Generational Garbage Collection Policies

Xufeng Zhao, Syouji Nakamura and Cunhua Qian

1 Introduction

In the computer science community, the technique of *garbage collection* [5] is an automatic process of memory recycling, which refers to those objects in the memory no longer referenced by programs are called *garbage* and should be thrown away. A *garbage collector* determines which objects are garbage and makes the heap space occupied by such garbage available again for the subsequent new objects. Garbage collection plays an important role in Java's security strategy, however, it adds a large overhead that can deteriorate the program performances. From related studies which are summarized in [5], a garbage collector spends between 25 and 40 percent of execution time of programs for its work in general, and delays caused by such a garbage collection are obtrusive.

In recent years, *generational garbage collection* [1, 17, 19, 20] has been popular with programmers as it can be made more efficiently. Compared with classical tracing collectors, e.g., reference counting collector, mark-sweep collector, mark-compact collector, and copying collector, a *generational garbage collector* is effective in computer programs with the characteristic that it is unnecessary to mark or copy all active data of the whole heap for every collection, i.e., the collector concentrates effort on

X. Zhao (✉)
Department of Business Administration, Aichi Institute of Technology,
1247 Yachigusa, Yakusa-cho, Toyota 470-0392, Japan
e-mail: g09184gg@aitech.ac.jp

S. Nakamura
Department of Human Life and Information, Kinjo Gakuin University,
1723 Omori 2-chome, Moriyama-ku, Nagoya 463-8521, Japan
e-mail: snakam@kinjo-u.ac.jp

C. Qian
School of Economics and Management, Nanjing University of Technology,
30 Puzhu Road, 211816 Nanjing, China
e-mail: qch64317@njut.edu.cn

T. Dohi and T. Nakagawa (eds.), *Stochastic Reliability and Maintenance Modeling*,
Springer Series in Reliability Engineering 9, DOI: 10.1007/978-1-4471-4971-2_15,

those objects that are most likely to be garbage. Based on the weak generational hypothesis [17] which asserts that most objects are short-lived after their allocation, a generational garbage collector segregates objects by age into two or more regions called *generations* or *multiple generations*. The survival rates of younger generations are always much lower than those of older ones, which means that younger generations are more likely to be garbage and can be collected more frequently than older ones. Although such generational collections cost much shorter time than that of a full collection, the problems of pointers from older generations to younger ones and the size of root sets for younger generations become more complicated. For these reasons, many generational collectors are limited to just two or three generations [5]. This generational technique is now in widespread use for memory management. For instance, the garbage collector, which is used in Sun's HotSpot Java Virtual Machine (JVM), manages heap space for both young and old generations [19]: New objects space *Eden*, two equal survivor spaces *SS♯1* and *SS♯2* for surviving objects, and tenured objects space *Old* (*Tenured*), where Eden, SS♯1 and SS♯2 are for younger generations, and Old (Tenured) is for older ones.

A generational garbage collector uses *minor collection* and *tenuring collection* [1] for younger generations and *major collection* for multi-generations [5]. Most generational garbage collectors are copying collectors, although it is possible to use mark-sweep collectors [2]. In this chapter, we concentrate on a generational garbage collector using copying collection. However, for every garbage collection, the manner of *stop and copy* pauses all application threads to collect the garbage. The duration of time for which the collector has worked is called *pause time* [5], which is an important parameter for interactive systems, and depends largely upon the volume of surviving objects and the type of collections. That is, pause time suffered for minor collection increases with the number of collections and is less than that of tenuring collection; major collection pause time is the longest among the three.

With regard to garbage collection modelings, there have been few research papers that studied analytical expressions of optimal policies for a generational garbage collector. Most problems were concerned with several ways to introduce garbage collection methods in techniques and how to tune the garbage collector by simulations, which is more complex and time-consuming due to the random accesses of programs in the memory in practice [4, 6, 7, 16, 18]. We propose that garbage collection is a *stochastic decision making process* and should be analyzed by the theory of stochastic processes from the viewpoints of management. As some applications of damage models, a garbage collection model for a database in the computer system [14] was studied, but the theoretical point of garbage collection was not considered essentially, and optimal policies for a generational garbage collector with tenuring threshold and major collection times according to practical working schemes [21, 22] were studied recently.

[1] Tenuring collection is also a kind of minor collection [5]. We define tenuring collection as distinct from minor collection because there may be some surviving objects tenured from survivor space into Old.

This chapter considers a pause time goal which is called time cost or cost for simplicity, and our problem is to obtain optimal collection times which minimize the expected cost rates. Using the techniques of cumulative processes and reliability theory [8–10], optimal tenuring collection times and major collection times are discussed. Furthermore, increase in objects might be unclear at discrete times for the high frequency of computer processes. According to [1, 19], it would be more practical to assume that surviving objects that should be copied increase with time continuously and roughly according to some mathematical laws. Applying the techniques of degradation processes [11, 15] and continuous wear processes [9], optimal tenuring collection times are discussed analytically and numerically.

2 Working Schemes

In general, the frequency of garbage collections depends on whether the computer processes are busy or not. Hence, it is practical to assume that garbage collections occur at a nonhomogeneous Poisson process with an intensity function $\lambda(t)$ and a mean-value function $R(t) \equiv \int_0^t \lambda(u)\mathrm{d}u$. Then, the probability that collections occur exactly j times in $(s, t]$ is [12]

$$H_j(s,t) \equiv \frac{[R(t)-R(s)]^j}{j!}\mathrm{e}^{-[R(t)-R(s)]} \quad (j = 0, 1, 2, \ldots).$$

Letting $F_j(s,t)$ $(j = 1, 2, \ldots)$ denote the probability that collections occur at least j times in the time interval $(s, t]$,

$$F_j(s,t) = \int_s^t H_{j-1}(s,u)\lambda(u)\mathrm{d}u = \sum_{i=j}^{\infty} H_i(s,t), \tag{1}$$

where $F_0(s,t) \equiv 1$ and

$$H_j(t) \equiv H_j(0,t) = \frac{[R(t)]^j}{j!}\mathrm{e}^{-R(t)},$$

$$F_j(t) \equiv F_j(0,t) = \sum_{i=j}^{\infty} H_i(t).$$

Further, the volume X_i of new objects in Eden at the ith collection has an identical distribution $G(x) \equiv \Pr\{X_i \le x\}$ $(i = 1, 2, \ldots)$, and survivor rate α_i $(0 \le \alpha_i < 1; i = 1, 2, \ldots)$, where $1 > \alpha_1 > \alpha_2 > \cdots > \alpha_i > \cdots \ge 0$, means that new objects will survive $100\alpha_i$ percent at the ith minor collection. That is, detailed working schemes of a generational garbage collector that have been introduced in [5, 19, 21, 22] are given as the following steps (Fig. 1):

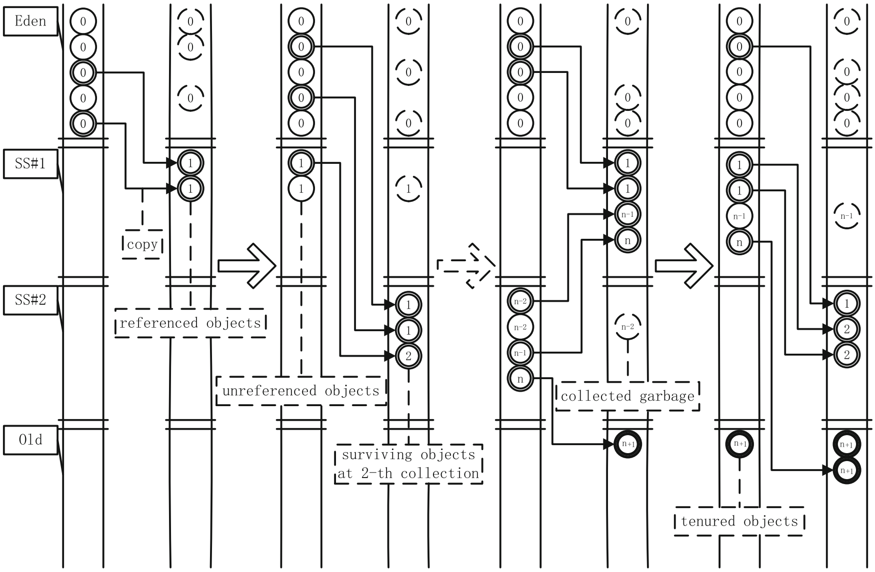

Fig. 1 Working schemes of a generational garbage collector

1. New objects X_1 are allocated in Eden.
2. When the first minor collection occurs, surviving objects $\alpha_1 X_1$ from Eden are copied into SS♯1.
3. When the second minor collection occurs, surviving objects $\alpha_1 X_2$ from Eden and $\alpha_2 X_1$ from SS♯1 are copied into SS♯2.
4. In the fashions of 1–3, minor collections copy surviving objects between SS♯1 and SS♯2 until they become tenured, *i.e.*, tenuring collection occurs when some parameter meets the tenuring threshold, and then, the older or the oldest objects are copied into Old.
5. When Old fills up, major collection of the whole heap occurs, and surviving objects from Old are kept in Old, while objects from Eden and survivor space are kept in survivor space.

