

ПОТОКИ СОБЫТИЙ, СИСТЕМЫ И СЕТИ МАССОВОГО ОБСЛУЖИВАНИЯ

УДК 519.283

Е.Н. Беккерман, С.Г. Катаев, С.С. Катаева

ОБ ОДНОМ АЛГОРИТМЕ РАСПОЗНАВАНИЯ МС-ПОТОКА СОБЫТИЙ

В работе предлагается алгоритм определения участков стационарности и оценивания соответствующих интенсивностей в распознаваемом МС-потоке событий.

В теории массового обслуживания широко исследуются системы массового обслуживания (СМО) с входящим МС-потоком заявок. Если на вход СМО поступает неизвестный поток заявок, то возникает задача распознавания в нем МС-потока для дальнейшего исследования этой системы уже разработанными методами.

МС-потоком называется поток, интенсивность которого представляет собой кусочно-постоянный марковский процесс $\lambda(t)$, принимающий значения из конечного множества констант $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, называемых состояниями процесса $\lambda(t)$. Длительности участков стационарности, т.е. участков, где $\lambda(t) = \text{const}$, есть случайные величины, распределенные по экспоненциальному закону с параметрами $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ соответственно. Переход из состояния i в состояние j осуществляется по вероятности p_{ij} из матрицы вероятностей переходов, которая обладает свойством $\sum_j p_{ij} = 1, p_{ii} = 0$.

Пусть дана выборка моментов наступления событий распознаваемого потока фиксированного размера n : $\{0, t_1, t_2, \dots, t_n\}$. Для любых пар событий с номерами i и j , где $i=1, \dots, n-1, j=2, \dots, n$, причем $i < j$, строим оценки интенсивности потока по формуле

$$\hat{\lambda}_{ij} = \frac{j-i+1}{t_{j+1} - t_{i-1}}, j > i. \quad (1)$$

По (1) получим $(n-1)(n-2)/2$ оценок интенсивности исследуемого потока по всевозможным группам событий из выборки. Представим полученные данные в виде верхнетреугольной матрицы D . Элемент (i, j) матрицы есть оценка интенсивности потока на интервале (t_{i-1}, t_{j+1}) по $(j-i+1)$ событиям. Предполагая, что исследуемый поток является МС-потоком, в дальнейшем для удобства будем использовать соответствующую терминологию. Заметим, что если исследуемый поток является МС-потоком, то для групп событий, принадлежащих одному и тому же интервалу стационарности, оценка (1) интенсивности потока на соответствующем временном интервале будет несмещенной, эффективной и асимптотически нормальной, так как получена методом максимального правдоподобия [1].

Рассмотрим свойства полученной матрицы и ее отдельных элементов. Будем называть элемент матрицы $D = \{\hat{\lambda}_{ij}\}$ чистой оценкой интенсивности потока на интервале стационарности, если все события, наступившие в моменты $\{t_i, t_{i+1}, \dots, t_{j+1}\}$, принадлежат одному и тому же интервалу стационар-

ности потока. Остальные элементы матрицы назовем ложными оценками или шумом. Если исследуемый поток пуассоновский, то все элементы матрицы D являются чистыми оценками его интенсивности. В случае МС-потока, если события, начиная с номера i_k и заканчивая номером j_k , относятся к одному интервалу стационарности, то для любых пар событий с номерами i и j , где $i=i_k, j_k=1, j=i_k+1, j_k$, причем $i < j$, элемент d_{ij} матрицы D будет чистой оценкой, так как на интервале стационарности МС-поток является пуассоновским. Эти оценки составят в матрице D треугольный блок размерности $(j_k - i_k + 1)$, расположенный на главной диагонали.

Сгруппируем данные, содержащиеся в матрице оценок, в гистограмму

$$\left\{ kh; \frac{n_k}{N}; N = \sum_k n_k \right\},$$

где h – шаг гистограммы; n_k – частоты, т.е. количество элементов матрицы D , значения которых принадлежат интервалу $((k-1)h, kh]$; N – количество элементов матрицы D , значения которых принадлежат интервалу $(0, mh]$; m – количество шагов гистограммы.

