УДК 519.21

ИССЛЕДОВАНИЕ РЕАЛИЗАЦИИ ТРАНСПОРТНОГО ПОТОКА БАРТЛЕТТА

А.М. Федоткин, А.А. Федоткин

Нижегородский госуниверситет им. Н.И. Лобачевского

fandr@vmk.unn.ru

Поступила в редакцию 31.01.2013

Рассмотрена задача исследования реализации конкретного случайного процесса. Показано, что интервалы между поступлениями заявок в данном потоке зависимы и имеют разное распределение. Предложен метод нелокального описания данного потока, и доказано, что он может быть аппроксимирован потоком Гнеденко—Коваленко. Разработан пакет программ, который позволяет проверить выдвинутые теоретические предположения нелокального описания.

Ключевые слова: случайный процесс, конечномерные распределения, маркированный точечный процесс, интенсивность.

В работах [1–3] впервые рассматривается теория транспортных потоков, которая частично учитывает свойства как пространственного, так и временного процесса. Такой подход, основанный на теории нелокального описания [5–8], позволяет исследовать процессы, в которых интервалы между событиями зависимы и имеют разное распределение. Был разработан метод выбора стробирующих моментов (разбивка потока на пачки) [2], который позволяет исследовать такие сложные неоднородные потоки. В данной статье решается конкретная задача, описанная в работе [4]. Для такого типа транспортного потока Бартлетту и другим исследователям не удалось найти подходящего закона распределения для промежутков времени между двумя последовательными пересечениями автомобилями виртуальной стоп-линии дорожной магистрали. Ниже приведена табл. 1 с экспериментальными данными из работы [4]. В табл. 1 указаны 128 интервалов между последовательными автомобилями. Интервалы измеряются в секундах. Читать табл. 1 следует по строкам. Из этой таблицы легко вычислить, что Бартлетт наблюдал процесс пересечения автомобилями виртуальной стоп-линии некоторой транспортной магистрали в течение времени около 2024 сек. При этом всех автомобилей было зафиксировано 129. На вероятностную структуру транспортного потока влияют физическая характеристика дороги, состав автомобилей, правила движения и погода. Первые три фактора можно считать постоянными. Изменения погоды всегда имеют случайный характер. При хороших метеорологических условиях движение машин по магистрали является пуассоновским.

При плохих погодных условиях обгон быстрыми машинами медленных занимает значительное время. В этих условиях на дороге будут образовываться группы автомобилей. Впереди каждой группы находится медленная машина. Таким образом, множество всех машин на магистрали разбиваем на быстрые и медленные. Экспериментальные данные табл. 1, по-видимому, отражают описанную ситуацию. До настоящего момента неизвестно подходящее теоретическое распределение для интервалов из табл. 1. Более того, применяя фазово-частотный критерий Валлиса и Мура [9] к статистическим данным табл. 1, получаем значение статистики $Z_1(n, X_1, X_2, ..., X_n)$ при $n = 128$, равное 3.378. Вид статистики Валлиса и Мура $Z_1(n, X_1, X_2, ...,X_n)$

Таблица 1

Значения интервалов $\tau'_{i+1} - \tau'_i$, $i = 1, 2, ..., 128$, при $\tau'_1 = 0$

2.8	3.4	1.4	14.5	1.9	2.8	2.3	15.3	1.8	9.5	2.5	9.4
1.1	88.6	1.6	1.9	1.5	33.7	2.6	12.9	16.2	1.9	20.3	36.8
40.1	70.5	2	8	2.1	3.2	1.7	56.5	23.7	2.4	21.4	5.1
7.9	20.1	14.9	5.6	51.7	87.1	1.2	2.7	1	1.5	1.3	24.7
72.6	119.8	1.2	6.9	3.9	1.6	3	1.8	44.8	5	3.9	125.3
22.8	1.9	15.9	6	20.6	12.9	3.9	13	6.9	2.5	12.3	5.7
11.3	2.5	1.6	7.6	2.3	6.1	2.1	34.7	15.4	4.6	55.7	2.2
6	1.8	1.9	1.8	42	9.3	91.7	2.4	30.6	1.2	8.8	6.6
49.8	58.1	1.9	2.9	0.5	1.2	31	11.9	0.8	1.2	0.8	4.7
8.3	7.3	8.8	1.8	3.1	0.8	34.1	3	2.6	3.7	41.3	29.7
17.6	1.9	13.8	40.2	10.1	11.9	11	0.2				