In practice, tenuring threshold mentioned in step 4 above is adaptive, which is called *adaptive tenuring* [5] and can be modified at any time. In this chapter, we propose two cases of working schemes according to the properties of adaptive tenuring:

Based on [17], new objects can be tenured only if they survive at least one minor collection, because objects that survive two minor collections are much less than those that survive just one. In other words, surviving objects are likely to reduce slightly with the number of minor collections beyond two. That is, for step 4:

4a. When tenuring collection occurs, surviving objects from Eden and survivor space are copied into the other survivor space and Old, respectively. That is, if tenuring collection is made at the jth ($j = 1, 2, \ldots$) collection, surviving objects $\alpha_1 X_j$

and $\alpha_2 X_{j-1} + \alpha_3 X_{j-2} + \cdots + \alpha_j X_1$ are copied into survivor space and Old, respectively.

4b. After tenuring collection, the same collection cycle begins with step 1. The collector works $1 \to 2 \to 3 \to 4a \to 4b \to 1 \to \cdots$. In this case, tenuring collections can be consider as renewal points of the collection processes, because Old will be filled with tenured objects slowly and major collection occurs rarely, especially when the tenuring threshold is high and the survivor rates are low. Modelings and optimizations of tenuring collection times are discussed in Sects. 3 and 5.

From [19], the oldest objects can be tenured from survivor space into Old at every collection time when tenuring collection begins. That is, for step 4:

4c. When tenuring collection occurs, the oldest objects from survivor space are copied into Old, and the other surviving objects from Eden and survivor space are copied into the other survivor space. That is, if tenuring collection is made at the jth ($j = 1, 2, \ldots$) collection, surviving objects $\alpha_1 X_j + \alpha_2 X_{j-1} + \cdots + \alpha_{j-1} X_2$ and $\alpha_j X_1$ are copied into survivor space and Old, respectively.

4d. When the next collection occurs, the collector works as the same rule as 4c. That is, when the second tenuring collection occurs, surviving objects $\alpha_1 X_{j+1} + \alpha_2 X_j + \cdots + \alpha_{j-1} X_3$ and $\alpha_j X_2$ are copied into the other survivor space and Old, respectively. The collector works $1 \to 2 \to 3 \to 4c \to 4d \to 5 \to 1 \to \cdots$. In this case, major collections can be consider as renewal points of the collection processes, because there are always some surviving objects tenured from survivor space into Old at every collection time when tenuring collection begins, especially when the tenuring threshold is low and the survivor rates are high. Related optimization problems of major collection times are discussed in Sect. 4.

From the above discussions, if tenuring collection is made at the jth ($j = 1, 2, \ldots$) collection, surviving objects that should be copied at the ith ($i = 0, 1, 2, \ldots, j-1$) minor collection, copied objects and tenured objects at the kth ($k = 1, 2, \ldots$) tenuring collection are, respectively,

$$\sum_{n=0}^{i-1} \alpha_{n+1} X_{i-n} < K, \qquad \sum_{n=1}^{j} \alpha_n X_{j+k-n} \geq K \quad \text{and} \quad \alpha_j X_k, \tag{2}$$

where $\sum_{n=0}^{-1} \equiv 0$, and K is tenuring threshold in step 4, which means that the total volume of surviving objects has exceeded level K. It could be easily seen that copied objects increase with the number of minor collections and are relatively stable with the number of tenuring collections. We define that the distribution of the total surviving objects at the ith minor collection is

$$G_i(x) \equiv \Pr\left\{ \sum_{n=0}^{i-1} \alpha_{n+1} X_{i-n} \leq x \right\} \quad (i = 0, 1, 2, \ldots), \tag{3}$$

where $G_i(x)$ decreases with i, and $G_0(x) \equiv 1$ means that there are no objects in the heap space at time 0. The probability that the total surviving objects exceed exactly a threshold level K at the $(i+1)$th $(i = 0, 1, 2, \ldots)$ minor collection is

$$p_i(K) \equiv \int_0^K \overline{G}(K-x)\mathrm{d}G_i(x) = G_i(K) - G_{i+1}(K), \tag{4}$$

where $\overline{V}(x) \equiv 1 - V(x)$ for any distribution $V(x)$.

Letting $c_S + c_M(x)$ be the cost suffered for every minor collection, where c_S is the constant cost of scanning surviving objects and x is the surviving objects that should be copied, $c_M(x)$ increases with x and $c_M(0) \equiv 0$. Then, the expected cost of the ith minor collection is

$$C(i, K) \equiv \frac{1}{G_i(K)} \int_0^K [c_S + c_M(x)]\,\mathrm{d}G_i(x) \qquad (i = 0, 1, 2, \ldots), \tag{5}$$

where $C(0, K) \equiv 0$ and $C(i, K)$ increases with i.

3 Tenuring Collection Times

Suppose that minor collections are made when the garbage collector begins to work, tenuring collection is made at a planned time T $(0 < T \leq \infty)$ or at the first collection time when surviving objects have exceeded a threshold level K $(0 < K \leq \infty)$, whichever occurs first. Then, the probability that tenuring collection is made at time T is

$$P_T = \sum_{j=0}^{\infty} H_j(T) G_j(K), \tag{6}$$

and the probability that tenuring collection is made at level K is

$$P_K = \sum_{j=0}^{\infty} F_{j+1}(T) p_j(K), \tag{7}$$

where note that $P_T + P_K \equiv 1$. The mean time to tenuring collection is

$$\begin{aligned} E_1(L) &= T \sum_{j=0}^{\infty} H_j(T) G_j(K) + \sum_{j=0}^{\infty} p_j(K) \int_0^T t\mathrm{d}F_{j+1}(t) \\ &= \sum_{j=0}^{\infty} G_j(K) \int_0^T H_j(t)\mathrm{d}t. \end{aligned} \tag{8}$$

The expected cost suffered for minor collections until tenuring collection is

$$C_M = \sum_{j=1}^{\infty}\sum_{i=1}^{j} C(i, K)H_j(T)G_j(K) + \sum_{j=1}^{\infty}\sum_{i=1}^{j} C(i, K)F_{j+1}(T)p_j(K)$$
$$= \sum_{j=1}^{\infty} C(j, K)F_j(T)G_j(K). \tag{9}$$

Then, the expected cost until tenuring collection is

$$E_1(C) = c_K - (c_K - c_T)\sum_{j=0}^{\infty} H_j(T)G_j(K) + \sum_{j=1}^{\infty} C(j, K)F_j(T)G_j(K), \tag{10}$$

where c_T and c_K $(c_T,\ c_K > c_S + c_M(K))$ are the costs suffered for tenuring collections at time T and when surviving objects have exceeded K, respectively. Therefore, from (8) to (10), by using the theory of renewal reward process [13], the expected cost rate is

$$C_1(T, K) = \frac{c_K - (c_K - c_T)\sum_{j=0}^{\infty} H_j(T)G_j(K) + \sum_{j=1}^{\infty} C(j, K)F_j(T)G_j(K)}{\sum_{j=0}^{\infty} G_j(K)\int_0^T H_j(t)\mathrm{d}t}. \tag{11}$$

3.1 Optimal Policies

1. Optimal T_1^*: When tenuring collection is made only at time T,

$$C_1(T) \equiv \lim_{K\to\infty} C_1(T, K) = \frac{1}{T}\left\{\sum_{j=1}^{\infty} F_j(T)\int_0^{\infty}[c_S + c_M(x)]\,\mathrm{d}G_j(x) + c_T\right\}. \tag{12}$$

Letting $f_j(t)$ be a density function of $F_j(t)$, i.e., $f_j(t) \equiv \mathrm{d}F_j(t)/\mathrm{d}t$. Then, differentiating $C_1(T)$ with respect to T and setting it equal to zero,

$$\sum_{j=1}^{\infty}\left[Tf_j(T) - F_j(T)\right]\int_0^{\infty}[c_S + c_M(x)]\,\mathrm{d}G_j(x) = c_T. \tag{13}$$

Letting $L_1(T)$ be the left-hand side of (13),

$$L_1(0) \equiv \lim_{T\to 0} L(T) = 0,$$

$$L_1'(T) = \lambda'(T)T \sum_{j=0}^{\infty} H_j(T) \int_0^{\infty} [c_S + c_M(x)]\, \mathrm{d}G_{j+1}(x)$$
$$+ \lambda(T)^2 T \sum_{j=0}^{\infty} H_j(T) \int_0^{\infty} p_{j+1}(x) \mathrm{d}c_M(x).$$

Thus, if $\lambda(t)$ increases with t and $L_1(\infty) > c_T$, there exists a finite and unique T_1^* $(0 < T_1^* < \infty)$ which satisfies (13), and the resulting cost rate is

$$C_1(T_1^*) = \lambda(T_1^*) \sum_{j=0}^{\infty} F_j(T_1^*) \int_0^{\infty} p_j(x) \mathrm{d}c_M(x).$$

In particular, when $H_j(t) = [(\lambda t)^j / j!]\mathrm{e}^{-\lambda t}$ $(j = 0, 1, 2, \ldots)$, i.e., garbage collections occur at a Poisson process with rate λ, (13) becomes

$$\sum_{j=1}^{\infty} j F_{j+1}(T) \int_0^{\infty} p_j(x) \mathrm{d}c_M(x) = c_T. \tag{14}$$

Differentiating the left-hand side of (14) with respect to T,

$$\lambda \sum_{j=1}^{\infty} j H_j(T) \int_0^{\infty} p_j(x) \mathrm{d}c_M(x) > 0.$$

Thus, if the left-hand side of (14) is greater than c_T, then there exists a finite and unique T_1^* $(0 < T_1^* < \infty)$ which satisfies (14).

