ОДНОЛИНЕЙНАЯ СИСТЕМА ОБСЛУЖИВАНИЯ С НЕНАДЕЖНЫМ ПРИБОРОМ В ДИСКРЕТНОМ ВРЕМЕНИ. НЕСТАЦИОНАРНЫЕ ХАРАКТЕРИСТИКИ

Боян Димитров, Чавдар Докев

1. Введение. В работе рассматривается дискретный вариант однолинейной системы обслуживания типа $M/G/1$ с неординарным входящим потоком и прибором, ненадежным в свободном состоянии. Такие системы были объектом изучения при непрерывном времени в работе [1] и в книге [2, гл. 8], где получены преобразования Лапласа (по временному аргументу) от производящей функции Ez^{ν_t} числа заявок ν_t в системе и от преобразования Лапласа — Стилтьеса $\omega(s, t)$ виртуального времени ожидания w_t. Знание таких функционалов от нестационарных характеристик процесса обслуживания позволяет лишь формально считать, что нам этот процесс известен для конечных значений момента t. Этот путь исследования дает возможность получить стационарные характеристики обслуживания путем применения разного рода тауберовых теорем, исследовать в лучшем случае скорость сходимости нестационарного процесса к стационарному, находить некоторые асимптотические выражения средних значений величин при больших значений параметра t. Исполнение обратных преобразований практически не удается даже в простейших случаях, так что неявно полученные нестационарные характеристики оставались за нашим полезрением. Впервые Дафермос и Нойтс в [5] обратили внимание на возможность использовать результаты для систем обслуживания в дискретном времени при эффективном получении приближенных значений нестационарных характеристик обслуживания для систем с непрерывном времени. У них же приведен ряд дополнительных доводов и соображений в пользу рассмотрения вероятностных систем (в частности системы обслуживания) в дискретном времени. Далее в серию статей под общим названием Numerical Analysis М. Нойтс и его сотрудники развили эту идею, исследовав характеристики системы $M/G/1$ с надежным прибором в дискретном времени, при некоторых ограничений, необходимых для численной реализации процесса на компютере. Мы следуем путь Нойтса [6], чтобы изучить однолинейных систем с ненадежным прибором. При этом сначала необходимо доказать возможность сведения изучения систем с ненадежным прибором вообще к системам с надежным прибором во время обслуживания и ненадежным только в свободном состоянии. Так возникает необходимость изучения времен блокировки обслуживающего ненадежного прибора в дискретном времени, подобно тому, что делается для

непрерывного случая, см. [2, гл. 5]. В работе найдены выражения для явного подсчета распределений времен блокировки для всех схем обслуживания при возможном прерывании обслуживания заявки, когда все участвующие величины независимы и дискретно распределены, при одной и той же единицы измерения времени. Далее рассмотрена система с прибором, ненадежным только в свободном состоянии. Получены выражения для совместного распределения длины очереди и остаточного времени (до конца обслуживания, восстановления или до выхода обслуживающего прибора из строя) в любой дискретный момент n. Найдены следующие характеристики: а) длина очереди; б) остаточное время; в) распределение возможного времени ожидания; г) распределение потока обслуженных заявок; д) распределение потока потерянных заявок. Введены ограничения на число отличных от нуля вероятностей, что удобно при использовании полученных результатов для определения вероятностных характеристик процесса обслуживания на цифровой вычислительной машине.

2. Описание системы. Рассматривается однолинейная система обслуживания, в которой все времена измеряются как кратными одной выбранной нами основной единицы измерения. Предположим, что число поступивших заявок в интервал времени $[n-1, n)$ есть случайная величина ν_n с разпределением $p_k=P\{\nu_n=k\}$, причем $k=0, 1, \ldots, k$ (для исключении тривиальностей предполагаем $0<p_0<1$). Прибывшие заявки в интервале $[n-1, n)$ присоединяются к очереди в моменте $n-0$. Случайные величины ν_n предполагаются независимыми между собой и независящими от остальных характеристик обслуживания. Последовательные времена обслуживания отдельных заявок предположим независимые одинаково распределенные целочисленные случайные величины со значениями в множестве $\{1, 2, \ldots, L_2\}$. Через r_ν, $\nu=1, 2, \ldots, L_2$, обозначим вероятность, что заявка потребует ровно ν единиц времени для своего обслуживания. Максимальная допустимая длина очереди в системе предположена конечна и равна L_1.

Обслуживающий прибор ненадежен как в свободном, так и в занятом состоянии. Отдельные времена безотказной работы обслуживающего прибора реализуются заново в каждом моменте начала или конца очередного обслуживания и в конце каждого восстановления после отказа. Отдельные реализации времен „жизни" прибора предположены независимые одинаково распределенные целочисленные случайные величины ζ_i с распределением

$$P(\zeta_i=j)=c_j,\ j=1, \ldots, L_2.$$

Последовательные времена восстановления прибора после отказа тоже независимы от остальных и одинаково распределенные целочисленные случайные величины с распределением $\{q_\nu\}$, $\nu=1, \ldots, L_2$, $L_2<\infty$.