На основании полученной гистограммы построим полигон частот. Из свойств оценки $\hat{\lambda}$ вытекает, что для пуассоновского потока при достаточной длине реализации полигон частот представляет собой δ -образную кривую, максимум которой совпадает со значением интенсивности потока λ с точностью до шага гистограммы [2]. В случае МС-потока это не так. Если бы в случае МС-потока удалось выделить из матрицы оценок только те элементы, которые являются чистыми оценками, полигон частот выглядел бы как сумма δ -образных кривых, максимумы которых совпадают с интенсивностями потока на интервалах стационарности [3].

В данной работе предлагается алгоритм выделения чистых оценок интенсивности МС-потока путем определения участков стационарности, причем в качестве исходной информации используются только времена наступления событий.

Поставим в соответствие каждому i -му событию выборки, $i=1, N$, вершину с тем же номером i некоторого графа $G(V, E)$, где V – множество мощности N вершин графа, а E – множество пар вершин, т.е. множество ребер графа. Стоимостью каждого ребра графа, связывающего i -ю и j -ю ($i < j$) вершины, будем считать величину d_{ij} , т.е. элемент матрицы D . Число ребер графа составит $N(N-1)/2$. Если группа событий с номерами

$i_1^p, i_2^p, \dots, i_k^p$ составляет участок стационарности потока с интенсивностью $\lambda = \lambda_p$, то это означает, что:

$$a) \frac{2}{k(k-1)} \sum_{(i,j) \in I} d_{ij} \cong \lambda_p, \text{ где } I - \text{множество упорядо-}$$

ченных по номерам пар из $\{i_1^p, i_2^p, \dots, i_k^p\}$;

б) номера $i_j^p, j=1, \dots, k$, идут подряд, т.е. $i_j^p = i_{j-1}^p + 1$;

в) все $d_{ij}, (i,j) \in I$, принадлежат в основном некоторому интервалу $(\lambda_p - \Delta\lambda, \lambda_p + \Delta\lambda)$.

Оценка интенсивности, полученная по всей группе событий, т.е. (i_1^p, i_k^p) -я компонента матрицы D , должна быть наиболее близкой к истинному λ_p .

Эти соображения определяют идею алгоритма выделения участков стационарности потока. При выборе удачного интервала $(\lambda_p - \Delta\lambda, \lambda_p + \Delta\lambda)$ и удалении из графа G всех ребер, стоимость которых не попадает в указанный интервал, в графе выделяются компоненты связности $G_k = (V_k, E_k)$, соответствующие участкам стационарности, относящимся к p -му состоянию МС-потока (G_k является подграфом графа G). Разумно ожидать, что участку стационарности в наибольшей степени отвечает такой подграф, в котором, во-первых, число ребер максимально, а во-вторых, каждая вершина должна быть соединена с наибольшим числом вершин, т.е. степени вершин должны быть примерно одинаковыми. Эти требования могут быть учтены введением следующих характеристик для выделенных компонент связности G_k :

а) степень компактности компоненты

$$\text{COMP}[k] = \frac{2 \sum_{i=1}^{N_k} s_i}{N_k(N_k - 1)},$$

где N_k – число вершин G_k ; s_i – степень i -й вершины. Величина $\text{COMP}[k]$ характеризует близость подграфа G_k к полному;

б) степень однородности компоненты

$$\text{ENTR}[k] = \frac{- \sum_{i=1}^{N_k} p_i \log p_i}{\log N_k},$$

где p_i – отношение степени i -й вершины компоненты G_k к величине $(N_k - 1)$. Характеристика $\text{ENTR}[k]$ по виду напоминает формулу энтропии случайной величины с распределением $p_i, i=1, N_k$. Величина энтропии, как известно [4], характеризует степень однородности множества.