Таблица 2

Значения векторов $(\tau_{i+1} - \tau_i, \eta_i)$, $i = 0, 1, ..., 87$, при $\tau_0 = 0$

(6.2, 2)	(1.4, 1)	(16.4, 2)	(2.8, 1)	(17.6, 2)	(1.8, 1)	(12, 2)	(9.4, 1)	(89.7, 2)
(1.6, 1)	(3.4, 2)	(33.7, 1)	(15.5, 2)	(16.2, 1)	(22.2, 2)	(36.8, 1)	(40.1, 1)	(70.5, 1)
(10.2, 2)	(2.1, 1)	(4.9, 2)	(56.5, 1)	(23.7, 1)	(23.8, 2)	(5.1, 1)	(28, 2)	(14.9, 1)
(57.3, 2)	(87.1, 1)	(3.9 2)	(1, 1)	(2.8, 2)	(24.7, 1)	(72.6, 1)	(119.8, 1)	(8.1, 2)
(3.9, 1)	(4.6, 2)	(1.8, 1)	(44.8, 1)	(8.9, 2)	(125.3, 1)	(24.7, 2)	(15.9, 1)	(26.6, 2)
(12.9, 1)	(16.9, 2)	(6.9, 1)	(14.8, 2)	(5.7, 1)	(13.8, 2)	(1.6, 1)	(9.9, 2)	(6.1, 1)
(36.8, 2)	(15.4, 1)	(60.3, 2)	(2.2, 1)	(7.8, 2)	(1.9, 1)	(43.8, 2)	(9.3, 1)	(91.7, 1)
(33, 2)	(1.2, 1)	(15.4, 2)	(49.8, 1)	(58.1, 1)	(4.8, 2)	(0.5, 1)	(32.2, 2)	(11.9, 1)
(2, 2)	(0.8, 1)	(13, 2)	(7.3, 1)	(10.6, 2)	(3.1, 1)	(34.9, 2)	(3, 1)	(6.3, 2)
(41.3, 1)	(29.7, 1)	(19.5, 2)	(13.8, 1)	(40.2, 1)	(22, 2)	(11, 1)	(*, 2)	

приведен в статье [3]. Пороговое значение величины C_α на 5-процентном уровне значимости равно 1.96. Заметим, что пороговое значение C_α при заданном уровне значимости α определяется по интегральной функции распределения $\Phi(x)$ стандартной нормальной случайной величины из условия $\Phi(-C_\alpha) = \alpha/2$. Так как значение статистики Валлиса и Мура $Z_1(128, x_1, x_2, ..., x_{128}) = 3.378 > 1.96$, то согласно фазово-частотному критерию выдвинутую гипотезу о независимости и одинаковом распределении интервалов между последовательными автомобилями следует отклонить.

Проведем статистический анализ реализации реального транспортного потока Бартлетта [1] с использованием способа его разбиения, приведенного впервые в работе [3].

Пусть при фиксированном $c = 0,1,...$ моменты $\tau_i^{(c)}$, $i \ge 0$, на оси времени $[0, \infty)$ совпадают с некоторыми моментами τ'_i, $i \ge 1$, поступления требований в систему. Другими словами, моменты $\tau_i^{(c)}$, $i \ge 0$, совпадают с некоторыми точками разрыва исходного считающего случайного процесса $\{\eta(t)\colon t \ge 0\}$, т. е. $\tau_i^{(c)} = \tau'_{k_{c,i}}$, $k_{c,i} \in \{1, 2, \ldots\}$. Тогда при каждом $c \ge 0$ величина $\eta_i^{(c)} = k_{c,\,i+1} - k_{c,\,i}$ задает число поступивших требований на промежутке $[\tau_i^{(c)}, \tau_{i+1}^{(c)})$ первоначального потока и является величиной i-й группы виртуального потока $\{(\tau_i^{(c)}, \eta_i^{(c)});\ i \ge 0\}$. Величина $\delta_i^{(c)} = \tau'_{k_{c,i+1}} - \tau'_{k_{c,i+1}-1}$ определяет временной интервал между последовательными группами с номерами i и $i + 1$ исходного входного потока $\{\eta(t)\colon t \ge 0\}$ при его нелокальном описании в виде последовательности $\{(\tau_i^{(c)}, \eta_i^{(c)});\ i \ge 0\}$. Тогда при каждом фиксированном $c \ge 0$ элементы $\tau_i^{(c)}$, $i \ge 0$, потока $\{(\tau_i^{(c)}, \eta_i^{(c)});\ i \ge 0\}$ будем строить с помощью рекуррентных соотношений вида:

$$k_{0,\,i+1} = \inf\{k\colon\ k > k_{0,\,i},\ \ \tau'_k - \tau'_{k-1} \ge h_0\},$$

$$s_c = \inf\{k\colon k \ge 0,\ \ \eta_k^{(c)} \le d,\ \ \eta_{k+1}^{(c)} \le d,\ \ \delta_{k+1}^{(c)} < h_1,\ \ \eta_k^{(c)} = \eta_{k-1}^{(c)}\},$$

$$\tau_i^{(c+1)} = \begin{cases} \tau_i^{(c)} & \text{при } i \le s_c, \\ \tau_{i+1}^{(c)} & \text{при } i > s_c. \end{cases}$$

В этих формулах $\eta_{-1}^{(c)} = 1$ при каждом $c \ge 0$, $k_{0,0} = 1$, d — некоторое натуральное число и постоянные величины h_0, h_1 удовлетворяют условию $h_0 < h_1$. Легко видеть, что множество

$$\{\omega\colon \lim_{c\to\infty}\tau_i^{(c)} \text{ существует}\} = \Omega$$

для любого $i \ge 0$. Теперь для любого $i \ge 0$ определим случайную величину $\tau_i = \lim_{c\to\infty}\tau_i^{(c)}$. При таком алгоритме выбора потока $\{(\tau_i, \eta_i);\ i \ge 0\}$ имеем: $\tau_i = \tau'_{k_i}$, $\eta_i = k_{i+1} - k_i$ для всех $i \ge 0$.

Подбирая теперь параметры h_0, h_1, d для данного способа, можно разбить первоначальный транспортный поток Бартлетта на группы автомобилей. Например, при $d = 1$, $h_0 = 0.1$ и $h_1 = 23.7$ получим 88 интервалов $\tau_{i+1} - \tau_i$, $i = 0, 1, ..., 87$, между медленными машинами и последовательность из 89 значений для размеров $\eta_0, \eta_1, ..., \eta_{88}$ транспортных пачек. При этом каждая транспортная пачка содержит одну или две машины. Эти данные приведены в табл. 2. Читать табл. 2 следует по строкам. В скобках первое число – временной интервал между пачками в секундах, второе число – количество автомобилей в пачках. Последний элемент этой таблицы содержит символ «*». Это означает, что момент появления пачки размера η_{89} не определен.

Компьютерная обработка статистических данных табл. 2 методами из работы [3] дает следующие результаты. Последовательности вида $\{(\tau_{i+1} - \tau_i);\ i \ge 0\}$, $\{\eta_i;\ i \ge 0\}$ составлены из независимых в совокупности случайных величин. Так, например, для последовательности $\{(\tau_{i+1} - \tau_i);\ i \ge 0\}$ значение статистики $Z_2(n, X_1, X_2, ..., X_n)$ инверсионного критерия при $n = 88$ приблизительно равно 0.004. При этом формула для статистики $Z_2(n, X_1, X_2, ..., X_n)$ приведена в [3] . Так как $0.004 < 1.96$, то гипотеза о независимости и одинаковом распределении интервалов между последовательными медленными автомобилями не отклоняется. Значения статистик $\Upsilon_2(n, X_1, X_2, ..., X_n)$ и $\Delta(n, X_1, X_2, ..., X_n)$ фазово-частотного критерия с учетом длин фаз для последовательности $\{\eta_i;\ i \ge 0\}$ при $n = 89$ равны 77 и 1 соответственно. Формулы для статистик $\Upsilon_2(n, X_1, X_2, ..., X_n)$ и $\Delta(n, X_1, X_2, ..., X_n)$ приведены в работе [3].

