Чернов В.П.

§ 6. СМО с отказами

Мы рассмотрим в этом параграфе систему обслуживания, удовлетворяющую следующим условиям:

(1) если в момент поступления требования имеется хотя бы один свободный узел обслуживания, то требование сразу начинает обслуживаться (любым из свободных узлов);
(2) каждый узел в любой момент времени обслуживает не более одного требования;
(3) каждое требование обслуживается одним узлом;
(4) обслуживание не прерывается;
(5) по окончании обслуживания требование покидает систему;
(6) входящий поток является пуассоновским;
(7) продолжительность обслуживания является случайной величиной, распределенной по экспоненциальному закону, одному и тому же для всех узлов обслуживания.
(8) Если в момент прихода требования все узлы в системе обслуживания оказываются занятыми, то требование получает отказ и сразу покидает систему.

Из условия (7) следует, что узлы обслуживания предполагаются одинаковыми (не по физическим свойствам, а по вероятностным характеристикам), причём вероятность того, что время обслуживания больше заданного времени t равна

$$P\left\{t_{обсл.} > t\right\} = e^{-\nu t}, \qquad (6.1)$$

где ν – интенсивность обслуживания, то есть среднее число требований, обслуживаемых узлом в единицу времени.

Такой системой обслуживания является, например, телефонная станция. Исторически теория массового обслуживания возникла из рассмотрения именно таких систем в исследованиях датского ученого А.К. Эрланга (первые работы относятся к 1909г.).

Общая схема СМО с отказами представлена на рис .6.1.

Пусть λ – параметр входящего пуассоновского потока требований. Посредством N обозначим число узлов обслуживания ($N \geq 1$). Опишем функционирование такой СМО в терминах процессов гибели и рождения. Для этого требуется ввести понятие состояния, обосновать то, что процесс перехода из состояния в состояние является марковским и доказать, что вероятности перехода удовлетворяют условиям процессов гибели и рождения.

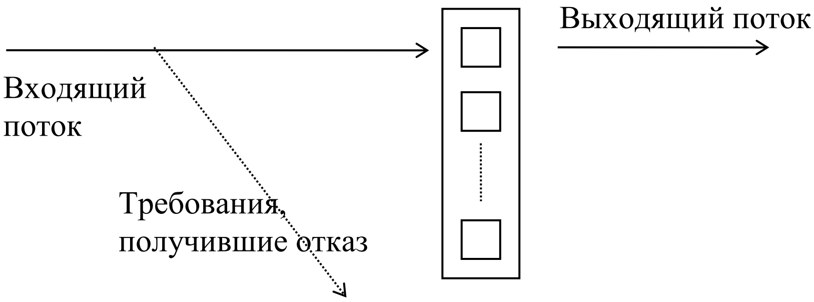

Рис.6.1.

Начнем с формализации понятия состояния. Состоянием СМО с отказами назовем число требований, находящихся в системе, то есть число требований, находящихся в узлах обслуживания. Состояниями являются целые неотрицательные числа в пределах от 0 до N,

{0, 1, … N}.

Процесс функционирования, то есть перехода из состояния в состояние, является марковским процессом. Действительно, пусть в момент t состоянием является i, а в момент $t+\tau$ состоянием является j. Изменение состояния, то есть изменение числа требований, находящихся в системе обслуживания, имеет две причины: поступление новых требований из входящего потока и уход обслуженных требований из системы.

Входящий поток является пуассоновским, поступление требований из него на любом отрезке времени не зависит от истории потока до начального момента этого отрезка (отсутствие последействия в пуассоновском потоке). Тем более оно не зависит от требований, покидающих систему. Таким образом, изменение состояния по первой причине обладает марковским свойством.

