 3.1, we employ a joint level-phase process for the analysis. In the following, we give a mathematical description of the level-phase process. Let $\boldsymbol{I}(t) \triangleq (I_1(t),\ldots,I_K(t))$. Then, we have $\boldsymbol{I}(t) \in \mathcal{F}$ (where $\mathcal{F}$ is defined in (7) at Step 1). Thus, we can apply the map $\mathcal{X}$ to $\boldsymbol{I}(t)$ as follows:

$$\mathcal{X}(\boldsymbol{I}(t)) = \big(L(\boldsymbol{I}(t)), J(\boldsymbol{I}(t))\big).$$

To simplify notation, let $L(t) \triangleq L(\boldsymbol{I}(t))$ and $J(t) \triangleq J(\boldsymbol{I}(t))$. Note that the level function $L(t)$ represents the total number of infected nodes at time t, and the phase function $J(t)$ specifies how $L(t)$ number of infected nodes are distributed to each class. We call $\{L(t); t \ge 0\}$ and $\{J(t); t \ge 0\}$ the *level process* and the *phase process*, respectively. In Lemma 1, we characterize the stochastic nature of joint level-phase process.

Lemma 1. The joint level-phase process $\{(L(t), J(t)); t \ge 0\}$ forms a two-dimensional CTMC.
Proof: See Appendix A. ∎

Since $\{\mathcal{X}(\boldsymbol{f}) : \boldsymbol{f} \in \mathcal{F}\} = \mathcal{U}$ (See Eq. (8)), the state space of the Markov chain in Lemma 1 is $\{\mathcal{X}(\boldsymbol{I}(t)) : \boldsymbol{I}(t) \in \mathcal{F}\} = \mathcal{U}$. Moreover, it has the following properties:

(P1) The level process $\{L(t); t \ge 0\}$ is a *counting process*.
(P2) The joint level-phase process $\{(L(t), J(t)); t \ge 0\}$ stops evolving when the level reaches N.

Hence, state transition of the Markov chain occurs only to the adjacent level from i to $i+1$, and then the Markov chain is eventually absorbed to level N. An example of a sample path is shown in Fig. 1. Thus, the state space $\mathcal{U}$ is decomposed into transient state space $\mathcal{T}$ and absorbing state space $\mathcal{A}$ as

$$\mathcal{T} \triangleq \big\{(i,j) \in \mathcal{U} : i < N\big\},$$
$$\mathcal{A} \triangleq \big\{(i,j) \in \mathcal{U} : i = N\big\}.$$

Note that $\mathcal{T}_\alpha$ defined at Step 2 reduces to $\mathcal{T}$ when $\alpha = 1$. Now we consider the α-completion time T_α for $\alpha = 1$. By Definition 1, T_1 is the time required for the level $L(t)$ to reach the state N. In terms of Lemma 1, T_1 is the time taken by the Markov chain $\{(L(t), J(t)); t \ge 0\}$ to be absorbed into the state space $\mathcal{A}$, i.e.,

$$T_1 = \inf\big\{t \ge 0 : (L(t), J(t)) \in \mathcal{A}\big\}. \tag{16}$$

Hence, we can understand T_1 as the *time until absorption*, which has been studied widely in Markov chain theory by e.g., Neuts [15]

and Bremaud [16]. In what follows, we summarize existing results that are directly related to our problem.

Lemma 2. Suppose that $\{X(t); t \geq 0\}$ is a CTMC with transient state space $\mathcal{E}$ and absorbing state space $\mathcal{E}^o$. Let $\boldsymbol{Q}$ be the infinitesimal generator of the CTMC $\{X(t); t \geq 0\}$. If $T \triangleq \inf\{t \geq 0 : X(t) \in \mathcal{E}^o\}$ denotes the time until absorption, then we have [15, Lemma 2.2.2]

$$\mathrm{P}(T > t) = \boldsymbol{h} \exp(\boldsymbol{F}t)\, \boldsymbol{e}, \tag{17}$$

where $\boldsymbol{h} \triangleq \big(\mathrm{P}(X(0) = x)\big)_{x \in \mathcal{E}}$ and $\boldsymbol{F} \triangleq \boldsymbol{Q}|_{\mathcal{E}\times\mathcal{E}}$. The nth moment of T $(n \in \mathbb{N})$ is given by [15, Eq. (2.2.7)]

$$\mathrm{E}[(T)^n] = n! \boldsymbol{h}(-\boldsymbol{F})^{-n} \boldsymbol{e}, \tag{18}$$

provided that the inverse of the matrix $\boldsymbol{F}$ exists.
Proof: Refer to [15]. ∎

Because of its importance, we call $\boldsymbol{F}$ in Lemma 2 the *fundamental matrix*. Lemma 2 says that both the initial distribution on transient state space and the fundamental matrix essentially govern the time until absorption. Accordingly, we first look into the initial distribution of the Markov chain $\{(L(t), J(t)); t \geq 0\}$ on transient state space $\mathcal{T}_1$. The initial value is determined by the seeder vector $\boldsymbol{s} = (s_k)$ and becomes $(L(0), J(0)) = \mathcal{X}(\boldsymbol{s})$. In the case when the number of seeders is selected in a deterministic manner as in our model, the initial distribution on $\mathcal{T}_1$ is obtained by

$$\big(\mathrm{P}(\mathcal{X}(\boldsymbol{s}) = (i,j))\big)_{(i,j)\in\mathcal{T}_1} = \big(\mathbf{1}_{\{\mathcal{X}(\boldsymbol{s})=(i,j)\}}\big)_{(i,j)\in\mathcal{T}_1},$$

which is identical to $\boldsymbol{h}_1(\boldsymbol{s})$ in (9) defined at Step 2.

We next look into the fundamental matrix of the Markov chain $\{(L(t), J(t)); t \geq 0\}$. Let $\boldsymbol{Q}$ denote the infinitesimal generator of the Markov chain. Then, it is of the following matrix form:

$$\boldsymbol{Q} = \big[\boldsymbol{R}_{i,i'}\big]_{1\leq i,i'\leq N}.$$

Here, $\boldsymbol{R}_{i,i'}$ is a $|\mathcal{F}_i| \times |\mathcal{F}_{i'}|$ matrix representing transition rate from level i to level i' (i.e., from state (i,j) to state (i',j') for $1 \leq j \leq |\mathcal{F}_i|$ and $1 \leq j' \leq |\mathcal{F}_{i'}|$). Suppose $i < N$. Then, from (P1) we have $\boldsymbol{R}_{i,i'} = \mathbf{0}$ unless $i' = i+1$ or $i' = i$. In the case $i' = i+1$, we denote $\boldsymbol{R}_{i,i'} (= \boldsymbol{R}_{i,i+1})$ by $\boldsymbol{R}_i$ and derive the closed-form expression for $\boldsymbol{R}_i$ below in Lemma 3. In the case $i' = i$, by (P1) again, $\boldsymbol{R}_{i,i}$ becomes a diagonal matrix. Its main diagonal is determined by the identity $\boldsymbol{Q}\boldsymbol{e} = \mathbf{0}$, and is thus given by

$$\boldsymbol{R}_{i,i}\boldsymbol{e} = -\sum_{i'\neq i} \boldsymbol{R}_{i,i'}\boldsymbol{e} = -\boldsymbol{R}_{i,i+1}\boldsymbol{e} = -\boldsymbol{R}_i\boldsymbol{e}.$$

Now suppose $i = N$. Then, from (P2), we have $\boldsymbol{R}_{i,i'} = \mathbf{0}$ for all i' (i.e., no further transition occurs from an absorbing state). Therefore, the infinitesimal generator $\boldsymbol{Q}$ is given by (19) at the bottom of the next page, in which $\boldsymbol{R}_i$ can be obtained by Lemma 3.

Lemma 3. The matrix $\boldsymbol{R}_i$ $(1 \leq i \leq N-1)$ is obtained by

$$\boldsymbol{R}_i = \big[\lambda(\mathcal{X}^{-1}(i,j), \mathcal{X}^{-1}(i+1,j'))\big]_{1\leq j\leq|\mathcal{F}_i|, 1\leq j'\leq|\mathcal{F}_{i+1}|},$$

where $\lambda(\cdot,\cdot)$ is defined in (10).
Proof: See Appendix B. ∎

Comparing (6) and (19), we have

$$\boldsymbol{Q}|_{\mathcal{T}_1\times\mathcal{T}_1} = \boldsymbol{F}_1. \tag{20}$$

That is, the matrix $\boldsymbol{F}_1$ defined in (6) at Step 3 represents transition rate from transient state space $\mathcal{T}_1$ to itself and becomes the fundamental matrix of the Markov chain $\{(L(t), J(t)); t \geq 0\}$.

We are now ready to prove Theorem 1 for the case $\alpha = 1$. By applying (17) in Lemma 2 to the completion time T_1, we have (12) in Theorem 1. It is clear from the formula in (12) that the function $H_1(\cdot)$ is strictly increasing and continuous. Hence, its inverse function $H_1^{-1}(\cdot)$ exists. Moreover, in accordance with Definition 1, we can obtain $G_{1,\beta}$ by solving $H_1(G_{1,\beta}) = \mathrm{P}(T_1 \leq G_{1,\beta}) = \beta$, i.e., $G_{1,\beta} = H_1^{-1}(\beta)$. This proves (13) in Theorem 1. In our spread model, the process $\{(L(t), J(t)); t \geq 0\}$ eventually enters the absorbing state space $\mathcal{A}$ with probability 1, which shows the existence of the inverse matrix of $\boldsymbol{F}_1$ [15, Lemma 2.2.1]. Hence, by applying (18) in Lemma 2, we have (14). This completes the proof of Theorem 1 for $\alpha = 1$.

3.3.2 *Proof for the case* $\alpha \in [0, 1)$

The key idea behind our derivation of this case is that T_α is the time taken by the process $L(t)$ to reach the state $\lceil \alpha N \rceil$. Expanding the idea, we judiciously redefine transient state space and absorbing state space depending on α, and construct a new Markov process on those redefined spaces. Then, by using an approach similar to the case $\alpha = 1$, we can derive formulas for $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$. In what follows, we present our derivation in detail.

First, we *truncate* the state space $\mathcal{U}$ to a smaller $\mathcal{U}_\alpha$ parameterized by α as follows:

$$\mathcal{U}_\alpha \triangleq \big\{(i,j) \in \mathcal{U} : i \leq \lceil \alpha N \rceil\big\}.$$

Next, we partition the space $\mathcal{U}_\alpha$ into transient state space $\mathcal{T}_\alpha$ and absorbing state space $\mathcal{A}_\alpha$ as

$$\mathcal{T}_\alpha \triangleq \big\{(i,j) \in \mathcal{U}_\alpha : i < \lceil \alpha N \rceil\big\},$$
$$\mathcal{A}_\alpha \triangleq \big\{(i,j) \in \mathcal{U}_\alpha : i = \lceil \alpha N \rceil\big\}.$$

Note that $\mathcal{T}_\alpha$ defined at Step 2 exactly refers to the above. On the state space $\mathcal{T}_\alpha \cup \mathcal{A}_\alpha$, we define a truncated level-phase process, denoted by $(L_\alpha(t), J_\alpha(t))$, from the process $(L(t), J(t))$ as follows: $(L_\alpha(t), J_\alpha(t))$ evolves identically to the process $(L(t), J(t))$ as long as $L(t) < \lceil \alpha N \rceil$ (i.e., $(L(t), J(t)) \in \mathcal{T}_\alpha$). If $L(t)$ enters the level $\lceil \alpha N \rceil$ (i.e., $(L(t), J(t)) \in \mathcal{A}_\alpha$), then the process $(L_\alpha(t), J_\alpha(t))$ stops evolving and is absorbed into the space $\mathcal{A}_\alpha$. Note that by Lemma 1, the truncated process $\{(L_\alpha(t), J_\alpha(t)); t \geq 0\}$ forms a two-dimensional CTMC with possibly multiple absorbing states in $\mathcal{A}_\alpha$. An example of a sample path for the truncated Markov chain is shown in Fig. 2.
Similarly to the case $\alpha = 1$, we have

$$T_\alpha = \inf\big\{t \geq 0 : (L_\alpha(t), J_\alpha(t)) \in \mathcal{A}_\alpha\big\}, \quad \alpha \in [0, 1).$$

That is, T_α for $\alpha \in [0,1)$ can be viewed as the *time until absorption*. Thus, by applying Lemma 2 with proper initial distribution and fundamental matrix, we can obtain the formulas for $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$ for $\alpha \in [0,1)$.

