Аналитико-численное моделирование работы распределенных информационных систем с самоподобным входным трафиком

А. Н. Скоба[1], В. К. Михайлов[1], С.А. Назаров[2], Айеш Ахмед Нафеа Айеш[1]

[1]Южно-Российский государственный политехнический университет (НПИ) им. М. И. Платова, Новочеркасск

[2]Каменский институт (филиал) Южно-Российского государственного политехнического университета (НПИ) им. М. И. Платова

Аннотация: в данной статье сформулированы основные этапы процесса конструирования аналитико-численных моделей функционирования распределенных информационных систем (РИС) с неэкспоненциальным (самоподобным) входным трафиком заявок. Каждый отдельный этап моделирования, на концептуальном уровне, сведен к решению определенного класса математических задач. В основе математического аппарата получения основных интегральных характеристик качества функционирования РИС лежит метод анализа средних значений сетей очередей. Приведены базовые соотношения для вычисления реактивности работы РИС.

Ключевые слова: распределенная информационная система, дискретный самоподобный процесс, параметр Хёрста, распределение Парето, сеть массового обслуживания, классификация Кендалла, формула Полячека-Хинчина, среднее время реакции системы.

Разработанные ранее авторами математические модели функционирования распрелделенных информационных систем (РИС) на базе файл-серверной, двухуровневой и трехуровневой клиент-серверной архитектур [1-3] основывались на предположении об экспоненциальном (пуассоновском) законе плотности вероятности распределения заявок на входе системы. В соответствии с этим, все аналитические выражения для нахождения интегральных характеристик качества функционирования РИС были получены для входных потоков, обладающих свойствами марковских случайных процессов [4].

Однако, проведенные в последние годы многочисленные исследования данной проблемы показали наличие во входных потоках реальных клиент-серверных информационных систем свойств самоподобия, которые были названы фрактальными процессами [5].

В случае самоподобного трафика входного потока, в силу его непредсказуемости, трудно получить аналитические выражения интегральных характеристик показателей качества функционирования РИС. Кроме этого, как показано в работах [6,7], наличие эффектов самоподобия во входном трафике заявок в системах передачи данных, как правило, ухудшает качество обслуживания (характеристики системы) по сравнению с тем, которое наблюдалось бы, например, в случае потока с экспоненциальным распределением заявок на входе. Все эти факторы указывают на важность решения задач, связанных с разработкой математических моделей функционирования РИС с неэкспоненциальными функциями плотности распределения длительности поступления и обслуживания заявок.

Математически дискретный самоподобный процесс можно определить следующим образом [8]: пусть задан случайный процесс $X = \{X_i = X(t_i)\}, i = 1,2,...$. Определим другой случайный процесс $X^{(m)} = \{X^{(m)}{}_k\}, k = 1,2,...$, который назовем агрегированным, полученным путем усреднения m значений исходного процесса на непересекающихся временных интервалах:

$$X_k^{(m)} = \frac{1}{m}\sum_{i=1}^{m} X_{(k+1)m+i}$$

Самоподобный случайный процесс характеризуется следующими основными свойствами:

1) убыванием дисперсии $D\{X^{(m)}\} = am^{-\beta}, 0 < \beta < 1$, при $m \to \infty$, т.е. дисперсия агрегированных процессов уменьшается медленнее, чем величина обратная размеру выборки $(1/m)$;

2) долгосрочной зависимостью, признаком которой является расходимость автокорреляционной функции:

$$\sum_{k=1}^{m} r(k, X^{(m)}) = \infty, r(k, X^{(m)}) = k^{-\beta},$$

т.е. функция автокорреляции не выраждается при $m \to +\infty$, в отличии от стохастических случайных процессов, для которых $r(k, X^{(m)}) \to 0$ при $m \to +\infty$;

3) флуктуационным характером спектра мощности: $f(\omega) = c\omega^{-\gamma}, \omega \to 0, \gamma = 1-\beta,$ где $f(\omega) = \sum_{k=1}^{m} r(k, X^{(m)}) e^{-j\omega k}$ – дисперсное преобразование Фурье автокорреляционной функции.