Пусть задан интервал значений оценок интенсивности $[\lambda', \lambda'']$, граф G разбит вышеупомянутым способом на компоненты связности и для каждой компоненты вычислены характеристики COMP и ENTR . Если при увеличении интервала $[\lambda', \lambda'']$ и соответствующем разбиении графа G на компоненты связности для k -й компоненты N_k остается неизменным, а число связей увеличивается, то благодаря увеличившемуся числителю значение $\text{COMP}[k]$ возрастает, и подграф G_k приближается к полному. Присоединение в новом разбиении к k -й компоненте изолированной в предыдущем разбиении вершины увеличивает знаменатель, а числитель может возрасти в меньшей степени, если, например, вершина присоединилась толь-

ко одной связью, и это приведет к уменьшению значения $\text{COMP}[k]$. С другой стороны, рост числа связей может достигаться увеличением степени только одной или нескольких вершин, поэтому само по себе большое значение $\text{COMP}[k]$ не гарантирует оптимальности разбиения. Характеристика $\text{ENTR}[k]$, отвечающая за однородность компоненты, достигает своего локального максимума, когда степени всех вершин компоненты равны. Для каждой компоненты G_k $\text{COMP}[k]$ и $\text{ENTR}[k]$ – слабо флуктуирующие многоэкстремальные функции от $[\lambda', \lambda'']$.

Опишем процедуру выделения связанных компонент графа.

1. Задан интервал $[\lambda', \lambda'']$. Из графа удаляются все ребра, стоимости которых не попадают в данный интервал. Граф G разбивается на компоненты связности $G_k, k=1, 2, \dots$ (в том числе изолированные вершины).

2. Интервал $[\lambda', \lambda'']$ увеличивается. Получаются новые разбиения графа G , в которых компоненты изменяются: расширяются за счет изолированных вершин, поглощают другие компоненты. Наконец, при достижении интервалом некоторой величины граф G не разобьется на компоненты, останется связным.

3. На множестве разбиений k -я компонента характеризуется последовательностями характеристик $\text{COMP}[k]$ и $\text{ENTR}[k]$. Выбирается то разбиение, для которого $\text{COMP}[k]$ достигает своего локального максимума, но при этом и $\text{ENTR}[k]$ близка к своему локальному максимуму. Те компоненты, которые этому условию удовлетворяют, считаются перспективными в смысле соответствия интервалам стационарности. Средняя стоимость ребра полученной компоненты, которая рассчитывается с учетом ранее отброшенных ребер, будет оценкой интенсивности потока для одного из состояний. Если эти оценки для разных компонент близки, то это интервалы стационарности одного состояния МС-потока.

Что касается выбора интервала $[\lambda', \lambda'']$ и принципов его расширения, то на этот счет можно привести следующие соображения. Можно, определив границы множества значений оценок по матрице D , зафиксировать левый край интервала λ' на нижней границе множества значений оценок и с заданным шагом увеличивать значение λ'' до тех пор, пока граф продолжает разбиваться на компоненты связности. Затем, зафиксировав λ'' на правой границе множества значений оценок, поступить аналогичным образом с λ' . С другой стороны, для выбора перспективных интервалов $[\lambda', \lambda'']$ можно воспользоваться информацией о местоположении пиков гистограммы.

Ясно, что если события с номерами $i_1, i_2, \dots, i_k$ относятся к одному интервалу стационарности, значит, (i, j) -е компоненты матрицы D при $i \in \{1, \dots, i_1 - 1\}$ и $j \in \{i_1, N\}$, а также при $i \in \{i_1, i_k\}$ и $j \in \{i_k + 1, N\}$ являются заведомо ложными оценками (шумом). Представляется целесообразным удалить из матрицы D ложные оценки и повторить процедуру выделения связанных компонент графа. Если состав компонент не изменяется, то интервалы стационарности для данных состояний считаются определенными.

Пример. Построена имитационная модель МС-потока с двумя состояниями и следующими параметра-

ми: $\lambda_1=20$; $\lambda_2=80$; $\alpha_1=0,1$; $\alpha_2=1,2$. Разбиение первых 200 событий по состояниям для данной реализации показано в табл. 1.

