

ПОСТРОЕНИЕ ФУНКЦИЙ РАСПРЕДЕЛЕНИЯ РЕАЛЬНОГО ТРАФИКА С ПОМОЩЬЮ КУМУЛЯНТНОГО АНАЛИЗА

Чупахина Л.Р., Куреева Н.В.

В данной статье рассматривается применение кумулянтного анализа для аппроксимации функций плотности распределения в звене мультисервисной сети. Возможность ее представления через кумулянты позволяет получить плотность распределения на выходе обрабатывающего устройства.

Ключевые слова: система массового обслуживания, трафик, моменты функций, кумулянты, распределения с «тяжелыми» хвостами.

Введение

Поток реального трафика в современных системах передачи обладает свойствами самоподобия, исследовать который невозможно методами, основанными на марковских моделях. В силу непуассоновского характера поведения реального трафика в качестве модели целесообразнее использовать системы массового обслуживания (СМО) типа G/G/1.

На практике при исследовании реальных систем редко встречаются известные законы распределения и обслуживания поступающего на вход системы трафика.

Невозможность в численном виде исследовать и оценить характеристики реального трафика позволяет рассматривать его в виде потока случайных событий, считая событием поступление пакета на обрабатывающее устройство. Поэтому, проводя кумулянтный анализ, вероятностные характеристики сети представим через моментные величины.

Аппроксимация функций плотности распределения

Одним из важных свойств кумулянтного описания случайных величин и процессов является то, что конечному набору кумулянтов всегда соответствует некоторая вещественная функция, аппроксимирующая вероятностное распределение. Это обстоятельство имеет особо важное значение в приближенном представлении вероятностных распределений тех случайных величин и процессов, для которых можно отыскать лишь конечные наборы кумулянтов и кумулянтных функций.

Решение данной задачи представим в виде алгоритма нахождения неизвестных плотностей вероятности $W(\tau)$ и $W(\xi)$ (см. рис. 1). На-

иболее распространена аппроксимация плотности вероятности в ряд Эджворта, дающего разложение произвольной плотности вероятности по производным гауссовского распределения [1].

$$W(x) = W_r(x) - \frac{\chi_3}{3!} W_r^{(3)}(x) + \frac{\chi_4}{4!} W_r^{(4)}(x) - \frac{\chi_5}{5!} W_r^{(5)}(x) + \frac{\chi_6 + 10\chi_3^2}{6!} W_r^{(6)}(x), \quad (1)$$

где $W_r^{(k)}(x)$ – производные плотности нормальной функции.

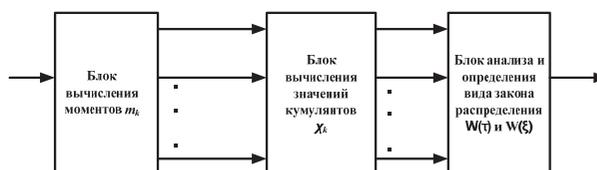


Рис. 1. Блок-схема реализации алгоритма

Для того чтобы описать процесс обслуживания в данном звене, необходимо определить:

- моменты входящего распределения поступления и обслуживания сообщения;
- кумулянты входящего распределения и распределения, по которому происходит процесс обслуживания в звене;
- плотности распределения $W(\tau)$ и $W(\xi)$ через их разложение в ряд Эджворта [2].

При решении задач в современных телекоммуникационных сетях мультисервисный трафик часто описывается с помощью распределений с «тяжелыми» хвостами, которые приводят к самоподобным процессам [3].

Выбирая наиболее часто встречающиеся распределения с «тяжелыми» хвостами: Вейбулла, Парето, логнормальное, можно смоделировать в качестве СМО систему типа G/G/1.

1. Рассмотрим аппроксимацию функции плотности распределения Вейбулла.

Функция распределения Вейбулла имеет вид:

$$F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha}, \quad x > 0, \alpha > 0, \beta > 0, \quad (2)$$

где α – параметр формы, β – масштабный параметр [4].