На уровне значимости α ($0.05 < \alpha < 0.0975$) выдвинутая гипотеза о независимости случайных величин η_i, $i \geq 0$, не отвергается, так как при

$\Upsilon_2(89, x_1, x_2, ..., x_n) = 77$, $\Delta(89, x_1, x_2, ..., x_n) = 1$,

$$\frac{2n-1}{3} - 1.96\frac{\sqrt{16n-29}}{\sqrt{90}} \approx 51.2834,\ C_{0,\alpha} = 6$$

выполняются неравенства

$$\Upsilon_2(89, X_1, X_2, ..., X_n) > \frac{2n-1}{3} - 1.96\frac{\sqrt{16n-29}}{\sqrt{90}},\quad \Delta(n, x_1, x_2, ..., x_n) < C_{0,\alpha}.$$

С помощью критерия хи-квадрат [10] проверим теперь гипотезу о том, что последовательность $\{(\tau_{i+1} - \tau_i);\ i \geq 0\}$ составлена из случайных величин с распределением типа (1)

$$P(\{\omega: \tau_{i+1} - \tau_i < t\}) = 1 - \exp\{-(t-h)/\sigma\},\ t > h;$$
$$P(\{\omega: \tau_{i+1} - \tau_i < t\}) = 0,\ t \leq h. \qquad (1)$$

В работе [3] приводится методика получения оценок $\tilde{h}$ и $\tilde{\sigma}$ для неизвестных параметров $h \geq 0$ и $\sigma > 0$ распределения (1). При этом значение случайной величины $\tau_{i+1} - \tau_i$ разбито на $s = 5$ непересекающихся частей с шагом $b = 25.5$ и с начальным полуинтервалом $[0, a)$, где $a = 1.4005$. Заметим, что последний полуинтервал имеет вид $[a + (s-2)b, \infty)$.

Из табл. 2 с помощью формул, приведенных в работе [3], вычисляем оценки $\tilde{h} \approx 0.0006$ и $\tilde{\sigma} \approx 23.9312$. Приближенное значение статистики хи-квадрат равно 2.2302. При двух степенях свободы и уровне значимости в 5% пороговое значение хи-квадрат-распределения равно 5.991. Отсюда видно, что принятое гипотетическое распределение (1) для интервалов $(\tau_{i+1} - \tau_i)$, $i \geq 0$, между последовательными медленными машинами хорошо соответствует экспериментальным данным из табл. 2. Таким образом, можно принять, что распределение интервалов между последовательными медленными машинами в наблюдаемом потоке Бартлетта имеет вид

$$P(\{\omega: \tau_{i+1} - \tau_i < t\}) = 1 - \exp\{-(t - 0.0006)/23.9312\},\ t > h;$$
$$P(\{\omega: \tau_{i+1} - \tau_i < t\}) = 0,\ t \leq 0.0006. \qquad (2)$$

Перейдем теперь к определению приемлемого гипотетического распределения случайных величин η_i, $i \geq 0$. Из табл. 1 вычислим оценку $\tilde{\lambda}_+ = 129/2024$ интенсивности пересечения всеми машинами поперечной линии магистрали. Так как число всех медленных машин равно числу 89 всех пачек, то число быстрых машин равно $129 - 89 = 40$. Отсюда получаем простую оценку $\tilde{\lambda}_0 = 40/2024$ для интенсивности λ_0 быстрых машин. На практике транспортники очень часто замечают, что интенсивность $\mu_{1,0}$ обгона быстрыми машинами медленной машины увеличивается с увеличением интенсивности β_1 потока пачек из одной машины. Обычно рассматривают линейную зависимость $\mu_{1,0} = K\beta_1$ между этими величинами, где число $K > 0$ является коэффициентом пропорциональности и определяется условием дорожного движения. Отсюда можно предложить оценку $\tilde{\mu}_{1,0} = K\tilde{\beta}_1$ для интенсивности $\mu_{1,0}$ обгона быстрыми машинами медленной, где оценка $\tilde{\beta}_1 = 49/2024$ для β_1. Если условия дорожного движения являются относительно хорошими, когда образуются пачки небольшого размера, то значение коэффициента K, как правило, приблизительно равно единице. Здесь можно предположить, что коэффициент $K = 0.9$.