Таблица 1

События	Состояние
1 – 109	1
110 – 182	2
183 – 190	1
191 – 200	2

Предполагается, что нам известны только моменты наступления событий. По этим моментам была построена (симметризованная для иллюстрации) матрица D оценок интенсивности потока, трехмерное изображение которой представлено на рис. 1, где по оси «оценка интенсивности» отложены значения соответствующих элементов матрицы D .

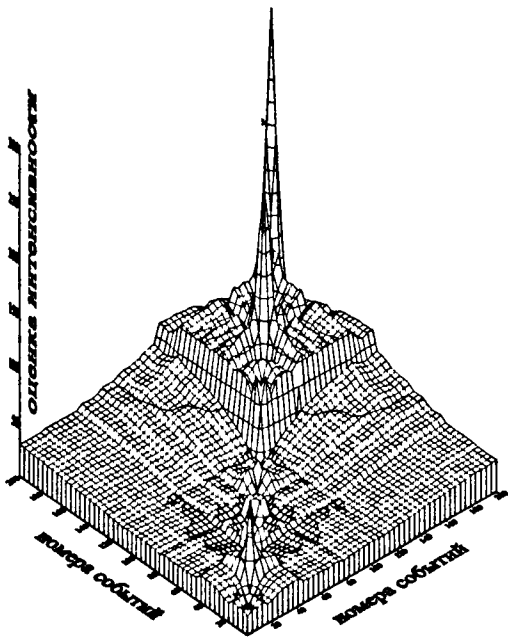


Рис. 1. Исходная матрица D оценок интенсивности потока событий

Гистограмма, соответствующая матрице D , изображена на рис. 2, из которого видно, что огибающая гистограммы заметно отличается от δ -образной кривой, а это говорит о наличии в рассматриваемом потоке как минимум двух состояний.

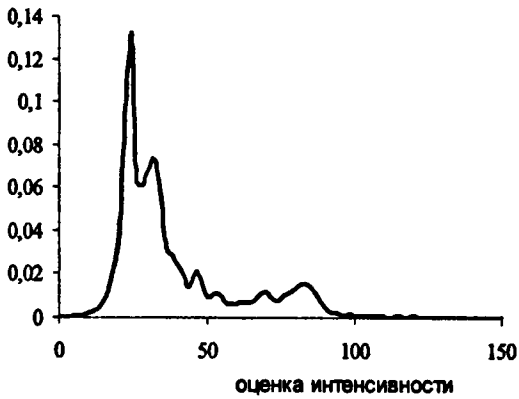


Рис. 2. Огибающая гистограммы оценок интенсивности

Анализ множества разбиений событий на группы по различным интервалам $[\lambda', \lambda'']$ в левой части множества значений оценок λ позволил выделить лучшее в смысле характеристик COMP и ENTR разбиение для наибольшей по численности группы событий. Полученное разбиение, упорядоченное по количеству событий в группе, относилось к интервалу [4, 19] и приведено в табл. 2.

Таблица 2

№ групп	Число событий	Начальное событие	Конечное событие	Ср. оценка интенсивности	Оценка по группе
1	60	28	87	19,648	21,316
2	27	89	115	21,888	23,259
3	23	1	23	21,334	23,818
4	5	181	185	21,007	23,652
5	3	188	190	18,170	18,216

По близости средних оценок интенсивности $\bar{\lambda}$ и оценок по группе $\hat{\lambda}$ можно сделать предположение о том, что все 5 групп событий относятся к одному состоянию МС-потока, а 1-я и 2-я группы, более того, принадлежат одному участку стационарности потока со случайно выпавшим событием под номером 88. Из матрицы D удаляются элементы d_{ij} при $i=1,27$ и $j=28,200$, а также при $i=28,115$ и $j=116,200$, и процедура разбиения событий на группы повторяется. Заметим, что поскольку разбиение фиксировалось для интервала, оптимального только для 1-й группы событий, то удаление шума из матрицы D для групп 3–5 не проводилось. Лучшее разбиение, относившееся к интервалу [4, 38], приведено в табл. 3.

