Research Reports on Mathematical and Computing Sciences

Subexponential interval graphs generated by immigration-death processes

Naoto Miyoshi, Mariko Ogura, Takeya Shigezumi and Ryuhei Uehara

December 2008, B–451

Department of Mathematical and Computing Sciences Tokyo Institute of Technology

series B: Operations Research

Subexponential interval graphs generated by immigration-death processes

Naoto Miyoshi¹^{*}, Mariko Ogura¹, Takeya Shigezumi¹ and Ryuhei Uehara²

¹Department of Mathematical and Computing Sciences Tokyo Institute of Technology

²School of Information Science Japan Advanced Institute of Science and Technology

Abstract

Scale-free graphs have recently attracted much attention since so-called scale-free phenomena have really appeared in various physical and social networks, where a graph is said to be scale-free if the distribution of degrees has a power-law tail. In the previous paper, the authors proposed a simple model of random interval graphs generated by immigration-death processes (also known as $M/G/\infty$ queueing processes) and showed that, when the interval lengths follow a power-law distribution, the generated interval graph is scale-free in the steady state. In this paper, we generalize this result to the model where the distribution of interval lengths is subexponential and provide a condition under which the stationary degree distribution is also subexponential. Furthermore, we consider the conditional expectation of the cluster coefficient of a vertex given its degree and derive its limit as the degree goes to infinity under the same condition as that for obtaining the tail asymptotics of the degree distribution.

Keywords. Scale-free graphs, interval graphs, immigration-death processes, $M/G/\infty$ queues, subexponential distributions, square-root insensitive distributions.

^{*}Corresponding author: 2-12-1-W8-52 Ookayama, Tokyo 152-8552, Japan. E-mail: miyoshi@is.titech.ac.jp

1 Introduction

Scale-free graphs have recently attracted much attention since so-called scale-free phenomena have really appeared in various physical and social networks, where we say that a graph is scale-free if the distribution of degrees (the numbers of edges incident to respective vertices) has a power-law tail. To throw light on such phenomena in real world, many models of random graphs realizing the scale-free property have so far been proposed and investigated since the early works by Watts & Strogatz ([16]) and Barabási & Albert ([3]). Among them, the authors' previous work [12] proposed a model of random interval graphs generated by immigration-death processes (also known as $M/G/\infty$ queueing processes; see, e.g., Cox & Isham [6, Section 5.6]) and showed that, when the interval lengths follow a power-law distribution, the generated interval graph is scale-free. Here, a graph G = (V, E) is said to be an interval graph when G has an interval representation \mathcal{I} , the set of intervals on the real line, such that each vertex $v \in V$ corresponds to an interval $I_v \in \mathcal{I}$ and there is an edge $(u, v) \in E$ connecting two vertices $u, v \in V$ if and only if $I_u \cap I_u \neq \emptyset$. In [12], each interval is then given as the period of a customer's stay in the $M/G/\infty$ queue; that is, the interval lengths correspond to the service (sojourn) times of customers. Interval graphs form one of the most important classes of graphs and have been studied thoroughly in the graph theory (see, e.g., Golumbic [10, Chapter 8]).

In the current paper, we generalize the result in [12] to the model where the distribution of interval lengths is subexponential (see, e.g., Embrechts et al. [8, Sections 1.3 & A3] or Rolski et al. [14, Section 2.5] for subexponential distributions). Namely, we consider random interval graphs generated by immigrationdeath processes with subexponential lifetime distributions, which we call subexponential interval graphs. We provide a condition on the lifetime (service time or interval length) distribution F under which the stationary degree distribution of the generated interval graph has an equivalent tail to that of $1 - F(x/\lambda)$; that is, the stationary degree distribution is also subexponential, where λ denotes the arrival rate of intervals. This derivation is based on the recent results on sampling of a stochastic process at random times according to subexponential distributions (see Asmussen et al. [2], Foss & Korshunov [9] and Jelenković et al. [11]). Furthermore, we consider the conditional expectation of the cluster coefficient of a vertex given its degree. In a given graph, the cluster coefficient of a vertex represents the fraction of couples of its neighbors such that the couple is connected by an edge, and it is observed that many scale-free graphs have high cluster coefficients (see, e.g., Newman [13]). We derive the limit of the conditionally expected cluster coefficient given the degree as the degree goes to infinity under the same condition as that for obtaining the tail asymptotics of the degree distribution. We will see that a simple example indeed exhibits the high cluster coefficient in such a limit.

The rest of the paper is organized as follows. In the next section, we describe the immigrationdeath process and present an algorithm constructing random interval graphs based on that process. In section 3, we analyze the subexponential interval graph generated by the immigration-death process in the steady state, where we discuss the tail asymptotics of the stationary degree distribution and the limit of the conditionally expected cluster coefficient given the degree of a vertex as the degree goes to infinity. Finally, Section 4 makes a concluding remark.

2 Interval graphs generated by immigration-death processes

In this section, we describe an immigration-death process (also known as an $M/G/\infty$ queueing process; see, e.g., [6, Section 5.6]) and construct a random interval graph based on that process. Let $\{T_n\}_{n\in\mathbb{Z}}$ denote a random sequence on \mathbb{R}_+ satisfying $0 = T_0 < T_1 < T_2 < \cdots$, at each of which an individual arrives and enters a system. We refer to the individual arriving at T_n as individual $n \in \mathbb{Z}_+$. The lifetime (service time) of individual n in the system is denoted by L_n (≥ 0), so that the individual n departs from the system at $T_n + L_n$. We assume that $\{T_n\}_{n \in \mathbb{N}}$ follows a homogeneous Poisson process with intensity $\lambda \in (0, \infty)$ and $\{L_n\}_{n \in \mathbb{Z}_+}$ is a sequence of mutually independent nonnegative random variables according to an identical distribution F, where $\{T_n\}_{n \in \mathbb{N}}$ and $\{L_n\}_{n \in \mathbb{Z}_+}$ are also independent each other. The distribution F is assumed to have its mean $\mu^{-1} = \int_0^\infty \overline{F}(x) \, dx < \infty$, where $\overline{F}(x) = 1 - F(x), x \ge 0$. Let $I_n = [T_n, T_n + L_n], n \in \mathbb{Z}_+$, and $Z(t) = \sum_{n \in \mathbb{Z}_+} \mathbf{1}_{I_n}(t), t \ge 0$, where $\mathbf{1}_A$ denotes the indicator function for set A. Note that Z(t) represents the number of individuals in the system at time $t \ge 0$ (the reason for the choice of $I_n = [T_n, T_n + L_n]$ rather than $[T_n, T_n + L_n)$ is clarified in Remark 1 below). It is well known that $\{Z(t)\}_{t\ge 0}$ has a stationary regime when both λ and μ are nonzero and finite (see, e.g., [6, Section 5.6] or Takács [15, Section 3.2]).