First, we look into the initial distribution of the Markov chain $\{(L_\alpha(t), J_\alpha(t)); t \geq 0\}$ on the space $\mathcal{T}_\alpha$. Similarly to the case $\alpha = 1$, the initial value is determined by the seeder vector as $(L_\alpha(0), J_\alpha(0)) = \mathcal{X}(\boldsymbol{s})$. Hence, the initial distribution on $\mathcal{T}_\alpha$ is obtained by

$$\big(\mathrm{P}(\mathcal{X}(\boldsymbol{s}) = (i,j))\big)_{(i,j)\in\mathcal{T}_\alpha} = \big(\mathbf{1}_{\{\mathcal{X}(\boldsymbol{s})=(i,j)\}}\big)_{(i,j)\in\mathcal{T}_\alpha},$$

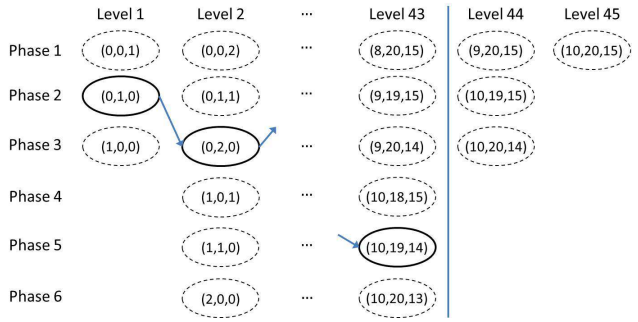


Fig. 2. Example of a sample path for $\{(L_\alpha(t), J_\alpha(t)); t \geq 0\}$ with $\alpha = 0.95$: we use the same parameters K, $\boldsymbol{s}$, and $\boldsymbol{N}$ as in Fig. 1. The Markov chain starts from $\mathcal{X}(\boldsymbol{s}) = (1, 2)$ and ends up with one of states having level $\lceil \alpha N \rceil = 43$.

which corresponds to the vector $\boldsymbol{h}_\alpha(\boldsymbol{s})$ in (9) defined at Step 2. This explains the necessity of Step 2.

Next, we look into the fundamental matrix. Since the level process $\{L(t); t \geq 0\}$ is a counting process and $L_\alpha(t)$ is the same with $L(t)$ until it reaches $\lceil \alpha N \rceil$, the fundamental matrix for the process $\{(L_\alpha(t), J_\alpha(t)); t \geq 0\}$ is obtained by truncating the infinitesimal generator $\boldsymbol{Q}$ as $\boldsymbol{Q}|_{\mathcal{T}_\alpha \times \mathcal{T}_\alpha}$. The resulting matrix corresponds to $\boldsymbol{F}_\alpha$ in (6) defined at Step 3.

Now we are ready to prove Theorem 1 for $\alpha \in [0, 1)$. Following the approach used for the case $\alpha = 1$, we can prove Theorem 1 for $\alpha \in [0, 1)$. The only difference is that we apply Lemma 2 with $\mathcal{T}_\alpha$, $\boldsymbol{h}_\alpha(\boldsymbol{s})$, and $\boldsymbol{F}_\alpha$ instead of $\mathcal{T}_1$, $\boldsymbol{h}_1(\boldsymbol{s})$, and $\boldsymbol{F}_1$ in the place of transient state space, initial distribution, and fundamental matrix, respectively. Due to similarity, we omit the details.

4 Analytical Characteristics and Applications

In this section, we identify analytical characteristics of the (α, β)-guaranteed time $G_{\alpha,\beta}$ and the ratio $R_{\alpha,\beta}$ through our temporal analysis framework. We then remark how we can utilize such characteristics in practical applications.

4.1 Impact of the Level of Infection Rates

Various spread factors control the behavior of the information spread time. We first answer the question on how the level of infection rates affect the distribution of the α-completion time T_α. To formalize, we suppose that the infection rate $\xi_{k,k'}$ from an infected node in class k to a susceptible node in class k' is scaled by $\gamma\,(>0)$ times for all $1 \leq k, k' \leq K$. Let $\hat{T}_\alpha$, $\hat{G}_{\alpha,\beta}$ and $\hat{R}_{\alpha,\beta}$ denote, respectively, the values of T_α, $G_{\alpha,\beta}$ and $R_{\alpha,\beta}$ that are computed with the scaled infection rate $\gamma\xi_{k,k'}$. Using our level-phase framework in Theorem 1, we can derive the relationship among those metrics, as presented in the following theorem.

Theorem 2. For any $\alpha \in [0, 1]$, we have

$$\hat{T}_\alpha \overset{\mathrm{d}}{=} \gamma^{-1} T_\alpha,$$

where $\overset{\mathrm{d}}{=}$ denotes "equal in distribution." As a consequence, we obtain the following for any $\alpha \in [0, 1]$ and $\beta \in [0, 1]$:

$$\hat{G}_{\alpha,\beta} = \gamma^{-1} G_{\alpha,\beta}, \quad \hat{R}_{\alpha,\beta} = R_{\alpha,\beta}.$$

Proof: See Appendix D. ∎

Theorem 2 says that the spread becomes faster *proportionally* to the level of infection rates in *distribution sense*. From the perspective of average analysis, the result in Theorem 2 also indicates the following. Let $\mathcal{M}(t) \triangleq \mathrm{E}[\sum_k I_k(t)]$ denote the average number of infected nodes in the network at time t. By Theorem 2, we have that $\hat{\mathcal{M}}(t)$, the value of $\mathcal{M}(t)$ that is computed with the scaled infection rate $\gamma\xi_{k,k'}$, satisfies the following *for all* $t \geq 0$:

$$\hat{\mathcal{M}}(t) = \mathcal{M}(\gamma t). \tag{21}$$

The proof of (21) is given in Appendix E. Likewise, the time derivatives $\mathcal{D}(t) \triangleq \mathcal{M}'(t)$ and $\hat{\mathcal{D}}(t) \triangleq \hat{\mathcal{M}}'(t)$ representing the speed of information propagation, should change over time t with the relation $\hat{\mathcal{D}}(t) = \gamma \mathcal{D}(\gamma t)$.

4.2 Impact of Network Size

We next investigate the impact of the network size N on the time for information spread. In our spread model, each non-informed node (i.e., susceptible node) can be considered as a workload to finish. In this respect, adding a node to the network might slow down the spread. However, once the node becomes informed (i.e., infected), it works in a similar manner as the seeder and is involved in spreading the information. In this respect, adding a node to the network might expedite the spread. Therefore, it is unclear whether the network size accelerates or slows down the speed of information propagation. Using our framework, we can give the answer, as shown in Theorem 3.

Theorem 3. Suppose $\alpha = 1$, $K = 1$, and $s_1 = 1$. Then, as the network size N increases, we have the following.

$$\boldsymbol{Q} = \begin{bmatrix} -diag(\boldsymbol{R}_1\boldsymbol{e}) & \boldsymbol{R}_1 & \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -diag(\boldsymbol{R}_2\boldsymbol{e}) & \boldsymbol{R}_2 & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & -diag(\boldsymbol{R}_3\boldsymbol{e}) & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \dots & -diag(\boldsymbol{R}_{N-2}\boldsymbol{e}) & \boldsymbol{R}_{N-2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{0} & -diag(\boldsymbol{R}_{N-1}\boldsymbol{e}) & \boldsymbol{R}_{N-1} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}. \tag{19}$$

(1) The average α-completion time $\mathrm{E}[T_\alpha]$ is strictly decreasing with N. Also, it asymptotically behaves as [6]

$$\mathrm{E}[T_\alpha] = \Theta\big(\xi^{-1}N^{-1}\ln N\big). \tag{22}$$

(2) The guaranteed time $G_{\alpha,\beta}$ is strictly decreasing with N if β is greater than a certain value $\beta_0 < 1$. Also, for any $\beta \in [0,1]$, it asymptotically behaves as

$$G_{\alpha,\beta} = \Theta\big(\xi^{-1}N^{-1}(\ln N - \ln(\ln \beta^{-1}))\big), \tag{23}$$

from which we have $R_{\alpha,\beta} = \Theta(1)$.

Proof: See Appendix F. ■

Theorem 3 indicates that adding a node to the network accelerates the information spread when per-pair infection rates are unchanged. This effect, which we call population effect, is also observed in Fig. 4 of [8] where the authors study the impact of heterogeneous human activities on epidemic spreading through simulations. Theorem 3 also says that the degree of acceleration is asymptotically *proportional* to the network size. Combining with Theorem 2, our analytic findings in Theorem 3 imply that information spread accelerated by the population effect shows a quantitatively similar behavior as if the level of infection rates is scaled up by the network size.

To assist understanding of Theorem 3, we consider a non-cooperative spread model in which only the seeder(s) chosen at the beginning of the spread is able to disseminate the information. In epidemiology, this non-cooperative model can be classified into a SIR model with zero recovery time from infection. Remark 4 summarizes our result with the proof given in Appendix G.

Remark 4. Suppose $\alpha = 1$, $K = 1$, and $s_1 = 1$. Then, as the network size N increases, we have the following for a non-cooperative spread model.

(1) The average α-completion time $\mathrm{E}[T_\alpha]$ is strictly increasing with N. Also, it asymptotically behaves as

$$\mathrm{E}[T_\alpha] = \Theta\big(\xi^{-1}\ln N\big).$$

(2) The guaranteed time $G_{\alpha,\beta}$ is strictly increasing with N for any $\beta \in [0,1]$. Also, it asymptotically behaves as

$$G_{\alpha,\beta} = \Theta\big(\xi^{-1}(\ln N - \ln(\ln \beta^{-1}))\big),$$

from which we have $R_{\alpha,\beta} = \Theta(1)$.

The higher order statistics of T_α for the non-cooperative spread model and our spread model (namely, cooperative model) are further compared in Table 1. In the table, $\zeta(c) \triangleq \sum_{n=1}^{\infty} n^{-c}$ denotes the Riemann zeta function. The proof of Table 1 is omitted due to similarity to the proofs of Theorem 3 and Remark 4. Our analysis showing that $G_{\alpha,\beta}$ behaves differently for the scaling of N and ξ tells that resource allocation for information spread should be carefully designed based on the willingness of cooperation in a spread process (i.e., infectivity in a spread process).

6. We adopt the following notations to describe asymptotic behaviors:
(i) $g(n) = O(h(n))$ if there exists a constant $c > 0$ and $\hat{n} \in \mathbb{N}$ such that $|g(n)| \le c|h(n)|$ for all $n \ge \hat{n}$.
(ii) $g(n) = \Omega(h(n))$ if $h(n) = O(g(n))$.
(iii) $g(n) = \Theta(h(n))$ if $g(n) = O(h(n))$ and $g(n) = \Omega(h(n))$.

TABLE 1
Comparison of Population Effect on the Spread Time

	Cooperative model	Non-cooperative model
Variance of T_α	Strictly decrease with N and scale as $\Theta(\xi^{-2}N^{-2})$	Strictly increase with N and converge to $\xi^{-2}\zeta(2)$
Skewness of T_α	Strictly decrease with N and scale as $\Theta(\xi^{-3}N^{-3})$	Strictly increase with N and converge to $\xi^{-3}\zeta(3)$
$\mathrm{E}[(T_\alpha)^n]$ $(n \ge 2)$	$\mathrm{E}[(T_\alpha)^n] < \infty$ for a fixed N	$\mathrm{E}[(T_\alpha)^n] < \infty$ for a fixed N
	$\lim_{N\to\infty}\mathrm{E}[(T_\alpha)^n] < \infty$	$\lim_{N\to\infty}\mathrm{E}[(T_\alpha)^n] = \infty$

4.3 Impact of Heterogeneity

The impact of heterogeneity in information or virus spread has been less explored. Using our framework, we analyze and understand the temporal spread behavior under a heterogeneous network with multiple classes compared with a homogeneous network. In particular, we focus on answering "Does heterogeneity persistently expedite the spread or not?", "Is there an optimal heterogeneity level for information spread?", and "Is there an upper or a lower bound on the gain from the heterogeneity over homogeneity?". In this section, we provide the answers to these questions by studying a dual community model ($K = 2$) compared with a single community model ($K = 1$). Note that our framework can be easily extended to study the cases when $K \ge 3$.

We first describe the system parameters for a heterogeneous network. We assume that (i) both classes are of the same size, i.e., $N_1 = N_2\,(= N/2)$. (ii) There is one seeder in the network. Without loss of generality, the seeder is chosen randomly from nodes in class 1, i.e., $\boldsymbol{s} = (1,0)$. (iii) The inter-class infection rates are the same in either direction, i.e., $\xi_{1,2} = \xi_{2,1}\,(\triangleq \xi_{\text{inter}})$. Hence, the rate matrix $\boldsymbol{\Lambda}$ can be expressed as

$$\boldsymbol{\Lambda} = \xi_{\text{inter}} \begin{pmatrix} r_1 & 1 \\ 1 & r_2 \end{pmatrix}, \tag{24}$$

where $r_1 \triangleq \xi_{1,1}/\xi_{\text{inter}}$ and $r_2 \triangleq \xi_{2,2}/\xi_{\text{inter}}$. The ratios r_1 and r_2 control the intra-class infection rates whose values are chosen freely in the range $0 \le r_1, r_2 < \infty$. Note that $(r_1, r_2) = (1,1)$ reduces to the homogeneous case, and the larger (r_1, r_2) is deviated from $(1,1)$, the more the heterogeneity is induced. To summarize, the heterogeneous network in this study is parameterized by the tuple $(N, \xi_{\text{inter}}, r_1, r_2)$.