Таблица 3

Номер группы	Число событий	Начальное событие	Конечное событие	Ср. оценка интенсивности	Оценка по группе
1	88	28	115	21,198	22,135
2	27	1	27	24,470	27,026
3	22	176	197	48,040	48,792
4	6	145	150	30,363	51,420
5	5	116	120	40,632	56,313
6	4	165	168	35,240	41,115
7	4	153	156	38,175	52,244

Вновь 1-я и 2-я группы объединяются в одну по близости $\bar{\lambda}$ и $\hat{\lambda}$, а также по последовательности номеров составляющих их событий (т.е. обе группы принадлежат одному участку стационарности потока). Группы 3–7 имеют заметно отличные от 1-й и 2-й оценки $\bar{\lambda}$ и $\hat{\lambda}$, поэтому их можно рассматривать либо как еще не сформировавшиеся участки стационарности для другого состояния потока, либо как совокупности событий, относящихся к разным участкам стационарности одного или различных состояний. Из матрицы D были удалены элементы d_{ij} при $i=1,115$, $j=116,200$, и процедура разбиения событий на группы была повторена. Лучшее разбиение, относившееся к интервалу [4, 27], приведено в табл. 4. Анализ этой таблицы позволяет сделать предварительный вывод о существовании двух участков стационарно-

сти, относящихся к состоянию МС-потока, величина интенсивности которого принадлежит интервалу $[22, 30]$. Первый участок содержит события с 1-го по 115-е, а второй – со 180-го по 191-е. Относительно 3-й и 4-й групп ничего определенного сказать нельзя.

Таблица 4

№ группы	Число событий	Начальное событие	Конечное событие	Ср. оценка интенсивности	Оценка по группе
1	115	1	115	22,567	23,192
2	12	180	191	25,576	30,672
3	4	146	149	28,037	35,700
4	3	117	119	29,218	36,444

Следующий этап состоял в выделении другого состояния МС-потока, в котором интенсивность наступления событий близка к верхней границе множества значений оценок интенсивности. Теперь разбиение событий на группы, аналогичное описанному, проводилось по различным интервалам $[\lambda', \lambda'']$ в правой части множества значений оценок интенсивности потока. Здесь окончательный результат относился к интервалу $[70, 76]$ и приведен в табл. 5.

Таблица 5

Номер группы	Число событий	Начальное событие	Конечное событие	Ср. оценка интенсивности	Оценка по группе
1	81	110	200	74,685	70,247
2	6	23	28	80,966	88,190

Трехмерное изображение матрицы D оценок интенсивности потока после удаления из нее всех ложных оценок (шума) представлено на рис. 3.

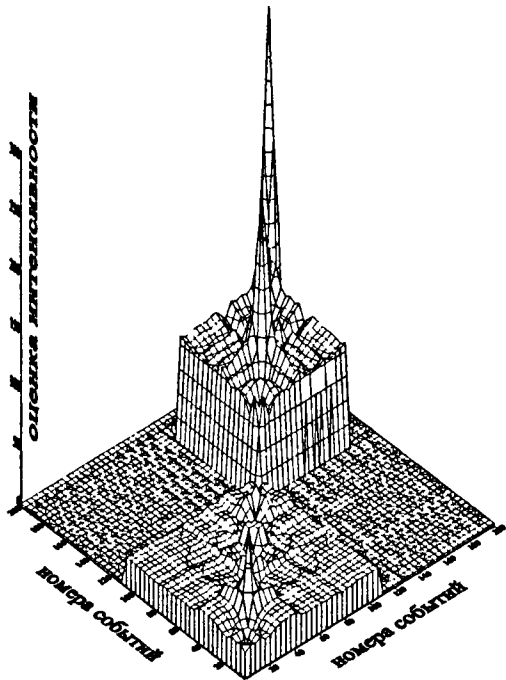


Рис. 3. Очищенная матрица D оценок интенсивности потока событий

Сравнивая табл. 4 и 5, видим, что все события оказались распределенными по состояниям. Это позволяет утверждать, что исследуемый поток есть МС-поток с двумя состояниями. Однако некоторые события оказались отнесенными как к 1-му, так и к 2-му состояниям. На рис. 4 эти «спорные» участки отмечены черными прямоугольниками. Надо принять решение об отнесении этих «спорных» участков к тому или иному состоянию.