3. Механизм обслуживания. Из-за дискретного измерения времени будем предполагать, что если обслуживание заявки начнется в n-ом интервале и требует ν единиц времени, мы будем рассматривать его как начавшемся в момент n и закончившемся в момент $n+\nu-0$ (при условии, что прибор не выйдет из строя во время обслуживания). Точно так же, если обслуживающий прибор выйдет из строя в интервале $[n-1, n)$,

будем считать, что это произошло в момент $n-0$ и если требовалось ν единиц времени для восстановления, прибор будет в исправности в момент $n+\nu-0$. Если в момент $n-0$ закончено обслуживание или восстановление и в системе нет больше заявок, следующее восстановление начнется ровно в момент $n+\nu-0$, если в моменте $n-0$ ему отведено ровно ν единиц времени „жизни".

Поскольку общее число мест в системе ограничено и принято равно L_1, то из любой группы в состав ν заявок будет потеряно $\nu-(L_1-k)$ заявок, если в системе уже имеется k заявок и не будет закончено обслуживание заявки.

4. Дисциплина обслуживания. При обслуживании с прибором, ненадежным в занятом состоянии, возможны следующие правила поведения системы при прерывании уже начатого обслуживания из-за выхода прибора из строя (см. [2, гл. 5, т. 2]).

Схема 1. Если во время обслуживания прибор выдет из строя, сразу начинается его восстановление, после чего заявка продолжает обслуживаться до окончании оставшихся (из потребованных) единиц времени.

Схема 2. При прерывании обслуживания из-за выхода прибора из строя заявка ждет до полного заканчивания восстановления, после чего уходит из системы недообслуженной.

Схема 3. Прерванное обслуживание начинается заново сразу после восстановления прибора. При этом различаются два варианта:

3а. Идентичное повторное обслуживание. Если при первом поступлении на обслуживании заявка потребовала ν единиц времени для своего обслуживания, тогда она требует с вероятностью единица ровно ν единиц при каждом следующим становлении на обслуживание.

3б. Неидентичное повторное обслуживание. При каждом следующем становлении на прибор после прерывания заявка обслуживается с новой реализацией времени обслуживания, независимо от предыстории процесса, т. е. каждый раз заявка обслуживается как новая.

Длительность интервала времени с момента первого становления заявки на прибор до первого следующего момента, когда прибор станет доступен для обслуживания других заявок, будем называть *временем блокировки*. Это одна из важнейших характеристик систем с ненадежным обслуживающем устройством, так как дает возможность им быть исследованы более простыми способами. Заменив в общем случае время обслуживание на время блокировки, мы можем рассматривать систему как обслуживающей с прибором, абсолютно надежным в занятом состоянии прибором и ненадежным лишь когда он свободен от обслуживания. Поскольку и в этом случае случайные процессы в систем обслуживания в дискретном времени можно изучить с помощью Марковской цепи, то использование дискретного времени блокировки особо удачно. Прежде чем приступить к определению времен блокировки и исследованию характеристик процесса обслуживания, мы введем обозначения и пояснения величин, участвующих в процессе.

5. Обозначения. Для простоты дальнейшего изложения введем следующие величины:

n — момент времени, в котором рассматривается система;

K — максимальное число заявок, поступивших в систему на интервале длины единица;

L_1 — максимальное число заявок в системе;

L_2 — максимальное число ненулевых вероятностей распределений;

ν_n — число заявок, поступивших в систему на интервале $[n-1, n)$,

$$P(\nu_n=i)=p_i,\ i=0,\ 1,\ \ldots,\ k;$$

ξ_k — время обслуживания k-ой очередной заявки,

$$P(\xi_k=j)=r_j,\quad j=1,\ \ldots,\ L_2;$$

β_i — время блокировки, связанное с i-ой схемой обслуживания,

$$P(\beta_i=j)=b_j,\ j=1,\ \ldots,\ L_2;$$

ω_k — время k-го восстановления обслуживающего прибора,

$$P(\omega_k=j)=q_j,\quad j=1,\ \ldots,\ L_2;$$

ζ_k — время „жизни" прибора после $k-1$-го его выхода из строя,

$$P(\zeta_k=j)=c_j,\quad j=1,\ \ldots,\ L_2;$$

X_n — число заявок в системе в момент времени n;

Y_n — остаточное время до конца обслуживания, текущего в момент n, до конца текущего восстановления обслуживающего прибора или до ожидаемого выхода свободного от обслуживания прибора из строя;

ε_n — состояние прибора в момент времени n: $\varepsilon_n=0$, если прибор восстанавливается; $\varepsilon_n=1$, если идет обслуживание некоторой заявки; $\varepsilon_n=2$, если прибор свободен, исправен (в таком случае ожидается выход прибора из строя).

Если $\alpha_1, \alpha_2, \ldots, \alpha_n$ независимые, одинаково распределенные целочисленные случайные величины с распределением $\{a_i\}$, через $\{a_i^{(k)}\}$ будем обозначать распределение суммы $\alpha^{(k)}=\alpha_1+\alpha_2+\ldots+\alpha_k$, т. е. $P\{\alpha_1+\alpha_2+\ldots+\alpha_k=i\}=a_i^{(k)}$;

N_j — число выходов обслуживающего устройства из строя за j единиц времени;

φ_i — часть времени блокировки (включающая интервал обслуживания и возможного интервала восстановления прибора) при очередном становлении заявки на прибор при условии, что заявка потребовала j единиц времени на обслуживание и будет обслужена без прерывания;

ψ_j — часть времени блокировки (включающая часть интервала обслуживания и интервал восстановления) при становлении заявки на прибор при условии, что заявка потребовала j единиц на обслуживание и обслуживание будет прервано;

φ — общая длительность времени блокировки при очередном становлении заявки на прибор при условии, что она будет обслужена до конца без прерывания;

ψ — длительность времени блокировки при очередном становлении заявки на прибор при условии, что обслуживание будет прервано.