Based on this immigration-death process, we consider a random interval graph $G_0 = (V_0, E_0)$ with interval representation $\mathcal{I}_0 = \{I_n\}_{n \in V_0}$, where $V_0 = \{0, 1, \ldots, n_0 - 1\}$ and n_0 is a predetermined positive integer. Namely, each individual in V_0 corresponds to a vertex of the graph and two vertices n and $m \in V_0$ are connected by an edge if and only if $I_n \cap I_m \neq \emptyset$. Note that such a graph has no multiedges or self-loops. Given n_0 , λ and distribution F, a simple algorithm constructing such random interval graphs is as follows, where Sample(F) denotes the sampled value extracted according to F and $Exp(\lambda)$ denotes the exponential distribution with parameter λ .

procedure generate_graph (n_0, λ, F)

 $T = 0, V = \{0\}, E = \emptyset, Q = \{0\}, U_0 = Sample(F), n = 1; \{Q: \text{Set of individuals in the system}; U_n: \text{departure time of individual } n\}$

while
$$n < n_0$$
 do

 $T \leftarrow T + Sample(\text{Exp}(\lambda)); V \leftarrow V \cup \{n\}; \quad \{\text{Individual } n \text{ arrives} \Rightarrow \text{Add vertex } n\}$ for *i* such that $i \in Q$ do
if $U_i < T$ then $Q \leftarrow Q \setminus \{i\};$ else $E \leftarrow E \cup \{(i,n)\}; \quad \{\text{Individual } i \text{ is still in the system at individual } n\text{'s arrival} \Rightarrow \text{Add edge } (i,n)\}$ end if
end for $U_n = T + Sample(F); Q \leftarrow Q \cup \{n\};$ $n \leftarrow n + 1;$ end while

Remark 1 When n_0 is large, the random interval graph constructed by the above algorithm ends up having many connected components with random but finite sizes and the size of any connected component does not tend to infinity even as $n_0 \to \infty$. Against such a feature, one may want to have one big connected graph. In such a case, it can be realized by adding extra intervals $J_n = [A_n, B_n]$, $n \in \mathbb{N}$, where $A_n = \inf\{t > B_{n-1} \mid Z(t) = 0\}$ and $B_n = \inf_{k \in \mathbb{N}}\{T_k > A_n\}$ with $B_0 = 0$; that is, $\{A_n\}_{n \in \mathbb{N}}$ and $\{B_n\}_{n \in \mathbb{N}}$ represent, respectively, the beginnings and the ends of idle periods for the corresponding $M/G/\infty$ queue. Two connected components in the original graph G_0 are then connected through a vertex with two edges in the modified graph \tilde{G}_0 , which has the interval representation $\{I_n\}_{n \in V_0} \cup \{J_m\}_{m \in U_0}$ with $U_0 = \{m \in \mathbb{N} : B_m < T_{n_0}\}$ (note that, for any $m \in U_0$, there exists an $n \in V_0$ such that $A_m = T_n + L_n$ and $B_m = T_{n+1}$).

3 Stationary analysis

In this section, we analyze the subexponential interval graphs; that is, the random interval graphs proposed in the preceding section such that the lifetime (interval length) distribution is subexponential. In the analysis, we extend the time range of the immigration-death process to the whole real line \mathbb{R} and consider it to be stationary. Namely, a sequence $\{T_n\}_{n\in\mathbb{Z}}$ follows a homogeneous Poisson process with intensity $\lambda \in (0, \infty)$ satisfying $\cdots < T_0 \leq 0 < T_1 < \cdots$ and $\{L_n\}_{n\in\mathbb{Z}}$ denotes a sequence of mutually independent nonnegative random variables according to the distribution F with mean $\mu \in (0, \infty)$, where $\{L_n\}_{n\in\mathbb{Z}}$ is also independent of $\{T_n\}_{n\in\mathbb{Z}}$. Let $Q(t), t \in \mathbb{R}$, denote the set of individuals in the system at time t; that is, $Q(t) = \{n \in \mathbb{Z} : t \in I_n\}$ for $I_n = [T_n, T_n + L_n]$. Then, clearly $|Q(t)| = Z(t) = \sum_{n\in\mathbb{Z}} 1_{I_n}(t),$ $t \in \mathbb{R}$, where |A| denotes the cardinality of set A. When Z(t) > 0, let $n_i(t), i = 1, \ldots, Z(t)$, denote the *i*th element of Q(t) satisfying $n_i(t) < n_j(t)$ when i < j. Let also $R_{(i)}(t) = T_{n_i(t)} + L_{n_i(t)} - t (\geq 0),$ $i = 1, \ldots, Z(t)$; that is, the residual lifetime of individual $n_i(t)$ at time $t \in \mathbb{R}$. It is then known that (see, e.g., [15, Section 3.2]), when both λ and μ are positive and finite, the stationary distribution of $\{Z(t), R_{(i)}(t), i = 1, \ldots, Z(t)\}_{t\in\mathbb{R}}$ is given by

$$P(Z(0) = l, R_{(1)}(0) \le x_1, \dots, R_{(l)}(0) \le x_l) = \frac{(\lambda/\mu)^l}{l!} e^{-\lambda/\mu} \prod_{i=1}^l F_e(x_i), \quad l \in \mathbb{Z}_+, \ x_1, \dots, x_l \in \mathbb{R}_+, \quad (1)$$

where F_e denotes the equilibrium residual lifetime distribution of F defined by $F_e(x) = \mu \int_0^x \overline{F}(y) \, dy$, $x \ge 0$, and when l = 0, the left-hand side just means P(Z(0) = 0) and conventionally $\prod_{i=1}^0 \cdot = 1$ on the right-hand side. Formula (1) states that, in the steady-state, the number of individuals in the system follows the Poisson distribution with mean λ/μ , and the residual lifetimes of the individuals in the system are mutually independent and identically distributed according to F_e . By the PASTA (Poisson arrivals see time averages) property (see Wolff [17]), the right-hand side of (1) also gives the distribution of $\{Z(T_n-), R_{(i)}(T_n-), i = 1, \ldots, Z(T_n-)\}_{n \in \mathbb{Z}}$ just before the arrivals of individuals.