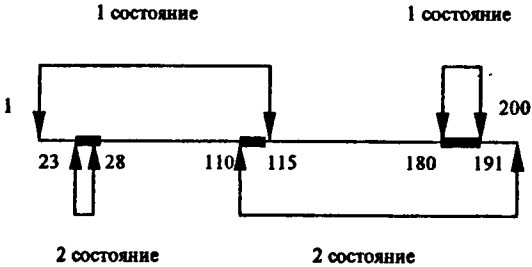


Рис. 4. Результирующее распределение событий по состояниям

В качестве критериев, которыми необходимо руководствоваться в этой ситуации, естественно использовать оценки интенсивностей. В табл. 6 приведены основные характеристики групп событий, составляющих «спорные» участки.

Таблица 6

Номер группы	Число событий	Начальное событие	Конечное событие	Ср. оценка интенсивности	Оценка по группе
1	6	23	28	80,966	88,190
2	6	110	115	60,502	64,410
3	12	180	191	25,576	30,672

Из табл. 6 видно, что участки 23–28, 110–115 следует отнести ко 2-му, а 180–191 – к 1-му состоянию. Окончательно полученное распределение событий по состояниям приведено в табл. 7.

Таблица 7

События	Состояние	Ср. оценка интенсивности
1–22	1	21,281
23–28	2	80,966
29–109	1	20,978
110–179	2	80,118
180–191	1	25,576
192–200	2	114,946

Результирующие оценки интенсивности МС-потока (их средние значения по участкам стационарности) равны: $\hat{\lambda}_1 = 21,58$, $\hat{\lambda}_2 = 83,87$. Рассмотренный пример показал, что:

- а) количество состояний определено точно;
- б) оценки интенсивности наступления событий для обоих состояний близки к истинным значениям интенсивности модельного потока;
- в) число интервалов стационарности потока отличается от истинного на 1;
- г) число событий, отнесенных не к «своему» интервалу, составляет 10 из 200.

В предложенном алгоритме не используется никакой информации о распознаваемом потоке, кроме времен наступления событий. При усовершенствовании процедуры принятия решений этот алгоритм позволит распознавать МС-потоки с различным числом состояний; оцени-

вать их интенсивности и длительности нахождения потока в каждом интервале стационарности, что при большем числе событий, нежели в приведенном примере, поможет оценить интенсивности перехода из одного состояния в другое.

ЛИТЕРАТУРА

1. Терпугов А. Ф. Математическая статистика. Томск: Изд-во ТГУ, 1974.
2. Катаева С. С. Об одном подходе к распознаванию МС-потока событий // Массовое обслуживание. Потоки, системы, сети: Материалы международной конференции «Математические методы исследования телекоммуникационных сетей». Минск, 1997. С. 43.
3. Беккерман Е. Н., Катаева С. С. Эвристический способ обнаружения информационного признака МС-потока и его исследование // Массовое обслуживание. Потоки, системы, сети: Материалы международной конференции «Математические методы исследования систем и сетей массового обслуживания». Минск, 1998. С. 5–9.
4. Шеннон К. Математическая теория связи // Работы по теории информации и кибернетике. М.: ИЛ, 1963.

Статья представлена кафедрой исследования операций факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию 19 февраля 2000 г.

УДК 519.8

Е. В. Глухова, А. С. Шкуркин

РАСЧЕТ ХАРАКТЕРИСТИК ПЕРИОДА ЗАНЯТОСТИ В ОДНОЛИНЕЙНОЙ СМО С ВЫТЕСНЕНИЕМ ЗАЯВОК

Находится преобразование Лапласа от плотности вероятностей длительности периода занятости в однолинейной СМО с вытеснением заявок, а также математическое ожидание и дисперсия длительности периода занятости. Строятся оценки интенсивности входящего потока заявок и среднего времени обслуживания по наблюдениям над моментами начала периодов занятости.