2. Optimal K_1^*: When tenuring collection is made only at level K,

$$C_1(K) \equiv \lim_{T \to \infty} C_1(T, K) = \frac{\sum_{j=1}^{\infty} \int_0^K [c_S + c_M(x)] \mathrm{d}G_j(x) + c_K}{\sum_{j=0}^{\infty} G_j(K) \int_0^{\infty} H_j(t) \mathrm{d}t}. \tag{15}$$

Letting $g_i(x)$ be a density function of $G_i(x)$ in (3), i.e., $g_i(x) \equiv \mathrm{d}G_i(x)/\mathrm{d}x$. Differentiating $C_1(K)$ with respect to K and setting it equal to zero,

$$Q_1(K) \sum_{j=0}^{\infty} G_j(K) \int_0^{\infty} H_j(t) \mathrm{d}t - \sum_{j=1}^{\infty} \int_0^K [c_S + c_M(x)]\, \mathrm{d}G_j(x) = c_K, \tag{16}$$

where

$$Q_1(K) \equiv \frac{[c_S + c_M(K)] \sum_{j=1}^{\infty} g_j(K)}{\sum_{j=1}^{\infty} g_j(K) \int_0^{\infty} H_j(t) \mathrm{d}t}.$$

Letting $L_1(K)$ be the left-hand side of (16),

$$L_1(0) \equiv \lim_{K\to 0} L_1(K) = Q_1(0)\int_0^\infty H_0(t)\mathrm{d}t,$$

$$L_1'(K) = Q_1'(K)\sum_{j=0}^{\infty} G_j(K)\int_0^\infty H_j(t)\mathrm{d}t.$$

Thus, if $Q_1(K)$ increases with K and $L_1(0) < c_K < L_1(\infty)$, then there exists a finite and unique K_1^* $(0 < K_1^* < \infty)$ which satisfies (16), and the resulting cost rate is

$$C_1(K_1^*) = Q_1(K_1^*).$$

In particular, when $H_j(t) = [(\lambda t)^j/j!]\mathrm{e}^{-\lambda t}$, (16) becomes

$$c_M(K) + \int_0^K [c_M(K) - c_M(x)]\,\mathrm{d}M(x) = c_K - c_S, \tag{17}$$

whose left-hand side increases with K from 0 to ∞, where $M(x) \equiv \sum_{j=1}^{\infty} G_j(x)$. Thus, there exists a finite and unique K_1^* $(0 < K_1^* < \infty)$ which satisfies (17).

3.2 Numerical Examples

When $\lambda(t) = \lambda$, X_i $(i = 1, 2, \ldots)$ has a normal distribution $N(\mu, \sigma^2)$, $\alpha_i = \alpha/i$ $(0 \le \alpha < 1; i = 1, 2, \ldots)$ and $c_M(x) = c_M x$. Then

$$F_j(t) = 1 - \sum_{i=0}^{j-1}\frac{(\lambda t)^i}{i!}\mathrm{e}^{-\lambda t}, \qquad G_j(x) = \Phi\left(\frac{x - \alpha\mu\nu_j}{\alpha\sigma\sqrt{\omega_j}}\right), \tag{18}$$

where $\Phi(x)$ is the standard normal distribution with mean 0 and variance 1, i.e., $\Phi(x) \equiv (1/\sqrt{2\pi})\int_{-\infty}^{x} \mathrm{e}^{-u^2/2}\mathrm{d}u$, and

$$\nu_j \equiv \sum_{n=1}^{j}\frac{1}{n}, \qquad \omega_j \equiv \sum_{n=1}^{j}\frac{1}{n^2}.$$

Tables 1 and 2 present λT_1^*, $C_1(T_1^*)/\lambda$, K_1^* and $C_1(K_1^*)/\lambda$ for $c_T = c_K = 20$, 30, 40, $\mu = 8$, 10 and $\alpha = 0.40, 0.45, 0.50, 0.55, 0.60$ when $c_S = 10$, $c_M = 1$ and $\sigma = 1$. These show that optimal tenuring collection times λT_1^* increase with cost c_T and decrease with both the volume of new objects in Eden at collection time μ and the survivor rate α, optimal tenuring collection times K_1^* increase with all of c_K, μ and α, and $C_1(T_1^*)/\lambda$ and $C_1(K_1^*)/\lambda$ increase with all of c_T or c_K, μ and α. We can explain all the results and obtain some interesting conclusions as follows:

Table 1 Optimal λT_1^* and $C_1(T_1^*)/\lambda$ when $c_S = 10$, $c_M = 1$ and $\sigma = 1$

μ	α	$c_T = 20$		$c_T = 30$		$c_T = 40$	
		λT_1^*	$C_1(T_1^*)/\lambda$	λT_1^*	$C_1(T_1^*)/\lambda$	λT_1^*	$C_1(T_1^*)/\lambda$
	0.40	8.99	19.24	12.48	20.18	15.84	20.89
	0.45	8.24	20.11	11.34	21.14	14.35	21.92
8	0.50	7.61	20.95	10.42	22.07	13.15	22.92
	0.55	7.08	21.77	9.66	22.98	12.17	23.90
	0.60	6.64	22.58	9.03	23.86	11.34	24.85
	0.40	7.61	20.95	10.42	22.07	13.14	22.91
	0.45	6.96	21.98	9.49	23.20	11.95	24.14
10	0.50	6.44	22.97	8.75	24.30	10.97	25.31
	0.55	6.01	23.95	8.13	25.37	10.17	26.47
	0.60	5.64	24.91	7.61	26.43	9.49	27.60

Table 2 Optimal K_1^* and $C_1(K_1^*)/\lambda$ when $c_S = 10$, $c_M = 1$ and $\sigma = 1$

μ	α	$c_K = 20$		$c_K = 30$		$c_K = 40$	
		K_1^*	$C_1(K_1^*)/\lambda$	K_1^*	$C_1(K_1^*)/\lambda$	K_1^*	$C_1(K_1^*)/\lambda$
	0.40	8.76	18.76	9.61	19.61	10.71	20.71
	0.45	9.25	19.25	11.04	21.04	11.57	21.57
8	0.50	9.71	19.71	12.03	22.03	12.39	22.39
	0.55	10.12	20.12	12.69	22.69	13.18	23.18
	0.60	10.49	20.49	13.32	23.32	13.94	23.94
	0.40	10.72	20.72	12.05	22.05	12.41	22.41
	0.45	11.24	21.24	12.87	22.87	13.39	23.39
10	0.50	11.64	21.64	13.64	23.64	14.33	24.33
	0.55	12.08	22.08	14.37	24.37	15.23	25.23
	0.60	12.44	22.44	15.07	25.07	16.09	26.09

- When tenuring collection cost c_T or c_K increases, it is not economical to make tenuring collections frequently, then T_1^* or K_1^* should be postponed.
- When μ or α increases, cost suffered for minor collections will increase in a shorter time, because of faster increase in copied objects. If cost c_T or c_K is constant in this case, T_1^* should be advanced. For K_1^*, it costs much shorter time to increase copied objects until level K, then K_1^* would increase suitably to decrease both the frequency of tenuring collections and the total minor collection cost.
- The resulting cost rates $C_1(T_1^*)$ or $C_1(K_1^*)$ increase with all μ, α and c_T or c_K, because the total expected cost of one cycle increases but the expected time decreases.
- It is interesting that $C_1(K_1^*)$ are always less than $C_1(T_1^*)$ for the same parameters, i.e., tenuring collections at level K are better than those at time T. In fact, from Tables 1 and 2, we know that the expected number of minor collections until tenuring collection for two models are almost the same. That is, from the assumption

of $\alpha_i = \alpha/i$, we can derive

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{[\lambda T_1^*]} < \frac{K_1^*}{\alpha\mu} < 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{[\lambda T_1^*] + 1}, \tag{19}$$

where $[x]$ denotes the greatest integer contained in x. For example, when $c_T = c_K = 20$, $\mu = 8$ and $\alpha = 0.4$, $\lambda T_1^* = 8.99$ and $K_1^* = 8.76$, and hence

$$1 + \frac{1}{2} + \cdots + \frac{1}{8} = 2.55 < \frac{8.76}{0.4 \times 8} = 2.74 < 1 + \frac{1}{2} + \cdots + \frac{1}{9} = 2.83.$$

We can estimate approximate values K_1^* from T_1^* using the relationship of the two policies in (19), and vice versa.