Введем еще и следующие случайные события:

A_j — заявка потребует ровно j единиц времени для своего обслуживания ($j=1, \ldots, L_2$);

B_j^k — во время обслуживания заявки, потребовавшей ровно j единиц времени для обслуживания, обслуживающий прибор выдет из строя ровно k раз;

u_j^0 — во время обслуживания заявки, потребовавшей j единиц времени для обслуживания, прибор не выйдет из строя ($j=1, \ldots, L_2$);

u_j^k — во время обслуживания заявки, потребовавшей j единиц времени, обслуживающий прибор выйдет из строя ровно через k единиц времени после начала обслуживания заявки ($k=1, \ldots, j$; $j=1, \ldots, L_2$);

G — заявка покинет систему обслуженной;

G_i — заявка будет обслуженной ровно при i-ом повторном становлении на обслуживание (после $i-1$ выходов прибора из строя).

6. Распределения времен блокировки. В этом параграфе выведем распределения времени блокировки для описанных в [4] схем возможных дисциплин обслуживания.

Схема 1. Пусть заявка потребовала ровно j единиц времени для своего обслуживания. Это произойдет с вероятностью r_j. Тогда (см. [2] стр. 93) имеем

$$P(N_j=0)=1-P(\zeta\leqq j)$$

$$P(N_j=k)=P(\zeta^{(k)}\leqq j)-P(\zeta^{(k+1)}\leqq j),\ k=1,\ 2,\ \ldots$$

Здесь отметим, что при $k>j$ имеем $P\{\zeta^{(k)}\leqq j\}=0$, откуда следует $P(N_j=k)=0$ при $k>j$, т. е. если обслуживание состоит из j единиц времени, прибор не может выйти из строя больше j раз. Если за время обслуживания заявки прибор выйдет из строя ровно k раз, то время блокировки β_1 будет представленно суммой из времени обслуживания j и k независимых одинаково распределенных времен восстановления $\omega_1, \omega_2, \ldots, \omega_k$. Следовательно

$$P(\beta_1=s/A_jB_j^k)=q_{s-j}^{(k)},$$

где, согласно введенным обозначениям $q_{s-j}^{(k)}$ есть $s-j$-ая координата свертки из k распределений времени восстановления прибора. Отсюда, применив формулу полной вероятности, получим

$$(1)\quad P(\beta_1=s)=\sum_{j=1}^{L_2}\sum_{k=0}^{j}P(A_jB_j^k)\,q_{s-j}^{(k)}=\sum_{j=1}^{L_2}\sum_{k=0}^{j}r_j(P\zeta^{(k)}\leqq j)-P(\zeta^{(k+1)}\leqq j))\,q_{s-j}^{(k)}.$$

Схема 2. Заявка потребует ровно j единиц времени обслуживания с вероятностью r_j. Пусть прибор выйдет из строя после того, как заявка обслуживалась k единиц времени, вероятность чего есть $P(\zeta=k)=c_k$.

Тогда время блокировки β_2 будет суммой из времени истекшего обслуживания k и одной реализации времени восстановления прибора ω. Если прибор не выйдет из строя до конца обслуживания заявки (вероятность чего равна $P(\zeta>j)$), то β_2 будет равно j. Следовательно

$$P(\beta_2=s/A_j u_j^0)=\delta_{sj}, \quad P(\beta_2=s/A_j u_j^k)=q_{s-j}.$$

Применив и теперь формулу полной вероятности, получим

$$P(\beta_2=s)=\sum_{j=1}^{L_2}\left[\sum_{k=1}^{j} P(A_j u_j^k)\, q_{s-j}+P(A_j u_j^0)\,\delta_{sj}\right] \tag{2}$$

$$=\sum_{j=1}^{L_2}\left[\sum_{k=1}^{j} r_j P(\zeta=k)\, q_{s-j}+r_j P(\zeta>j)\,\delta_{sj}\right]$$

$$=\sum_{j=1}^{L_2} r_j \sum_{k=1}^{j} P(\zeta=k)\, q_{s-j}+r_s P(\zeta>s).$$

Из выражения (2) видно, что

$$P(G)=\sum_{j=1}^{L_2} r_j P(\zeta\geqq j)=\sum_{j=1}^{L_2} r_j \sum_{k=j}^{L_2} c_k, \tag{3}$$

$$P(\{\beta_2=s\}\, G)=\sum_{j=1}^{L_2} r_j P(\zeta=j)\, q_{s-j}+r_s P(\zeta>s), \tag{4}$$

$$P(\{\beta_2=s\}\, \bar{G})=\sum_{j=1}^{L_2} r_j \sum_{k=1}^{j-1} P(\zeta=k)\, q_{s-j}. \tag{5}$$

Схема 3. По существу и в обоих случаях обслуживание по этой схеме представляет собой многократное повторение схемы 2, выражающееся в следующем: заявка делает попытку обслужится как в схеме 2. Если заявка будет обслужена, она покидает обслуживающего устройства как только последнее окажется в исправности; если прибор выйдет из строя во время обслуживания заявки, она ждет до конца последующего восстановления и заново поступает на обслуживание. Это повторяется до тех пор, когда заявка успеет обслужится до выхода прибора из строя.