Величину H характеризующую «степень самоподобия» случайного процесса называют параметром Хёрста, при этом $0{,}5 < H < 1{,}0$. Многочисленные измерения показывают наличие существенно самоподобных свойств входного трафика в клиент-серверных информационных системах реализованных на базе различных архитектур – от «классических» двухуровневых, трехуровневых, до многоуровневых с *Web*-доступом с параметром Хёрста $H \approx 0{,}65 \div 0{,}8$ и $\beta = 2(1-H)$ [8]. А кроме этого для практического описания входного трафика современных информационных систем, реализованных на базе клиент-серверных сетей, обладающего свойствами самоподобного потока, используются распределения Парето, Вейбулла, логнормальное распределения, а так же модели типа *fMB* – фрактальные броуновские движения и *fGN* – фрактальный гауссовский шум [8].

Таким образом, процесс моделирования функционирования современных РИС для получения их характеристик *QoS* (*Quality of Service*) можно представить в виде следующей последовательности этапов:

1) моделирования самоподобного случайного процесса с заданной плотностью распределения вероятности случайной величины, описывающего входной трафик заявок на входе системы;

2) аналитическое представление (аппроксимацию) полученного самоподобного процесса;

3) применение математического аппарата: СМО, сетей массового обслуживания (СеМО), тензорного анализа, сетей Петри, для конструирования моделей функционирования РИС с целью получения их основных интегральных характеристик;

4) решения оптимизационных задач на разработанных математических моделях РИС с целью выбора оптимальной топологической структуры их конструирования, оптимизации распределения информационных ресурсов (баз данных) по узлам системы по различным критериям (минимума среднего времени реакции на запросы пользователей, на группу запросов пользователей, функциональной надежности передаваемой информации и т.д.);

5) программную реализацию полученных математических моделей и алгоритмов оптимизации;

6) выбор рекомендаций по наилучшей организации вычислительного процесса в РИС.

В настоящее время существуют два основных подхода к моделированию самоподобных процессов:

1) метод Мандельброта, предусматривающий использование нескольких независимых *ON-OFF* источников, у которых закон чередования включенного и выключенного состояний распределен по закону Парето [8];

2) методы на основе Броуновского движения [8].

Рассмотрим более подробно основные технические моменты, связанные с реализацией первого подхода.

Плотность распределения Парето задается функцией распределения вида:

$$W(a) = \frac{\alpha k^{\alpha}}{x^{\alpha+1}},\ \alpha > 0,\ x \geq k\ ,$$

где, α – параметр формы; k – параметр, определяющий нижнюю границу случайной величины. Путем несложных преобразований можно получить выражение для моделируемой случайной величины: $x = \dfrac{k}{\sqrt[\alpha]{1-y}}$, где y – случайное число с равномерной плотностью распределения вероятности на $(0;1)$, $\alpha = 3 - 2H$.

Следующим этапом является аппроксимация полученных численных значений плотности распределения вероятности случайной величины x аналитическим выражением. Основными подходами к решению данной задачи являются: аппроксимация суммой затухающих экспонент; метод Прони; коммулятивный анализ [9]. Так, для случая, когда значения функции плотности распределения вероятности случайной величины задаются в численной форме наиболее подходящим методом является метод аппроксимации суммой затухающих экспонент [9]. (Метод Прони является частным случаем предыдущего метода, когда аппроксимация осуществляется линейной комбинацией экспоненциальных функций с положительными и отрицательными значениями пользователей экспонент). Таким образом, наиболее простым решением данной задачи является аппроксимация полученных значений случайной величины выражением вида: $f(x) = \sum_{k=1}^{n} a_k e^{-\alpha_k x}$,где $\alpha_k = 1,2,...$. Выбор α_k обусловлен требуемой погрешностью аппроксимации. Для вычисления неизвестных значений a_k можно использовать, например, метод наименьших квадратов, сводящий задачу нахождения коэффициентов a_k к решению системы линейных алгебраических уравнений.

