

Применение аппарата сетей массового обслуживания для аналитико-численного моделирования работы информационной системы без учета влияния блокировок

А.Н. Скоба, Е.В. Состина

Южно-Российский государственный политехнический университет (НПИ)

им. М. И. Платова, Новочеркасск

Аннотация: В данной статье с использованием аппарата замкнутых экспоненциальных сетей массового обслуживания (СеМО) разработана математическая модель функционирования корпоративной информационной системы в режиме информационного обслуживания без учета влияния блокировок. Представлены аналитические выражения для вычисления интегральных характеристик системы: закона распределения количества запросов пользователей, среднего времени пребывания запросов пользователей в узлах системы, её быстродействия.

Ключевые слова: База данных, блокировка, сеть массового обслуживания, дисциплина обслуживания заявок, экспоненциальный закон распределения, концептуальная модель, пространство состояний сети массового обслуживания, уравнение глобального баланса, стационарная вероятность, матрица переходных вероятностей, закон распределения количества заявок, среднее время пребывания заявки в узле.

Особенность рассматриваемой информационной системы (ИС) состоит в том, что число используемых баз данных превышает число пользователей, что позволяет пренебречь влиянием блокировок. В качестве разделяемых ресурсов рассматривается процессор и объединенный ресурс «канал-внешняя память», все запросы пользователей полагаются однократными и однородными, а для определения последовательности их выполнения используется дисциплина обслуживания FCFS [1]. («Первый пришел – первый обслужен»). Однородность запросов предполагает совпадение их вероятностных характеристик, а их однократность состоит в том, что пользователь формирует очередной запрос только после получения ответного сообщения на предыдущий [2].

В качестве исходных данных предполагается, что длительность активного состояния пользователей распределена по экспоненциальному закону с плотностью $f(t)=\lambda e^{-\lambda t}$, время обработки заявок процессором – по

закону $\varphi(t) = \mu e^{-\mu t}$, время обслуживания заявок ресурсом «канал – внешняя память» - по закону $q(t) = \eta e^{-\eta t}$, где $\lambda^{-1}, \mu^{-1}, \eta^{-1}$ – средние значения соответственно длительности активного состояния, времени обслуживания заявок процессором и ресурсом «канал – внешняя память» [2].

На концептуальном уровне функционирование данной ИС можно представить в виде замкнутой однородной экспоненциальной сети массового обслуживания (СеМО) [3], состоящей из взаимодействующих систем массового обслуживания - $СМО_1$, $СМО_2$ и $СМО_3$. Концептуальная модель рассматриваемой ИС представлена на рис 1. При этом, $СМО_1$ - имитирует процесс работы m пользователей за терминалами. Каждый из пользователей может находиться в двух состояниях: активном – формирование запросов к базам данных и пассивном – ожидании ответа на сформированный запрос. Учитывая, что каждый из пользователей до получения ответа может сформировать только одну заявку, то общее количество заявок, циркулирующих в сети равно m и, следовательно, отсутствует очередь к $СМО_1$. Следующая $СМО_2$ имитирует работу СУБД и её программного обеспечения по обработке запросов, поступивших с терминалов m пользователей, и по формированию обращений к базам данных на считывание информации. Система включает прибор P_1 и накопитель M_1 емкостью $m-1$, предназначенные для моделирования работы процессора и буферной зоны памяти. Последняя $СМО_3$ имитирует доступ к данным и состоит из прибора P_2 и накопителя N_2 емкостного $m-1$, моделирующих работу ресурса “канал - внешняя память” и его входные очереди.

Для получения интегральных характеристик данной информационной системы целесообразно использовать математический аппарат СеМО [3,4].

Для идентификации состояний сети вводится векторное пространство состояний:

$$\bar{E} = \{E_i(n_1, n_2, n_3), i = \overline{1, \rho}\}, \text{ где}$$

$\{n_1, n_1 = \overline{0, m}\}$ – описывает количество активных пользователей системы (терминалов); $\{n_2, n_2 = \overline{0, m}\}, \{n_3, n_3 = \overline{0, m}\}$ – количество заявок находящихся в очереди и на обработке соответственно процессором и ресурсом “канал-внешняя память”.