We next describe the system parameters for a homogeneous network. For a fair comparison between homogeneous and heterogeneous networks, we impose the following constraint:

$$\xi = \frac{\sum_{a=1}^{N}\sum_{b=1,b\neq a}^{N}\xi_{k(a),k(b)}}{N(N-1)}. \tag{25}$$

Here, the left-hand side is the infection rate in a homogeneous network, which equals the per-pair averaged infection rate in a heterogeneous network on the right-hand side. Thus, a heterogeneous network with $(N, \xi_{\text{inter}}, r_1, r_2)$ and a homogeneous network with (N, ξ) are fairly comparable if

$$\xi = \frac{\xi_{\text{inter}}}{2(N-1)}\left\{\left(\frac{N}{2}-1\right)(r_1+r_2) + N\right\}. \tag{26}$$

With the help of Theorems 2 and 3 concerning the scaling of ξ and N, we can characterize and generalize the impact of heterogeneity by only observing a specific set of (ξ, N). Hence,

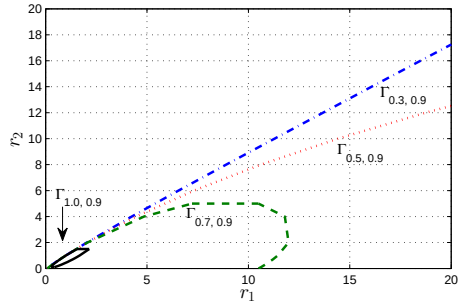


Fig. 3. Comparison of the guaranteed time $G_{\alpha,\beta}$ between heterogeneous and homogeneous networks for $\beta = 0.9$ and $\alpha \in \{0.3, 0.5, 0.7, 1.0\}$: if $(r_1, r_2) \in \Gamma_{\alpha,\beta}$, then heterogeneity with the level (r_1, r_2) accelerates the information spread (i.e., reduces the guaranteed time $G_{\alpha,\beta}$). If $(r_1, r_2) \notin \Gamma_{\alpha,\beta}$, then heterogeneity slows down the information spread.

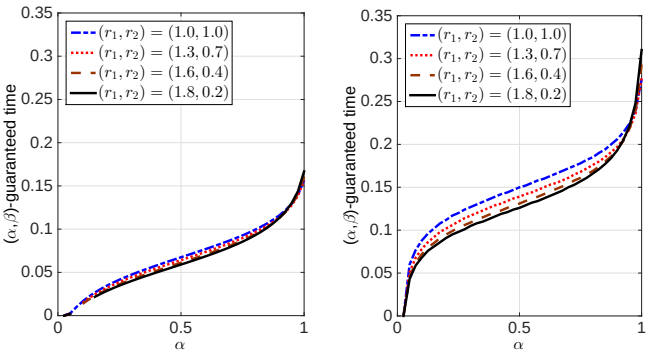


Fig. 4. The guaranteed time $G_{\alpha,\beta}$ for $\beta = 0.1$ (left) and $\beta = 0.9$ (right): $(r_1, r_2) = (1, 1)$ reduces to the homogeneous case, and the larger (r_1, r_2) is deviated from $(1, 1)$, the more the heterogeneity is induced.

we fix $(\xi, N) = (1, 40)$ in this study, and let (r_1, r_2) vary in the range $0 \le r_1, r_2 \le 20$.[7]

Using our framework, we obtain the guaranteed time $G_{\alpha,\beta}$ for each (r_1, r_2) and compare it with the homogeneous counterpart. Fig. 3 shows the result. In the figure, $\Gamma_{\alpha,\beta}$ is the region such that $(r_1, r_2) \in \Gamma_{\alpha,\beta}$ if and only if heterogeneity with the level (r_1, r_2) yields reduced guaranteed time $G_{\alpha,\beta}$ compared to the homogeneous case. Hence, the region $\Gamma_{\alpha,\beta}$ can be interpreted as the area where heterogeneity accelerates the information spread. From the figure, we can have the following observations and interpretations:

(1) For a fixed β, the region $\Gamma_{\alpha,\beta}$ becomes reduced as α increases, i.e., $\Gamma_{\alpha_1,\beta} \subset \Gamma_{\alpha_2,\beta}$ for $\alpha_1 \ge \alpha_2$. Hence, for a fixed (r_1, r_2), there exists a threshold $\alpha_{\text{th}} \in [0, 1]$ such that $(r_1, r_2) \in \Gamma_{\alpha,\beta}$ if $\alpha \le \alpha_{\text{th}}$ and $(r_1, r_2) \notin \Gamma_{\alpha,\beta}$ if $\alpha > \alpha_{\text{th}}$. This implies that heterogeneity in infection rates accelerates the spread at the beginning phase of the spread process (i.e., $\alpha \le \alpha_{\text{th}}$), whereas slowing down the spread at the ending phase (i.e., $\alpha > \alpha_{\text{th}}$). In addition, the threshold α_{th} decreases as (r_1, r_2) deviates from (1,1) meaning that the time portion of the acceleration becomes shorter with more heterogeneity.

(2) For $\alpha \in \{0.3, 0.5, 0.7, 1.0\}$, the region $\bigcap_\alpha \Gamma_{\alpha,\beta}$ is non-empty. This shows that there is an *optimal heterogeneity level* for information spread which always accelerates the spread entirely from the beginning to the end of the spread process (i.e., $\alpha_{\text{th}} = 1$). Such an optimal heterogeneity level can be found in the intersection $\bigcap_\alpha \Gamma_{\alpha,\beta} (= \Gamma_{1,\beta})$.

(3) For any α, we have $\Gamma_{\alpha,\beta} \bigcap \{(r_1, r_2) : r_1 < r_2\} = \varnothing$. Hence, in the region $\{(r_1, r_2) : r_1 < r_2\}$, heterogeneity always slows down the information spread from the beginning to the end of the spread process. This observation indicates that if the seeder is chosen from a less infective community, then heterogeneity never accelerates the information spread.

In the following, we support the above observations for a special case when the inter-class infection rate is determined from the intra-class infection rates by $\xi_{\text{inter}} = (\xi_{1,1} + \xi_{2,2})/2$. In this case, the rate matrix $\mathbf{\Lambda}$ in (24) is further simplified as

$$\mathbf{\Lambda} = \xi_{\text{inter}} \begin{pmatrix} r_1 & 1 \\ 1 & 2 - r_1 \end{pmatrix}.$$

7. Then, from the constraint (26), we can compute ξ_{inter} for each (r_1, r_2), which in turn determines uniquely the rate matrix $\mathbf{\Lambda}$ in (24).

In this study, we fix $(\xi, N) = (1, 40)$ as above and vary r_1 in the range $\{1.0, 1.3, 1.6, 1.8\}$. Here, we set $r_1 \ge 1$ (i.e., $\xi_{1,1} \ge \xi_{2,2}$) so that the seeder is chosen from a more infective community. Under this system setup and parameters, we obtain the (α, β)-guaranteed time $G_{\alpha,\beta}$ for $\alpha \in [0, 1]$, $\beta \in \{0.1, 0.9\}$ and show the result in Fig. 4. From the figure, we confirm that heterogeneity indeed accelerates the spread for smaller penetration (i.e., for low α) but slows down it for higher penetration.

4.4 Heterogeneity Advantage on Spread

To assist understanding of our numerical observations in Section 4.3, we investigate the rates at which the spread process leaves each level i ($i = 1, 2, \ldots, N-1$), i.e., the rate of informing one additional node provided that i nodes have been informed. Let μ_i be the rate of leaving the level i in a homogeneous network. Then, we have

$$\mu_i = \xi i (N - i).$$

Different from the homogeneous case, we have multiple phases for each level in a heterogeneous network, and the rate of leaving a level varies depending upon which phase realizes the level. Hence, in the case of a heterogeneous network, we define $\hat{\mu}_{i,X}$ to be the rate of leaving the level i when the associated phase is X.

Let i be fixed. Then, we have $X = X_i \in \{1, \ldots, |\mathcal{F}_i|\}$. For each $X_i = j \in \{1, \ldots, |\mathcal{F}_i|\}$, the rate $\hat{\mu}_{i,j}$ can be obtained by adding the transition rate from level-phase state (i, j) to $(i+1, j')$ (i.e., adding the (j, j')th element of the matrix $\boldsymbol{R}_i$) for all $1 \le j' \le |\mathcal{F}_{i+1}|$ as follows:

$$\hat{\mu}_{i,j} = \sum_{1 \le j' \le |\mathcal{F}_{i+1}|} [\boldsymbol{R}_i]_{j,j'} = (\boldsymbol{N} - \boldsymbol{f})(\boldsymbol{f}\mathbf{\Lambda})^T, \qquad (27)$$

where $\boldsymbol{f} = \mathcal{X}^{-1}(i, j)$ and A^T denotes the transpose of a matrix (or a vector) A, and where the second equality comes from (10). Hence, the collection of $\hat{\mu}_{i,j}$ $(j = 1, 2, \ldots, |\mathcal{F}_i|)$ for a given i can be obtained from $\boldsymbol{R}_i$ as

$$(\hat{\mu}_{i,j})_{j=1,2,\ldots,|\mathcal{F}_i|} = (\boldsymbol{R}_i \boldsymbol{e})^T.$$

It follows from the theory of absorbing Markov chains [15] that the probability distribution of X_i on $\{1, 2, \ldots, |\mathcal{F}_i|\}$ is

$$(\mathrm{P}(X_i = j))_{j=1,2,\ldots,|\mathcal{F}_i|} = \boldsymbol{h}_\alpha (-\boldsymbol{F}_\alpha)^{-1} \begin{bmatrix} \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \\ \boldsymbol{R}_{\lceil \alpha N \rceil - 1} \end{bmatrix}, \qquad (28)$$

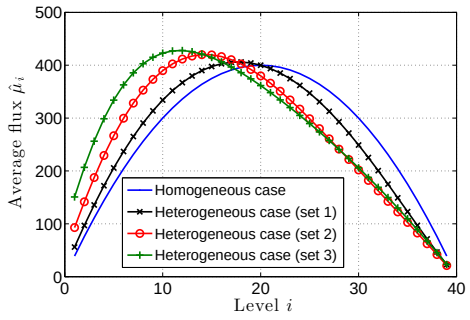


Fig. 5. Average rate $\hat{\mu}_i$ of leaving each level i for the spread parameters in (30).

where $\alpha = \frac{i}{N}$.[8] Hence, the expectation of $\hat{\mu}_{i,X}$ accommodating all the possibilities of the phase $X(= X_i)$ for a given level i is

$$\hat{\mu}_i \triangleq \mathrm{E}[\hat{\mu}_{i,X}] = \sum_{j=1}^{|\mathcal{F}_i|} \mathrm{P}(X_i = j)\hat{\mu}_{i,j}$$
$$= \boldsymbol{h}_{\frac{i}{N}} \left(-\boldsymbol{F}_{\frac{i}{N}}\right)^{-1} \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \boldsymbol{R}_{i-1} \end{bmatrix} \boldsymbol{R}_i \boldsymbol{e}. \qquad (29)$$

The metric $\hat{\mu}_i$ represents the average rate of information flow (or dispersion) passing through susceptible nodes from i informed nodes, which we call flux. Based on (29), we compare the flux $\hat{\mu}_i$ with μ_i for $i = 1, 2, \ldots, N-1$ and show the result in Fig. 5. In the figure, we use $N = 40$ and $\xi = 1$ for homogeneous case. For heterogeneous case, we use the following three sets of spread parameters:

$$\begin{aligned} &\text{(Set 1)} \quad \boldsymbol{N} = (20, 20), \quad \boldsymbol{\Lambda} = \begin{bmatrix} 1.9 & 1 \\ 1 & 0.1 \end{bmatrix}, \\ &\text{(Set 2)} \quad \boldsymbol{N} = (10, 30), \quad \boldsymbol{\Lambda} = 1.63 \begin{bmatrix} 1.9 & 1 \\ 1 & 0.1 \end{bmatrix}, \\ &\text{(Set 3)} \quad \boldsymbol{N} = (5, 35), \quad \boldsymbol{\Lambda} = 2.74 \begin{bmatrix} 1.9 & 1 \\ 1 & 0.1 \end{bmatrix}. \end{aligned} \qquad (30)$$

For a fair comparison, those parameters are chosen to satisfy the constraint in (26). In all three cases, we choose one seeder randomly from nodes in a more infective community (i.e., class 1).

We can observe that at the beginning of the spread, the average flux $\hat{\mu}_i$ in the heterogeneous network is larger than that of the homogeneous network. This trend becomes flipped after a certain moment as shown in Fig. 5. This phenomenon can be understood intuitively given that heterogeneity allows a wider range of mobility and contact patterns resulting in speedy spread together with procrastinatory spread.