Следующим этапом является конструирование математической модели функционирования РИС. В настоящее время, при конструировании концептуальных и математических моделей такого рода систем широко используется аппарат теории СМО и ее подраздела сетей массового обслуживания (СеМО). Моделями, наиболее адекватно описывающими процесс функционирования РИС с самоподобным входным трафиком заявок, является СМО, в условных обозначениях в соответствии с классификацией Кендалла [4] вида: G/M/m, fMB/M/m, fGN/M/m, Г/M/m (Г – гамма распределения), W/M/m (W- распределение Вейбулла). Однако, для большинства представленных моделей СМО расчет характеристик QoS, например, по формуле Полачека-Хинчина [4] затруднен. Причиной этому является слабая формализованность модели самоподобных потоков, в следствие чего и невозможно получить аналитически обоснованные результаты характеристик СМО, в которых и обслуживаются данные потоки. В частности, из-за отсутствия строгой теоретической базы, способной дополнить классическую теорию систем массового обслуживания при проектировании СМО с самоподобным трафиком, на сегодняшний день не существует достоверной методики расчета параметров и показателей качества РИС, в условиях эффекта самоподобия, кроме как аппарата имитационного моделирования. При этом, также, при разработке математических моделей функционирования РИС, в большинстве случаев остается открытым вопрос определения матриц переходных вероятностей, описывающих процесс функционирования заявок по узлам сети; не разработаны методики перехода от реальных параметров сети к таким параметрам концептуальных моделей как, например, интенсивности обработки информации в ее узлах при различных режимах её работы.

В работе [10] авторами была представлена разработанная модель функционирования РИС с произвольной, неэкспоненциалной, длительностью обслуживания заявок в узлах системы с использованием метода декомпозиционной аппроксимации. В работе приведены базовые соотношения для вычисления среднего времени реакции системы на запросы пользователей. Приведенные численные эксперименты с разработанной моделью показали ее достойно высокую эффективность при сравнительно небольшой загрузке обслуживающих центров, что характерно для РИС, работающих в реальном масштабе времени.

В данной статье, в рамках конструирования математической модели функционирования РИС с самоподобным входным трафиком заявок предлагается подход на основе итерационного метода анализа средних значений [11], который допускает решения для расчета СеМО большой размерности с произвольным распределением обслуживания заявок в центрах. Так, согласно [11], для СеМО с произвольным (неэкспоненциальным) потоком заявок и экспоненциальной длительностью обслуживания в центрах сети, выражение для среднего времени пребывания сообщения класса $r(r=\overline{1,R})$ в центре $i(i=\overline{1,M})$ будет иметь вид:

$$T_{ir}(\overline{N}_R)=\tau_{ir}+\sum_{j=1}^{R}\tau_{ij}L_{ij}(\overline{N}_R-1_r),$$

где $\overline{N}_R=(1,...,1)$; τ_{ij} – средняя длительность обслуживания сообщений класса $j(j=\overline{1,R})$ в i-м центре $(i=\overline{1,M})$; $L_{ij}(\overline{N}_R)$ – среднее число сообщений j-го класса в i-м центре; $\bar{1}_r=(0,...,0,1,0,...,0)$ – вектор, r-ая координата которого равна 1, а все остальные равны 0; $\tau_{ij}=1/\mu_{ij}$, где μ_{ij} – интенсивность обслуживания j-го сообщения в i-м центре. При этом

$$L_{ij}(\overline{N}_R-\bar{1}_r)=\sum_{r=1}^{N_R}\sum_{\bar{n}_R=\bar{1}_j}^{\overline{N}_R}P_i(\bar{n}_j,\overline{N}_R-\bar{1}_r)\,n_j,$$