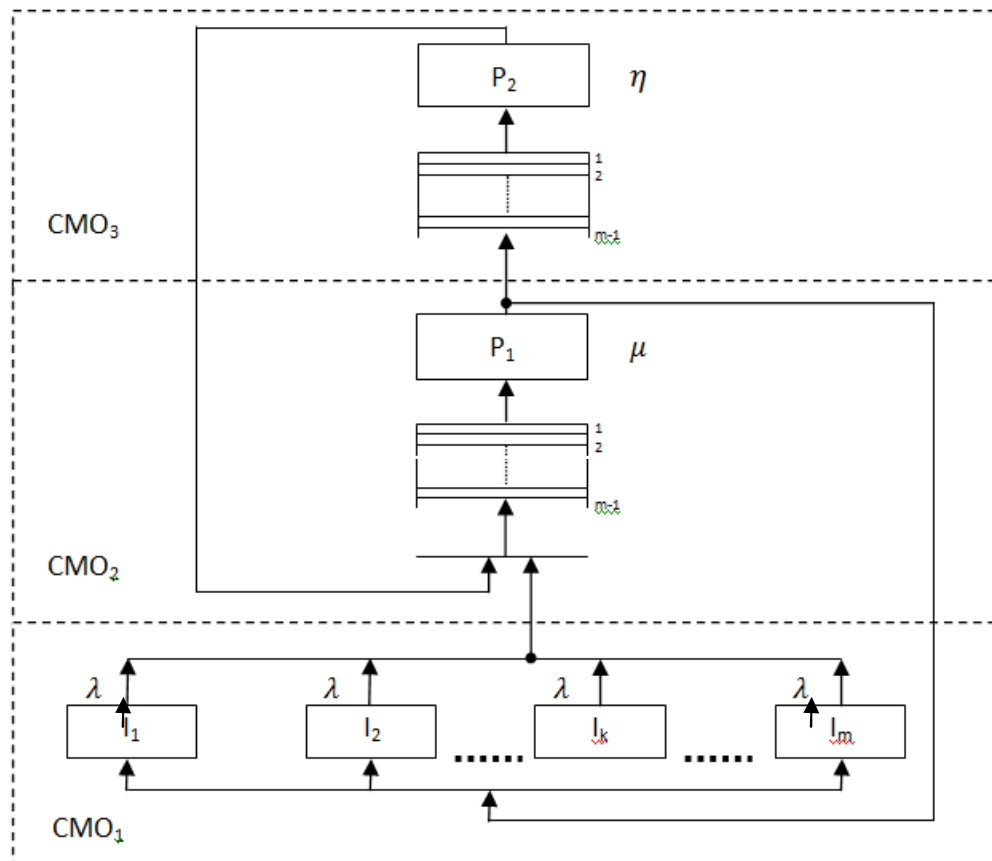


Рис.1. Концептуальная модель функционирования информационной системы

Учитывая, что общее количество заявок в сети конечно и равно m , то компоненты векторов состояний удовлетворяют ограничению:

$$n_1 + n_2 + n_3 = m.$$

Представляющие интерес характеристики СеМО определяются стационарными вероятностями состояний сети. Пусть $P(\bar{n})$ - стационарная вероятность того, что сеть находится в состоянии $\bar{n}=(n_1, n_2, n_3)$. Можно показать, что процесс изменения состояний сети описывается регулярным марковским процессом [1,5] и тогда уравнение глобального баланса относительно $P(\bar{n})$ для стационарного режима функционирования сети имеет вид [6,7]:

$$\sum_{k=1}^3 \mu_k P(\bar{n}) = \sum_{i=1}^3 \sum_{k=1}^3 P(\bar{n} - \bar{1}_k + \bar{1}_i) \mu_i P_{ik}, \quad (1)$$

где μ_k - интенсивность обслуживания заявок в k -ом узле (СМО _{k} , $k=\overline{1,3}$); P_{ik} - вероятность того, что заявка, после обслуживания в i -ом узле (СМО _{i} , $i=\overline{1,3}$), попадет в k -ый узел (СМО _{k} , $k=\overline{1,3}$); $\bar{1}_k$ - вектор, k - я координата которого равна 1 ($k=\overline{1,3}$), а все остальные координаты равны нулю.