Постановка задачи

Однолинейные СМО с вытеснением заявок встречаются при описании так называемого продлевающегося мертвого времени [1]. Математическая модель таких систем СМО выглядит следующим образом. Имеется однолинейная СМО, на которую поступает рекуррентный поток заявок с плотностью вероятностей для интервалов времени τ между заявками вида $p_\tau(\tau) = \lambda p_1(\lambda\tau)$, где $p_1(z)$ – функция, обладающая свойствами:

- а) $p_1(z) \geq 0$,
- б) $\int_0^\infty p_1(z) dz = 1$, $\int_0^\infty z p_1(z) dz = 1$.

Тогда λ имеет смысл интенсивности потока заявок.

Если каждая заявка обслуживается независимо от остальных, то время её обслуживания t имеет плотность

вероятностей вида $p_t(t) = \frac{1}{\theta_0} p_0\left(\frac{t}{\theta_0}\right)$, где функция

$p_0(z)$ имеет те же свойства, что и функция $p_1(z)$. В этом случае параметр θ_0 – среднее время обслуживания заявки. Термин «вытеснение заявок» означает следующее: если в период обслуживания какой-то заявки придет следующая заявка, то она вытесняет с обслуживающего прибора находящуюся там заявку и сама занимает её место. Вытесненная заявка теряется и на обслуживание не возвращается. Обозначим через ξ длительность периода занятости в такой системе и через η – интервал времени, проходящий между началами периодов занятости. Нас будет интересовать распределение вероятностей этих величин. В частном случае пуассоновского входящего потока эта задача решена в [1].

Распределение вероятностей величин ξ и η

Найти плотности вероятностей величин ξ и η затруднительно, а преобразование Лапласа от них

вычислить можно. Пусть $p_\xi(\xi)$ и $p_\eta(\eta)$ – плотности вероятностей ξ и η , а

$$\begin{aligned} g_\xi(s) &= \int_0^\infty e^{-s\xi} p_\xi(\xi) d\xi = M\{e^{-s\xi}\}, \\ g_\eta(s) &= \int_0^\infty e^{-s\eta} p_\eta(\eta) d\eta = M\{e^{-s\eta}\} \end{aligned} \quad (2)$$

– их преобразования Лапласа. Найдём выражение для $g_\eta(s)$. Пусть в пустую СМО поступила заявка, требующая для своего обслуживания времени t . Тогда возможны два варианта.

1. За время обслуживания этой заявки не поступит никакой другой заявки, т.е. наступит событие $t > \tau$. В этом случае $\eta = \tau$, так как через время t период занятости окончится и следующий период начнётся в момент τ .

2. За время обслуживания этой заявки в систему поступит новая заявка, т.е. наступит событие $\tau < t$. В этом случае $\eta = \tau + \eta'$, где η' – интервал времени до начала нового периода занятости, отсчитываемый от поступления этой новой заявки. η и η' имеют одно и то же распределение вероятностей, так как новая заявка вытесняет старую. Получившаяся ситуация ничем не отличается от исходной и имеет место соотношение:

$$\begin{aligned} g_\eta(s) &= \int_0^\infty \frac{1}{\theta_0} p\left(\frac{t}{\theta_0}\right) dt \int_0^\infty e^{-s\tau} \lambda p_1(\lambda\tau) d\tau + \\ &+ M\{e^{-s\eta'}\} \int_0^\infty \frac{1}{\theta_0} p\left(\frac{t}{\theta_0}\right) dt \int_0^\infty e^{-s\tau} \lambda p_1(\lambda\tau) d\tau. \end{aligned}$$

Но, по сказанному выше, $M\{e^{-s\eta'}\} = g_\eta(s)$ и

$$g_\eta(s) = \frac{\int_0^\infty p\left(\frac{t}{\theta_0}\right) \frac{dt}{\theta_0} \int_0^\infty \lambda p_1(\lambda\tau) e^{-s\tau} d\tau}{1 - \int_0^\infty p\left(\frac{t}{\theta_0}\right) \frac{dt}{\theta_0} \int_0^\infty \lambda p_1(\lambda\tau) e^{-s\tau} d\tau}, \quad (3)$$

что и даёт явное выражение для $g_\eta(s)$.