In the following two subsections, we consider the infinite size of random interval graph G = (V, E), $V = \mathbb{Z}$, with interval representation $\mathcal{I} = \{I_n\}_{n \in \mathbb{Z}}$ and the subexponential interval length (lifetime) distribution F. We discuss the tail asymptotics of the stationary degree distribution and the limit of the conditionally expected cluster coefficient of a vertex given its degree as the degree goes to infinity. In the analysis, we use the standard notation that, for any two real functions f(x) and g(x) on \mathbb{R} , $f(x) \sim g(x)$ as $x \to a$ stands for $\lim_{x \to a} f(x)/g(x) = 1$, where a is possibly infinity.

3.1 Degree distribution

A random graph G = (V, E) is said to be scale-free if its degree distribution has a power-law tail; that is, for some constants C > 0 and $\gamma > 0$,

$$P(D_0 = k) \sim \frac{C}{k^{\gamma}} \quad \text{as } k \to \infty,$$
 (2)

where $D_n = \sum_{i \in V} 1_E(n, i)$ denotes the degree of vertex $n \in V$. Note that D_0 satisfying (2) has the *m*th moment if $\gamma > m + 1$. The authors [12] showed that, in a discrete-time model setting, the random interval graph G = (V, E) with interval representation $\mathcal{I} = \{I_n\}_{n \in \mathbb{Z}}$ is scale-free in the steady state when the interval length distribution F has a power-law tail. We here extend this by applying the recent results on sampling of a stochastic process at random times according to subexponential distributions (see [2, 9, 11]) and provide a more general condition on F under which the stationary degree distribution satisfies

$$P(D_0 > k) \sim \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty.$$
 (3)

We will see that the power-law distribution F such that $\overline{F}(x) \sim c/x^{\alpha}$ as $x \to \infty$ with c > 0 and $\alpha > 1$ fulfills the provided condition, so that (3) leads to $P(D_0 > k) \sim c (\lambda/k)^{\alpha}$ as $k \to \infty$, which implies (2) with $C = c \alpha \lambda^{\alpha}$ and $\gamma = \alpha + 1$.

To provide the condition on the lifetime distribution under which (3) holds, we first give the definition of subexponential distributions. A distribution F and the corresponding random variable are said to be subexponential (see, e.g., Chistyakov [4] or [8, Sections 1.3 & A3], [14, Section 2.5]) if $\overline{F}(x) > 0$ for all $x \ge 0$ and

$$\lim_{x \to \infty} \frac{\overline{F^{*2}}(x)}{\overline{F}(x)} = 2, \tag{4}$$

where F^{*n} denotes the *n*th-fold convolution of F with itself. Note that, if F is subexponential, then $\overline{F}(x+a) \sim \overline{F}(x)$ as $x \to \infty$ for any $a \in \mathbb{R}$; that is, subexponential distributions are long-tailed. The following is a well-known and basic property of the subexponential distributions.

Lemma 1 (see, e.g., Cline [5]) Let F denote a subexponential distribution and let G_i , i = 1, 2, denote distributions on $[0, \infty)$ such that $\lim_{x\to\infty} \overline{G_i(x)}/\overline{F}(x) = c_i \in [0, \infty)$. Then, $\lim_{x\to\infty} \overline{G_1 * G_2(x)}/\overline{F}(x) = c_1 + c_2$, where $G_1 * G_2$ denotes the convolution of G_1 and G_2 .

Another important class of heavy-tailed distributions is recently introduced by [11] in problems of random time sampling and reduced load equivalence (see also [2, 9]). A distribution F and the corresponding random variable are said to be *square-root insensitive* if $\overline{F}(x) > 0$ for all $x \ge 0$ and

$$\lim_{x \to \infty} \frac{\overline{F}(x - \sqrt{x})}{\overline{F}(x)} = 1.$$
(5)

Note that, if F is square-root insensitive, then $\overline{F}(x - a\sqrt{x}) \sim \overline{F}(x)$ as $x \to \infty$ for any $a \in \mathbb{R}$. Also, if a random variable X is square-root insensitive, then $P(\sqrt{X} > x + a) \sim P(\sqrt{X} > x)$ as $x \to \infty$ for any $a \in \mathbb{R}$. It is known that distribution F is square-root insensitive when its tail is heavier than $\exp(-x^{\beta})$ with $\beta < 1/2$, whereas any distribution with a tail lighter than $e^{-\sqrt{x}}$ is not square-root insensitive ([2]).

Lemma 2 (see [2, 9, 11]) Let N denote a (delayed or non-delayed) renewal process with inter-renewal sequence $\{\tau_i\}_{i\in\mathbb{Z}_+}$ satisfying $E(\tau_1^2) < \infty$ and let L denote a nonnegative random variable independent of N. If L follows a square-root insensitive distribution F, then $P(N((0, L]) > k) \sim P(\lambda L > k) = \overline{F}(k/\lambda)$ as $k \to \infty$, where $\lambda = 1/E\tau_1$.

The proof of Lemma 2 is given in Appendix only for the non-delayed case. The tail-equivalence of the delayed and non-delayed cases is shown in Lemma 2.3 of [9]. Assume et al. [2] and Jelenković et al. [11] consider a more general case including that N in Lemma 2 is a regenerative process. Foss & Korshunov [9] also consider another general case where $E(\tau_1^{\beta}) < \infty$ for $\beta \in [1, 2)$. In this paper, however, the above form of the lemma is sufficient to show the following.