4 Major Collection Times

4.1 Model 1

Suppose that minor collections are made before surviving objects exceed a threshold level K $(0 < K < \infty)$, and when they have exceeded K, tenuring collections are always made. Further, major collection is made at time T $(0 < T \leq \infty)$ or at the Nth $(N = 1, 2, \ldots)$ collection including minor and tenuring collections, whichever occurs first. Furthermore, Letting c_{kT} $(k = 1, 2, \ldots)$ be the cost suffered for the kth tenuring collection, where $c_S + c_M(K) < c_{1T} < c_{2T} < \cdots$, and c_F $(c_F > c_{kT})$ be the cost suffered for major collection. Then, the probability that major collection is made at time T is

$$P_T = \sum_{j=0}^{N-1} H_j(T) G^{(j)}(K) + \sum_{j=1}^{N-1} \sum_{i=1}^{j-1} H_j(T) p_i(K) = 1 - F_N(T), \tag{20}$$

and the probability that major collection is made at collection N is

$$P_N = F_N(T) G^{(N)}(K) + \sum_{j=0}^{N-1} F_N(T) p_j(K) = F_N(T), \tag{21}$$

where note that $P_T + P_N \equiv 1$. The mean time to major collection is

$$E_2(L) = \int_0^T t \, \mathrm{d}F_N(t) + T \sum_{j=0}^{N-1} H_j(T) = \int_0^T [1 - F_N(t)] \, \mathrm{d}t. \tag{22}$$

The expected costs suffered for minor collections and tenuring collections when major collection is made at time T are, respectively,

$$C_{TM} = \sum_{j=1}^{N-1} H_j(T)\left[\sum_{i=1}^{j} C(i,K)G^{(j)}(K) + \sum_{i=1}^{j-1}\sum_{k=1}^{i} C(k,K)p_i(K)\right]$$

$$= \sum_{j=1}^{N-1} H_j(T)\sum_{i=1}^{j} C(i,K)G^{(i)}(K), \tag{23}$$

$$C_{TT} = \sum_{j=1}^{N-1} H_j(T)\sum_{i=0}^{j-1}\sum_{k=1}^{j-i} c_{kT}\, p_i(K)$$

$$= \sum_{j=1}^{N-1} H_j(T)\sum_{i=0}^{j-1}\left[c_{(i+1)T} - c_{(j-i)T}G^{(i+1)}(K)\right], \tag{24}$$

and the expected costs suffered for minor collections and tenuring collections when major collection is made at collection N are, respectively,

$$C_{NM} = F_N(T)\left[\sum_{j=1}^{N} C(j,K)G^{(N)}(K) + \sum_{j=1}^{N-1}\sum_{i=1}^{j} C(i,K)p_j(K)\right]$$

$$= F_N(T)\sum_{j=1}^{N} C(j,K)G^{(j)}(K), \tag{25}$$

$$C_{NT} = F_N(T)\sum_{j=0}^{N-1}\sum_{i=1}^{N-j} c_{iT}\, p_j(K)$$

$$= F_N(T)\sum_{j=0}^{N-1}\left[c_{(j+1)T} - c_{(N-j)T}G^{(j+1)}(K)\right]. \tag{26}$$

Thus, the total expected cost until major collection is, summing up from (23) to (26) and adding the cost c_F of major collection,

$$E_2(C) = c_F + \sum_{j=1}^{N} C(j,K)F_j(T)G^{(j)}(K)$$

$$+ \sum_{j=1}^{N} F_j(T)\left[c_{jT} - \sum_{i=0}^{j-1} G^{(j-i)}(K)(c_{(i+1)T} - c_{iT})\right]. \tag{27}$$

Therefore, the expected cost rate is, from (22) and (27),

$$C_2(T, N) = \frac{c_F + \sum_{j=1}^{N} F_j(T) A_j}{\int_0^T [1 - F_N(t)]\mathrm{d}t}, \tag{28}$$

where

$$A_j \equiv c_{jT} + \int_0^K [c_S + c_M(x)]\, \mathrm{d}G^{(j)}(x) - \sum_{i=0}^{j-1} G^{(j-i)}(K)(c_{(i+1)T} - c_{iT}).$$

It can be easily proved that A_j increases with j because

$$A_{j+1} - A_j = (c_{1T} - c_S - c_M(K))p_j(K) + \int_0^K p_j(x)\mathrm{d}c_M(x)$$
$$+ \sum_{i=1}^{j} p_{j-i}(K)(c_{(i+1)T} - c_{iT}) > 0.$$

1. Optimal T_2^*: When major collection is made only at time T,

$$C_2(T) \equiv \lim_{N\to\infty} C_2(T, N) = \frac{1}{T}\left[\sum_{j=1}^{\infty} F_j(T) A_j + c_F\right]. \tag{29}$$

Differentiating $C_2(T)$ in (29) with respect to T and setting it equal to zero,

$$\sum_{j=1}^{\infty} A_j \left[T\lambda(T) H_{j-1}(T) - F_j(T)\right] = c_F,$$

that is,

$$\sum_{j=1}^{\infty} A_j \int_0^T t\mathrm{d}\left[\lambda(t) H_{j-1}(t)\right] = c_F. \tag{30}$$

Letting $L_2(T)$ be the left-hand side of (30),

$$L_2'(T) = \sum_{j=0}^{\infty} A_{j+1} \int_0^T t\lambda'(t) H_j(t)\mathrm{d}t + \sum_{j=0}^{\infty} (A_{j+2} - A_{j+1}) \int_0^T t[\lambda(t)]^2 H_j(t)\mathrm{d}t,$$

$$L_2(\infty) = \sum_{j=1}^{\infty} A_j \int_0^{\infty} t\mathrm{d}[\lambda(t) H_{j-1}(t)].$$

Thus, if $\lambda(t)$ increases with t and $L_2(\infty) > c_F$, then there exists a finite and unique T_2^* $(0 < T_2^* < \infty)$ which satisfies (30).

In particular, when $\lambda(t) = \lambda$,

$$L_2'(T) = \sum_{j=0}^{\infty}(j+1)F_{j+2}(T)(A_{j+2} - A_{j+1}),$$

$$L_2(\infty) = \sum_{j=1}^{\infty}(A_\infty - A_j).$$

Therefore, if $\sum_{j=1}^{\infty}(A_\infty - A_j) > c_F$, then there exists a finite and unique T_2^* $(0 < T_2^* < \infty)$, and the resulting cost rate is

$$\frac{C_2(T_2^*)}{\lambda} = \sum_{j=0}^{\infty} H_j(T_2^*)A_{j+1}.$$

2. Optimal N_2^*: When major collection is made only at collection N,

$$C_2(N) \equiv \lim_{T\to\infty} C_2(T, N) = \frac{\sum_{j=1}^{N} A_j + c_F}{\int_0^\infty [1 - F_N(t)]\mathrm{d}t} \quad (N = 1, 2, \ldots). \tag{31}$$

From the inequality $C_2(N+1) - C_2(N) \geq 0$,

$$\sum_{j=0}^{N-1}\left[\frac{A_{N+1}}{\int_0^\infty H_N(t)\mathrm{d}t}\int_0^\infty H_j(t)\mathrm{d}t - A_{j+1}\right] \geq c_F. \tag{32}$$

Letting $L_2(N)$ be the left-hand side of (32),

$$L_2(N+1) - L_2(N) = \left[\frac{A_{N+2}}{\int_0^\infty H_{N+1}(t)\mathrm{d}t} - \frac{A_{N+1}}{\int_0^\infty H_N(t)\mathrm{d}t}\right]\int_0^\infty \left[1 - F_{N+1}(t)\right]\mathrm{d}t. \tag{33}$$

Thus, if $A_{N+1}/\int_0^\infty H_N(t)\mathrm{d}t$ increases with N and $L_2(\infty) > c_F$, then there exists a finite and unique minimum N_2^* $(1 \leq N_2^* < \infty)$ which satisfies (32).

In particular, when $\lambda(t) = \lambda$,

$$L_2(N) = \sum_{j=1}^{N}(A_{N+1} - A_j),$$

$$L_2(N+1) - L_2(N) = (N+1)(A_{N+2} - A_{N+1}) > 0.$$

It is assumed that $A_\infty \equiv \lim_{j\to\infty} A_j < \infty$. Then,

$$L_2(\infty) = \sum_{j=1}^{\infty}(A_\infty - A_j).$$

Further, because $\sum_{j=1}^{N}(A_{N+1} - A_j) \geq A_{N+1} - A_1$ $(N = 1, 2, \ldots)$, if $A_\infty = \infty$, then $L_2(\infty) = \infty$. Therefore, if $\sum_{j=1}^{\infty}(A_\infty - A_j) > c_F$, then there exists a finite and unique minimum N_2^* $(1 \leq N_2^* < \infty)$, and the resulting cost rate is

$$A_{N_2^*} \leq \frac{C_2(N_2^*)}{\lambda} < A_{N_2^*+1}.$$

It is of interest that when collections occur at a Poisson process with rate λ, if $\sum_{j=1}^{\infty}(A_\infty - A_j) > c_F$, then both finite and unique T_2^* and N_2^* exist.