Плотность распределения Вейбулла имеет следующий вид (см. также рис. 2):

$$f(x) = \alpha \beta^{-\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha} \quad (3)$$

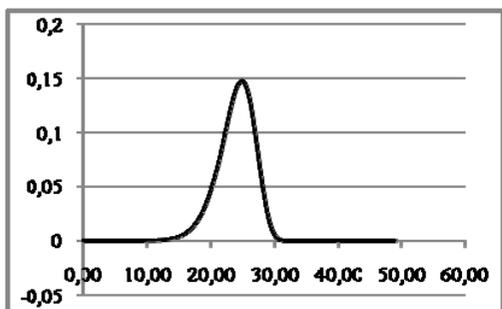


Рис. 2. Функция плотности распределения Вейбулла при $\alpha = 10$ и $\beta = 25$

Известно [5], что моменты m_k случайной величины x распределения Вейбулла имеют вид:

$$m_k = \beta^\alpha \Gamma\left(1 + \frac{k}{\alpha}\right), \quad (4)$$

где $\Gamma(z)$ – гамма функция. Для полного описания $W(x)$ найдем шесть первых моментов из выражения (4).

Вычисляем кумулянты χ_k через формулы [1]:

$$\begin{aligned} \chi_1 &= m_1; \quad \chi_2 = m_2 - m_1^2; \\ \chi_3 &= m_3 - 3m_1m_2 + 2m_1^3; \\ \chi_4 &= m_4 - 3m_2^2 - 4m_1m_3 + 12m_1^2m_2 - 6m_1^4; \\ \chi_5 &= m_5 - 5m_1m_4 - 10m_2m_3 + 20m_1^2m_3 + \\ &+ 30m_1m_2^2 - 60m_1^3m_2 + 24m_1^5; \\ \chi_6 &= m_6 - 6m_1m_5 - 15m_2m_4 + 30m_1^2m_4 - \\ &- 10m_3^2 + 120m_1m_3 - 120m_1^3m_3 + 30m_2^3 - \\ &- 270m_1^2m_2^2 + 360m_1^4m_2 - 120m_1^6. \end{aligned} \quad (5)$$

В соответствии с (5) получим численные значения кумулянтов, представленные в таблице 1. Исходя из вышесказанного, можем разложить функцию в ряд Эджворта (1).

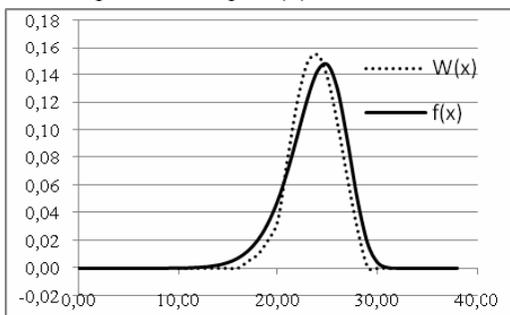


Рис. 3. Сравнение двух плотностей распределения Вейбулла

Из выражения (1) непосредственно видна особая ценность кумулянтов при оценке отклонения плотности вероятности от гауссовского распределения. Для выбранных значений α и β с учетом (4)-(5) можно получить аппроксимацию распределения (3) в виде, представленном на рис. 3.

При построении (см. рис. 3) учтено, что аппроксимирующее выражение для плотности должно удовлетворять условию нормировки:

$$\int_0^\infty W(x) dx = 1. \quad (6)$$

Из (6) следует выражение нормирующей константы:

$$const = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \left(\frac{\chi_5}{40} - \frac{\chi_3}{6} \right). \quad (7)$$

Аналогично вышесказанному проводятся вычисления для распределения Парето и логнормального распределения.

2. Плотность распределения Парето имеет следующий вид (см. также рис. 4):

$$f(x) = \frac{\alpha}{\beta} \left(\frac{\beta}{x}\right)^{\alpha+1}, \quad x > \beta, \quad \alpha > 0, \quad \beta > 0, \quad (8)$$

где α – параметр формы, β – масштабный параметр [4].

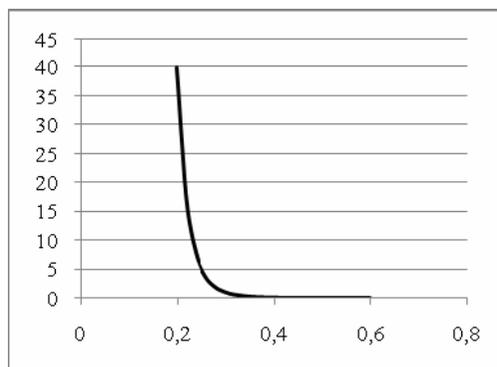


Рис. 4. Функция плотности распределения Парето при $\alpha = 8$ и $\beta = 0,2$

Моменты m_k случайной величины x распределения Парето имеют вид [5]:

$$m_k = \frac{\alpha \beta^k}{\alpha - k}, \quad k < \alpha \quad (9)$$

Численные значения кумулянтов с учетом (9) представлены в таблице 1. Так как аппроксимирующее выражение остается неизменным (1), неизменным остается и выражение для нормирующей константы (7).

