

ИССЛЕДОВАНИЕ ПОТОКОВ ЗАЯВОК, ПОСТУПАЮЩИХ НА ВЕБ - СЕРВЕР ОРГАНИЗАЦИИ

Тукубаев З.Б ¹, Умаров А.А. ²

*(Международный Казахско-Турецкий университет имени
Ясауи, г.Туркестан, Казахстан)*

В статье на основе экспериментальных измерений входящего потока требований в веб-сервер организации, установлен закон его распределения как пуассоновский.

Ключевые слова: входящий поток вызовов, статистическая аппроксимация, статистический ряд распределения

В современных цифровых сетях интегрального обслуживания в условиях перегрузки сети (в час пик) качество обслуживания заметно ухудшается; увеличивается вероятность отказов и среднее время обслуживания [1,2]. В таких случаях целесообразно управлять организацией сервера узла связи; оптимально выбирать кластерную структуру сервера, рассчитывая на максимальный трафик [3,4,5,6]. При этом, процесс оптимизации сервера основывается на модели входящего потока вызовов (заявок).

Поэтому разработка методов автоматизации обработки входных потоков сообщений, аппроксимации и построения модели входных потоков сервера является актуальной.

В настоящей статье для построения модели входящего потока вызовов в центральном сервере МКТУ им. Ясауи. Были получены аналогичные данные и в сервере компьютерной фирмы [9,10].

Измерение производилось в течение трех месяцев; объем измеренных данных превышал 1000 часов. В табл. 1 приведен фрагмент измерений моментов поступлений заявок в сервер входящего потока вызовов (в часах); интервалом анализа выбран 1 час, а на рисунке 1 и 2 приведены гистограммы интервалов

времени между поступающими заявками с интервалами 1 и 10 заявок.

Статистический ряд (статистика) распределения количества К заявок в час приведен в табл. 2. При этом, объем выборки превышает 1000 часов, а количество вызовов превышает 761.

Результаты обработки также приведены в табл. 2.

Этапы обработки измеренных значений потоков приведены ниже.

Таблица 1. Статистика моментов поступлений заявок

№ заявки	Моменты времени, (в час.)						
1	0,1469	26	4,3671	51	8,7047	76	13,6344
2	0,2624	27	4,5318	52	8,8769	77	13,8628
3	0,4040	28	4,7395	53	9,0110	78	13,9911
4	0,6673	29	4,9202	54	9,1680	79	14,1692
5	0,9246	30	5,0226	55	9,3642	80	14,2953
6	1,1489	31	5,1476	56	9,5450	81	14,5293
7	1,3402	32	5,2901	57	9,7966	82	14,6911
8	1,4612	33	5,5218	58	10,014	83	14,8331
9	1,6245	34	5,7499	59	10,194	84	14,9612
10	1,7434	35	5,9168	60	10,454	85	15,1084
11	1,8605	36	6,0872	61	10,662	86	15,2578
12	1,9737	37	6,2558	62	10,892	87	15,4839
13	2,2053	38	6,5236	63	11,075	88	15,6351
14	2,3851	39	6,6701	64	11,282	89	15,7653
15	2,5138	40	6,7832	65	11,434	90	15,9705
16	2,751511	41	6,980729	66	11,684	91	16,184
17	2,853625	42	7,15494	67	11,794	92	16,342
18	3,054369	43	7,379783	68	12,050	93	16,591
19	3,165383	44	7,548423	69	12,231	94	16,693
20	3,3043	45	7,7369	70	12,469	95	16,922
21	3,441945	46	7,910319	71	12,697	96	17,102
22	3,566849	47	8,020537	72	12,878	97	17,245

23	3,676935	48	8,219601	73	13,114	98	17,456
24	3,933401	49	8,389272	74	13,251	99	17,630
25	4,179725	50	8,5512	75	13,431	10 0	17,887

Рис. 1. и 2. Гистограмма интервалов времени между заявками потока.

Из табл.1 определяется статистический ряд распределения K заявок в интервале 1 час, при этом $K=4-15$ (с шагом 1) статистический ряд n распределения ровно K заявок в интервале 1 час. Крайние четыре разряда объединены по два; при этом $(4+5)/2=4,5$ и $(14+15)/2=14,5$ (см. 1-столбец табл.2).

Далее, 1 - столбце определяется значение выборочного среднего \bar{K} по формуле:

$$\bar{K} = \sum_i n_i \cdot k_i / \sum_i n_i. \quad (1)$$

Таблица 1. Статистическая аппроксимация экспериментальных измерений потоков требований

k_i	λ^{k_i}	$P_n(K)$	$(n_i - n'_i)^2$	k'_i	n'_i
4,5	7417,938315	24	24,73541	4,9970	0,202
6	68842,08607	120	45,91134	15,2985	0,333
7	638888,1402	720	70,36006	0,1296	0,0018
8	5929193,594	5040	93,28224	27,9021	0,299
9	55025808,84	40320	108,2131	0,0454	0,4*E-3
10	510666347,9	362880	111,5856	2,5140	0,0225
11	4739232814	3628800	103,5569	20,7650	0,2005
12	43982392335	39916800	87,36893	11,3497	0,1299
13	4,08178E+11	479001600	67,56886	6,5990	0,0976
14,5	3,78809E+12	6227020800	48,2363	588,727	12,205
	3,51553E+13		760,8187		13,492

При этом, интенсивность потока $\lambda = \bar{K}$. В рассматриваемом примере $\lambda = \bar{K} = 9,28$. Далее, в 3-м столбце применяя формулу (1) получается общее количество обработанных заявок 761.