4.2 Model 2

Suppose that minor collections are made before surviving objects exceed a threshold level K, and after they have exceeded K, tenuring collections are always made. Further, major collection is made at time T $(0 < T \leq \infty)$ or at collection N $(N = 1, 2, \ldots)$ including tenuring collections, whichever occurs first. Then, the probability that major collection is made at time T is

$$P_T = \sum_{j=0}^{\infty}\sum_{i=0}^{N-2} p_j(K)\int_0^\infty H_i(u, u+T)\mathrm{d}F_{j+1}(u), \tag{34}$$

and the probability that major collection is made at collection N is

$$P_N = \sum_{j=0}^{\infty}\sum_{i=N-1}^{\infty} p_j(K)\int_0^\infty H_i(u, u+T)\mathrm{d}F_{j+1}(u). \tag{35}$$

The mean time to major collection is

$$E_3(L) = \sum_{j=0}^{\infty} p_j(K)\int_0^\infty \left[\int_0^T (u+t)\mathrm{d}F_{N-1}(u, u+t)\right]\mathrm{d}F_{j+1}(u)$$
$$+ \sum_{j=0}^{\infty}\sum_{i=0}^{N-2} p_j(K)\int_0^\infty (u+T)H_i(u, u+T)\mathrm{d}F_{j+1}(u)$$

$$= \sum_{j=0}^{\infty} p_j(K) \int_0^{\infty} u \mathrm{d}F_{j+1}(u)$$
$$+ \sum_{j=0}^{\infty} p_j(K) \int_0^{\infty} \left\{ \int_0^T \left[1 - F_{N-1}(u, u+t)\right] \mathrm{d}t \right\} \mathrm{d}F_{j+1}(u). \quad (36)$$

The expected costs suffered for minor collections and tenuring collections when major collection is made at time T are, respectively,

$$C_{TM} = \sum_{j=0}^{\infty} \sum_{i=1}^{j} c_{iM} p_j(K) \int_0^{\infty} \left[1 - F_{N-1}(u, u+T)\right] \mathrm{d}F_{j+1}(u), \quad (37)$$
$$C_{TT} = \sum_{j=0}^{\infty} \sum_{i=0}^{N-2} \sum_{k=1}^{i+1} c_{kT} p_j(K) \int_0^{\infty} H_i(u, u+T) \mathrm{d}F_{j+1}(u), \quad (38)$$

and the expected costs suffered for minor collections and tenuring collections when major collection is made at collection N are, respectively,

$$C_{NM} = \sum_{j=0}^{\infty} \sum_{i=1}^{j} c_{iM} p_j(K) \int_0^{\infty} F_{N-1}(u, u+T) \mathrm{d}F_{j+1}(u), \quad (39)$$
$$C_{NT} = \sum_{j=0}^{\infty} \sum_{i=1}^{N} c_{iT} p_j(K) \int_0^{\infty} F_{N-1}(u, u+T) \mathrm{d}F_{j+1}(u). \quad (40)$$

Thus, the total expected cost until major collection is, summing up from (37) to (40) and adding the cost c_F of major collection,

$$E_3(C) = c_F + \sum_{j=1}^{\infty} \sum_{i=1}^{j} c_{iM} p_j(K)$$
$$+ \sum_{j=0}^{\infty} \sum_{i=1}^{N} c_{iT} p_j(K) \int_0^{\infty} F_{i-1}(u, u+T) \mathrm{d}F_{j+1}(u). \quad (41)$$

Therefore, from (36) to (41), the expected cost rate is

$$C_3(T, N) = \frac{E_3(C)}{E_3(L)}. \quad (42)$$

1. Optimal T_3^*: When major collection is made only at time T,

$$C_3(T) \equiv \lim_{N\to\infty} C_3(T,N) = \frac{\begin{array}{l} c_F + \sum_{j=1}^{\infty}\sum_{i=1}^{j} c_{iM}\, p_j(K) \\ + \sum_{j=0}^{\infty}\sum_{i=1}^{\infty} c_{iT}\, p_j(K) \\ \times \int_0^{\infty} F_{i-1}(u, u+T)\mathrm{d}F_{j+1}(u) \end{array}}{\sum_{j=0}^{\infty} p_j(K)\int_0^{\infty} u\mathrm{d}F_{j+1}(u) + T}. \tag{43}$$

Differentiating $C_3(T)$ with respect to T and setting it equal to zero,

$$\sum_{j=0}^{\infty} p_{j+1}(K)\int_0^{\infty} Q_3(u,T)\mathrm{d}F_{j+1}(u) = c_F + \sum_{j=1}^{\infty} c_{jM} G^{(j)}(K), \tag{44}$$

where

$$\begin{aligned} Q_3(u,T) &\equiv \sum_{i=1}^{\infty} c_{iT}\int_0^{\infty} (l+x)\,\mathrm{d}\left[\lambda(u+x)H_{i-2}(u,u+x)\right] \\ &= \sum_{i=1}^{\infty} c_{iT}\int_0^{\infty} (l+x)\,\lambda'(u+x)H_{i-2}(u,u+x)\mathrm{d}x \\ &\quad + \sum_{i=1}^{\infty}(c_{(i+3)T} - c_{(i+2)T})\int_0^{\infty} (l+x)\,[\lambda(u+x)]^2 H_i(u,u+x)\mathrm{d}x, \end{aligned}$$

and

$$l \equiv \sum_{j=1}^{\infty} p_j(K)\int_0^{\infty} t\mathrm{d}F_j(t),$$

which represents the mean time until surviving objects have exceeded K. Letting $L_3(T)$ be the left-hand side of (44). Thus, if $\lambda(t)$ increases with t, $L_3(T)$ increases with T. Therefore, if $L_3(\infty) > c_F + \sum_{j=1}^{\infty} c_{jM} G^{(j)}(K)$, then there exists a finite and unique T_3^* $(0 < T_3^* < \infty)$ which satisfies (44).

In particular, when $\lambda(t) = \lambda$, then $l = [1 + M(K)]/\lambda$, and

$$\begin{aligned} Q_3(u,T) = [1+M(K)]\sum_{j=1}^{\infty} F_j(T)(c_{(j+2)T} - c_{(j+1)T}) \\ + \sum_{j=1}^{\infty} jF_{j+1}(T)(c_{(j+2)T} - c_{(j+1)T}), \end{aligned}$$

$$L_3(\infty) = \sum_{j=1}^{\infty}(c_{\infty T} - c_{(j+1)T}) + [1+M(K)](c_{\infty T} - c_{2T}).$$

Therefore, if

$$\sum_{j=1}^{\infty}(c_{\infty T}-c_{(j+1)T})+[1+M(K)](c_{\infty T}-c_{2T})>c_F+\sum_{j=1}^{\infty}c_{jM}G^{(j)}(K),$$

then there exists a finite and unique T_3^* $(0<T_3^*<\infty)$, and the resulting cost rate is

$$\frac{C_3(T_3^*)}{\lambda}=\sum_{j=0}^{\infty}H_j(T_3^*)c_{(j+2)T}.$$

2. Optimal N_3^*: When major collection is made only at collection N,

$$C_3(N)\equiv\lim_{T\to\infty}C_3(T,N)=\frac{c_F+\sum_{j=1}^{\infty}\sum_{i=1}^{j}c_{iM}p_j(K)+\sum_{j=1}^{N}c_{jT}}{\sum_{j=0}^{\infty}p_j(K)\int_0^{\infty}[1-F_{j+N}(t)]\mathrm{d}t}$$
$$(N=1,2,\ldots). \quad (45)$$

From the inequality $C_3(N+1)-C_3(N)\geq 0$,

$$Q_3(N)c_{(N+1)T}-\sum_{j=1}^{N}c_{jT}\geq c_F+\sum_{j=1}^{\infty}c_{jM}G^{(j)}(K), \quad (46)$$

where

$$Q_3(N)\equiv\frac{\sum_{j=0}^{\infty}p_j(K)\int_0^{\infty}[1-F_{j+N}(t)]\mathrm{d}t}{\sum_{j=0}^{\infty}p_j(K)\int_0^{\infty}H_{j+N}(t)\mathrm{d}t.}$$

Letting $L_3(N)$ be the left-hand side of (46),

$$L_3(N+1)-L_3(N)=\left[\widetilde{Q}_3(N+1)-\widetilde{Q}_3(N)\right]\sum_{j=0}^{\infty}p_j(K)\int_0^{\infty}\left[1-F_{j+N+1}(t)\right]\mathrm{d}t,$$

where

$$\widetilde{Q}_3(i)\equiv\frac{c_{(i+1)T}}{\sum_{j=0}^{\infty}p_j(K)\int_0^{\infty}H_{j+i}(t)\mathrm{d}t}.$$

Thus, if $\widetilde{Q}_3(i)$ increases with i, $L_3(N)$ increases with N. Therefore, if $L_3(\infty)>c_F+\sum_{j=1}^{\infty}c_{jM}G^{(j)}(K)$, then there exists a finite and unique minimum N_3^* $(1\leq N_3^*<\infty)$ which satisfies (46).