Проверим гипотезу о том, что интенсивность λ_0 потока быстрых машин равна 40/2024 и интенсивность $\mu_{1,0}$ обгона быстрыми машинами медленной равна 44.1/2024. В работе [2] было получено, что для стационарного режима вероятность появления в транспортной пачке одной машины равна $p = \mu_{1,0}(\lambda_0 + \mu_{1,0})^{-1} = 44.1/84.1$ и вероятность появления в транспортной пачке двух машин равна $q = \lambda_0(\lambda_0 + \mu_{1,0})^{-1} = 40/84.1$. Первоначальную гипотезу теперь можно заменить эквивалентной гипотезой о том, что гипотетическое распределение величины транспортной пачки является распределением Бернулли с постоянным параметром $p = 44.1/84.1$. Из табл. 2 получаем, что случайная величина η_i приняла значение единица $\nu = 49$ раз. Используя формулу (30.2.1) из работы [9], вычисляем значение статистики хи-квадрат $(\nu - np)^2/(npq)$ при $n = 89$. Значит, мера расхождения между выборочным распределением величины транспортной пачки и ее гипотетическим бернуллиевским распределением приближенно равна 0.2446. Это значение статистики хи-квадрат меньше 5-процентного значения величины χ_1^2 с одной степенью свободы, равного 3.841. Поэтому статистические данные табл. 2 можно считать совместимыми с гипотезой.

В заключение сделаем важное замечание. Исследование статистических данных Бартлетта позволяет сделать следующий вывод. Во-первых, последовательность $\{(\tau_{i+1} - \tau_i);\ i \geq 0\}$ составлена из независимых случайных величин с распределением (2), в котором параметр смещения $h = 0.0006$, т. е. приближенно равен нулю. Во-вторых, последовательность $\{\eta_i;\ i \geq 0\}$ составлена из независимых случайных величин, каждая из которых распределена по закону Бернулли. Поэтому можно считать, что наблюдаемый Бартлеттом вблизи Лондона в 1963 году транспортный поток можно нелокально описать потоком типа Гнеденко–Коваленко [11].

Список литературы

1. Федоткин А.М. Механизм образования неоднородной транспортной пачки и изучение динамики распределения ее величины // Вестник Нижегородского университета им. Н.И. Лобачевского. 2012. № 2. С. 171–177.

2. Федоткин А.М. Нелокальное описание входных потоков неоднородных требований // Вестник Нижегородского университета им. Н.И. Лобачевского. 2012. № 4. С. 211–217.

3. Федоткин А.М. Управляющие конфликтные системы и аппроксимация потока Гнеденко–Коваленко // Проблемы теоретической кибернетики. Материалы Международной конференции. Н. Новгород, 2011. С. 354–361.

4. Bartlett M.S. The spectral analysis of point processes // J. R. Statist. Soc. B. 1963. V. 25. № 2. P. 264–296.

5. Федоткин М.А. Процессы обслуживания и управляющие системы // Математические вопросы кибернетики. М.: Наука, 1996. Вып. 6. С. 51–70.

6. Федоткин М.А. Неполное описание потоков неоднородных требований. Теория массового обслуживания. М.: МГУ, ВНИИСИ, 1981. С. 113–118.

7. Федоткин М.А. Основы прикладной теории вероятностей и статистики. М.: Высшая школа, 2006. 368 с.

8. Федоткин А.М., Федоткин М.А. Анализ и оптимизация выходных процессов при циклическом управлении конфликтными транспортными потоками Гнеденко–Коваленко // Автоматика и телемеханика. 2009. № 12. С. 92–108.

9. Закс Л. Статистическое оценивание. М.: Статистика, 1976. 600 с.

10. Крамер Г. Математические методы статистики. М.: Мир, 1975. 648 с.

11. Гнеденко Б.В., Коваленко И.Н. Введение в теорию массового обслуживания. 3-е изд. испр. и доп. М.: Ком Книга, 2005. 400 с.

STUDY OF BARTLETT TRAFFIC FLOW IMPLEMENTATION

A.M. Fedotkin, A.A. Fedotkin

The implementation of a specific random process, which was given in Bartlett's article [1], is examined. It is shown that the intervals between demands in this traffic flow are dependent and have a different distribution. A method is proposed of non-local description of the flow, which is proved to be approximated by the Gnedenko-Kovalenko stream. A software package has been developed to test the proposed theoretical assumptions of the non-local description.

Keywords: random process, finite-dimensional distributions, marked point process, intensity.