Результат аппроксимации распределения Парето представлен на рис. 5.

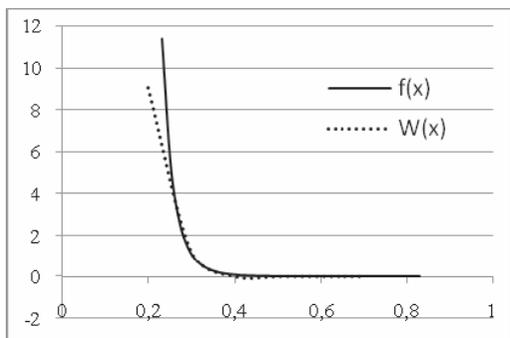


Рис. 5. Сравнение двух плотностей распределения Парето

3. Плотность логнормального распределения имеет следующий вид (см. также рис. 6):

$$f(x) = \frac{e^{-0,5\left(\frac{\ln x - \mu}{\sigma}\right)^2}}{\sqrt{2\pi\sigma^2 x^2}}; \quad x > 0; -\infty \leq \mu \leq \infty; \sigma \geq 0; \quad (10)$$

где σ – параметр формы, μ – масштабный параметр [4].

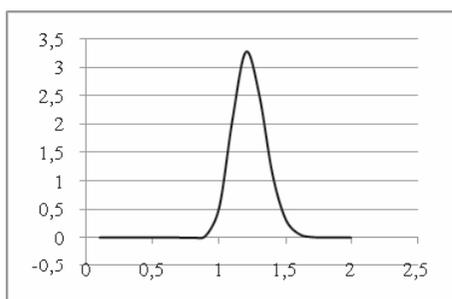


Рис. 6. Функция плотности логнормального распределения при $\mu = 0,2$ и $\sigma = 0,1$

Моменты m_k случайной величины x логнормального распределения имеют вид [5]:

$$m_k = e^{k\mu + 0,5(k\sigma)^2}. \quad (11)$$

Численные значения кумулянтов с учетом (11) представлены в таблице 1.

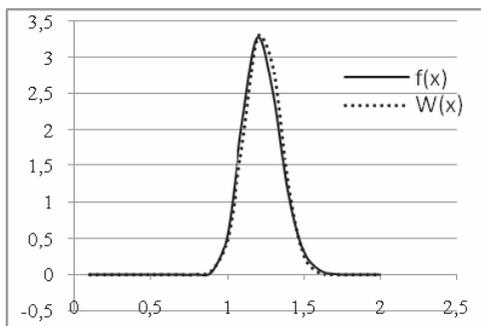


Рис. 7. Сравнение двух плотностей логнормального распределения

Результат аппроксимации логнормального распределения представлен на рис. 7. Рассмотрим случай, когда на вход звена случайно и непрерывно с определенным объемом информации поступает поток пакетов. Неизвестны величины длительности поступления τ и длительности обслуживания пакетов ξ (см. рис. 8).



Рис. 8. Параметры звена мультисервисной сети

Разработка моделей, учитывающих реальный характер потоков сообщений, и детали обслуживания мультисервисного трафика разных коммуникационных приложений является приоритетным вопросом при анализе и прогнозировании трафика.

Получив с помощью программы Fractan, статистические данные реального потока, поступающие на вход сетевого элемента, обрабатывающего самоподобный трафик, и представленные параметрами τ_i и ξ_i , можно аппроксимировать функцию плотности вероятности $W(\tau)$.

Для примера возьмем выборку величин длительности поступления τ_i , где $i = 1 \dots 3176$.

Согласно алгоритму (см. рис. 1) находим моменты произвольного распределения, поступающего на вход системы обслуживания. Моменты произвольной длительности имеют вид:

$$m_k = \frac{1}{N} \sum_{i=1}^N \tau_i^k. \quad (12)$$

Численные значения кумулянтов с учетом (12) представлены в таблице 1.