Следовательно, если заявка будет обслуженной ровно при i-ом очередном своем становлении на прибор, то время блокировки будет суммой из $i-1$ незаконченных времен обслуживания, $i-1$ времен восстановлений прибора и из времени последнего (успешного) обслуживания, удлиненного еще возможно с одним интервалом восстановления прибора, в случаях, когда он выдет из строя на последнем единичном интервале обслуживания.

Рассмотрим теперь оба случая в отдельности:

3а. Пусть заявка потребовала при первом становлении на обслужи-

вании j единиц времени. Вероятность этого события равна r_j. Согласно формул (4) и (5) схемы 2 имеем при $P(\zeta \geqq j) > 0$

$$P(\varphi_j = s) = P(\beta_2 = s/A_j G_1) = \frac{P(\{\beta_2 = s\} A_j G_1)}{P(A_j G_1)}$$
$$= \frac{P(\zeta = j)\, q_{s-j} + P(\zeta > j)\, \delta_{sj}}{P(\zeta \geqq j)}$$

и при $P(\zeta \geqq j) = 0$

$$P(\varphi_j = s) = \delta_{s0}.$$

Для распределения величины ψ_j получаем при $P(\zeta < j) > 0$

$$P(\psi_j = s) = P(\beta_2 = s/A_j \bar{G}_1) = \frac{P(\{\beta_2 = s\} A_j \bar{G}_1)}{P(A_j \bar{G}_1)} = \frac{\sum_{k=1}^{j-1} P(\zeta = k)\, q_{s-k}}{P(\zeta < j)}$$

и при $P(\zeta < j) = 0$

$$P(\psi_j = s) = \delta_{s0}.$$

Поэтому имеем

$$P(\beta_3' = s/A_j G_i) = P(\psi_j^{(i-1)} + \varphi_j = s),$$

откуда следует

$$P(\beta_3' = s/A_j) = \sum_{i=1}^{\infty} P(G_i/A_j)\, P(\psi_j^{(i-1)} + \varphi_j = s).$$

Так как

$$P(G_i/A_j) = [P(\bar{G}/A_j)]^{i-1} P(G/A_j) = [1 - P(G/A_j)]^{i-1} \cdot P(G/A_j)$$
$$= [1 - P(\zeta \geqq j)]^{i-1} P(\zeta \geqq j),$$

то

$$P(\beta_3' = s) = \sum_{j=1}^{L_2} r_j \sum_{i=1}^{\infty} [1 - P(\zeta \geqq j)]^{i-1} P(\zeta \geqq j)\, P(\psi_j^{(i-1)} + \varphi_j = s).$$

3б. Легко сообразить, что

$$P(\varphi = s) = P(\beta_2 = s/G) = \frac{P(\{\beta_2 = s\} G)}{P(G)} = \frac{1}{P(G)} \sum_{j=1}^{L_2} P(\{\alpha_2 = s\} A_j G).$$

Приняв ввиду, что

$$P(G) = \sum_{j=1}^{L_2} P(G/A_j) P(A_j) = \sum_{j=1}^{L_2} r_j P(\zeta \geqq j),$$

и кроме того

$$P(\{\beta_2 = s\} A_j G) = P(A_j)\, P(G/A_j)\, P(\{\beta_2 = s\}/A_j G) = r_j P(\zeta \geqq j)\, P(\varphi_j = s),$$

получаем для распределения величины φ

$$P(\varphi=s)=\left[\sum_{j=1}^{L_2} r_j P(\zeta\geqq j)\right]^{-1}\sum_{j=1}^{L_2} r_j P(\zeta\geqq j)\,P(\varphi_j=s).$$

Аналогичным образом находим распределение величины ψ:

$$P(\psi=s)=P(\beta=s/\overline{G})=\left[\sum_{j=1}^{L_2} r_j P(\zeta<j)\right]^{-1}\sum_{j=1}^{L_2} r_j P(\zeta<j)\,P(\psi_j=s).$$

Теперь из равенства

$$P(\beta_3''=s/G_2)=P(\psi^{(i-1)}+\varphi=s)$$

следует

$$P(\beta_3''=s)=\sum_{i=1}^{\infty} P(\beta_3''=s)G_i P(G_i)$$

$$=\sum_{i=1}^{\infty}\left[\sum_{j=1}^{L_2} r_j P(\zeta<j)\right]^{i-1}\left[\sum_{j=1}^{L_2} r_j P(\zeta\geqq j)\right] P(\psi^{(i-1)}+\varphi=s).$$

Видно, что в схеме 3 время блокировки с положительной вероятностью может принимать сколь угодно большие значения. При этом, если (хотя бы формально) существует такое j, что $P(\zeta<j)=1$ и $r_j>0$, то в случае 3а время блокировки будет стохастически неограниченной величины с вероятностью $r_j+r_{j+1}+\cdots+r_{L_2}$ будет принимать значение безконечность).

Чтобы исключить этот тривиальный случай, мы ввели предположение, что $c_{L_2}>0$. Поэтому можем утверждать, что какого бы не было $\varepsilon>0$ существует такое целое число N, что вероятность заявки быть обслуженной через N или большее число попытков будет меньше ε. Для случая 3а достаточно выбрать в качестве N наименьшее целое число, удовлетворяющее условию

$$\sum_{j=1}^{L_2} r_j\,[P(\zeta<j)]^N<\varepsilon,$$

а в случае 3б N определяется условием

$$1-\sum_{j=1}^{L_2} r_j\,P(\zeta\geqq j))^N<\varepsilon.$$

Замечание 1. Если в системе с непрерывном времени время „жизни" прибора экспоненциально распределено, то в соответствующей дискретной модели величина ζ будет иметь геометрическое распределение

$$P(\zeta=j)=q^{j-1}(1-q). \tag{6}$$

Этот случай, рассмотренный в [3], получается как следствие вышеприведенных результатов, если распределение ζ задается равенствами (6).