In particular, when $\lambda(t)=\lambda$, then $Q_3(N)=M(K)+N$, where $M(x)\equiv\sum_{j=1}^{\infty}G^{(j)}(x)$ is the expected number of minor collections before surviving objects exceed x, and

$$L_3(N) = \sum_{j=1}^{N}(c_{(N+1)T} - c_{jT}) + M(K)c_{(N+1)T},$$

$$L_3(N+1) - L_3(N) = [M(K) + N + 1]\,(c_{(N+2)T} - c_{(N+1)T}) > 0.$$

It is assumed that $c_{\infty T} \equiv \lim_{j\to\infty} c_{jT} < \infty$. Then,

$$L_3(\infty) = \sum_{j=1}^{\infty}(c_{\infty T} - c_{jT}) + M(K)c_{\infty T}.$$

Clearly, if $c_{\infty T} = \infty$, then $L_2(\infty) = \infty$. Therefore, if

$$\sum_{j=1}^{\infty}(c_{\infty T} - c_{jT}) + M(K)c_{\infty T} > c_F + \sum_{j=1}^{\infty} c_{jM} G^{(j)}(K),$$

then there exists a finite and unique minimum N_3^* $(1 \le N_3^* < \infty)$ which satisfies (46), and the resulting cost rate is

$$c_{N_3^* T} \le \frac{C_3(N_3^*)}{\lambda} < c_{(N_3^*+1)T}.$$

4.3 Numerical Examples

It is assumed that $c_{kT} = c_T + k\beta$ $(\beta > 0; k = 1, 2, \ldots)$, and other assumptions are the same as in Sect. 3.2. We give numerical examples of each model as follows:

Tables 3–6 present optimal λT_i^* and $C_i(T_i^*)/\lambda$ $(i = 2, 3)$, N_i^* and $C_i(N_i^*)/\lambda$ $(i = 2, 3)$, when $c_F = 100$, $c_T = c_N = 20$, $c_S = 10$, $c_M = 1$, $\mu = 10$ and $\sigma = 1$ for different α and β. These show that both λT_2^* and N_2^* decrease with α or β, both λT_3^* and N_3^* increase with α and decrease with β, all $C_i(T_i^*)/\lambda$ $(i = 2, 3)$ and $C_i(N_i^*)/\lambda$ $(i = 2, 3)$ increase with α or β.

It can be explained as follows:

- When α or β increases, it means that the total cost suffered for minor collections or tenuring collections increases, then optimal major collection times should be advanced, but even then the expected cost rates increase.
- The differences between Tables 3 and 5, Tables 4 and 6, are that when α increases, $M(K)$ decreases, then optimal major collection times should be postponed, because it is not economic to make major collection frequently.
- Compared Tables 3 with 4, Tables 5 with 6, these show that $C_2(T_2^*) > C_2(N_2^*)$ and $C_3(T_3^*) > C_3(N_3^*)$ for the same parameters, that is, major collections made at N_2 or N_3 are better than those at T_2 or T_3. It is interesting that $C_2(N_2^*) \approx C_3(N_3^*)$ and

Table 3 Optimal λT_2^* and $C_2(T_2^*)/\lambda$ when $c_F = 100$, $c_T = 20$, $c_S = 10$, $c_M = 1$, $\mu = 10$ and $\sigma = 1$

α	$\beta = 1$		$\beta = 2$		$\beta = 5$	
	λT_2^*	$C_2(T_2^*)/\lambda$	λT_2^*	$C_2(T_2^*)/\lambda$	λT_2^*	$C_2(T_2^*)/\lambda$
0.3	17.98	24.4699	15.51	25.1134	13.14	26.0229
0.4	14.09	28.1939	10.66	31.5677	7.77	35.1867
0.5	13.80	31.6197	9.95	35.5256	6.60	42.2212
0.6	12.86	32.8499	9.86	37.6922	6.33	46.6393
0.7	12.86	33.6762	9.86	39.2143	6.27	50.0259

Table 4 Optimal N_2^* and $C_2(N_2^*)/\lambda$ when $c_F = 100$, $c_N = 20$, $c_S = 10$, $c_M = 1$, $\mu = 10$ and $\sigma = 1$

α	$\beta = 1$		$\beta = 2$		$\beta = 5$	
	N_2^*	$C_2(N_2^*)/\lambda$	N_2^*	$C_2(N_2^*)/\lambda$	N_2^*	$C_2(N_2^*)/\lambda$
0.3	17	24.1785	16	24.3642	15	24.7538
0.4	14	28.6941	11	30.5818	8	32.8485
0.5	14	31.1212	10	34.5257	7	39.7804
0.6	14	32.3506	10	36.6942	6	44.1853
0.7	14	33.1763	10	38.2160	6	47.5548

Table 5 Optimal λT_3^* and $C_3(T_3^*)/\lambda$ when $c_F = 100$, $c_T = 20$, $c_S = 10$, $c_M = 1$, $\mu = 10$ and $\sigma = 1$

α	$\beta = 1$		$\beta = 2$		$\beta = 5$	
	λT_3^*	$C_3(T_3^*)/\lambda$	λT_3^*	$C_3(T_3^*)/\lambda$	λT_3^*	$C_3(T_3^*)/\lambda$
0.3	1.95	23.9629	0.06	24.1471	0.01	24.3401
0.4	6.92	28.9179	3.42	30.8389	0.57	32.8532
0.5	9.45	31.4559	5.54	35.0679	2.10	40.4869
0.6	10.75	32.7357	6.69	37.3694	3.06	45.3683
0.7	11.60	33.5878	7.48	38.9639	3.80	49.0298

Table 6 Optimal N_3^* and $C_3(N_3^*)/\lambda$ when $c_F = 100$, $c_N = 20$, $c_S = 10$, $c_M = 1$, $\mu = 10$ and $\sigma = 1$

α	$\beta = 1$		$\beta = 2$		$\beta = 5$	
	N_3^*	$C_3(N_3^*)/\lambda$	N_3^*	$C_3(N_3^*)/\lambda$	N_3^*	$C_3(N_3^*)/\lambda$
0.3	3	23.9071	2	24.1387	1	24.3365
0.4	8	28.6662	5	30.5026	2	32.5947
0.5	10	31.1223	7	34.4997	3	39.6811
0.6	12	32.3455	8	36.6799	4	44.1223
0.7	13	33.1731	9	38.2105	5	47.4478

$C_2(T_2^*) \approx C_3(T_3^*)$, that is, although the two policies are different, the resulting expected cost rates are almost the same.

- We can derive the relationship of the two polices, that is,

$$\lambda T_2^* \approx 1 + M(K) + \lambda T_3^*,$$
$$N_2^* \approx M(K) + N_3^*.$$

For example, when $\alpha = 0.3$ and $\beta = 1$, $M(K) = 14.8$, then

$$\lambda T_2^* = 17.98, \qquad 1 + M(K) + \lambda T_3^* = 1 + 14.8 + 1.95 = 17.75,$$
$$N_2^* = 17, \qquad M(K) + N_3^* = 14.8 + 3 = 17.8.$$

Therefore, the concrete performances of the two kinds of policies would depend on the program engineers and software system structures at the beginning, and so on.

5 Continuous Models

From the related studies in Sect. 2, we know that the volume of surviving objects that should be copied increases with the number of minor collections and is relatively stable with the number of tenuring collections. However, it may be difficult to inspect the survivor rates exactly at collection times. Hence, in this section, we assume that the total volume of surviving objects in Eden and survivor space at time t is $Z(t) = A(t)t + \sigma B(t)$ with distribution $\Pr\{Z(t) \le x\} = W(t, x)$, where both $A(t)$ and $B(t)$ are random variables of time t. Then, the expected cost of minor collection at time t is

$$C(t, K) = \frac{1}{W(t, K)} \int_0^K [c_S + c_M(x)]\, \mathrm{d}W(t, x), \tag{47}$$

where $C(0, K) \equiv 0$. Letting $r(t, x)$ be the failure rate of $W(t, x)$, i.e., $r(t, x) \equiv -[\mathrm{d}W(t, x)/\mathrm{d}t]/W(t, x)$ [3]. It is clear that if $r(t, x)$ increases with t for any $x \ge 0$, $C(t, K)$ increases with t for any $K \ge 0$.