Theorem 1 If the lifetime distribution F is subexponential and square-root insensitive; that is, F fulfills (4) and (5), then the stationary degree distribution of the random interval graph G = (V, E) satisfies (3).

Theorem 1 states that, if the lifetime distribution F is subexponential and square-root insensitive, then so is the stationary degree distribution of the obtained random interval graph. The power-law distributions are subexponential and square-root insensitive, so that Theorem 1 covers the previous result in [12]. In the proof below and thereafter, N denotes the counting measure corresponding to $\{T_n\}_{n\in\mathbb{Z}}$; that is, N(A) represents the number of points of $\{T_n\}_{n\in\mathbb{Z}}$ in $A \in \mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ -field on \mathbb{R} . *Proof:* We here consider the Palm version satisfying $T_0 = 0$; that is, an arrival occurs at the origin. It is then known that $\{T_n\}_{n\neq 0}$ is also the Poisson process with the same intensity λ (see, e.g., Daley & Vere-Jones [7, Example 13.1(c)]). We can observe that the degree of vertex 0 consists of the number of individuals in the system just before the arrival of individual 0 and the number of new arrivals during the lifetime of individual 0; that is,

$$D_0 = \sum_{n<0} 1_E(0,n) + \sum_{n>0} 1_E(0,n) = Z(0-) + N(I_0) \quad \text{a.s.}$$
(6)

Since a Poisson process has independent increments and the lifetimes of individuals are mutually independent, Z(0-) and $N(I_0)$ are also independent each other, so that the distribution of D_0 is given as the convolution of those of Z(0-) and $N(I_0)$. Since F is square-root insensitive, Lemma 2 implies that $P(N(I_0) > k) \sim \overline{F}(k/\lambda)$ as $k \to \infty$. By (1), on the other hand, Z(0-) follows the Poisson distribution with mean λ/μ , so that $P(Z(0-) > k)/\overline{F}(k/\lambda) \to 0$ as $k \to \infty$ since F is subexponential. Hence, we have by Lemma 1 that

$$\mathcal{P}(D_0 > k) = \mathcal{P}(Z(0-) + N(I_0) > k) \sim \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty.$$

3.2 Cluster coefficient

In a given graph, the cluster coefficient of a vertex represents the fraction of couples of its neighbors such that the couple is connected by an edge. The cluster coefficient of vertex 0 of graph $G = (V, E), V = \mathbb{Z}$, is then given by

$$C_{0} = \frac{\sum_{n \in \mathbb{Z}} \sum_{m > n} 1_{E}(0, n) 1_{E}(0, m) 1_{E}(n, m)}{\binom{D_{0}}{2}}.$$
(7)

We here evaluate the limit of the conditional expectation $E(C_0 \mid D_0 > k)$ as $k \to \infty$ under the same condition as in Theorem 1.

Theorem 2 If the lifetime distribution F is subexponential and square-root insensitive; that is, F fulfills (4) and (5), then $\lim_{k\to\infty} E(C_0 \mid D_0 > k) = \eta$ exists and is given by

$$\eta = \int_0^\infty \frac{2}{x^2} \int_0^x \int_0^{x-y} \overline{F}(z) \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}F(x) = 1 - \mathrm{E}\Big[\Big\{\Big(1 - \frac{L_1}{L_0}\Big)^+\Big\}^2\Big],\tag{8}$$

where L_i , i = 0, 1, are mutually independent random variables according to F and $x^+ = \max(x, 0)$, $x \in \mathbb{R}$.

For example, if the distribution F is Pareto distribution such that $\overline{F}(x) = x^{-\alpha}$, $x \ge 1$, with $\alpha > 1$, then we have $\eta = 1 - 1/[(\alpha+1)(\alpha+2)]$. When $\alpha = 1.5$ ($\gamma = 2.5$ in (2)), it takes $\eta = 31/35 = 0.885714 \cdots$.

Proof: We consider the Palm version satisfying $T_0 = 0$ as in the proof of Theorem 1 and verify that $E(C_0 1_{\{D_0 > k\}}) \sim \eta P(D_0 > k)$ as $k \to \infty$. For simplicity of the notation, we write the event $A(0, n, m) = \{(0, n) \in E, (0, m) \in E, (n, m) \in E\}$. Recall that $D_0 = Z(0-) + N(I_0)$ a.s. as seen in (6). We then have from (7) that

$$\mathcal{E}(C_0 \, 1_{\{D_0 > k\}}) = \sum_{l=k+1}^{\infty} \frac{2}{l\,(l-1)} \, \mathcal{E}\left[\sum_{n \in \mathbb{Z}} \sum_{m>n} 1_E(0,n) \, 1_E(0,m) \, 1_E(n,m) \, 1_{\{D_0 = l\}}\right]$$

$$=\sum_{l=k+1}^{\infty} \frac{2}{l(l-1)} \sum_{j=0}^{l} \sum_{n \in \mathbb{Z}} \sum_{m>n} P(A(0,n,m) \mid Z(0-) = j, N(I_0) = l-j) \\ \times P(Z(0-) = j, N(I_0) = l-j).$$
(9)

In the following, we consider the last expression above by separating the sum over $-\infty < n < m < +\infty$ into three cases; i) n < m < 0, ii) n < 0 < m and iii) 0 < n < m. We will see that the first two cases leads to the terms which are $o(P(D_0 > k))$ and the third case yields the term which is tail-equivalent to $P(D_0 > k)$ as $k \to \infty$.

i) Case of n < m < 0. Whenever $(0, n) \in E$ and $(0, m) \in E$ for n, m < 0, it is necessary that $(n, m) \in E$ since individuals n and m are in the system when individual 0 arrives, so that,

$$\sum_{n=-\infty}^{-2} \sum_{m=n+1}^{-1} P(A(0,n,m) \mid Z(0-) = j, N(I_0) = l-j) = {j \choose 2}.$$