From the discussion above, we form a conjecture that there exists $\hat{N} \leq N$ such that

$$\begin{aligned} \hat{\mu}_i &> \mu_i \quad \text{if} \quad i < \hat{N}, \\ \hat{\mu}_i &< \mu_i \quad \text{if} \quad i > \hat{N}. \end{aligned} \qquad (31)$$

8. In terms of the level-phase process $\{(L(t), J(t)); t \geq 0\}$, the probability $\mathrm{P}(X_i = j)$ in (28) can be rewritten as $\mathrm{P}(J(t) = j | L(t) = i)$.

Due to technical difficulty in handling the inverse matrix $(\boldsymbol{F}_{i/N})^{-1}$ analytically, we prove this conjecture in a simplified case where $K = 2, N_1 = N_2$ and the rate matrix $\boldsymbol{\Lambda}$ is given by

$$\boldsymbol{\Lambda} = \xi_{\text{inter}} \begin{pmatrix} r_1 & 1 \\ 1 & 2 - r_1 \end{pmatrix}. \qquad (32)$$

This is the model underlying Fig. 4 of Section 4.3. The heterogeneity parameter r_1 is in the range $[1, 2)$. When $r_1 = 1$, it reduces to the homogeneous case. As r_1 increases, the more the heterogeneity is induced. The following theorem presents the relation between $\hat{\mu}_i$ and μ_i.

Theorem 4. Suppose $K = 2, N_1 = N_2$, $\boldsymbol{s} = (1, 0)$, and the rate matrix $\boldsymbol{\Lambda}$ is given by (32). Then, the average rate $\hat{\mu}_i$ of leaving the level i satisfies

$$\hat{\mu}_i = \mu_i + c(r_1 - 1)\left(\frac{N}{2} - i\right), \qquad i = 1, 2, \ldots, N-1,$$

where $c > 0$ is a constant.

Proof: See Appendix H. ■

The result in Theorem 4 implies that (i) due to variability in r_1, the average flux $\hat{\mu}_i$ can be deviated from μ_i. (ii) The more the heterogeneity is induced, the larger $\hat{\mu}_i$ is deviated from μ_i. (iii) Concerning the conjecture, we have $\hat{\mu}_i > \mu_i$ if $i < \frac{N}{2}$ and $\hat{\mu}_i < \mu_i$ if $i > \frac{N}{2}$, showing that $\hat{N}$ in (31) is $\frac{N}{2}$.

4.5 Implication

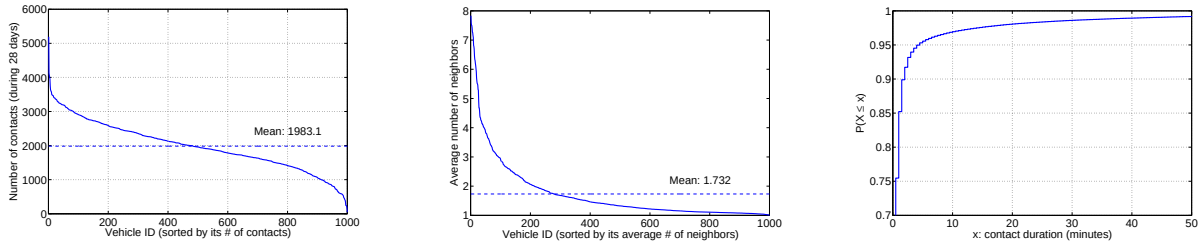
How to optimally distribute given resources to nodes in a network to minimize the time for information spread is of an important research question. Our results on heterogeneity provide key understanding to this question. We note that there exists a small region of $\boldsymbol{\Lambda}$ with heterogeneous contact rates, which always make the spread faster than a homogeneous network for a target β as shown in Fig. 3. This implies that when utilizing a vehicular network for information delivery (e.g., DieselNet [17]) or a social network for advertising a product, $\boldsymbol{\Lambda}$ can be manipulated to be heterogeneous by allocating uneven fuel to vehicles or providing distinguished incentive to users. How to realize such $\boldsymbol{\Lambda}$ from fuel or incentive distribution needs experimental study that is beyond the scope of this paper.

5 Simulation Study

5.1 Contact Statistics of a Vehicular Network

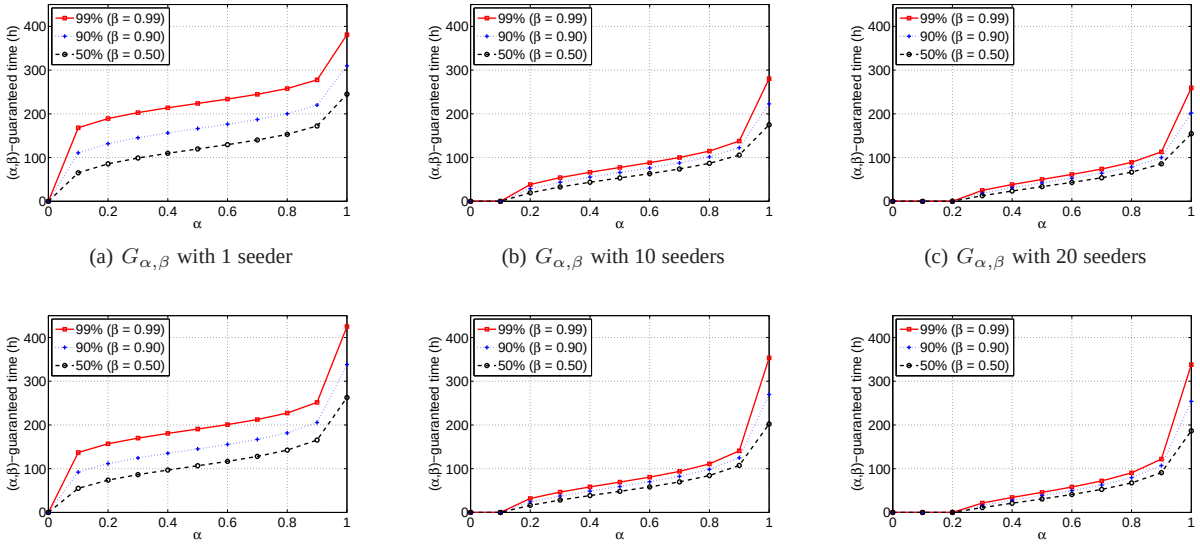
We study the efficacy of our framework and characterizations using a vehicular mobility trace obtained from more than a thousand taxies in Shanghai, China [18]. The experimental trace tracked GPS coordinates of taxies at every 30 seconds during 28 days in Shanghai. The trace was previously analyzed in [19], and it was shown that the taxies have exponentially distributed pairwise inter-contact time, which is well aligned with our CTMC-based framework.

Figs. 6 (a), (b), and (c) characterize the statistics of the taxi network with 1000 randomly chosen taxies in the aspect of number of contacts, number of neighbors in a communication range (50 meter in our analysis), and contact duration, respectively. We apply these three factors for evaluating the infection rates $\lambda_{a,b}^{\text{eff}} = \lambda_{a,b}\varphi_a\psi_b$ in (2), where the infectivity φ_a is 1 and the susceptibility ψ_b is derived from how many infection (i.e., data transmission) opportunities a contact duration can hold out of all neighboring nodes who are willing to infect others. The results



(a) Number of contacts for each vehicle (b) Average number of neighbors (c) CDF of aggregated contact durations

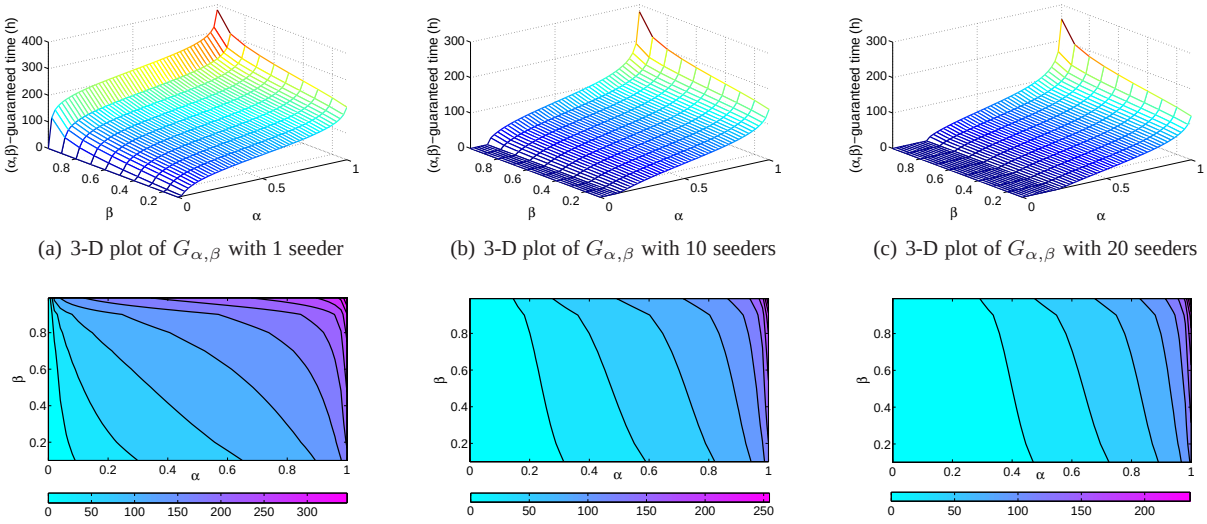
Fig. 6. (a) Number of contacts of a vehicle with all other vehicles during 28 days. (b) Average number of neighbors when a vehicle is in a contact with another vehicle. (c) CDF of aggregated contact durations between all taxi pairs.



(a) $G_{\alpha,\beta}$ with 1 seeder (b) $G_{\alpha,\beta}$ with 10 seeders (c) $G_{\alpha,\beta}$ with 20 seeders

(d) $G_{\alpha,\beta}$ with 1 seeder (e) $G_{\alpha,\beta}$ with 10 seeders (f) $G_{\alpha,\beta}$ with 20 seeders

Fig. 7. Distribution of (α,β)-guaranteed time for $\alpha \in [0,1]$ and $\beta = \{0.5, 0.9, 0.99\}$ with (a) 1 seeder, (b) 10 seeders, and (c) 20 seeders in a homogeneous network and with (d) 1 seeder, (e) 10 seeders, and (f) 20 seeders in a heterogeneous network with two classes.



(a) 3-D plot of $G_{\alpha,\beta}$ with 1 seeder (b) 3-D plot of $G_{\alpha,\beta}$ with 10 seeders (c) 3-D plot of $G_{\alpha,\beta}$ with 20 seeders

(d) Contour map of $G_{\alpha,\beta}$ with 1 seeder (e) Contour map of $G_{\alpha,\beta}$ with 10 seeders (f) Contour map of $G_{\alpha,\beta}$ with 20 seeders

Fig. 8. (α,β)-guaranteed time for $\alpha \in [0,1]$ and $\beta \in [0.1, 0.99]$ with (a) 1 seeder, (b) 10 seeders, and (c) 20 seeders in a homogeneous network and contour maps of $G_{\alpha,\beta}$ with (d) 1 seeder, (e) 10 seeders, and (f) 20 seeders.

are summarized in Table 2 for a homogeneous network and a heterogeneous network with $K = 2$. Note that the infection rates in Table 2 satisfy the constraint in (25) that was introduced for a fair comparison between a homogeneous model and a heterogeneous model.

TABLE 2
Infection rates for a homogeneous network and for a heterogeneous network with two classes of taxies

Homogeneous case	Heterogeneous case		
ξ	$\xi_{1,1}$	$\xi_{2,2}$	$\xi_{1,2}\,(=\xi_{2,1})$
4.14×10^{-4}	7.17×10^{-4}	1.93×10^{-4}	3.72×10^{-4}

Based on the statistics in Table 2, it is possible to predict the information spread time and to find out efficient methods for properly allocating resources to the taxi network. To this end, we simulate probabilistic guarantees for the completion time in a homogeneous and a heterogeneous network, each with 100 taxies. We assume an application scenario of a firmware update to be distributed for mobile devices, which will take around 90 seconds demanding 1.15 number of contacts on average. In our simulation, the number of taxies is limited to be 100 due to computational complexity involved in the matrix operations.

5.2 Information Spread Time in a Vehicular Network

We first study the spread time in a homogeneous network. Figs. 7 (a), (b), and (c) show the (α, β)-guaranteed time for $\alpha \in [0, 1]$ and $\beta \in \{0.5, 0.9, 0.99\}$ with the number of seeders given by 1, 10, and 20, respectively. The figures tell that if we target 90% penetration with 99% confidence (i.e., $(\alpha, \beta) = (0.9, 0.99)$), then the taxi network with a single seeder is estimated to take about 11.6 days (i.e., 278 hours) to achieve the target level of information spread. This estimation largely differs from the existing estimation of the average time to achieve 90% of penetration, which is close to 7 days. This clarifies that designing plans associated with the successful spread to 90% of nodes should incorporate about 4.6 additional days. If someone wants to avoid the associated plans being delayed, our framework is able to suggest adding more seeders to the network as shown in Figs. 7 (b) and (c). As the number of seeders increases to 10 or 20, the time required for 90% penetration with 99% confidence reduces from 278 hours to 137 hours (5.7 days) and 113 hours (4.7 days), respectively.