Уход требований за данный отрезок времени также не связан с тем, что происходило раньше. В частности, если в момент времени t требование уже обслуживалось, то вероятность окончания обслуживания до момента $t+\tau$ не зависит от того, сколько длилось обслуживание до момента t. Это обстоятельство связано с экспоненциальностью распределения длительности обслуживания. Соответствующее свойство было нами подробно рассмотрено ранее при изучении свойств пуассоновских потоков; здесь можно повторить те же рассуждения, рассматривая вместо интервала времени между моментами поступления требований интервал между моментами начала и окончания обслуживания. Таким образом, изменение состояния по второй причине также обладает марковским свойством и, следовательно, процесс функционирования является марковским.

Найдем вероятности переходов из одного состояния в другое. Переход $i \to i+1$ за время t связан с поступлением одного требования в систему за это время и уходом нуля требований (то есть с тем, что ни один из занятых узлов за это время обслуживания не закончил). Число занятых узлов равно i.

$$P_{i,i+1}^{(t)} = \lambda t \cdot (1 - \lambda t + o(t)) \cdot (1 - i\nu t + o(t)) = \lambda t + o(t). \quad (6.2)$$

Переход $i \to i-1$ связан с поступлениям нуля требований и уходом одного (то есть с тем, что ровно один из занятых узлов окончил обслуживание).

$$P_{i,i-1}^{(t)} = e^{-\lambda t} \cdot (e^{-\nu t})^{i-1} \cdot (1 - e^{-\nu t}) \cdot C_i^1 = i\nu t + o(t). \quad (6.3)$$

Далее, переход $i \to i$ связан с отсутствие поступления требований в систему и ухода требований из системы. Такой переход имеет вероятность

$$P_{i,i}^{(t)} = e^{-\lambda t} \cdot (e^{-\nu t})^i = 1 - \lambda t - i\nu t + o(t). \quad (6.4)$$

Все остальные вероятности переходов суть величины бесконечно малые. Например,

$$P_{i,i+2}^{(t)} = \frac{(\lambda t)^2}{2} \cdot e^{-\lambda t} \cdot (e^{-\nu t})^i = o(t) \quad (6.5)$$

$$P_{i,i-2}^{(t)} = e^{-\lambda t} \cdot (e^{-\nu t})^{i-2} \cdot (1 - e^{-\nu t})^2 \cdot C_i^2 = o(t) \quad (6.6)$$

При определении формул для вероятностей (6.2) – (6.4) мы учли не все возможные случаи. Так, например, переход $i \to i+1$ возможен не только при поступлении одного, но и при поступлении n требований и уходе $n+1$ требований. Легко можно подсчитать вероятности таких событий и убедиться, что при $n \geq 2$ они (и их сумма) являются бесконечно малыми, так что единственный существенный вклад в эту вероятность даст учтенный выше случай $n = 0$.

Мы убедились, что вероятности перехода удовлетворяют условиям процессов гибели и рождения. При этом коэффициенты λ_i и ν_i, характеризующие такие процессы, определяются формулами

$$\lambda_i = \lambda,$$

$$\nu_i = \begin{cases} i\nu & \text{при } i \le N \\ N\nu & \text{при } i \ge N \end{cases} \qquad (6.7)$$

Подставив (6.7) в (5.22), мы получим формулы для финальных вероятностей состояний. При $k \le N$

$$P_k = \frac{\lambda}{k\nu} \cdot \frac{\lambda}{(k-1)\nu} \cdot \ldots \cdot \frac{\lambda}{1 \cdot \nu} \cdot P_0 = \frac{\lambda^k}{\nu^k \cdot k!} \cdot P_0 \qquad (6.8)$$

Введем обозначение

$$\rho = \frac{\lambda}{\nu} \qquad (6.9)$$

Величина ρ является отношением интенсивности входящего потока к интенсивности обслуживания и называется <u>загрузкой</u> системы. Это величина безразмерная и, в отличие от λ и ν, она не зависит от того, какой интервал времени выбран в качестве единицы. Как мы увидим далее, ρ является естественной и важной характеристикой систем обслуживания и через нее выражаются многие другие характеристики.