Substituting this into (9), we have

$$\sum_{l=k+1}^{\infty} \sum_{j=2}^{l} \frac{j(j-1)}{l(l-1)} P(Z(0-) = j, N(I_0) = l-j)$$

$$= \left(\frac{\lambda}{\mu}\right)^2 \sum_{j=k+1}^{\infty} P(Z(0-) = j-2) \sum_{l=j}^{\infty} \frac{P(N(I_0) = l-j)}{l(l-1)}$$

$$+ \left(\frac{\lambda}{\mu}\right)^2 \sum_{j=2}^{k} P(Z(0-) = j-2) \sum_{l=k+1}^{\infty} \frac{P(N(I_0) = l-j)}{l(l-1)},$$
(10)

where we use the fact that Z(0-) follows the Poisson distribution with mean λ/μ and $j(j-1) P(Z(0-) = j) = (\lambda/\mu)^2 P(Z(0-) = j-2), j = 2, 3, ...$ For the first term on the right-hand side above, since $1/[l(l-1)] \leq 1$ for l > 1 and D_0 is subexponential by Theorem 1, we have

(1st term on RHS of (10))
$$\leq \left(\frac{\lambda}{\mu}\right)^2 P(Z(0-) > k-2) = o(P(D_0 > k))$$
 as $k \to \infty$.

We now consider the second term on the right-hand side of (10). For any $\epsilon > 0$, there exists a $k_{\epsilon} > 0$ such that $1/[l(l-1)] < \epsilon$ for $l > k_{\epsilon}$. Thus, we have for $k \ge k_{\epsilon}$,

(2nd term on RHS of (10))
$$\leq \epsilon \left(\frac{\lambda}{\mu}\right)^2 P(D_0 > k - 1).$$

Since ϵ is arbitrarily small, this implies that the second term on the right-hand side of (10) is $o(P(D_0 > k))$ as $k \to \infty$.

ii) Case of n < 0 < m. Given that Z(0-) = j, we have by (1) that the residual lifetimes of these j individuals at time 0 are independent and identically distributed according to F_e . Also, given that $N(I_0) = l - j$ and $L_0 = x$ (> 0), the property of Poisson processes implies that the arrival times of these l - j individuals are independent and uniformly distributed on [0, x] (see, e.g., [7, Section 2.1]). Note that interval I_n , n < 0, has an overlap with interval I_m , m > 0, which has its left endpoint at $y \in [0, x]$, when the residual lifetime of individual n at time 0 is longer than y. Therefore,

$$\sum_{n=-\infty}^{-1} \sum_{m=1}^{+\infty} P(A(0,n,m) \mid Z(0-) = j, N(I_0) = l-j) = j(l-j) \int_0^\infty \frac{1}{x} \int_0^x \overline{F_e}(y) \, \mathrm{d}y \, \mathrm{d}F(x).$$
(11)

We write η_1 for the integral on the right-hand side, which can also be expressed as $\eta_1 = \mathbb{E}[1 \wedge (R_{(1)}/L_0)]$, where L and $R_{(1)}$ are mutually independent and according to the distributions F and F_e , respectively, and $x \wedge y = \min(x, y)$ for $x, y \in \mathbb{R}$. Substituting (11) into (9), we have

$$\sum_{l=k+1}^{\infty} \sum_{j=1}^{l-1} \frac{2\eta_1 j (l-j)}{l(l-1)} P(Z(0-) = j, N(I_0) = l-j)$$

$$= \frac{2\eta_1 \lambda}{\mu} \sum_{j=k}^{\infty} P(Z(0-) = j-1) \sum_{l=j+1}^{\infty} \frac{(l-j) P(N(I_0) = l-j)}{l(l-1)}$$

$$+ \frac{2\eta_1 \lambda}{\mu} \sum_{j=1}^{k-1} P(Z(0-) = j-1) \sum_{l=k+1}^{\infty} \frac{(l-j) P(N(I_0) = l-j)}{l(l-1)}.$$
(12)

For the first term on the right-hand side above, since $(l-j)/[l(l-1)] \leq 1$ for $l > j \geq 1$,

(1st term on RHS of (12))
$$\leq \frac{2\eta_1 \lambda}{\mu} P(Z(0-) > k-2) = o(P(D_0 > k))$$
 as $k \to \infty$.

For the second term on the right-hand side of (12), we have for any $\epsilon > 0$, there exists a $k_{\epsilon} > 0$ such that $(l-j)/[l(l-1)] \leq \epsilon$ for $l > k_{\epsilon}$ and $j \geq 1$. Thus, for $k \geq k_{\epsilon}$,

(2nd term on RHS of (12))
$$\leq \frac{2 \epsilon \eta_1 \lambda}{\mu} P(D_0 > k - 1),$$

where ϵ is arbitrarily small, so that this leads to $o(P(D_0 > k))$ as $k \to \infty$.

iii) Case of 0 < n < m. Given that $N(I_0) = l - j$ and $L_0 = x$ (> 0), the arrival times of these l - j individuals are independent and uniformly distributed on [0, x]. The event that interval I_n whose left endpoint is at $y \in [0, x]$ has an overlap with interval I_m whose left endpoint is at $z \in [y, x]$ is realized when $L_n > z - y$, so that

$$\sum_{n=1}^{+\infty} \sum_{m=n+1}^{+\infty} \mathbb{P}\left(A(0,n,m) \mid Z(0-) = j, N(I_0) = l-j\right) = \binom{l-j}{2} \int_0^\infty \frac{2}{x^2} \int_0^x \int_y^x \overline{F}(z-y) \, \mathrm{d}z \, \mathrm{d}y \, \mathrm{d}F(x),$$
(13)

and the integral on the right-hand side above is given by η in (8). We now substitute (13) into (9) and show that both the upper and lower bounds are asymptotically equal to $\eta P(D_0 > k)$ as $k \to \infty$. For the upper bound, since $(l-j)(l-j-1)/[l(l-1)] \le 1$ for l > 1 and $0 \le j < 1$, we have clearly

$$\sum_{l=k+1}^{\infty} \sum_{j=0}^{l-2} \frac{\eta \left(l-j\right) \left(l-j-1\right)}{l \left(l-1\right)} \operatorname{P}(Z(0-)=j, \ N(I_0)=l-j) \le \eta \operatorname{P}(D_0>k)$$