Далее, в 4-м столбце вычисляется значение $e^{-\lambda}$, которое остается постоянным в столбце. В 5-м столбце вычисляется λ^{k_i} соответствующих значений k_i . Далее, в 6-м столбце вычисляется значение $P_n(K)$ по закону Пуассона по формуле

$$P_n(K) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{при различных } \bar{K} = 4,5 \div 14,5.$$

Сумма содержимого 9-го столбца (сумма частных вероятностей $P_n(K)$) должна быть равна 1. В нашем случае

$$\sum_{i=1}^{10} P_n(K) = 1,008.$$

Далее, перемножив содержимое столбцов 2 (n_i) и 6 ($P_n(K)$) заполняется 7-й столбец значениями n_i' , то есть $n_i' = n_i \cdot P_n(k)$. В примере сумма этого столбца равна 760,8. Суммы столбцов 3 и 7 должны быть одинаковыми.

В 8-й столбец записываются разности столбцов 2 и 7, а 9-й столбец заполняется значениями $(n_i - n_i')^2$.

В 10-й столбец записываются значения $(n_i - n_i')^2 / n_i'$, сумма 9-го столбца дает величину χ^2 - квадрат. В рассматриваемом примере сумма 10-го столбца равна 13,49.

Далее по табл. χ^2 – распределения [9] для степеней свободы $\nu = 8$ и для уровня значимости $\alpha = 0,05$ находится критическое значение $\chi^2_{кр} = 15,5$, которое больше чем $\chi^2_{набл}$, что дает основание не опровергать гипотезу о пуассоновости входящих потоков заявок в узел связи.

Приведенную методику можно использовать для анализа потоков вызовов в узлы связи, сервера и порталы.

Методику также можно использовать в учебном процессе по курсам “Вычислительные системы, сети и телекоммуникации”, “Компьютерные сети” и др.

Литература

1. ЛИФШИЦ А. Л., МАЛЫЦ Э. А. *Статистическое моделирование систем массового обслуживания*// -М.: Сов. Радио, 1978, -248 с.
2. КЛЕЙНРОК Л. *Вычислительные системы с очередями*// -М., Мир. 1979.

3. БРОЙДО В. Л. *Вычислительные системы, сети и телекоммуникации*//2- изд. М.,С.-Пб.,Н.-Новгород-Воронеж. 2004.-703 с.
4. КУЛЬГИН М. *Компьютерные сет*// Практика построения для профессионалов. 2-изд.-М., С.-Пб., 2003.-462 с.
5. Н. РЕНДАЛЛ. *Кластеризация серверов*// PC Magazine, №2. 1998.
6. КУЗНЕЦОВ С. Д. *Основы современных баз данных*// Информационно-аналитические материалы. www.citforum.ru
7. ТУКУБАЕВ З. Б. и др. *Моделирование и исследование алгоритмов распределения нестационарных приоритетных потоков в сети СМО*// Сб.“Алгоритмы”АН РУ, Ташкент, Вып. 60, 1986.
8. ТУКУБАЕВ З. Б. и др. *Модель эффективной загрузки узлов сет*// Тезисы докладов XI – Всесоюзного совещания по проблемам управления, г.Ташкент, 1989.
9. ТУКУБАЕВ З. Б. *Обобщенные модели в системах массового обслуживания*// “Вестник МКТУ”, №2, г.Туркестан, 2006.
10. УМАРОВ А. А. *Разработка веб узла фирмы “АББА*// V Всероссийская школа-семинар молодых ученых, “Управление большими системами”. Сб. трудов. Т.2, г.Липецк, ЛГТУ, 2008. с.151-157 .
11. ГМУРМАН В. Е. *Руководство к решению задач по теории вероятности и математической статистики*// -М.:Высшая школа, 1975.

RESEARCH METHOD OF TRANSACTIONS STREAMS ON THE CORPORATION'S WEB-SERVER.

Zukirkhan Beisekovich Tukubayev, International Kazakh-Turkish University named H.A.Yassavi Turkistan city. Kazakhstan (zuhr@pochta.ru).

Amantur Amangekdievich Umarov, International Kazakh-Turkish University named H.A.Yassavi Turkistan city. Kazakhstan (Unix77@yandex.ru).

Abstract: The article deals with the research method of transactions streams on the corporation's web-server.

Keywords: incoming stream of origination, statistic approximation, statistic row distribution.