Suppose that garbage collections occur at a nonhomogeneous Poisson process in Sect. 2, minor collections are made when the garbage collector begins to work, tenuring collection is made at a planned time T $(0 < T \le \infty)$, or when surviving objects have exceeded a threshold level K $(0 < K \le \infty)$, whichever occurs first. Then, the mean time to tenuring collection is

$$E_4(L) = TW(T, K) + \int_0^T t\mathrm{d}\overline{W}(t, K) = \int_0^T W(t, K)\mathrm{d}t, \tag{48}$$

where $\overline{V}(t, x) \equiv 1 - V(t, x)$ for any distribution $V(t, x)$.

The expected cost suffered for minor collections until tenuring collection is

$$C_M = W(T, K) \sum_{j=1}^{\infty} \int_0^T C(t, K) \mathrm{d}F_j(t) + \int_0^T \left[\sum_{j=1}^{\infty} \int_0^t C(u, K) \mathrm{d}F_j(u) \right] \mathrm{d}\overline{W}(t, K)$$
$$= \int_0^T \lambda(t) C(t, K) W(t, K) \mathrm{d}t. \tag{49}$$

Then, the expected cost until tenuring collection is

$$E_4(C) = c_K - (c_K - c_T) W(T, K) + \int_0^T \lambda(t) C(t, K) W(t, K) \mathrm{d}t. \tag{50}$$

Therefore, from (48) to (50), the expected cost rate is

$$C_4(T, K) = \frac{c_K - (c_K - c_T) W(T, K) + \int_0^T \lambda(t) C(t, K) W(t, K) \mathrm{d}t}{\int_0^T W(t, K) \mathrm{d}t}. \tag{51}$$

5.1 *Optimal Policies*

It can be seen that $C_4(T, K)$ includes the following collection polices:

- Tenuring collection is made at time T for a given K, the reason why making such a policy is $c_T < c_K$.
- Tenuring collection is made at level K for a given T. In this case, $c_K < c_T$.
- Tenuring collection is made only at time T or only at level K. In these two cases, $c_K = c_T$.

1. Optimal T_4^*: When $c_T < c_K$, we find an optimal T_4^* which minimizes $C_4(T, K)$ in (51) for a given K. Differentiating $C_4(T, K)$ with respect to T and setting it equal to zero,

$$(c_K - c_T) \left[r(T, K) \int_0^T W(t, K) \mathrm{d}t - \overline{W}(T, K) \right] + \int_0^T \left[\lambda(T) C(T, K) - \lambda(t) C(t, K) \right] W(t, K) \mathrm{d}t = c_T. \tag{52}$$

Letting $L_4(T)$ be the left-hand side of (52),

$$L_4(0) \equiv \lim_{T \to 0} L_4(T) = 0,$$

$$L_4'(T) = (c_K - c_T)r'(T,K)\int_0^T W(t,K)\mathrm{d}t + \left[\lambda'(T)C(T,K) + \lambda(T)C'(T,K)\right]\int_0^T W(t,K)\mathrm{d}t.$$

Thus, if both $r(t, K)$ and $\lambda(t)$ increase with t, then the left-hand side of (52) increases with t from 0. Therefore, there exists a unique optimal T_4^* $(0 < T_4^* \le \infty)$ which satisfies (52), and the resulting cost rate is

$$C_4(T_4^*, K) = (c_K - c_T)r(T_4^*, K) + \lambda(T_4^*)C(T_4^*, K).$$

2. Optimal K_4^*: When $c_K < c_T$, we find an optimal K_4^* which minimizes $C_4(T, K)$ in (51) for a given T. Letting $w(t, x)$ be a density function of $W(t, x)$, i.e., $w(t, x) \equiv \mathrm{d}W(t, x)/\mathrm{d}x$. Then, differentiating $C_4(T, K)$ with respect to K and setting it equal to zero,

$$(c_T - c_K)\left[Q_4(T,K)\int_0^T W(t,K)\mathrm{d}t - W(T,K)\right] + \int_0^T \left[\widetilde{Q}_4(T,K) - \lambda(t)C(t,K)\right]W(t,K)\mathrm{d}t = c_K, \tag{53}$$

where

$$Q_4(T,K) \equiv \frac{w(T,K)}{\int_0^T w(t,K)\mathrm{d}t}, \qquad \widetilde{Q}_4(T,K) \equiv \frac{[c_S + c_M(K)]\int_0^T \lambda(t)w(t,K)\mathrm{d}t}{\int_0^T w(t,K)\mathrm{d}t}.$$

Letting $L_4(K)$ be the left-hand side of (53),

$$L_4(0) \equiv \lim_{K\to 0} L_4(K) = 0,$$

$$L_4'(K) = (c_T - c_K)Q_4'(T,K)\int_0^T W(t,K)\mathrm{d}t + \widetilde{Q}_4'(T,K)\int_0^T W(t,K)\mathrm{d}t.$$

Thus, if both $Q_4(T, K)$ and $\widetilde{Q}_4(T, K)$ increase with K, then the left-hand side of (53) increases with K from 0. Therefore, there exists a unique optimal K_4^* $(0 < K_4^* \le \infty)$ which satisfies (53), and the resulting cost rate is

$$C_4(T, K_4^*) = (c_T - c_K)Q_4(T, K_4^*) + \widetilde{Q}_4(T, K_4^*).$$

3. Optimal $\widetilde{T}_4^*$: When $c_K = c_T$, putting that $K = \infty$ in (51), the expected cost rate is

$$\widetilde{C}_4(T) \equiv \lim_{K\to\infty} C_4(T,K) = \frac{1}{T}\left[\int_0^T \lambda(t)C(t,\infty)\mathrm{d}t + c_T\right], \tag{54}$$

where

$$C(t,\infty) \equiv \int_0^\infty [c_S + c_M(x)]\,\mathrm{d}W(t,x) = c_S + \int_0^\infty \overline{W}(t,x)\mathrm{d}c_M(x).$$

From (52), if $\lambda(t)$ increases with t, then an optimal tenuring collection time $\widetilde{T}_1^*$ which minimizes (54) is given by a unique solution of the equation

$$\int_0^T [\lambda(T)C(T,\infty) - \lambda(t)C(t,\infty)]\,\mathrm{d}t = c_T, \tag{55}$$

and the resulting cost rate is

$$\widetilde{C}_4(\widetilde{T}_4^*) = \lambda(\widetilde{T}_4^*)C(\widetilde{T}_4^*,\infty).$$

In particular, when $\lambda(t) = \lambda$, (55) becomes

$$\int_0^\infty \left\{\int_0^T [W(t,x) - W(T,x)]\mathrm{d}t\right\} \mathrm{d}c_M(x) = \frac{c_T}{\lambda}, \tag{56}$$

which increases with T, and the resulting cost rate is

$$\frac{\widetilde{C}_4(\widetilde{T}_4^*)}{\lambda} = c_S + \int_0^\infty \overline{W}(\widetilde{T}_4^*, x)\mathrm{d}c_M(x).$$

4. Optimal $\widetilde{K}_4^*$: When $c_K = c_T$, putting that $T = \infty$ in (51), the expected cost rate is

$$\widetilde{C}_4(K) = \lim_{T\to\infty} C_4(T,K) = \frac{\int_0^\infty \lambda(t)C(t,K)W(t,K)\mathrm{d}t + c_K}{\int_0^\infty W(t,K)\mathrm{d}t}. \tag{57}$$

From (53), if $\widetilde{Q}_4(\infty, K)$ increases with K, then an optimal tenuring collection time $\widetilde{K}_4^*$ which minimizes (57) is given by a unique solution of the equation

$$\int_0^\infty \left[\widetilde{Q}_4(\infty,K) - \lambda(t)C(t,K)\right] W(t,K)\mathrm{d}t = c_K, \tag{58}$$

and the resulting cost rate is

$$\widetilde{C}_4(\widetilde{K}_4^*) = \widetilde{Q}_4(\infty, \widetilde{K}_4^*).$$

In particular, when $\lambda(t) = \lambda$, (58) becomes

$$\int_0^\infty \left[\int_0^K W(t,x)\mathrm{d}c_M(x)\right]\mathrm{d}t = \frac{c_K}{\lambda}, \tag{59}$$

Table 7 Optimal T_4^* and $C_4(T_4^*, K)$ when $c_T = 10$ and $c_S = \lambda = \mu = \sigma = 1$

K	c_K	$c_M = 0.1$		$c_M = 0.5$		$c_M = 1.0$	
		T_4^*	$C_4(T_4^*, K)$	T_4^*	$C_4(T_4^*, K)$	T_4^*	$C_4(T_4^*, K)$
	20	4.73	0.4723	4.09	0.5462	3.57	0.6324
	30	3.25	0.5473	3.06	0.6087	2.87	0.6822
5	40	2.80	0.5891	2.69	0.6453	2.56	0.7133
	50	2.57	0.6191	2.49	0.6718	2.40	0.7361
	20	7.43	0.2990	5.77	0.4285	4.50	0.5586
	30	6.37	0.3160	5.40	0.4339	4.42	0.5596
10	40	5.91	0.3252	5.18	0.4378	4.36	0.5604
	50	5.63	0.3320	5.03	0.4408	4.31	0.5612

which increases with K and the resulting cost rate is

$$\frac{\widetilde{C}_4(\widetilde{K}_4^*)}{\lambda} = c_S + c_M(\widetilde{K}_4^*).$$

5.2 *Numerical Examples*

We compute numerical examples of the models discussed above for $Z(t) = \mu t + \sigma B(t)$ when $B(t)$ is normally distributed with mean 0 and variance t or for $Z(t) = A(t)t$ when $A(t)$ is normally distributed with mean μ and variance σ^2/t, that is,

$$W(t, x) = \Phi\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right), \tag{60}$$

where $\Phi(x)$ is the standard normal distribution with mean 0 and variance 1, *i.e.*, $\Phi(x) \equiv (1/\sqrt{2\pi}) \int_{-\infty}^{x} \mathrm{e}^{-u^2/2}\mathrm{d}u$.