On the other hand, for the asymptotic lower bound,

$$\sum_{l=k+1}^{\infty} \sum_{j=0}^{l-2} \frac{\eta \left(l-j\right) \left(l-j-1\right)}{l \left(l-1\right)} \operatorname{P}(Z(0-)=j, \ N(I_0)=l-j)$$

$$\geq \eta \sum_{j=0}^{k-1} \operatorname{P}(Z(0-)=j) \sum_{l=k+1}^{\infty} \frac{\left(l-j\right) \left(l-j-1\right)}{l \left(l-1\right)} \operatorname{P}(N(I_0)=l-j)$$

Here, for any $\epsilon > 0$, there exists a $k_{\epsilon} > 0$ such that $P(Z(0-) < k_{\epsilon}) \ge 1 - \epsilon$. Furthermore, there exists an $l_{\epsilon} > k_{\epsilon}$ such that $(l-j)(l-j-1)/[l(l-1)] \ge 1 - \epsilon$ for $l > l_{\epsilon}$ and $0 \le j < k_{\epsilon}$. Thus, for $k > l_{\epsilon}$, the right-hand side above is bounded below by

$$(1-\epsilon)\,\eta\sum_{j=0}^{k_{\epsilon}-1} P(Z(0-)=j)\,P(N(I_0)>k-j) \ge (1-\epsilon)^2\,\eta\,P(N(I_0)>k).$$

Since Theorem 1 states that $P(N(I_0) > k) \sim P(D_0 > k)$ as $k \to \infty$ and ϵ is arbitrary, this case yields the term which is tail-equivalent to $\eta P(D_0 > k)$ as $k \to \infty$.

Remark 2 In considering the connected interval graph \tilde{G} in Remark 1, we have to modify slightly the result on the stationary degree distribution in Theorem 1. The fact that $P(Z(0-)=0) = e^{-\lambda/\mu}$ by (1) states that $\{A_n\}_{n\in\mathbb{Z}}$ is a stationary point process with intensity $\lambda e^{-\lambda/\mu}$. Thus, since the superposed point process $\{T_n\}_{n\in\mathbb{Z}} \cup \{A_n\}_{n\in\mathbb{Z}}$ has intensity $\lambda (1+e^{-\lambda/\mu})$, the probability that an arbitrary chosen vertex is not the one which is extraneously added in Remark 1 is given by $(1+e^{-\lambda/\mu})^{-1}$, so that the tail asymptotics of the stationary degree distribution in the modified graph \tilde{G} becomes $P(\tilde{D}_0 > k) \sim (1+e^{-\lambda/\mu})^{-1} \overline{F}(k/\lambda)$ as $k \to \infty$. The limit of the conditionally expected cluster coefficient, on the other hand, remains the same as that given on the right-hand side of (8) in Theorem 2.

4 Concluding remark

In this paper, we have analyzed the stationary subexponential interval graphs generated by immigrationdeath processes. Namely, we have derived the tail asymptotics of the stationary degree distribution when the lifetime distribution of the immigration-death process is subexponential and square-root insensitive. Furthermore, we have derived the limit of the conditionally expected cluster coefficient given the degree as the degree goes to infinity under the same condition as that for obtaining the tail asymptotics of the stationary degree distribution. In future works, we can consider problems like evaluating the stationary distribution of the sizes of connected components and the diameter of a connected component, which represents the length of the shortest path connecting any pair of vertices in the connected component.

Acknowledgment

The authors wish to thank Osamu Watanabe for valuable discussions. The third author's work was supported in part by JSPS Global COE program "Computation as a Foundation for the Sciences."

References

- [1] S. Asmussen. Applied Probability and Queues. Springer-Verlag, 2nd edition, 2003.
- S. Asmussen, C. Klüppelberg, and K. Sigman. Sampling at subexponential times, with queueing applications. Stochastic Process. Appl., 79:265–286, 1999.
- [3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509–512, 1999.
- [4] V. P. Chistyakov. A theorem on sums of independent positive random variables and its applications to branching random processes. *Theory Probab. Appl.*, 9:640–648, 1964.
- [5] D. B. H. Cline. Convolution tails, product tails and domains of attraction. Probab. Theory Related Fields, 72:529–557, 1986.
- [6] D. R. Cox and V. Isham. Point Processes. Chapman & Hall, 1980.
- [7] D. J. Daley and D. Vere-Jones. An Introduction to the Theory of Point Processes, volumes I, II. Springer-Verlag, 2nd edition, 2003, 2008.
- [8] P. Embrechts, C. Klüppelberg, and T. Mikosch. Modelling Extremal Events; for Insurance and Finance. Springer-Verlag, Berlin, 1997.
- [9] S. Foss and D. Korshunov. Sampling at a random time with a heavy-tailed distribution. Markov Process. Related Fields, 6:543–568, 2000.
- [10] M. C. Golumbic. Algorithmic Graph Theory and Perfect Graphs. Elsevier, 2nd edition, 2004.
- [11] P. Jelenković, P. Momčilović, and B. Zwart. Reduced load equivalence under subexponentiality. Queueing Syst., 46:97–112, 2004.

- [12] N. Miyoshi, T. Shigezumi, R. Uehara, and O. Watanabe. Scale free interval graphs. Submitted for publication, 2008.
- [13] M. Newman. The structure and function of complex networks. SIAM Review, 45:167–256, 2003.
- [14] T. Rolski, H. Schmidli, V. Schmidt, and J. Teugels. Stochastic Processes for Insurance and Finance. John Wiley & Sons, Chichester, 1999.
- [15] L. Takács. Introduction to the Theory of Queues. Oxford University Press, 1962.
- [16] D. J. Watts and D. H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393:440–442, 1998.
- [17] R. W. Wolff. Poisson arrivals see time averages. Oper. Res., 30:223–231, 1982.

A Proof of Lemma 2

We here provide the proof of Lemma 2 only for the case where N is a non-delayed renewal process. The tail equivalence of P(N(0,L]) > k) as $k \to \infty$ for delayed and non-delayed N is shown in Lemma 2.3 of [9]. The proof mainly follows [11]. In the following, for any two real functions f(x) and g(x) on \mathbb{R} , $f(x) \leq g(x)$ and $f(x) \geq g(x)$ as $x \to a$ stand for, respectively, $\limsup_{x\to a} f(x)/g(x) \leq 1$ and $\liminf_{x\to a} f(x)/g(x) \geq 1$, where a is possibly infinity. We first verify the following preliminary lemma.