Similarly, we can study a heterogeneous network with two classes. Figs. 7 (d), (e), and (f) show the (α, β)-guaranteed time for $\alpha \in [0, 1]$ and $\beta \in \{0.5, 0.9, 0.99\}$ with 1, 10, and 20 seeders, respectively. Direct comparison between Figs. 7 (a), (b), (c) and Figs. 7 (d), (e), (f) confirms our claims in Section 4.3. The claims tell that the (α, β)-guaranteed time in a heterogeneous network is faster than a homogeneous network for lower α, and is slower for higher α close to 1. This implies that if it is mandatory to achieve 100% penetration, making the properties of nodes in a network to be more homogeneous (e.g., by providing more resources to inactive groups of nodes) can be helpful in spreading information.

5.3 Effective Penetration of Information

We further analyze detailed cases of a homogeneous network with α and β parameters in the ranges of $[0, 1]$ and $[0.1, 0.99]$, respectively. Figs. 8 (a), (b), and (c) show the 3-D plots of $G_{\alpha,\beta}$ with different number of seeders. These 3-D plots commonly show

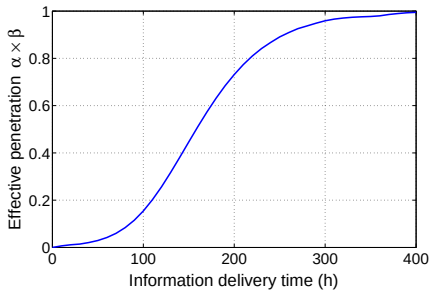
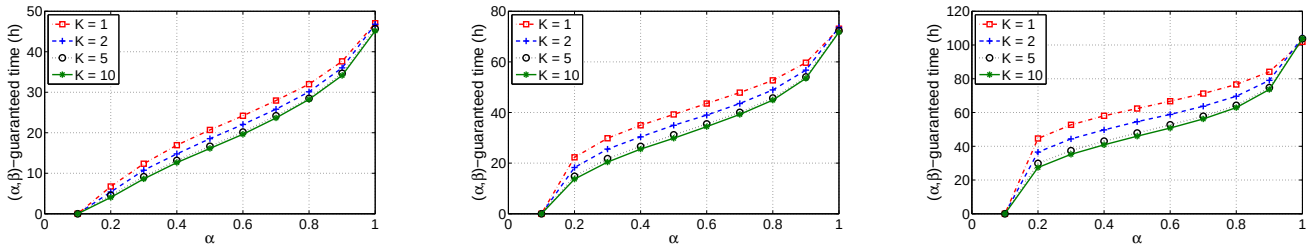


Fig. 9. Effective penetration of information, $\alpha \times \beta$, over time from 1 seeder in a homogeneous network of infection rate $\xi = 4.14 \times 10^{-4}$.

that approaching to $\alpha = 1$ or $\beta = 1$ significantly extends the corresponding (α, β)-guaranteed time. Especially, achieving both $\alpha = 1$ and $\beta = 0.99$ is shown to require much longer time. It is intriguing to note that the ratio of guaranteed times of achieving $(\alpha, \beta) = (1, 0.99)$ and achieving $(\alpha, \beta) = (0.5, 0.5)$ gets larger as the number of seeders increases. This tells that full penetration of information to a network is relatively hard to improve whereas a moderate penetration (e.g., half penetration with 50% guarantee) is much easier to speed up by adding more seeders.

Figs. 8 (d), (e), and (f) show contour maps of Figs. 8 (a), (b), and (c), respectively. The lines in the contour maps connect the combinations of α and β parameters that yield the same guaranteed time $G_{\alpha,\beta}$. The resulting lines provide us useful guidance on designing an information spread system. For instance, Fig. 8 (d) shows that the contour lines are changing from concave to convex as the time of spread proceeds. In addition, the convexity of the lines gets steeper as the time proceeds. These patterns that are also observed in Figs. 8 (e) and (f) clarify that targeting to balance α and β is more effective than aiming at an extreme α or β, especially when maximizing the *effective penetration of information* defined as $\alpha \times \beta$. More intuitively speaking in the context of Fig. 8 (d), spreading to 60% of population with 60% of guarantee takes the same amount of time with spreading to 82% of population with 30% of guarantee, but the former is more efficient in terms of the effective penetration. The choice of α and β surely depends on the application scenarios and the system designer's goals. However, given time budget, it is highly recommended for the system designers to adjust α or β slightly to see if they can achieve higher effective penetration without deviating too much from their original goals.

Fig. 9 details the changing pattern of the maximum effective penetration at a given information delivery time where the maximum of $\alpha \times \beta$ is taken from the contour line of the same time. As it is aforementioned, the pattern of changing from concavity to convexity makes the maximum effective penetration look like a sigmoid function. This sigmoid function well captures the information spreading behavior in a network as the speed of penetration goes up and then down as the numbers of infected and uninfected nodes become balanced and then again unbalanced (with only few remaining uninfected). When balanced (e.g., 50:50), the possibility of infection is surely maximized. Fig. 9 can also work as a quick reference for checking the minimum required time for any $\alpha \times \beta$. For instance, achieving effective penetration $\alpha \times \beta = 0.25$ would take at least 120 hours no matter how we combine α and β.



(a) $G_{\alpha,\beta}$ with $\beta = 0.50$ for $K = 1, 2, 5, 10$. (b) $G_{\alpha,\beta}$ with $\beta = 0.90$ for $K = 1, 2, 5, 10$. (c) $G_{\alpha,\beta}$ with $\beta = 0.99$ for $K = 1, 2, 5, 10$.

Fig. 10. (α, β)-guaranteed time with $\beta \in \{0.50, 0.90, 0.99\}$ for various number of classes K. Larger K (i.e., stronger heterogeneity) helps spreading at the beginning, but delays spreading at the end.

5.4 Impact of Multi-classes

In this section, we study the impact of multi-classes on the spread time with increasing number of classes K. In practice, how to classify nodes in a network may be of a difficult question. Given the notion and the definition of a class (i.e., all nodes in a class keep homogeneity), segregating humans or vehicles into multiple exclusive classes is infeasible since even the most similar behaviors of humans or vehicles in a class cannot statistically guarantee their homogeneity. This implies that classifying nodes is mainly for improving mathematical tractability and reducing the computational complexity in predicting spreading patterns. We here test the impact of grouping nodes with various number of classes from $K = 1$ (i.e., totally homogeneous) to $K = 10$ (i.e., totally heterogeneous) for a network consisting of 10 nodes. Infectivity upon a contact is set to be 0.25 for capturing a wide variety of spreading patterns. Fig. 10 shows how $G_{\alpha,\beta}$ changes for different number of classes. Note that the rate matrix $\mathbf{\Lambda}$ for 10 nodes are captured from the most active taxies in the Shanghai trace. Inter and intra contact rates for different classes are assigned by taking the average behavior of nodes in each class, since all nodes in a class are assumed to be homogeneous. From the figure, we can observe that larger K leads to quicker spread at the beginning, but procrastinatory spread at the end. This behavior is aligned with the heterogeneity advantage on the spread time presented in Section 4.3, as larger K induces more heterogeneity. For those 10 taxies, we run a trace driven simulation using the trajectories for those taxies included in the Shanghai trace. In order to make the statistical property of the trace driven simulation close to that of numerical analysis, we applied the same infectivity and collected 50,000 sample paths. Fig. 11 shows the 50%, 90%, and 99% guaranteed spread time for various α values. Note that β guaranteed time is equivalent to β quantile spread time from all sample paths. The comparison between Fig. 10 and Fig. 11 supports that the totally heterogeneous case ($K = 10$) that went through minimum approximation captures the real spreading patterns in the most accurate manner for all tested β values.

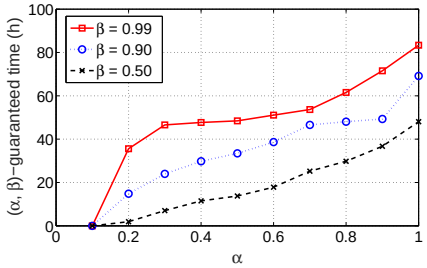


Fig. 11. Trace driven simulation results from 50,000 sample paths of information spread with 10 most active Shangai taxies.

6 Discussion

A computational problem can be encountered when applying (12), (13), and (14) in Theorem 1, especially when computing the matrix exponential $\exp(\boldsymbol{F}_\alpha t)$. The matrix $\boldsymbol{F}_\alpha$ in the theorem is a square matrix of order n (i.e., $\boldsymbol{F}_\alpha \in \mathbb{R}^{n\times n}$), where n is the number of states in our Markov chain and scales as $O(\alpha(1 + \frac{N}{K})^K)$. Hence, for each fixed N, the computational complexity increases with K and α, and the worst case occurs when $\alpha = 1$ and $K = N$ (i.e., targeting 100% penetration in a totally heterogeneous case) yielding $n = 2^N$.

Nonetheless, the matrix $\boldsymbol{F}_\alpha$ has an extra property that we can take advantage of: $\boldsymbol{F}_\alpha$ is a highly *sparse* matrix. According to (10), it allows non-zero transition rates only between state vectors that differ in one component. Consequently, $\boldsymbol{F}_\alpha$ becomes a banded upper triangular matrix and more than half of its elements are zero, as shown in (6). This property enables us to exploit the existing technique, called *Krylov approximation* (e.g., [20]–[22] and references therein), and a ready-to-use software package, called EXPOKIT [23], to resolve the computational problem as described below.

There have been extensive studies on the numerical algorithms for computing matrix exponentials in mathematics and physics [24]. The case where the matrix is of moderate dimension has benefited from the classical methods such as Padé approximation and Taylor series approximation. For a large sparse matrix, Krylov subspace projection technique has been shown to provide robust reduction of computational burden in exponentiating such matrix, especially for the one arising in Markov chains [20]. Its substantial performance gain has been justified by theoretical characterizations as well as practical studies [21], [22], [25].

Let us explain briefly the underlying principle of Krylov subspace projection technique. The gist of Krylov approximation is that the original large sparse problem (of size n) is converted to a small dense problem (of size $m \ll n$) by incorporating the well-known Arnoldi algorithm [26]. Practically, $m \leq 50$ whereas n can exceed thousands. The underlying technique is to approximate

$$w(t) \triangleq \exp(\boldsymbol{F}_\alpha t)\boldsymbol{e} = \boldsymbol{e} + \frac{(\boldsymbol{F}_\alpha t)}{1!}\boldsymbol{e} + \frac{(\boldsymbol{F}_\alpha t)^2}{2!}\boldsymbol{e} + \dots$$

by an element of the Krylov subspace $\mathcal{K}_m(\boldsymbol{F}_\alpha t, \boldsymbol{e}) = \mathrm{Span}\{\boldsymbol{e}, (\boldsymbol{F}_\alpha t)\boldsymbol{e}, \dots, (\boldsymbol{F}_\alpha t)^{m-1}\boldsymbol{e}\}$, where m is the dimension

of Krylov subspace. Thus, Krylov approximation gives

$$\tilde{w}(t) = \gamma \boldsymbol{V}_m \exp(\boldsymbol{H}_m t)\boldsymbol{e}_1,$$

where $\boldsymbol{V}_m \in \mathbb{R}^{n\times m}$ and $\boldsymbol{H}_m \in \mathbb{R}^{m\times m}$ are, respectively, the orthonormal basis of the Krylov subspace $\mathcal{K}_m(\boldsymbol{F}_\alpha t, \boldsymbol{e})$ and the upper Hessenberg matrix resulting from Arnoldi process, and $\gamma \triangleq ||\boldsymbol{e}||_2$.[9] Usually, the dimension m of $\boldsymbol{H}_m$ is much smaller than that n of $\boldsymbol{F}_\alpha$. As m increases, Krylov approximation becomes more accurate. It is proven that the error in the approximation behaves like [21]

$$||w(t) - \tilde{w}(t)||_2 = O\left(e^{-||\boldsymbol{F}_\alpha||_2 t}\left(||\boldsymbol{F}_\alpha||_2 t\frac{e}{m}\right)^m\right),$$

for $m \geq 2t||\boldsymbol{F}_\alpha||_2$. EXPOKIT is a software package that provides a set of ready-to-use routines (in Matlab and Fortran 77) for computing matrix exponentials and is available online [27]. One of the component, EXPOKIT/dmexpv(m), implements Krylov subspace technique to cope with a large sparse matrix in Markov chains. A full description can be found in [23]. The computational complexity involved in our work can be reduced with help of this software.

7 Conclusion

In this paper, we characterize the probabilistic guarantee on the time for information spread in opportunistic networks by developing a two-dimensional CTMC-based analytical framework and introducing the metric $G_{\alpha,\beta}$. Our characterization includes understanding of the temporal scaling behavior of information spread through a set of various spread measures. We also introduce various examples of application scenarios and demonstrate with Shanghai taxi traces that our framework enables us to estimate proper amount of resource for information spread by providing the detailed statistics of the guaranteed time for given penetration targets. We believe our framework can be viewed as an important first step in the design of highly sophisticated acceleration methods for information spread or prevention methods for epidemics.