согласно [1,2]:

$$P_i(\bar{n}_j,\overline{N}_R-\bar{1}_r)=\frac{Z_i(\overline{n_R})}{G(\overline{N}_R-\bar{1}_r)}[G(\overline{N}_R-\overline{n_R}-\bar{1}_r)-\sum_{j=1}^{R}\sum_{i=1}^{R}x_{ij}G(\overline{N}_R-\bar{1}_j-\bar{1}_r-\bar{n}_j)],$$ где $x_{ij}=\dfrac{e_{ij}}{\mu_{ij}}$,

$e_{ij}=\sum_{S=1}^{M}e_{Sj}P_{Si}(i),\ i=\overline{1,M},\ j=\overline{1,R}$; $G(\overline{N}_R)$ – нормализующая константа для вычисления которой может быть использован рекуррентный метод Бузена

[12]. Методики вычисления интенсивностей обслуживания в узлах сети $\mu_{ij}(i=\overline{1,M},\ j=\overline{1,R})$ и матриц переходных вероятностей $\|P_{si}(j)\|\ (S=\overline{1,M}, i=\overline{1,M},\ j=\overline{1,R})$ для различных архитектур построения РИС изложено в работах [1-3].

Расчет величины $\overline{T}$ – среднюю времени реакции системы на запросы пользователей может быть выполнен по формуле $\overline{T}=(1\Big/\sum\limits_{i=1}^{M}\lambda_i)\sum\limits_{i=1}^{M}\lambda_i\overline{T}_i$, где

$\overline{T}_i=\sum\limits_{r=1}^{R}T_{ir}(\overline{N}_R)\cdot r.$

Дальнейшие этапы моделирования работы РИС могут быть выполнены с помощью разработанного оптимизационного алгоритма [13] и программного продукта [14].

Литература

1. Скоба А.Н., Состина Е.В. Математическая модель оптимального размещения распределённой базы данных по узлам ЛВС на базе файл-серверной архитектуры // Инженерный вестник Дона. 2015. №2. URL: ivdon.ru/ru/magazine/archive/n2y2015/2881.

2. Скоба А.Н., Состина Е.В. Математическая модель оптимального размещения распределённой базы данных по узлам ЛВС на базе двухуровневой клиент-серверной архитектуры // Инженерный вестник Дона, 2015, №2. URL:ivdon.ru/ru/magazine/archive/n2y2015/2882.

3. Скоба А.Н., Михайлов В.К., Айеш Ахмед Нафеа Айеш. Модель оптимального размещения информационных ресурсов по узлам распределённой системы обработки информации предприятия на базе трёхуровневой архитектуры «клиент-сервер» с учётом влияния блокировок. // Изв. вузов. Электромеханика, 2018. Т. 61. №3. С. 68-75.

4. Клейнрок Л. Вычислительные системы с очередями. // Перевод с англ. под ред. Б.С. Цыбакова. – М.: Мир. – 1979. – 600с.

5. Соколов Д. Е., Треногин Н.Г. Характер сетевого трафика на клиентском участке распределенной клиент-серверной системы. // Материалы Международной научно-практической конференции «Информатика и проблемы телекоммуникации». – Новосибирск: СибГУТИ, 2001, с.34-36.

6. Нейман В.И. Самоподобные процессы и их применение в теории телетрафика. // Труды Международной академии связи. – 1999. – №3 – С.11-15.

7. Bo Ryu, S.B. Lowen. Point Process Models for Self-Similar Network Traffic, with Applications. // Stochastic Models, – 1998. – pp. 735-761.

8. Наместников С.М., Служивый М.Н., Украинцев Ю.Д. Основы теории телетрафика: учебное пособие. // Ульяновск, УлГТУ. – 2016. – 154с.

9. Киреева Н.В., Чупахина Л.Р. Сравнение возможностей использования различных методов аппроксимации для анализа трафика с самоподобным распределением. // Международный журнал прикладных и фундаментальных исследований. – 2016. – № 12 (часть 7) – С. 1287-1289.