Lemma 3 Let N denote a non-delayed renewal process with inter-renewal sequence $\{\tau_i\}_{i\in\mathbb{N}}$ satisfying $E(\tau_1^2) < \infty$.

(i) For any $\delta > 0$, there exists a $c_{\delta} > 0$ such that

$$P\Big(N((0,t]) - \frac{t}{E\tau_1} > u\Big) \le e^{-c_{\delta}u^2/t}, \quad t > 0, \ 0 \le u \le \delta t.$$

(ii) There exist $c_1 > 0$ and $c_2 > 0$ such that

$$\mathbf{P}\Big(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > u\Big) \le e^{-c_1 u^2/t} + e^{-c_2 t}, \quad t > 0, \ u \ge 0.$$

Proof: We first verify (i). Markov's inequality yields for s > 0,

$$\begin{split} \mathbf{P}\Big(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > u\Big) &= \mathbf{P}\Big(N((0,t]) \ge \left\lfloor u + \frac{t}{\mathbf{E}\tau_1} \right\rfloor + 1\Big) \\ &= \mathbf{P}\Big(\sum_{i=1}^{\lfloor u+t/\mathbf{E}\tau_1 \rfloor + 1} \tau_i \le t\Big) \le e^{st} \, (\mathbf{E}e^{-s\tau_1})^{(u+t/\mathbf{E}\tau_1)} \end{split}$$

where $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$, the maximal integer not greater than $x \in \mathbb{R}$. Here, applying $e^{-x} \leq 1 - x + x^2, x \in \mathbb{R}$, and then $1 + y \leq e^y, y \in \mathbb{R}$, into the last expression above, we have

$$\mathbf{P}\left(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > u\right) \le \exp\left\{-s \, u \, \mathbf{E}\tau_1 + s^2 \, \mathbf{E}(\tau_1^{\ 2}) \left(u + \frac{t}{\mathbf{E}\tau_1}\right)\right\}$$

Now, we choose $s = u (E\tau_1)^2 / [2t(1 + \delta E\tau_1) E(\tau_1^2)]$. Then, the inside of the braces on the right-hand side above leads to

$$-\frac{(\mathrm{E}\tau_1)^3}{4(1+\delta\,\mathrm{E}\tau_1)\,\mathrm{E}(\tau_1{}^2)}\left(2-\frac{1+(u/t)\,\mathrm{E}\tau_1}{1+\delta\,\mathrm{E}\tau_1}\right)\frac{u^2}{t} \le -\frac{(\mathrm{E}\tau_1)^3}{4(1+\delta\,\mathrm{E}\tau_1)\,\mathrm{E}(\tau_1{}^2)}\frac{u^2}{t}$$

where the inequality follows from $u/t \leq \delta$.

We next show (ii). Fix some $\delta > 0$. If $u \leq \delta t$, then (i) provides

$$\mathbf{P}\Big(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > u\Big) \le e^{-c_\delta u^2/t}.$$

If $u > \delta t$, on the other hand, then we have

$$\mathbf{P}\Big(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > u\Big) \le \mathbf{P}\Big(N((0,t]) - \frac{t}{\mathbf{E}\tau_1} > \delta t\Big) \le e^{-c_\delta \delta^2 t},$$

where the second inequality follows from (i) with $u = \delta t$.

We are now at the position to verify Lemma 2. We evaluate the asymptotic upper bound and lower bound separately for P(N(0,L]) > k) as $k \to \infty$.

A.1 Asymptotic upper bound

We first show the asymptotic upper bound,

$$P(N((0,L]) > k) \lesssim \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty.$$
 (14)

We have for $a \in (0, \sqrt{\lambda k}), b \in (0, 1/\lambda)$, and $k > a^2/[\lambda (1 - \lambda b)^2]$,

$$P(N((0,L]) > k) \le P\left(L > \frac{k}{\lambda} - \frac{a}{\lambda}\sqrt{\frac{k}{\lambda}}\right) + P\left(N((0,L]) > k, \ b \ k < L \le \frac{k}{\lambda} - \frac{a}{\lambda}\sqrt{\frac{k}{\lambda}}\right) + P\left(N((0,b \ k]) > k\right).$$
(15)

Since F is square-root insensitive, the first term on the right-hand side above leads to

$$P\left(L > \frac{k}{\lambda} - \frac{a}{\lambda}\sqrt{\frac{k}{\lambda}}\right) \sim P\left(L > \frac{k}{\lambda}\right) = \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty.$$

Thus, one needs to show that the last two terms on the right-hand side of (15) are $o(\overline{F}(k/\lambda))$ as $k \to \infty$. We start with the third term. Since $b \in (0, 1/\lambda)$, there exists a $\delta > 0$ such that $0 \le 1 - \lambda b \le \delta b$, so that by Lemma 3(i),

$$P(N((0, b\,k]) > k) = P(N((0, b\,k]) - \lambda \,b \,k > (1 - \lambda \,b) \,k) \le e^{-c_{\delta}(1 - \lambda b)^{2}k/b} = o\left(\overline{F}\left(\frac{k}{\lambda}\right)\right) \quad \text{as } k \to \infty.$$

To deal with the second term on the right-hand side of (15), note that in view of Lemma 3(ii),

$$(\text{2nd term on RHS of } (15)) = \int_{bk}^{k/\lambda - (a/\lambda)\sqrt{k/\lambda}} \mathbb{P}\left(N((0, x]) > k\right) \mathrm{d}F(x)$$

$$\leq \int_{bk}^{k/\lambda - (a/\lambda)\sqrt{k/\lambda}} e^{-c_1(k-\lambda x)^2/x} \,\mathrm{d}F(x) + \int_{bk}^{k/\lambda - (a/\lambda)\sqrt{k/\lambda}} e^{-c_2x} \,\mathrm{d}F(x).$$
(16)

Here, it is easy to see that

(2nd term on RHS of (16))
$$\leq e^{-c_2bk} = o\left(\overline{F}\left(\frac{k}{\lambda}\right)\right)$$
 as $k \to \infty$.