Acknowledgments

This research has been supported in part by grants from the NSF CNS-1421576, CNS-1065136, Army Research Lab MURI W911NF-12-1-0385, Defense Thrust Reduction Agency HDTRA1-14-1-0058, and Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2014R1A1A2057793) and the Ministry of Science, ICT & Future Planning (NRF-2014R1A1A1006330).

References

[1] M. J. Keeling, M. E. J. Woolhouse, D. J. Shaw, L. Matthews, M. Chase-Topping, D. T. Haydon, S. J. Cornell, J. Kappey, J. Wilesmith, and B. T. Grenfell, "Dynamics of the 2001 uk foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape," *Science*, vol. 294, no. 5543, pp. 813–817, October 2001.

[2] S. H. Sellke, N. B. Shroff, and S. Bagchi, "Modeling and automated containment of worms," *IEEE Transactions on Dependable and Secure Computing*, vol. 5, pp. 71–86, April 2008.

[3] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance modeling of epidemic routing," *Elsevier Computer Networks*, vol. 51, no. 10, pp. 2867–2891, July 2007.

[4] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*. Springer, 2000.

[5] M. J. Keeling and K. T. Eames, "Networks and epidemic models," *Journal of Royal Society Interface*, vol. 2, no. 4, pp. 295–307, September 2005.

[6] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, "Small but slow world: How network topology and burstiness slow down spreading," *Physical Review E*, vol. 83, no. 2, pp. 025102-1–025102-4, February 2011.

[7] P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, February 2009.

[8] Z. Yang, A.-X. Cui, and T. Zhou, "Impact of heterogeneous human activities on epidemic spreading," *Physica A*, vol. 390, pp. 4543–4548, 2011.

[9] C. C. Zou, W. Gong, and D. Towsley, "Code red worm propagation modeling and analysis," in *Proceedings of the 9th ACM Conference on Computer and Communications Security (CCS)*, 2002.

[10] V. Conan, J. Leguay, and T. Friedman, "Characterizing pairwise inter-contact patterns in delay tolerant networks," in *Proceedings of Intl. Conference on Autonomic Computing and Communication Systems*, 2007.

[11] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *Proceedings of ACM Intl. Symposium on Mobile Ad Hoc Networking and Computing*, 2009.

[12] C.-H. Lee, J. Kwak, and D. Y. Eun, "Characterizing link connectivity for opportunistic mobile networking: Does mobility suffice?" in *Proceedings of IEEE INFOCOM*, 2013.

[13] S. Ioannidis, A. Chaintreau, and L. Massoulie, "Optimal and scalable distribution of content updates over a mobile social network," in *Proceedings of IEEE INFOCOM*, 2009.

[14] Y. Kim, K. Lee, N. B. Shroff, and I. Rhee, "Providing probabilistic guarantees on the time of information spread in opportunistic networks," in *Proceedings of IEEE INFOCOM*, 2013.

[15] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Courier Dover Publications, 1981.

[16] P. Bremaud, *Markov Chains*. Springer, 2008.

[17] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine, "Routing for vehicle-based disruption tolerant networks," in *Proceedings of IEEE INFOCOM*, 2006.

[18] S. J. U. Traffic Information Grid Team, Grid Computing Center, "Shanghai taxi trace data," http://wirelesslab.sjtu.edu.cn/.

[19] K. Lee, Y. Yi, J. Jeong, H. Won, I. Rhee, and S. Chong, "Max-contribution: On optimal resource allocation in delay tolerant networks," in *Proceedings of IEEE INFOCOM*, 2010.

[20] R. B. Sidje and W. J. Stewart, "A numerical study of large sparse matrix exponentials arising in Markov chains," *Computational Statistics & Data Analysis*, vol. 29, no. 3, pp. 345–368, 1999.

[21] M. Hochbruck and C. Lubich, "On Krylov subspace approximations to the matrix exponential operator," *SIAM Journal on Numerical Analysis*, vol. 34, no. 5, pp. 1911–1925, 1997.

[22] Y. Saad, "Analysis of some Krylov subspace approximations to the matrix exponential operator," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 209–228, 1992.

[23] R. B. Sidje, "EXPOKIT. A software package for computing matrix exponentials," *ACM Trans. Math. Softw.*, vol. 24, no. 1, pp. 130–156.

[24] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM Review*, vol. 45, no. 1, pp. 3–49, 2003.

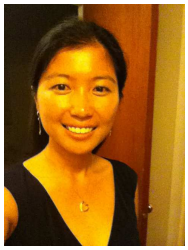
[25] E. Gallopoulos and Y. Saad, "Efficient solution of parabolic equations by Krylov approximation methods," *SIAM Journal on Scientific and Statistical Computing*, vol. 13, no. 5, pp. 1236–1264, 1992.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*. The Johns Hopkins University Press, 2013.

[27] http://www.maths.uq.edu.au/expokit/.

[28] G. Strang, *Linear Algebra and its Applications*, 3rd ed. Philadelphia, PA: Saunders, 1988.

9. $||\cdot||_2$ denotes the Euclidean norm or its induced matrix norm.



Yoora Kim (S'05-M'09) received her B.S., M.S. and Ph.D. degrees in Mechanical Engineering, Applied Mathematics and Mathematical Sciences from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2003, 2005 and 2009, respectively.

She is currently an Assistant Professor with the Department of Mathematics at University of Ulsan, Ulsan, South Korea. Prior to joining University of Ulsan, she has been with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, as a Post-Doctoral Research Associate. Her research interests include modeling, design, and performance evaluation of communication systems and scheduling and resource allocation problems in wireless networks under various stochastic dynamics.

Kyunghan Lee (S07-A10) received his B.S., M.S., and Ph.D. degrees in Electrical Engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2002, 2004, and 2009, respectively. He is currently an assistant professor in the school of electrical and computer engineering at UNIST (Ulsan National Institute of Science and Technology), Ulsan, South Korea. Prior to joining UNIST, he has been with the Department of Computer Science, North Carolina State University, Raleigh, US as a senior research scholar. His research interests include the areas of human mobility modeling, delay tolerant networking, information centric networking, context-aware system design, and cloud-powered network system design.

Ness B. Shroff (S91-M93-SM01-F07) received his Ph.D. degree from Columbia University, NY in 1994 and joined Purdue university immediately thereafter as an Assistant Professor. At Purdue, he became Professor of the school of Electrical and Computer Engineering in 2003 and director of CWSA in 2004, a university-wide center on wireless systems and applications. In July 2007, he joined Ohio State University where he holds the Ohio Eminent Scholar Chaired Professorship in Networking and Communications. His research interests span the areas of wireless and wireline communication networks. He is especially interested in fundamental problems in the design, performance, control, and security of these networks. Dr. Shroff is a past editor for IEEE/ACM Trans. on Networking and the IEEE Communications Letters. He currently serves on the editorial board of the Computer Networks Journal. He has served on the technical and executive committees of several major conferences and workshops. He was the TPC co-chair of IEEE INFOCOM03, ACM Mobihoc08, and general chair of IEEE CCW99 and WICON08. Dr. Shroff is a fellow of the IEEE. He has received numerous awards for his work, including two best paper awards at IEEE INFOCOM (in 2006 and 2008), the flagship conference of the field. He has also received the IEEE WiOPT 2013, IEEE WiOpt 2012, and the IWQoS best student paper award, the 2005 best paper of the year award for the Journal of Communications and Networking, the 2003 best paper of the year award for Computer Networks, and the NSF CAREER award in 1996 (his IEEE INFOCOM 2013 and IEEE INFOCOM 2005 papers ware selected as runner-up papers). Dr. Shroff is among the list of highly cited researchers from Thomson Reuters ISI and in Thomson Reuters Book on The Worlds Most Influential Scientific Minds in 2014. In 2014, he received the IEEE INFOCOM achievement award for seminal contributions to scheduling and resource allocation in wireless networks.

APPENDIX A
PROOF OF LEMMA 1

Suppose that at time t_0, the system enters state $(L(t_0), J(t_0)) = (i, j)$. Let $\mathcal{I}_k(i,j)$ and $\mathcal{S}_k(i,j)$ denote, respectively, the index sets of infected nodes and susceptible nodes in class k for the given level-phase (i,j). If $i = N$, then $\mathcal{S}_k(i,j) = \varnothing$ for all $k = 1, \ldots, K$, and thus the system stops evolving. From now on, we consider $i \leq N-1$. Define

$$\Omega(i,j) \triangleq \big\{m : \mathcal{X}^{-1}(i+1,m) = \mathcal{X}^{-1}(i,j) + \boldsymbol{e}_k \text{ for some } k\big\},$$

that denotes the collection of phases to which the very next transition from state (i,j) can occur.

For $j' \in \{1, 2, \ldots, |\mathcal{F}_{i+1}|\}$, let $Y_{(i,j)\to(i+1,j')}$ denote the time required to jump from (i,j) to $(i+1,j')$. If $j' \notin \Omega(i,j)$, then the transition never happens, i.e.,

$$Y_{(i,j)\to(i+1,j')} = \infty, \qquad \text{if } j' \notin \Omega(i,j). \tag{33}$$

If $j' \in \Omega(i,j)$, then there exists k' such that $\mathcal{X}^{-1}(i+1,j') = \mathcal{X}^{-1}(i,j) + \boldsymbol{e}_{k'}$, and $Y_{(i,j)\to(i+1,j')}$ becomes the time to have one more infected node in class k'. Note that at time t_0, there are $\mathcal{I}_k(i,j)$ infected nodes in each class k and $\mathcal{S}_{k'}(i,j)$ susceptible nodes in class k'. Hence, we obtain

$$Y_{(i,j)\to(i+1,j')} = \min_{a \in \bigcup_{k=1}^K \mathcal{I}_k(i,j),\ b \in \mathcal{S}_{k'}(i,j)} \{M^{\text{eff}}_{a,b}\}, \qquad \text{if } j' \in \Omega(i,j). \tag{34}$$

Hence, by (2), $Y_{(i,j)\to(i+1,j')}$ becomes an exponential random variable. The sojourn time of state (i,j) is the minimum among $Y_{(i,j)\to(i+1,j')}$ for all $j' \in \{1, 2, \ldots, |\mathcal{F}_{i+1}|\}$. Thus, by (33) and (34), it follows an exponential distribution. Therefore, the joint level-phase process is a CTMC.

APPENDIX B
PROOF OF LEMMA 3

In this proof, we use the same notation as in the proof of Lemma 1, unless otherwise mentioned. Let $[\boldsymbol{R}_i]_{j,j'}$ denote the (j,j')th element of the matrix $\boldsymbol{R}_i$. Then, it represents the transition rate from (i,j) to $(i+1,j')$. From (33), we obtain $[\boldsymbol{R}_i]_{j,j'} = 0$ if $j' \notin \Omega(i,j)$. In addition, from (34), the transition rate $[\boldsymbol{R}_i]_{j,j'}$ for the case $j' \in \Omega(i,j)$ can be derived as

$$\begin{aligned}[\boldsymbol{R}_i]_{j,j'} &= \sum_{a \in \bigcup_{k=1}^K \mathcal{I}_k(i,j),\ b \in \mathcal{S}_{k'}(i,j)} \xi_{k(a),k(b)} \\ &= |\mathcal{S}_{k'}(i,j)| \sum_{k=1}^K |\mathcal{I}_k(i,j)| \xi_{k,k'}\end{aligned} \tag{35}$$

Let $\boldsymbol{f} \triangleq \mathcal{X}^{-1}(i,j)$ and $\boldsymbol{g} \triangleq \mathcal{X}^{-1}(i+1,j')$. Then, we can rewrite the condition $j' \in \Omega(i,j)$ as $\boldsymbol{g} = \boldsymbol{f} + \boldsymbol{e}_{k'}$. In addition, $|\mathcal{I}_k(i,j)|$ and $|\mathcal{S}_k(i,j)|$ can be expressed as $|\mathcal{I}_k(i,j)| = (\boldsymbol{f})_k$ and $|\mathcal{S}_k(i,j)| = N_k - |\mathcal{I}_k(i,j)| = (\boldsymbol{N} - \boldsymbol{f})_k$, respectively. Hence, (35) is simplified as

$$\begin{aligned}|\mathcal{S}_{k'}(i,j)| \sum_{k=1}^K |\mathcal{I}_k(i,j)| \xi_{k,k'} &= (\boldsymbol{N} - \boldsymbol{f})_{k'} \sum_{k=1}^K (\boldsymbol{f})_k \xi_{k,k'} \\ &= (\boldsymbol{N} - \boldsymbol{f})_{k'} (\boldsymbol{f}\boldsymbol{\Lambda})_{k'}.\end{aligned} \tag{36}$$