From Tables 7– 9, we can obtain the following results:

- Optimal tenuring collection times increase with the initial parameters and decrease with minor or tenuring collection cost, however, the resulting cost rates have the opposite tendencies, that is, they decrease with the initial parameters and increase with minor or tenuring collection cost. Take T_4^* and $C_4(T_4^*, K)$ in Table 7 for an example: T_4^* increase with K and decrease with c_K or c_M. Increasing in K, c_K or c_M means that tenuring collection time made at a given level K is postponed, tenuring or minor collection cost is increased, respectively, so that tenuring collection times should be postponed for K or be advanced for c_K or c_M to decrease the frequency of tenuring collections or to decrease the total minor collection cost. $C_4(T_4^*, K)$ decrease with K and increase with c_K or c_M for the reason that the

Table 8 Optimal K_4^* and $C_4(T, K_4^*)$ when $c_K = 10$ and $c_S = \lambda = \mu = \sigma = 1$

T	c_T	$c_M = 0.1$		$c_M = 0.5$		$c_M = 1.0$	
		K_4^*	$C_4(T, K_4^*)$	K_4^*	$C_4(T, K_4^*)$	K_4^*	$C_4(T, K_4^*)$
	20	4.35	0.4677	3.77	0.5342	3.23	0.6063
	30	3.45	0.5500	3.14	0.6048	2.80	0.6661
5	40	3.01	0.6153	2.79	0.6633	2.54	0.7183
	50	2.67	0.6719	2.55	0.7155	2.36	0.7659
	20	7.43	0.2973	5.30	0.4209	3.92	0.5328
	30	6.37	0.3202	4.93	0.4221	3.80	0.5390
10	40	5.80	0.3372	4.67	0.4318	3.71	0.5448
	50	5.42	0.3512	4.48	0.4405	3.62	0.5503

Table 9 Optimal $\widetilde{T}_4^*$, $\widetilde{C}_4(\widetilde{T}_4^*)$, $\widetilde{K}_4^*$ and $\widetilde{C}_4(\widetilde{K}_4^*)$ when $c_T = c_K = 10$ and $c_S = \lambda = \mu = \sigma = 1$

c_M	$\widetilde{T}_4^*$	$\widetilde{C}_4(\widetilde{T}_4^*)$	$\widetilde{K}_4^*$	$\widetilde{C}_4(\widetilde{K}_4^*)$
0.1	14.24	0.2424	14.15	0.2413
0.2	10.11	0.3021	9.99	0.2997
0.3	8.28	0.3486	8.15	0.3444
0.4	7.20	0.3882	7.05	0.3820
0.5	6.46	0.4233	6.30	0.4151
0.6	5.91	0.4554	5.75	0.4449
0.7	5.49	0.4851	5.32	0.4722
0.8	5.15	0.5130	4.97	0.4976
0.9	4.86	0.5394	4.68	0.5114
1.0	4.62	0.5546	4.44	0.5239

frequency of tenuring collections is decreased and tenuring or minor collection cost is increased.

- Compared with Tables 7 and 8, we can derive that $T_4^* \approx K_4^*$, in fact, this means that $\mu T_4^* \approx K_4^*$, which corresponds to the assumption of $Z(t)$. $C_4(T_4^*, K) \approx C_4(T, K_4^*)$, however, $C_4(T_4^*, K)$ are sometimes greater than and sometimes less than $C_4(T, K_4^*)$. That is, we can not compare them exactly.
- $\widetilde{C}_4(T)$ and $\widetilde{C}_4(K)$ are the particular cases of $C_4(T, K)$. Take T_4^* and $\widetilde{T}_4^*$ in Tables 7 and 9 for an example, when $c_M = 0.1, 0.5, 1.0$, $\widetilde{T}_4^*$ should be greater than T_4^* and $C_4(T_4^*, K)$ should be less than $\widetilde{C}_4(\widetilde{T}_4^*)$ when $K = 10$ and $c_K = 20$.

6 Conclusions

This chapter has discussed the problems of when to make tenuring and major collections for a generational garbage collector to meet the pause time goal. According to the properties of adaptive tenuring, two cases of working schemes have been introduced first, where tenuring and major collections have been considered as renewal points of the collection processes, respectively. Second, analyses of the costs suffered for collections, including minor, tenuring and major collections, have been given. Third, using the techniques of cumulative processes and degradation processes or continuous wear processes, expected cost rates for the two cases have been derived, and optimal tenuring collection times and major collection times are discussed analytically. Fourth, numerical examples have been given and some comparisons of the policies have been made. Such theoretical analyses would be applied to actual garbage collections by suitable modifications and extensions.

Acknowledgments This work is partially supported by the Grant-in-Aid for Scientific Research (C) of Japan Society for the Promotion of Science (22500897, 21530318), National Natural Science Foundation of China (70471017, 70801036) and Humanities and Social Science Research Foundation of China (05JA630027).

References

1. Appel AW (1989) Simple generational garbage collection and fast allocation. Softw Pract Exper 19:171–183
2. Armstrong J, Virding R (1995) One-pass real-time generational mark-sweep garbage collection. In: Proceedings of international workshop on memory managementx (Lecture notes in computer science), vol 986. Springer, Berlin, pp 313–322
3. Barlow RE, Proschan F (1965) Mathematical theory of reliability. Wiley, New York
4. Clinger WD, Rojas FV (2006) Linear combinations of radioactive decay models for generational garbage collection. Sci Comput Program 62:184–203
5. Jones R, Lins R (1996) Garbage collection: algorithms for automatic dynamic memory management. Wiley, Chichester
6. Kaldewaij A, Vries L (2001) Optimal real-time garbage collection for acyclic pointer structures. Inf Process Lett 77:151–157
7. Lee WH, Chang JM (2004) A garbage collection policy based on empirical behavior. Inf Sci 167:129–146
8. Nakagawa T (2005) Maintenance theory of reliability. Springer, London
9. Nakagawa T (2007) Shock and damage models in reliability theory. Springer, London
10. Nakamura S, Nakagawa T (2010) Stochastic reliability modeling, optimization and applications. World Scientific, Singapore
11. Nikulin MS, Balakrishnan N (2010) Advances in degradation modeling: applications to reliability, survival analysis, and finanace. Birkhöuser, Boston
12. Osaki S (1992) Applied stochastic system modeling. Springer, Berlin
13. Ross SM (1983) Stochastic processes. Wiley, New York
14. Satow T, Yasui K, Nakagawa T (1996) Optimal garbage collection policies for a database in a computer system. RAIRO Oper Res 30:359–372
15. Sato K (2001) Basic results on Lévy processes. In: Bandorff-Nielsen O, Mikosch T, Resnick S (eds) Lévy processes, theory and applications. Birkhöuser, Boston
16. Soman S, Krintz C (2007) Application-specific garbage collection. J Syst Softw 80:1037–1056
17. Ungar D (1984) Generation scavenging: A non-disruptive high performance storage reclamation algorithm. ACM Sigplan Not 19:157–167
18. Ungar D, Jackson F (1992) An adaptive tenuring policy for generation scavengers. ACM Trans Program Lang Syst 14:1–27

19. Vengerov D (2009) Modeling, analysis and throughput optimization of a generational garbage collector. Technical Report, Sun Labs, TR-2009-179
20. Wilson PR (1992) Uniprocessor garbage collection techniques. In: International workshop on memory management, (Lecture notes in computer science), vol 637. Springer, London, pp 1–42
21. Zhao XF, Nakamura S, Nakagawa T (2010) Optimal policies for random and periodic garbage collections with tenuring threshold. In: Tomar GS, Chang RS, Gervasi O, Kim T, Bandyopadhyay SK (eds) vol 74 Communications in computers and information science. Springer, Berlin, pp 125–135
22. Zhao XF, Nakamura S, Nakagawa T (2011) Two generational garbage collection models with major collection time. IEICE transactions on fundamentals of electronics communications and computer sciences, E94-A:1558–1566