Замечание 2. Простота формул для нахождения дискретного распределения времени блокировки удивительна. Она дает возможность для эффективного определения функции распределения блокировки и ее моментов с помощью вычислительной машины. В этом и особая ценность найденных результатов. Чтобы действительно оценить их по достойнству, следует привести для сравнения любого из результатов [2, гл. 5]. К примеру мы приведем формулу [2, стр. 179].

$$\beta(s)=\int_0^\infty \frac{e^{-st}[1-c(t)]}{1-q(s)\int_0^t e^{-sx}\,dc(x)}\,dG(t), \tag{7}$$

котороя определяет пробразование Лапласа — Стилтьеса $\beta(s)$ распределения $B(x)$ времени блокировки прибора в случае 3а. В (7) $c(t)$ есть ф. р. времени „жизни" прибора, $G(t)$ — ф. р. абсолютного времени обслуживания заявки, $q(s)$ — преобразование Лапласа — Стилтьеса от ф. р. $Q(x)$ времени восстановления прибора. Сама ф. р. $B(x)$ задана формулой (7) неявно. Чтобы найти ее, необходимо выполнить операцию обращения преобразования Лапласа — Стилтьеса

$$B(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\frac{\beta(s)}{s}e^{st}\,ds,$$

что нам кажется практически невыполнимо даже в простейших случаях Более того, (7) не дает возможность определить хотя бы первый момент распределения $B(x)$, когда в дискретном случае мы можем определять практически все интересные характеристики времени блокировки. Наши результаты можно рассматривать еще как своеобразные численные методы для обращения формул типа (7) (и остальных из [2, гл. 5]). Если $x>0$, то приближением $B(x)$ можно служить сумма $\sum_{s=0}^{n}\beta_s$, где β_s получены путем дискретизации распределений величин ζ, η, ω, при основной единицей времени $\delta>0$ (и $\delta\approx 0$), и n определено как наименьшее целое число, большее величини t/δ.

7. Представление процесса обслуживания с помощью Марковской цепи. Далее будем рассматривать систему обслуживания как абсолютно надежной во время обслуживания любой заявки и с прибором, ненадежным только в свободном состоянии. Это оправдано, если за время обслуживания заявки берется время блокировки прибора при первом поступлении заявки на прибор в соответствии с принятой дисциплины обслуживания. Поэтому всюду ниже мы будем писать β_i величиной обслуживания i-ой заявки, подразумевая при этом (если это необходимо) время блокировки. Для остальных величин сохраняются обозначения и смысл, введенных в разделе 4.

Легко заметить, что тройка $(X_n, Y_n, \varepsilon_n)$ полностью задает состояние системы в момент n. Между величинами $(X_n, Y_n, \varepsilon_n)$ и $(X_{n+1}, Y_{n+1}, \varepsilon_{n+1})$ и ν_{n+1} имеют место следующие соотношения:

$$X_{n+1}=\min(X_n+\nu_{n+1}-1, L_1),\ Y_{n+1}=Y_n-1,\ \varepsilon_{n+1}=\varepsilon_n,$$
$$\text{если } X_n\geqq 1,\ Y_n>1,\ \varepsilon_n=0,\ 1,$$
$$X_{n+1}=\min(X_n+\nu_{n+1}-1, L_1),\ Y_{n+1}=\beta,\ \varepsilon_{n+1}=1,$$
$$\text{если } X_n+\nu_{n+1}>1,\ Y_n=1,\ \varepsilon_n=1,$$
$$X_{n+1}=0,\ Y_{n+1}=\zeta,\ \varepsilon_{n+1}=2, \text{ если } X_n=1,\ \nu_{n+1}=0,\ Y_n=1,\ \varepsilon_n=1,$$
$$X_{n+1}=\min(X_n+\nu_{n+1}, L_1),\ Y_{n+1}=\beta,\ \varepsilon_{n+1}=1, \text{ если } X_n+\nu_{n+1}>0,\ Y_n=1,\ \varepsilon_n=0,$$
$$X_{n+1}=0,\ Y_{n+1}=\zeta,\ \varepsilon_{n+1}=2, \text{ если } X_n=\nu_{n+1}=0,\ Y_n=1,\ \varepsilon_n=0,$$
$$X_{n+1}=\min(\nu_{n+1}, L_1),\ Y_{n+1}=\beta,\ \varepsilon_{n+1}=1, \text{ если } x_n=0,\ \nu_{n+1}>0,\ Y_n\geqq 1,\ \varepsilon_n=2,$$
$$X_{n+1}=0,\ Y_{n+1}=Y_n-1,\ \varepsilon_n=2, \text{ если } X_n=0,\ \nu_{n+1}=0,\ Y_n>1,\ \varepsilon_n=2,$$
$$X_{n+1}=\min(\nu_{n+1}, L_1),\ Y_{n+1}=\omega,\ \varepsilon_n=0, \text{ если } X_n=0,\ \nu_{n+1}=0,\ Y_n=1,\ \varepsilon_n=2.$$