Combining (10) and (36) gives

$$[\boldsymbol{R}_i]_{j,j'} = \lambda(\boldsymbol{f}, \boldsymbol{g}) = \lambda(\mathcal{X}^{-1}(i,j), \mathcal{X}^{-1}(i+1,j')).$$

APPENDIX C
PROOF OF REMARK 1

Since $\boldsymbol{F}_\alpha$ is an upper triangular matrix, its diagonal entries become the eigenvalues of $\boldsymbol{F}_\alpha$ [28]. If the diagonal entries of $\boldsymbol{F}_\alpha$ are all distinct (that is, $\boldsymbol{F}_\alpha$ has no repeated eigenvalues), then the matrix $\boldsymbol{F}_\alpha$ has M linearly independent eigenvectors, and consequently it is diagonalizable as $\boldsymbol{F}_\alpha = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{-1}$ where $\boldsymbol{D} \triangleq diag\big((-\zeta_1, \ldots, -\zeta_M)\big)$ [28]. Thus, we obtain

$$\exp(\boldsymbol{F}_\alpha t) = \exp(\boldsymbol{V}(\boldsymbol{D}t)\boldsymbol{V}^{-1}) = \boldsymbol{V}\exp(\boldsymbol{D}t)\boldsymbol{V}^{-1},$$

where the second equality follows from the property of the matrix exponential that $\exp(\boldsymbol{Y}\boldsymbol{X}\boldsymbol{Y}^{-1}) = \boldsymbol{Y}\exp(\boldsymbol{X})\boldsymbol{Y}^{-1}$ for an invertible matrix $\boldsymbol{Y}$. Furthermore, $\exp(\boldsymbol{D}t)$ reduces to

$$\exp(\boldsymbol{D}t) = diag\big((e^{-\zeta_1 t}, \ldots, e^{-\zeta_M t})\big).$$

APPENDIX D
PROOF OF THEOREM 2

In this proof, we introduce a symbol $\wedge$ over a variable, say $\hat{x}$, to denote the value of x computed with the scaled infection rate $\lambda\xi_{k,k'}$. Since the infection rate $\xi_{k,k'}$ is multiplied by γ for all $1 \leq k, k' \leq K$, the rate matrix satisfies $\hat{\boldsymbol{\Lambda}} = \gamma\boldsymbol{\Lambda}$. Hence, by (35) and (36) in the proof of Lemma 3, we obtain $\hat{\boldsymbol{R}}_i = \gamma\boldsymbol{R}_i$ for all $1 \leq i \leq N-1$, which in turn gives $\hat{\boldsymbol{F}}_\alpha = \gamma\boldsymbol{F}_\alpha$ by (6). Applying $\hat{\boldsymbol{F}}_\alpha = \gamma\boldsymbol{F}_\alpha$ to (12) in Theorem 1, we obtain

$$\begin{aligned}\mathrm{P}(\hat{T}_\alpha > t) &= \boldsymbol{h}_\alpha(\boldsymbol{s})\exp(\hat{\boldsymbol{F}}_\alpha t)\boldsymbol{e} = \boldsymbol{h}_\alpha(\boldsymbol{s})\exp(\boldsymbol{F}_\alpha(\gamma t))\boldsymbol{e} \\ &= \mathrm{P}(T_\alpha > \gamma t) = \mathrm{P}(\gamma^{-1}T_\alpha > t).\end{aligned} \tag{37}$$

Since (37) holds for any $t \geq 0$, we have $\hat{T}_\alpha \overset{\mathrm{d}}{=} \gamma^{-1}T_\alpha$. Let $\hat{H}_\alpha(t) \triangleq \mathrm{P}(\hat{T}_\alpha \leq t)$ denote the CDF of $\hat{T}_\alpha$. Then, (37) gives $\hat{H}_\alpha(t) = H_\alpha(\gamma t)$, and thus we have $\hat{H}^{-1}_\alpha(t) = \gamma^{-1}H^{-1}_\alpha(t)$. From (13) of Theorem 1, $\hat{G}_{\alpha,\beta}$ is then obtained by $\hat{G}_{\alpha,\beta} = \hat{H}^{-1}_\alpha(\beta) = \gamma^{-1}H^{-1}_\alpha(\beta) = \gamma^{-1}G_{\alpha,\beta}$. Since $\hat{T}_\alpha \overset{\mathrm{d}}{=} \gamma^{-1}T_\alpha$, we have $\mathrm{E}[\hat{T}_\alpha] = \gamma^{-1}\mathrm{E}[T_\alpha]$. Hence, $\hat{R}_{\alpha,\beta} = \frac{\hat{G}_{\alpha,\beta}}{\mathrm{E}[\hat{T}_\alpha]} = R_{\alpha,\beta}$.

APPENDIX E
PROOF OF EQUATION (21)

Since $\sum_k I_k(t)$ takes on only natural numbers from 1 to N, $\mathcal{M}(t)$ can be obtained by $\mathcal{M}(t) = \sum_{i=1}^N \mathrm{P}(\sum_k I_k(t) \geq i)$. Note that by Definition 1, the event $\{\sum_k I_k(t) \geq i\}$ is equivalent to $\{T_{i/N} \leq t\}$. Hence, we have $\mathcal{M}(t) = \sum_{i=1}^N \mathrm{P}(T_{i/N} \leq t)$. Similarly, $\hat{\mathcal{M}}(t)$ is given by $\hat{\mathcal{M}}(t) = \sum_{i=1}^N \mathrm{P}(\hat{T}_{i/N} \leq t) = \sum_{i=1}^N \mathrm{P}(T_{i/N} \leq \gamma t)$, where the second equality comes from Theorem 2. Therefore, we obtain $\hat{\mathcal{M}}(t) = \mathcal{M}(\gamma t)$.

APPENDIX F
PROOF OF THEOREM 3

In this proof, we add a subscript N or $N+1$ to T_α and $G_{\alpha,\beta}$ in order to explicitly denote the underlying network size.

Proof of Theorem 3. (1): We first consider a network with N nodes. From (15) with $\alpha = 1$ and $s_1 = 1$, we have

$$\mathrm{E}[T_{\alpha,N}] = \sum_{i=1}^{N-1} \frac{1}{\boldsymbol{R}_i} = \xi^{-1} \sum_{i=1}^{N-1} \frac{1}{i(N-i)}. \tag{38}$$

Similarly, for a network with $N+1$ nodes, we have $\mathrm{E}[T_{\alpha,N+1}] = \xi^{-1}\sum_{i=1}^{N}\frac{1}{i(N+1-i)}$. Hence, for any $N \in \mathbb{N}$, we obtain

$$\begin{aligned}\mathrm{E}[T_{\alpha,N+1}] - \mathrm{E}[T_{\alpha,N}] &= \xi^{-1}\bigg(\frac{1}{N} + \sum_{i=1}^{N-1}\frac{1}{N-i}\Big(\frac{1}{i+1}-\frac{1}{i}\Big)\bigg)\\ &< \xi^{-1}\bigg(\frac{1}{N} + \frac{1}{N-1}\sum_{i=1}^{N-1}\Big(\frac{1}{i+1}-\frac{1}{i}\Big)\bigg)\\ &= \xi^{-1}\bigg(\frac{1}{N}-\frac{1}{N}\bigg) = 0. \end{aligned} \tag{39}$$

That is, the average α-completion time $\mathrm{E}[T_{\alpha,N}]$ is strictly decreasing with an increasing N.

Next, we prove (22). Suppose that N is an odd number. When N is an even number, we can prove similarly and thus omit the detailed proof. For an odd N, $\mathrm{E}[T_{\alpha,N}]$ in (38) can be rewritten as

$$\mathrm{E}[T_{\alpha,N}] = 2\xi^{-1}\sum_{i=1}^{\frac{N-1}{2}}\frac{1}{i(N-i)}. \tag{40}$$

Let f be a function defined by $f(x) \triangleq \frac{1}{x(N-x)}$ for $0 < x < N$. Since f is a strictly decreasing convex function for $0 < x \le \frac{N}{2}$, the expectation in (40) is bounded above as

$$\mathrm{E}[T_{\alpha,N}] = 2\xi^{-1}\sum_{i=1}^{\frac{N-1}{2}} f(i) \le 2\xi^{-1}\bigg\{f(1) + \int_1^{\frac{N-1}{2}} f(x)\,\mathrm{d}x\bigg\}.$$

It is straightforward to obtain $f(1) + \int_1^{\frac{N-1}{2}} f(x)\,\mathrm{d}x = \frac{1}{N-1} + \frac{1}{N}\ln\frac{(N-1)^2}{N+1}$, which gives $\mathrm{E}[T_{\alpha,N}] = O(\xi^{-1}N^{-1}\ln N)$. By the same reason, the expectation in (40) is bounded below as

$$\mathrm{E}[T_{\alpha,N}] \ge 2\xi^{-1}\int_1^{\frac{N+1}{2}} f(x)\,\mathrm{d}x = 2(\xi N)^{-1}\ln(N+1),$$

which gives $\mathrm{E}[T_{\alpha,N}] = \Omega(\xi^{-1}N^{-1}\ln N)$. Therefore, we have $\mathrm{E}[T_{\alpha,N}] = \Theta(\xi^{-1}N^{-1}\ln N)$.

Proof of Theorem 3. (2): For two random variables A and B, let $A \preceq B$ denote $\mathrm{P}(A > x) \le \mathrm{P}(B > x)$ for all $x \in \mathbb{R}$. From (15), it is clear that $T_{\alpha,N} \succeq Z_1$, i.e.,

$$\mathrm{P}(T_{\alpha,N} > t) \ge \mathrm{P}(Z_1 > t) = \exp(-\xi(N-1)t). \tag{41}$$

Similarly, we have $T_{\alpha,N+1} \preceq \mathrm{Erlang}(N,\xi N)$, where $\mathrm{Erlang}(k,\lambda)$ denotes the Erlang random variable with shape $k \in \mathbb{N}$ and rate $\lambda > 0$. Hence,

$$\begin{aligned}\mathrm{P}(T_{\alpha,N+1} > t) &\le \mathrm{P}(\mathrm{Erlang}(N,\xi N) > t)\\ &= \exp(-\xi N t)\sum_{i=0}^{N-1}\frac{(\xi N t)^i}{i!}.\end{aligned} \tag{42}$$

From (42) and (41), we obtain

$$\lim_{t\to\infty}\frac{\mathrm{P}(T_{\alpha,N+1} > t)}{\mathrm{P}(T_{\alpha,N} > t)} \le \lim_{t\to\infty}\exp(-\xi t)\sum_{i=0}^{N-1}\frac{(\xi N t)^i}{i!} = 0.$$

Accordingly, there exists $t_0\,(\ge 0)$ such that $\mathrm{P}(T_{\alpha,N+1} > t) < \mathrm{P}(T_{\alpha,N} > t)$ for all $t \ge t_0$. Let $\beta_0 \triangleq \mathrm{P}(T_{\alpha,N} \le t_0)$. Then, $\beta_0 < 1$ and, as shown in Fig. 12, we have $G_{\alpha,\beta,N} > G_{\alpha,\beta,N+1}$ for $\beta \ge \beta_0$. That is, the guaranteed time $G_{\alpha,\beta,N}$ is strictly decreasing with an increasing N for $\beta \ge \beta_0$.

Next, we prove (23). Suppose that N is an odd number. When N is an even number, we can prove similarly and thus omit the detailed proof. For an odd N, the α-completion time $T_{\alpha,N}$ in (15) can be rewritten as

$$T_{\alpha,N} = \sum_{i=1}^{\frac{N-1}{2}} Z_i + \sum_{i=\frac{N+1}{2}}^{N-1} Z_i = \sum_{i=1}^{\frac{N-1}{2}} Z_i + \sum_{i=1}^{\frac{N-1}{2}} Z_{N-i}. \tag{43}$$

Then, we can derive upper and lower bounds on $T_{\alpha,N}$ as

$$T_{\text{lower}} \preceq T_{\alpha,N} \preceq T_{\text{upper}}, \tag{44}$$

where $T_{\text{upper}} \triangleq \sum_{i=1}^{N-1} Z_i^u$ and $T_{\text{lower}} \triangleq \sum_{i=1}^{N-1} Z_i^l$, and where Z_i^u and Z_i^l are independent exponential random variables with rates $\xi N i/4$ and $\xi N i$, respectively. Note that the rate of Z_i is greater than that of Z_{2i}^u if and only if $i \le \frac{N}{2}$. Hence, we have

$$Z_i \preceq Z_{2i}^u \quad \text{for } i = 1,2,\ldots,(N-1)/2. \tag{45}$$

Since $Z_{N-i} \overset{\mathrm{d}}{=} Z_i$ and $Z_{2i}^u \preceq Z_{2i-1}^u$ for all i, we further have

$$Z_{N-i} \preceq Z_{2i-1}^u \quad \text{for } i = 1,2,\ldots,(N-1)/2. \tag{46}$$

Combining (45) and (46) yields $\sum_{i=1}^{\frac{N-1}{2}} Z_i + \sum_{i=1}^{\frac{N-1}{2}} Z_{N-i} \preceq \sum_{i=1}^{N-1} Z_i^u$. This proves $T_{\alpha,N} \preceq T_{\text{upper}}$. By using a similar approach, we can prove $T_{\text{lower}} \preceq T_{\alpha,N}$. Due to similarity, we omit the details. Let $t_{\text{upper}} \triangleq 4(\xi N)^{-1}\{\ln(N-1) - \ln(\ln\beta^{-1})\}$ and $t_{\text{lower}} \triangleq \frac{1}{4}t_{\text{upper}}$. In the following, we will show that

$$\lim_{N\to\infty}\mathrm{P}(T_{\text{lower}} > t_{\text{lower}}) = 1-\beta, \tag{47a}$$

$$\lim_{N\to\infty}\mathrm{P}(T_{\text{upper}} > t_{\text{upper}}) = 1-\beta. \tag{47b}$$

The results in (44), (47a), and (47b) imply that there exists $N_0 \in \mathbb{N}$ such that $t_{\text{lower}} \le G_{\alpha,\beta,N} \le t_{\text{upper}}$ for all $N \ge N_0$, which gives $G_{\alpha,\beta,N} = \Theta\big(\xi^{-1}N^{-1}(\ln N - \ln(\ln\beta^{-1}))\big)$.