10. Скоба А.Н., Михайлов В.К., Назаров С.А., Скорик Н.С. Математическая модель оценки реактивности распределенной системы обработки информации с использованием метода декомпозиционной аппроксимации. // Инженерный вестник Дона, 2018, №4. URL:ivdon.ru/ru/magazine/archive/n4y2018/5291.

11. Вишневский В.М. Теоретические основы проектирования компьютерных сетей. – М.: Техносфера, 2003. – 512с.

12. Buzen J.P. Computational Algorithms for Closed Queueing Networks with Exponential Servers. Commun. ACM. 1983. Vol.16, №9. – pp. 527-531.

13. Скоба А.Н., Айеш Ахмед Нафеа Айеш. Эвристический алгоритм решения задачи оптимального размещения информационных ресурсов. // Инженерный вестник Дона, 2018, №1. URL:ivdon.ru/ru/magazine/archive/n1y2018/4744.

14. Михайлов В.К., Скоба А.Н., Айеш Ахмед Нафеа Айеш. Программный комплекс для решения задач оптимального размещения информационных ресурсов и моделирования влияния основных интегральных показателей на реактивность распределенных систем обработки информации. // Изв. вузов. Электромеханика. 2018. Т. 61. №6. С.102-109.

References

1. Skoba A.N., Sostina E.V. Inženernyj vestnik Dona (Rus), 2015. №2. URL:ivdon.ru/ru/ magazine/archive/n2y2015/2881.

2. Skoba A.N., Sostina E.V. Inženernyj vestnik Dona (Rus), 2015. №2. URL:ivdon.ru/ru/magazine/archive/n2y2015/2882.

3. Skoba A.N., Mikhaylov V.K., Ayesh Achmed Nafea Ayesh. Izv. vuzov. E`lektromexanika, 2018. T. 61. №3. pp. 68-75.

4. Klejnrok L. Vy`chislitel`ny`e sistemy` s ocheredyami. [Computing systems with queues]. M.: Mir, 1979. 600p.

5. Sokolov D. E., Trenogin N.G. Xarakter setevogo trafika na klientskom uchastke raspredelennoj klient-servernoj sistemy`. [The nature of network traffic on the client side of a distributed client-server system]. Novosibirsk: SibGUTI, 2001, pp. 34-36.

6. Nejman V.I. Samopodobny`e processy` i ix primenenie v teorii teletrafika. [Self-similar processes and their application in the theory of teletraffic]. 1999. №3. pp. 11-15.

7. Bo Ryu, S.B. Lowen. Stochastic Models. 1998. pp. 735-761.

8. Namestnikov S.M., Sluzhivy`j M.N., Ukraincev Yu.D. Osnovy` teorii teletrafika: uchebnoe posobie. [Fundamentals of Teletraffic Theory: Tutorial]. Ul`yanovsk, UlGTU. 2016. 154p.

9. Kireeva N.V., Chupaxina L.R. Mezhdunarodnyj zhurnal prikladnyh i fundamental'nyh issledovanij, 2016. № 12. pp. 1287-1289.

10. Skoba A.N., Mikhaylov V.K., Nazarov S.A., Skoryk N.S. Inženernyj vestnik Dona (Rus), 2018. №4. URL:ivdon.ru/ru/magazine/archive/n4y2018/5291.

11. Vishnevskij V.M. Teoreticheskie osnovy` proektirovaniya komp`yuterny`x setej [Theoretical foundations of computer network design]. M.: Texnosfera, 2003. 512p.

12. Buzen J.P. Computational Algorithms for Closed Queueing Networks with Exponential Servers. Commun. ACM. 1983. Vol.16, №9. pp. 527-531.

13. Skoba A.N., Ayesh Achmed Nafea Ayesh. Inženernyj vestnik Dona (Rus), 2018. №1. URL:ivdon.ru/ru/magazine/archive/n1y2018/4744.

14. Mikhaylov V.K., Skoba A.N., Ayesh Achmed Nafea Ayesh. Izv. vuzov. E`lektromexanika, 2018. T. 61. №6. pp. 102-109.