Тройка $(X_n, Y_n, \varepsilon_n)$ при $n\geqq 0$ представляет собой неразложимую, апериодическую цепь Маркова со значениями (i, j, e), где $i=1, \ldots, L_1$, $j=1, \ldots, L_2$, $e=0, 1, 2$. Пусть

$$P_n(i, j, \varepsilon)=P\{X_n=i,\ Y_n=j,\ \varepsilon_n=\varepsilon/X_0=i,\ Y_0=j_0,\ \varepsilon_0=e_0\},$$
$$i=1, \ldots, L_1,\ j=1, \ldots, L_2,\ \varepsilon=0, 1, 2.$$

Теперь просто записать основные рекуррентные соотношения, связывающие вероятности $P_{n+1}(i, j, e)$ и $P_n(i, j, e)$:

(8) $$P_{n+1}(0, j, 0)=p_0 P_n(0, j+1, 0)+p_0 q,\ P_n(0, 1, 2),\ j=1, \ldots, L_2,$$

(9) $$P_{n+1}(i, j, 0)=\sum_{\nu=0}^{i} p_\nu^* P_n(i-\nu, j+1, 0),\ i=1, \ldots, L_1,\qquad j=1, \ldots, L_2$$

(10) $$P_{n+1}(i, j, 1)=\sum_{\nu=0}^{i+1} p_\nu^* P_\nu(i-\nu, j+1, 1)+\sum_{\nu=0}^{i} p_\nu^* b_j P_n(i-\nu+1, 1, 1)$$
$$+b_j\sum_{\nu=0}^{i} p_\nu^* P_n(i-\nu, 1, 0)+\sum_{k=1}^{L_2} p_i^* b_j P_n(0, k, 2),\ i=1, \ldots, L_1,\ j=1, \ldots, L$$

(11) $$P_{n+1}(0, j, 2)=p_0 c_j P_n(0, 1, 0)+p_0 c_j P_n(1, 1, 1)+p_n P_n(0, j+1, 2),$$
$$j=1, \ldots, L_2,$$

где положено

12) $$p_\nu^*=p_\nu \text{ при } i=0, 1, \ldots, L_1-1,$$

$$p_\nu^* = \sum_{k=\nu}^{L_1} p_k \text{ при } i = L_1. \tag{13}$$

Эти уравнения очень удобны для численных подсчетов, так как область изменения значений аргументов конечна. Вероятности $P_n(i, j, \varepsilon)$ находятся последовательно при заданном начальном условии. Если в начальный момент $n=0$ система находится в состоянии (i_0, j_0, e_0), то необходимо удовлетворить начальным условиям $P_0(i_0, j_0, e_0)=1$ и $P_0(i, j, e)=0$ для всех остальных троек $(i, j, e) \neq (i_0, j_0, e_0)$. Когда начальные условия даны в виде распределения величин x_0, Y_0 и ε_0, это сводится (при вычислений) к другому определению исходного массива.

7. Распределение длины очереди. Вероятност того, что в момент n длина очереди равна нулю, задается выражением

$$m_0 = \sum_{j=1}^{L_1} [P_n(0, j, 0) + P_n(0, j, 2)].$$

Для $i=1, \ldots, L_1$ легко получить

$$m_i = P\{X_n = i / X_0 = i_0, \quad Y_0 = i_0, \ \varepsilon_0 = e_0\} = \sum_{j=1}^{L_1} (P_n(i, j, 1\} + P_n(i, j, 0)].$$

Располагая с распределением длины очереди, легко можем найти моменты этого распределения.

Замечание. И в данном случае стоит провести сравнение полученного результата с тем, что известно для системы $M|G|1$ с непрерывным временем. Если ν_t — число заявок в системе в момент времени $t>0$, то известна лишь функция [2, стр. 265]

$$p(z, s) = \int_0^\infty e^{-st} E z^{\nu_t} \, dt,$$

заданная выражением

$$p(z, s) = \left\{1 + \frac{\lambda}{s+\lambda}[1-\varphi(s+\lambda)] \ \pi(s) - \varphi(s+\lambda) g(s+\lambda-\lambda\pi(s))\right\}^{-1}$$

$$\times \left\{\frac{1-\varphi(s+\lambda)}{s+\lambda} + \frac{\lambda}{s+\lambda}[1-\varphi(s+\lambda)] \frac{1-\beta(s+\lambda-\lambda z)}{s+\lambda-\lambda z} \cdot \frac{z-\pi(s)}{1-z^{-1}\beta(z+\lambda-\lambda z)}\right.$$

$$\left. + \varphi(s+\lambda) \frac{1-\beta(s+\lambda-\lambda z)}{s+\lambda-\lambda z} \frac{g(s+\lambda-\lambda z) - g(s+\lambda-\lambda\pi(s))}{1-z^{-1}\beta(s-\lambda-\lambda z)} + \varphi(s+\lambda) \frac{1-g(s+\lambda-\lambda z)}{s+\lambda-\lambda z}\right\},$$

где положено $k(s, z) = s+\lambda-\lambda z$ и $\beta(s)$, $\varphi(s)$, $g(s)$ — преобразования Лапласа — Стилтьеса от ф. р. времени обслуживания, „жизни“ прибора и его восстановления, λ — параметр входящего потока, а $\pi(s)$ — решение функционального уравнения

$$\pi(s) = \beta(s+\lambda-\lambda\pi(s)).$$

Обращение этого выражения, уже ясно каждому, врядь ли выполнимо, когда путем дискретизации процесса по времени мы можем найти просто считаемые приближения вероятностного распределения длины очереди в момент t.