Согласно [3,4] выражения для стационарных вероятностей состояний сети, описываемой уравнением (1), имеют мультипликативную форму и могут быть представлены в виде:

$$P(\bar{n}) = \frac{\prod_{i=1}^3 \left(\frac{e_i}{\mu_i} \right)^{n_i}}{\sum_{\bar{n} \in E(m,3)} \prod_{i=1}^3 \left(\frac{e_i}{\mu_i} \right)^{n_i}}, \quad (2)$$

где $E(m,3)$ - множество состояний сети, а количество этих состояний равно мощности множества $|E(m,3)|$ и представляет собой число различных распределений m заявок по 3 система (узлам) и равно числу состояний $|E(m,3)| = C_m^{m+3-1}$ [3]. В выражении (2) входят величины e_k ($k=\overline{1,3}$), которые находятся из решения системы линейных алгебраических решений:

$$e_k = \sum_{k=1}^3 e_i P_{ik}, \quad k=\overline{1,3}. \quad (3)$$

Число независимых уравнений в системе (3) на единицу меньше количества переменных, так что её решение единственно с точностью до мультипликативной константы. Для отыскания однозначного решения системы (3) достаточно произвольно задать значение e_i , например, положить $e_i = 1, (i = \overline{1,3})$. В этом случае величины e_i можно интерпретировать как среднее число посещений сообщением i -го узла между двумя последовательными посещениями им первого узла.

При расчете величин $\mu_i (i = \overline{1,3})$ можно положить $\mu_1 = \lambda, \mu_2 = \mu, \mu_3 = \eta$.

Элементы матрицы переходных вероятностей $\| P_{ik} \| (i, k = \overline{1,3})$ определим следующим образом:

$$P_{ik} = \begin{cases} 1, & \text{если } i=1, k=2 \text{ и } i=3, k=2; \\ \frac{1}{\alpha}, & \text{если } i=2, k=1; \\ \frac{\alpha-1}{\alpha}, & \text{если } i=2, k=3; \\ 0, & \text{иначе.} \end{cases}$$

Здесь α - среднее число обращений к базам данных при выполнении одного запроса, сформированного пользователем.

Для расчета величины $G(m,3) = \sum_{\bar{n} \in E(m,3)} \prod_{i=1}^3 \left(\frac{e_i}{\mu_i} \right)^{n_i}$ - нормализующей

константы, может быть использован рекуррентный метод Бузена [3-4,8].

Основные интегральные характеристики сети находим из следующих выражений [3,9,10] :

- закон распределения количества заявок в S -м узле:

$$P_{s,l} = \sum_{n_s=l} P(n_1, n_2, n_3), \quad s = \overline{1,3}, l = \overline{0,m},$$

где $P_{s,l}$ – вероятность того, что в s -м узле находится точно l – заявок пользователей, а заявки среди других узлов распределяются любыми возможными сочетаниями;

- среднее количество заявок находящихся в S -м узле :

$$\bar{N}_s = \sum_{l=0}^m P_{sl} \cdot l, \quad s = \overline{1,3};$$

- среднее время пребывания заявки в S -м узле :

$$\bar{T}_s = \frac{\bar{N}_s}{\mu_s}, \quad (s = \overline{1,3}).$$

- среднее время пребывания заявки в системе:

$$\bar{T} = \frac{\bar{N}_a}{\bar{\lambda}},$$

где \bar{N}_a - среднее количество активных пользователей, $\bar{\lambda}$ - средняя интенсивность формирования заявок пользователями.

$$\bar{N}_a = \sum_{l=0}^m P_{1,l} \cdot l, \quad \text{где } P_{1,l} = \sum_{n_s=l} P(l, n_2, n_3) \quad s = \overline{2,3}, l = \overline{0,m};$$

$$\bar{\lambda} = \lambda (m - \bar{N}_a).$$

В работах [3,4] показано, что расчет выше приведенных интегральных характеристик сети сводится, по существу, к вычислению нормализующей константы $G(m,3)$, для расчета которой может быть использован рекуррентный метод Бузена.