Equations (47a) and (47b) remain to be proven. It is well-known that the sum S of n independent exponential random variables with rates $r_i\,(i = 1,\ldots,n)$ follows the generalized Erlang distribution. When $r_i \ne r_j$ for all $i \ne j$, e.g., in the case $r_i = \xi N i$, the generalized Erlang distribution is given by

$$\mathrm{P}(S > t) = \sum_{i=1}^{n}\bigg(\prod_{j=1,j\ne i}^{n}\frac{r_j}{r_j - r_i}\bigg)\exp(-r_i t). \tag{48}$$

Replacing r_i by $\xi N i$ and simplifying (48) yields

$$\mathrm{P}(T_{\text{lower}} > t) = \sum_{i=1}^{N-1}(-1)^{i-1}\binom{N-1}{i}\exp(-\xi N i t).$$

Hence, we obtain

$$\begin{aligned}\mathrm{P}(T_{\text{lower}} > t_{\text{lower}}) &= \sum_{i=1}^{N-1}(-1)^{i-1}\binom{N-1}{i}\Big(\frac{\ln\beta^{-1}}{N-1}\Big)^i\\ &= 1 - \sum_{i=0}^{N-1}\binom{N-1}{i}\Big(\frac{\ln\beta}{N-1}\Big)^i\\ &= 1 - \Big(1+\frac{\ln\beta}{N-1}\Big)^{N-1}.\end{aligned} \tag{49}$$

By taking N to ∞ in (49), we have $\lim_{N\to\infty}\mathrm{P}(T_{\text{lower}} > t_{\text{lower}}) = 1 - \exp(\ln\beta) = 1-\beta$, which proves (47a). Similarly as above, we can prove (47b) and omit the details.

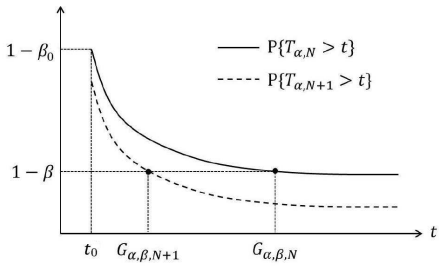


Fig. 12. Proof of Theorem 3. (2): $G_{\alpha,\beta,N} > G_{\alpha,\beta,N+1}$ for $\beta \geq \beta_0$

APPENDIX G
PROOF OF REMARK 4

Our framework is applicable for the analysis of the non-cooperative spread model. The only change occurs at Step 3 in the computation of $\boldsymbol{R}_i$. The function $\lambda(\boldsymbol{f},\boldsymbol{g})$ in (10) now should be replaced by $\lambda^o(\boldsymbol{f},\boldsymbol{g})$ as follows:

$$\lambda^o(\boldsymbol{f},\boldsymbol{g}) \triangleq \begin{cases} (\boldsymbol{N}-\boldsymbol{f})_k(\boldsymbol{s}\boldsymbol{\Lambda})_k & \text{if } \boldsymbol{g}=\boldsymbol{f}+\boldsymbol{e}_k \text{ for some } k, \\ 0 & \text{otherwise.} \end{cases}$$

With this replacement, we can prove Remark 4 by applying the technique used in the proof of Theorem 3. Due to similarity, we give only a sketch of the proof and omit the details.

Proof of Remark 4. (1): Let $T^o_{\alpha,N}$ be the α-completion time in a network consisting of N non-cooperative nodes. Then,

$$T^o_{\alpha,N} = \sum_{i=1}^{N-1} Z^o_i, \tag{50}$$

where Z^o_i are independent exponential random variables with rates ξi. Hence, we have $\mathrm{E}[T^o_{\alpha,N}] = \xi^{-1}\sum_{i=1}^{N-1}\frac{1}{i}$, which is strictly increasing with N and scales as $\Theta(\xi^{-1}\ln N)$.

Proof of Remark 4. (2): Since $T^o_{\alpha,N+1} \overset{\mathrm{d}}{=} T^o_{\alpha,N} + Z^o_N$ by (50), we have $\mathrm{P}(T^o_{\alpha,N} > t) < \mathrm{P}(T^o_{\alpha,N+1} > t)$ for all $t > 0$. Hence, $G_{\alpha,\beta,N}$ is strictly increasing with an increasing N for any $\beta \in [0,1]$. Also, by applying the formula (48), we can derive $\mathrm{P}(T^o_{\alpha,N} > t) = \sum_{i=1}^{N-1}(-1)^{i-1}\binom{N-1}{i}\exp(-\xi i t)$. Let $t^o \triangleq \xi^{-1}\{\ln(N-1) - \ln(\ln\beta^{-1})\}$. Then, we have

$$\mathrm{P}(T^o_{\alpha,N} > t^o) = 1 - \left(1 + \frac{\ln\beta}{N-1}\right)^{N-1},$$

which gives $\lim_{N\to\infty}\mathrm{P}(T^o_{\alpha,N} > t^o) = 1-\exp(\ln\beta) = 1-\beta$. Therefore, we have $G_{\alpha,\beta,N} = \Theta(\xi^{-1}(\ln N - \ln(\ln\beta^{-1})))$.

APPENDIX H
PROOF OF THEOREM 4

In the case $K=2$, we can set the phase X_i to be the number of infected nodes in class 1 for the level i, as it can specify uniquely which sample instance realizes the level. Then, we have $\mathcal{X}^{-1}(i,j) = (j, i-j)$, and Eq. (27) leads to

$$\hat{\mu}_{i,j} = \mu_i + 2\xi_{\text{inter}}\left(j - \frac{i}{2}\right)(r_1 - 1)\left(\frac{N}{2} - i\right).$$

The expectation $\hat{\mu}_i$ is thus obtained by

$$\hat{\mu}_i = \mu_i + c(r_1-1)\left(\frac{N}{2} - i\right),$$

where $c \triangleq 2\xi_{\text{inter}}(\mathrm{E}[X_i] - \frac{i}{2})$. Using mathematical induction, we prove that $c > 0$ for all $i = 1, 2, \ldots, N-1$, as follows.

We first rewrite $\mathrm{E}[X_{i+1}]$ by conditioning on X_i as

$$\mathrm{E}[X_{i+1}] = \sum_j \mathrm{P}(X_i = j)\mathrm{E}[X_{i+1}|X_i = j]. \tag{51}$$

Since X_i represents the number of infected nodes in class 1 when there are i infected nodes in the network, we have either $X_{i+1} = X_i + 1$ (i.e., newly infected node belongs to class 1) or $X_{i+1} = X_i$ (i.e., newly infected node belongs to class 2) at the very next level. Hence,

$$\begin{aligned}\mathrm{E}[X_{i+1}|X_i = j] &= (j+1)\mathrm{P}(X_{i+1} = j+1|X_i = j) \\ &\quad + j\mathrm{P}(X_{i+1} = j|X_i = j) \\ &= j + \mathrm{P}(X_{i+1} = j+1|X_i = j).\end{aligned} \tag{52}$$

Substituting (52) into (51) gives

$$\mathrm{E}[X_{i+1}] = \mathrm{E}[X_i] + \sum_j \mathrm{P}(X_i = j)\mathrm{P}(X_{i+1} = j+1|X_i = j). \tag{53}$$

In the following, we will show that the second probability on the right-hand side of (53) is bounded by

$$\mathrm{P}(X_{i+1} = j+1|X_i = j) \geq \frac{\frac{N}{2} - j}{N - i}. \tag{54}$$

Let $Y_{(i,j)\to(i+1,j')}$ denote the time required to jump from level-phase state (i,j) to $(i+1,j')$ (as we defined in Appendix A). Then, as shown in Appendices A and B, $Y_{(i,j)\to(i+1,j')}$ follows an exponential distribution whose rate is $[\boldsymbol{R}_i]_{j,j'}$ and can be obtained from (10) as

$$[\boldsymbol{R}_i]_{j,j'} = \begin{cases} (\frac{N}{2} - j)\begin{bmatrix} j & i-j \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}_{1,1} \\ \boldsymbol{\Lambda}_{2,1} \end{bmatrix} & \text{if } j' = j+1, \\ (\frac{N}{2} - i + j)\begin{bmatrix} j & i-j \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}_{1,2} \\ \boldsymbol{\Lambda}_{2,2} \end{bmatrix} & \text{if } j' = j, \end{cases}$$

where $\boldsymbol{\Lambda}_{i,j}$ denotes the (i,j)th element of the rate matrix $\boldsymbol{\Lambda}$. Note that $\boldsymbol{\Lambda}_{1,1} \geq \boldsymbol{\Lambda}_{2,1} = \boldsymbol{\Lambda}_{1,2} \geq \boldsymbol{\Lambda}_{2,2}$, and thus the rate above is bounded by

$$\begin{aligned}[\boldsymbol{R}_i]_{j,j'} &\geq (\frac{N}{2} - j)\begin{bmatrix} j & i-j \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}_{2,1} \\ \boldsymbol{\Lambda}_{2,1} \end{bmatrix} = (\frac{N}{2} - j)i\boldsymbol{\Lambda}_{2,1} \\ &\qquad \text{if } j' = j+1, \\ [\boldsymbol{R}_i]_{j,j'} &\leq (\frac{N}{2} - i + j)\begin{bmatrix} j & i-j \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}_{2,1} \\ \boldsymbol{\Lambda}_{2,1} \end{bmatrix} = (\frac{N}{2} - i + j)i\boldsymbol{\Lambda}_{2,1} \\ &\qquad \text{if } j' = j.\end{aligned}$$

Note that the event $\{X_{i+1} = X_i + 1\}$ occurs if and only if $Y_{(i,j)\to(i+1,j+1)} < Y_{(i,j)\to(i+1,j)}$. Hence,

$$\begin{aligned}&\mathrm{P}(X_{i+1} = j+1|X_i = j) \\ &\quad = \mathrm{P}(Y_{(i,j)\to(i+1,j+1)} < Y_{(i,j)\to(i+1,j)}) \\ &\quad = \frac{[\boldsymbol{R}_i]_{j,j+1}}{[\boldsymbol{R}_i]_{j,j+1} + [\boldsymbol{R}_i]_{j,j}} \\ &\quad \geq \frac{(\frac{N}{2} - j)i\boldsymbol{\Lambda}_{2,1}}{(\frac{N}{2} - j)i\boldsymbol{\Lambda}_{2,1} + (\frac{N}{2} - i + j)i\boldsymbol{\Lambda}_{2,1}} \\ &\quad = \frac{\frac{N}{2} - j}{N - i}.\end{aligned}$$

This proves (54). Combining (53) and (54) yields

$$\begin{aligned} \mathrm{E}[X_{i+1}] &\geq \mathrm{E}[X_i] + \sum_j \mathrm{P}(X_i = j)\frac{\frac{N}{2} - j}{N - i} \\ &= \mathrm{E}[X_i] + \frac{\frac{N}{2} - \mathrm{E}[X_i]}{N - i} \\ &= \frac{(N - i - 1)\mathrm{E}[X_i] + \frac{N}{2}}{N - i}. \end{aligned} \tag{55}$$

If $\mathrm{E}[X_i] > \frac{i}{2}$, then the last equation in (55) is further bounded by

$$\mathrm{E}[X_{i+1}] > \frac{(N - i - 1)\frac{i}{2} + \frac{N}{2}}{N - i} = \frac{i + 1}{2}.$$

That is, if $\mathrm{E}[X_i] > \frac{i}{2}$, then $\mathrm{E}[X_{i+1}] > \frac{i+1}{2}$. Moreover, since the seeder is selected from class 1, we have $\mathrm{E}[X_1] = 1 > \frac{1}{2}$. By mathematical induction, we therefore have $\mathrm{E}[X_i] > \frac{i}{2}$ for all $i = 1, 2, \ldots, N - 1$. This competes the proof.