8. Распределение остаточного времени. В любой момент времени n величина ε_n задает состояние прибора, а Y_n — остаточное время пребывания в этом состоянии (при $\varepsilon_n=0,2$) или до конца текущего обслуживания (при $\varepsilon_n=1$). Таким образом получаем

а) распределение остаточного времени „жизни" прибора

$$P\{Y_n=j,\ \varepsilon_n=2\}=\sum_{j=1}^{L_2}P_n(0,\ j,\ 2);$$

б) распределение остаточного времени до конца восстановления прибора

$$P\{Y_n=j,\ \varepsilon_n=0\}=\sum_{i=0}^{L_1}P_n(i,\ j,\ 0);$$

в) распределение времени до ухода обслуживаемой заявки из системы

$$P\{Y_n=j,\ \varepsilon_n=1\}=\sum_{i=1}^{L_1}P_n(i,\ j,\ 1).$$

9. Распределение времени ожидания. Время ожидания в момент n определим как интервал времени, необходимый для сведения прибора в свободном и исправном состоянии при условии, что после момента n в системе не придут новые заявки. Из этого определения следует, что время ожидания в момент n равно нулю только тогда, когда $\varepsilon_n=2$, т. е.

$$w_0=\sum_{j=1}^{L_2}P_n(0,\ j,\ 2).$$

При $\varepsilon_n=0$ или $\varepsilon_n=1$ время ожидания будет сформироваться из остаточного времени Y_n и из времен обслуживания (полностью) ожидающих $X_n-\varepsilon_n$ заявок, т. е.

$$W_n=Y_n+\beta_1+\cdots+\beta X_n-\varepsilon_n.$$

Ввиду зависимости величин X_n, Y_n и ε_n и независимость следующих времен обслуживания от состояния процесса в момент n, получим

$$P\{W_n=s\}=\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}\sum_{e=0}^{1}P\{W_n=s/X_n=i,\ Y_n=j,\ \varepsilon_n=e\}\,P_n(i,\ j,\ e)$$

$$=\sum_{i=1}^{L_1}\sum_{j=1}^{L_2}\sum_{e=0}^{1}P\{j_1+\beta_1+\cdots+\beta_{i-e}=s\}\ P_n(i,\ j,\ e)$$

$$\sum_{i=2}^{L_1}\sum_{j=1}^{L_2}\sum_{e=0}^{1}\beta_{s-j}^{(i-e)}P_n(i,\ j,\ e).$$

Замечание. И в этом случае приведем (для сравнения с выше найденных формул) выражение для виртуального времени ожидания w_t в системе $M|G|1$ с непрерывным временем, чей аналог является величина w_n. Известно лишь преобразование Лапласа—Стилтьеса $w(s,\ t)=Ee^{-sw_t}$ распределения величины w_t [2, стр. 269]:

$$w(s,\ t)=e^{(s-\lambda+\lambda\beta(s))t}\left\{1-s\int_0^t e^{-(s-\lambda+\lambda\beta(s))x}P_0(x)\,dx\right.$$

$$\left.-(1-g(s))\int_0^t e^{-(s-\lambda+\lambda\beta(s))x}P_1(x)\,dx\right\}.$$

Здесь функции $P_0(x)$ и $P_1(x)$ известны только своими преобразуванниями Лапласа $p_i(s)=\int_0^\infty e^{-st}P_i(t)\,dt,\ i=0,\ 1$:

$$p_0(s)=\frac{1}{s+\lambda}\Big[1-\varphi(s+\lambda)\Big]/\{1-\varphi(s+\lambda)g(s+\lambda-\lambda\pi(s))$$

$$-\lambda[1-\varphi(s+\lambda)]\pi(s)/(s+\lambda)\};$$

$$p_1(s)=[1-g(s+\lambda-\lambda\pi(s)]^{-1}\{1-[s+\lambda-\lambda\pi(s)]\}p_0(s)\},$$

где функция $\pi(s)$ определена как указано в замечании к разделу 7.

10. Распределение числа обслуженных и потерянных заявок на данном интервале времени. Пусть α_n—число обслуженных заявок на интервале $[n-1,\ n)$. Величина α_n может принимать лишь значения 0 и 1. Зная совместную плотность распределения величин X_{n-1}, Y_{n-1} и ε_{n-1} легко подсчитать, что вероятность моменту n быть моментом конца обслуживания некоторой заявки есть

$$P(\alpha_n=1)=\sum_{i=1}^{L_1}P_{n-1}(i,\ 1,\ 1).$$

Тогда

$$P(\alpha_n=0)=1-\sum_{i=1}^{L_1}P_{n-1}(i,\ 1,\ 1).$$

Распределение числа потерянных заявок на интервале $[n-1,\ n)$ тоже определяется знанием вероятностей $P_{n-1}(i,\ j,\ e)$. Принимая ввиду, что

событие $\{\gamma_n=j\}$ за $j=1, 2, \cdots, k$ реализируется в следующих случаях:

а) когда $\varepsilon_{n-1}=0$, $X_{n-1}=L_1-k+j$, $\nu_n=k$, $k\geqq j$;

б) когда $\varepsilon_{n-1}=1$, $X_{n-1}=L_1-k-j$, $Y_{n-1}>1$, $\nu_n=k$, $k\geqq j$;

в) когда $\varepsilon_{n-1}=1$, $X_{n-1}=L_1-k+j$, $Y_{n-1}=1$, $\nu_n=k+1$, $k\geqq j$;

г) когда $\varepsilon_{n-1}=2$, $\nu_2=L+j$.