Таблица 1. Значения кумулянтов χ_k

| Кумулянты χ_k | Численные значения кумулянтов | | | | | |
|---------------------------------|-------------------------------|----------|----------|----------|----------|----------|
| | χ_1 | χ_2 | χ_3 | χ_4 | χ_5 | χ_6 |
| Распределение Вейбулла | 1,3127 | 0,0248 | -0,0023 | 0,00010 | 12,488 | -277,725 |
| Распределение Парето | 0,2286 | 0,0011 | 0,0001 | 0,00002 | -0,2576 | 0,332 |
| Логнормальное распределение | 1,2275 | 0,0151 | 0,0006 | 0,00004 | 7,2631 | -146,545 |
| Распределение реального трафика | 0,1841 | 0,2017 | 0,2346 | 0,23042 | 0,4178 | 5,424 |

Исходя из вышесказанного, можем разложить функцию в ряд Эджворта (1). Рассчитанная плотность вероятности распределения $W(\tau)$ реального трафика представлена на рис. 9.

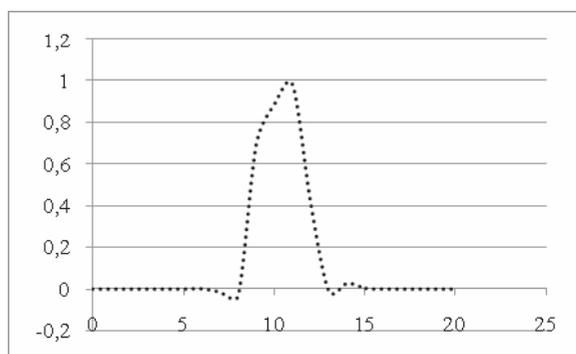


Рис. 9. Искомая функция $W(\tau)$, выраженная через ряд Эджворта

Таким образом, полученные аппроксимирующие выражения для функций плотностей распределений с «тяжелыми» хвостами с помощью кумулянтов позволяют сравнить их с теоретическими данными, а для реального трафика получить представление плотности распределения на выходе обрабатывающего устройства.

Заключение

На данный момент анализ систем типа G/G/1 не решен, с точки зрения того, чтобы можно было описать эту систему с помощью любого математического аппарата либо любого другого программно-аппаратного комплекса для решения данной задачи.

Аппроксимация функций распределения плотности с помощью кумулянтного анализа мо-

жет быть использована для прогнозирования поведения потоков трафика в сети.

В основном для анализа характеристик реального трафика используются «зашитые» программы и блоки, которые подходят к исследованию узких задач. Однако для массового решения проблемы, которая бы интегрировалась с разными приложениями и при появлении новизны не теряла своей актуальности, эти методы теряют свою полезность.

В связи с постоянным развитием и обновлением мультисервисных сетей на первое место выходит исследование и аппроксимация функции плотности распределения, которое реализовано в представленном алгоритме.

Литература

1. Малахов А.Н. Кумулянтный анализ случайных негауссовских процессов и их преобразований. М., Сов. радио, 1978. – 376 с.
2. Гарипова Л.Р., Киреева Н.В. Кумулянтный подход к исследованию системы G/G/1// XIV МК и В «Цифровая обработка сигналов и ее применение». Москва, 2012. – С. 120.
3. Шелухин О.И., Тенякшев А.М., Осин А.В. Фрактальные процессы в телекоммуникациях. М.: Радиотехника, 2003. – 480с.
4. Лоу А.М., Кельтон В.Д. Имитационное моделирование. Классика CS. М.: 2004. – 848 с.
5. Королюк В.С., Портенко Н.И., Скороход А.В. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985. – 640 с.

CREATION OF FUNCTIONS OF DISTRIBUTION OF THE REAL TRAFFIC BY MEANS OF THE CUMULANT ANALYSIS

Chupakhina L.R., Kireeva N.V.

This article discusses the use of cumulant analysis to approximate the density functions in the chain of multi-service network. And can be displayed via the cumulants can obtain the density distribution at the output processing unit.

Keywords: queuing system, the traffic, the moments of functions, cumulants, distribution with «heavy» tails.

Чупахина Лилия Равилевна, аспирант Кафедры мультисервисных сетей и информационной безопасности (МСИБ) Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). Тел. 8-927-710-67-17. E-mail: garip4ik555@mail.ru

Киреева Наталья Валерьевна, к.т.н., доцент Кафедры МСИБ ПГУТИ. Тел. (8-846) 339-11-67; 8-992-77-74. E-mail: zeppelinSN@yandex.ru