Consider the first term on the right-hand side of (16). Note that, for any $x \in (b\,k, k/\lambda - (a/\lambda)\sqrt{k/\lambda})$, we have $e^{-c_1(k-\lambda x)^2/x} \leq e^{-c_1\lambda(k-\lambda x)^2/k}$, so that integration by parts and change of variables $(y = \sqrt{\lambda/k} (k - \lambda x))$ result in

$$\begin{aligned} (1\text{st term on RHS of } (16)) &\leq \int_0^{k/\lambda - (a/\lambda)\sqrt{k/\lambda}} e^{-c_1\lambda(k-\lambda x)^2/k} \,\mathrm{d}F(x) \\ &\leq e^{-c_1\lambda k} + \int_0^{k/\lambda - (a/\lambda)\sqrt{k/\lambda}} \frac{2\,c_1\,\lambda^2\,(k-\lambda\,x)}{k}\,e^{-c_1\lambda(k-\lambda x)^2/k}\,\overline{F}(x)\,\mathrm{d}x \end{aligned}$$

$$= e^{-c_1\lambda k} + \int_a^{\sqrt{\lambda k}} 2 c_1 y e^{-c_1 y^2} \overline{F}\left(\frac{k}{\lambda} - \frac{y}{\lambda}\sqrt{\frac{k}{\lambda}}\right) \mathrm{d}y,\tag{17}$$

and the first term on the right-hand side above is $o(\overline{F}(k/\lambda))$ as $k \to \infty$. For the integrand above, since \sqrt{L} is long-tailed when L is square-root insensitive (see [11]), we have for $y \leq \sqrt{\lambda k}$, $\epsilon > 0$, and sufficiently large k,

$$\overline{F}\left(\frac{k}{\lambda} - \frac{y}{\lambda}\sqrt{\frac{k}{\lambda}}\right) = P\left(L > \frac{k}{\lambda} - \frac{y}{\lambda}\sqrt{\frac{k}{\lambda}}\right) \le P\left(\sqrt{L} > \sqrt{\frac{k}{\lambda}} - \frac{y}{\lambda}\right) \le c_{\epsilon} e^{\epsilon y} \overline{F}\left(\frac{k}{\lambda}\right),$$

where the first inequality follows from $\sqrt{x - u\sqrt{x}} \ge \sqrt{x} - u$ for $u \le \sqrt{x}$ and the last inequality follows from the property of long-tailed distributions; that is, for any long-tailed random variable X and any $\epsilon > 0$, there exist $c_{\epsilon} > 0$ and $x_{\epsilon} > 0$ such that $P(X > x - u) \le c_{\epsilon} e^{\epsilon u} P(X > x)$ for all $x - u > x_{\epsilon}$. Thus, we obtain

$$(\text{2nd term on RHS of } (17)) \le c_{\epsilon} \overline{F}\left(\frac{k}{\lambda}\right) \int_{a}^{\infty} e^{\epsilon y} \, 2 \, c_{1} \, y \, e^{-c_{1} y^{2}} \, \mathrm{d}y = c_{\epsilon} \, \overline{F}\left(\frac{k}{\lambda}\right) \mathbb{E}[e^{\epsilon Y} \, 1_{\{Y>a\}}],$$

where Y denotes a random variable according to Weibull distribution $P(Y > y) = e^{-c_1 y^2}$, $y \ge 0$. Since $E[e^{\epsilon Y}] < \infty$, there exists an $a_{\epsilon} > 0$ such that the second term on the right-hand side of (17) is bounded by $\epsilon \overline{F}(k/\lambda)$ for $a \ge a_{\epsilon}$; that is, it is $o(\overline{F}(k/\lambda))$ as $k \to \infty$, and eventually we have (14).

A.2 Asymptotic lower bound

We next show the asymptotic lower bound,

$$P(N((0,L]) > k) \gtrsim \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty.$$
 (18)

We have for a > 0,

$$\begin{split} \mathbf{P}(N((0,L]) > k) &\geq \mathbf{P}\Big(N((0,L]) > k, \ L > \frac{k}{\lambda} + a\sqrt{\frac{k}{\lambda}}\Big) \\ &\geq \mathbf{P}\Big(N\Big(\Big(0,\frac{k}{\lambda} + a\sqrt{\frac{k}{\lambda}}\Big]\Big) > k\Big) \ \overline{F}\Big(\frac{k}{\lambda} + a\sqrt{\frac{k}{\lambda}}\Big) \end{split}$$

Here, one obtains for $k \ge \lambda a^2$,

$$\begin{split} \mathbf{P}\Big(N\Big(\Big(0,\frac{k}{\lambda}+a\sqrt{\frac{k}{\lambda}}\,\Big]\Big) > k\Big) &= \mathbf{P}\Big(\frac{N((0,k/\lambda+a\sqrt{k/\lambda}\,])-(k+a\sqrt{\lambda k})}{\sqrt{k/\lambda+a\sqrt{k/\lambda}}} > -\frac{\lambda\,a}{\sqrt{1+a\sqrt{\lambda/k}}}\Big) \\ &\geq \mathbf{P}\Big(\frac{N((0,k/\lambda+a\sqrt{k/\lambda}\,])-(k+a\sqrt{\lambda k})}{\sqrt{k/\lambda+a\sqrt{k/\lambda}}} > -\frac{\lambda\,a}{\sqrt{2}}\Big) \end{split}$$

Therefore, the square-root insensitivity and the central limit theorem for renewal processes (see, e.g., Asmussen [1, Chap. V, Theorem 6.3]) result in for an appropriate $\sigma > 0$,

$$\mathbf{P}(N(0,L]) > k) \gtrsim \Phi\left(\frac{\lambda a}{\sigma\sqrt{2}}\right) \overline{F}\left(\frac{k}{\lambda}\right) \quad \text{as } k \to \infty,$$

where Φ denotes the standard normal distribution. Finally, letting $a \to \infty$ leads to (18).