Теперь формула полной вероятности даст

$$P(\gamma_n=j)=\sum_{k=j}^{K} p_k \sum_{\nu=1}^{L_2} P_{n-1}(L_1-k+j,\ \nu,\ 0)+\sum_{k=j}^{K} p_k \sum_{\nu=2}^{L_2}(L_1-k+j,\ \nu,\ 1)$$

$$+\sum_{k=j}^{K} p_{k+1} P_{n-1}(L_1-k+j,\ 1,\ 1)+p_{L_1+j}\sum_{\nu=1}^{L_2} P_{n-1}(0,\ \nu,\ 2),\ j=1,2,\cdots,K.$$

И когда $j=0$, очевидно

$$P(\gamma_n=0)=1=\sum_{j=1}^{K} P(\gamma_n=j).$$

Точно так же можно определить интенсивности переходов процесса между его состояниями в любом фиксированном моменте n:

из занятого состояния прибора в свободном, это равно вероятности $p_0 P_{n-1}(1,\ 1,\ 1)$;

из состояния ремонта в состоянии свободном $p_0\ P_{n-1}\ (0,\ 1,\ 0)$;

из состояния ремонта в состоянии обслуживания

$$\sum_{i=1}^{L_1} P_{n-1}(i,\ 1,\ 0)+(1-p_0)\,P_n(0,\ 1,\ 0);$$

из свободного состояния в состоянии ремонта $p_0 P_{n-1}(0,\ 1,\ 2)$;

из свободного состояния в состоянии обслуживания

$$(1-p_0)\sum_{j=1}^{L_2} P_{n-1}(0,\ j,\ 2).$$

11. Заключение. Целесообразность рассмотрения систем обслуживания в дискретном времени нам кажется безспорна. Под конец мы всего лишь отметим некоторые новые возможности для решения практических проблем, на базе полученных дополнительных характеристик нестационарного процесса обслуживания. Во первых, ясно, что уравнения для вероятностей состояний процесса останутся теми же, когда рассматривается и система с нестационарным входящим потоком без последействия. Для этого нужно добавить индекс n на вероятностное распределение числа пришедших заявок в системе на интервале $[n-1,\ n)$. Далее представляется интересной воз

можность считать распределения потоков обслуженных и потерянных заявок при проэктировании и изучении комплексов обслуживания. Таким образом можно рассматривать обслуженных (как и потерянных) (заявок в качестве входящего потока к другим обслуживающим приборам комплекса. Качество обслуживания комплексом можно определят например суммой из времен ожидания и времен обслуживания заявки у отдельных приборов. Примером такого обслуживающего комплекса может служит сеть электронно вычислительных машин. Заявки в системе могут быть отправленны либо к одной определенной ЭВМ, либо проявлять необходимость от многоэтапного обслуживания несколькими машинами. Кроме того можно располагать несколькими еднотипными устройствами, на любом из которых заявка может обслуживаться на данном этапе. Тогда, используя метод разделения обслуживающего комплекса на подсистем, можно ставить задачу об оптимальном распределении ресурса и заявок по узлам обслуживающего комплекса.

ЛИТЕРАТУРА

1. Даниелян, Е. А., Димитров, Б. Н.: О системе $M/G/1$ с ожиданием и двумя типами отказов. Math. Balcanica, 2 (1972), 21—37.
2. Обретенов, А., Димитров, Б., Даниелян, Е.: Масово обслужване и приоритетни системи на обслужване, С., 1973.
3. Докев, Ч.: Време на блокировката при обслужване на заявки в дискретно време. Пета пролетна конференция на БМД, Габрово, 1976.
4. Докев, Ч.: Еднолинейна система с прост входящ поток и ненадеждно обслужващо устройство в дискретно време. Шеста пролетна конференция на БМД, Варна, 1977.
5. Dafermos, S., Neuts, M. F: A Single Server Queue in Discrete Time. Cahiers du Centre de Recherche Opérationalle, 13 (1971), 23—40.
6. Neuts, M. F.: The Single Server Queue in Discrete Time—Numerical Analysis I. Naval Res. Log. Quart., 20, No. 2, 1973, 297—304.

Постъпила на 19. I. 1977 г.

THE SINGLE SERVER QUEUEING SYSTEM WITH UNRELIABLE SERVER IN DISCRETE TIME. NON-STATIONARY CHARACTERISTICS

B. Dimitrov, Ch. Dokev

(SUMMARY)

We consider a discrete variant of the well known single server queue of type $M/G/1$ with batch arrivals and unreliable server. The blocking time is studied in the case when the server can break by service for (i) supplementary service; (ii) interruption service; (iii) pre-emptive repeat identical and (iv) pre-emptive different service disciplines. We obtain explicit expressions for the discrete distributions of blocking times under assumptions of independence of the serving times and

intervals of renewal and of life times. After change of service time by the blocking one we study the system with unreliable server only when the queue is empty. We find expressions for the joint distribution of the queue size and the residual time (of the service, of the renewal or of the time until the break) in any time n. Also the following non-stationary distributions are given: i) of the queue size; ii) of the residual time; iii) of the virtual waiting ime; iv) of the losed and of the served flow. A comparison of the present results and the known results for the corresponding system in continuous time is discussed.