Подставив (6.9) в (6.8), получим

$$P_k = \frac{\rho^k}{k!} \cdot P_0 \qquad (6.14)$$

Поскольку

$$\sum_{k=0}^{N} P_k = 1, \qquad (6.15)$$

то из (6.14 – 6.15) получаем

$$P_0 = \frac{1}{\sum_{k=0}^{N} \frac{\rho^k}{k!}} \qquad (6.16)$$

Отсюда имеем

$$P_j = \frac{\frac{\rho^j}{j!}}{\sum_{k=0}^{N} \frac{\rho^k}{k!}} \qquad (6.17)$$

Эти формулы носят название <u>формул Эрланга</u>. Можно доказать, что последние формулы верны для СМО с отказами при любом (не обязательно экспоненциальном) законе распределения длительности обслуживания, при интенсивности обслуживания, равной ν.

Одной из важнейших характеристик работы системы является вероятность отказа

Требование, поступающее в систему, получает отказ в том и только том случае, когда все узлы заняты. Поэтому вероятность отказа $P_{отк}$ равна

$$P_{отк} = P_N = \frac{\dfrac{\rho^N}{N!}}{\sum\limits_{k=0}^{N} \dfrac{\rho^k}{k!}} \qquad (6.18)$$

Вероятность $P_{отк}$ характеризует долю требований получающих отказ. Величина

$$\alpha = 1 - P_{отк} = 1 - P_N \qquad (6.19)$$

определяет долю обслуженных требовании. Она называется относительной пропускной способностью СМО. Среднее число требований, поступающих в узлы обслуживания в единицу времени есть

$$A = \lambda \cdot (1 - P_N). \qquad (6.20)$$

Эта характеристика называется абсолютной пропускной способностью системы.

Среднее число занятых узлов обслуживания $M_{зан}$ равна

$$M_{зан} = \sum_{k=0}^{N} k \cdot P_k = \sum_{k=1}^{N} k \cdot \frac{\rho^k}{k!} \cdot P_0 = \rho \cdot \sum_{m=0}^{N-1} \frac{\rho^m}{m!} \cdot P_0 = \rho \cdot (1 - P_N) \qquad (6.21)$$

Среднее число свободных узлов $M_{св}$ равно, естественно, разности

$$M_{св} = N - M_{зан} \qquad (6.22)$$

Для случая, когда в СМО с отказами имеется всего один узел обслуживания, мы получаем

$$P_0 = \frac{1}{1+\rho} \qquad (6.23)$$

$$P_{отк} = P_1 = \frac{\rho}{1+\rho} \qquad (6.24)$$

$$M_{зан} = \frac{\rho}{1+\rho} \qquad (6.25)$$

Упражнения

1. Найдите формулы для характеристик работы системы с отказами, имеющей два узла обслуживания.
2. Если число узлов велико по сравнению с загрузкой системы, то его удобно иногда считать равным бесконечности. Докажите, что для систем с $N = \infty$ верны следующие формулы:

$$P_0 = e^{-\rho}, \tag{6.26}$$

$$P_k = \frac{\rho^k}{k!} \cdot e^{-\rho}, \tag{6.27}$$

$$M_{зан} = \rho. \tag{6.28}$$

3. Покажите, что среднее число занятых узлов есть отношение интенсивности потока тех требований, которые не получают отказ, к интенсивности обслуживания.
4. На телефонной станции имеются три соединительных линии. Вызов, поступающий, когда все линии заняты, получает отказ. Поток вызовов является пуассоновским с параметром $\lambda = 0{,}5$ вызовов в минуту; время обслуживания распределено по экспоненциальное закону. Средняя продолжительность разговора составляет 3 мин.

Найдите вероятность отказа, относительную и абсолютную пропускные способности и долю свободного времени, приходящегося в среднем на какую линию. При каком наименьшем числе линий доля получающих отказ требований составит меньше 5%?