Литература

1. Клейнрок Л. Вычислительные системы с очередями: Пер. с англ.- М.Мир,1979.-600с.
 2. Черноморов Г.А. Теория принятия решений: Учебное пособие / Юж. - Рос. гос. техн. ун-т.-3-е изд. перераб. и доп. - Новочеркасск : Ред. журн. - «Изв. Вузов. Электроомеханика», 2005.-448с.
 3. Жожикашвили В.А.,Вишневский В.М. Сети массового обслуживания. Теория и применение к сетям ЭВМ. - М.:Радио и связь, 1988.-192с.
 4. Герасимов А.И. Теория и практическое применение стохастических сетей. - М.:Радио и связь.,1994.-175с.
 5. Chakka R., Harrison P.G. A Markov modulated multi-server queue with negative customers –The MM CPP/GE/c/LG-queue // Acta Informatika/-2001.- v.37.pp.785-799.
 6. 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.
 7. 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.
 8. Buzen J.P. Computational Algorithms for Closed Queueing Networks with Exponential Servers // Commun. ACM. -1983. –Vol.16, №9.pp.527-531.
 9. Вишневский В.М. Теоретические основы проектирования компьютерных сетей.- М.: Техносфера, 2003.- 512 с.
 10. Antunes C.H. et al. A Multiple Objective Routing Algorithm for Integrated Communication Network // Proc. ITC-16.-1999.V.3b.-PP.1291-1300.
-

References

1. Kleynrok L. Vychislitel'nye sistemy s ocheredyami [Queueing Systems]: Per. s angl. M.Mir, 1979. 600p.
 2. Chernomorov G.A. Teoriya prinyatiya resheniy [decision making theory]: Uchebnoe posobie. Yuzh. Ros.gos.tekhn. un-t. 3-e izd. pererab. I dop.Novocherkassk : Red. zhurn."Izv.vuzov.Elektromekhanika", 2005. 448p.
 3. Zhozhikashvili V.A., Vishnevskiy V.M. Seti massovogo obsluzhivaniya. Teoriya i primeneniye k setyam EVM [Queueing networks. Theory and its network application].M.: Radio i svyaz', 1988. 192p.
 4. Gerasimov A.I. Teoriya i prakticheskoye primeneniye stokhasticheskikh setey [Theory and practical application of stochastic networks].M.: Radio i svyaz',1994. 175p.
 5. Chakka R., Harrison P.G. A Markov modulated multi-server queue with negative customers .The MM CPP/GE/c/LG-queue . Acta Informatika.2001. v.37.pp.785-799.
 6. Skoba A.N., Sostina E.V. Inženernyj vestnik Dona (Rus), 2015. №2.URL:ivdon.ru/ru/magazine/archive/n2y2015/2881.
 7. Skoba A.N., Sostina E.V. Inženernyj vestnik Dona (Rus), 2015. №2.URL:ivdon.ru/ru/magazine/archive/n2y2015/2882.
 8. Buzen J.P. Computational Algorithms for Closed Queueing Networks with Exponential Servers. Commun. ACM. 1983. Vol.16, №9.pp.527-531.
 9. Vishnevskiy V.M. Teoreticheskie osnovy proektirovaniya komp'yuternykh setey [Theoretical foundations of computer network design].M.: Tekhnosfera , 2003. 512p.
 10. Antunes C. H. et al. A Multiple Objective Routine Algorithm for Integrated Communication Network. Proc ITC-16. 1999. V. 3b. pp. 1291-1300.
-

