УДК 519.872

РАСЧЕТ СЕТИ ОБСЛУЖИВАНИЯ С ОГРАНИЧЕНИЕМ ВРЕМЕНИ ЖИЗНИ ЗАЯВОК

Ю.И. Рыжиков
Санкт-Петербургский институт информатики и автоматизации РАН
Россия, 199178, Санкт-Петербург, 14 линия В.О., 39
E-mail: ryzhbox@yandex.ru

А.В. Уланов
Военно-космическая академия им. А.Ф. Можайского
Россия, 197082, Санкт-Петербург, Ждановская ул., 13
E-mail: ulanov246@rambler.ru

Ключевые слова: сети массового обслуживания, потокоэквивалентная декомпозиция, ограничение времени пребывания заявки в сети

Аннотация: Предложен алгоритм приближенного расчета разомкнутой сети массового обслуживания, в которой время прохождения маршрута (время жизни заявки) ограничено случайной величиной. Узлы сети предложено рассчитывать как многоканальные системы массового обслуживания с простейшим входящим потоком и фазовыми распределениями времен обслуживания и пребывания заявок в них. Получено выражение для моментов распределения времени прохождения «успешных» заявок по сети.

1. Введение

Одной из актуальных задач теории очередей является расчет сети массового обслуживания (СеМО), в которой накладывается ограничение на *полное* время пребывания заявок в ней. Примеры таких сетей:

- в логистике – цепи поставок товаров с ограниченным сроком годности;
- в учреждениях здравоохранения – оказание помощи тяжелобольным, требующим проведения ряда медицинских мероприятий;
- в военном деле – перехват крылатых ракет противника эшелонированной системой противовоздушной обороны (ПВО) с ограничением времени нахождения цели в зоне поражения средствами ПВО;
- в управлении – когда влияющая на принятие решений информация имеет ограниченное время актуальности, и т.д.

Исследованием *систем* массового обслуживания (СМО) с ограничением времени ожидания и пребывания заявок начали заниматься достаточно давно, а по расчету *сетей* авторам встретилась лишь статья [1]. В ней предложен метод расчета разомкнутой сети массового обслуживания, состоящей из двух одноканальных узлов с показательным распределением времени обслуживания в каждом, время пребывания заявки в сети (время жизни) ограничено постоянной величиной.

В данной статье предложен метод, основанный на потокоэквивалентной декомпозиции сети [2, 3], и позволяющий рассчитывать временные характеристики разомкнутой сети обслуживания без ограничения на количество узлов и распределение времени жизни заявок. При этом узлы сети рассчитываются как многоканальные СМО типа $M/G/n–G$ (в дополнение к стандартной классификации Кендалла после тире указан тип распределения времени пребывания) с простейшим входящим потоком и произвольными распределениями времен обслуживания и пребывания заявки в системе. Расчет таких СМО можно выполнить через аппроксимацию распределений времен обслуживания и пребывания распределениями фазового типа, построение диаграмм переходов для полученной цепи Маркова и решение систем уравнений Чепмена-Колмогорова. Подобные задачи, только при ограничении на время ожидания начала обслуживания, решались в [4, 5]. В качестве фазового распределения перспективно использовать гиперэкспоненциальное второго порядка (обозначается H_2). Оно позволяет выровнять три начальных момента исходного распределения, что обеспечивает разумно достаточную точность при расчете СМО [6].

Для численной реализации представленного в статье метода расчета исследуемой СеМО предполагается выполнить следующие подзадачи.

1) Разработка имитационной модели (ИМ) узла сети как системы $M/G/n–G$.
2) Верификация ее на модели $M/M/n–M$, для которой существует аналитическое решение.
3) Разработка численного метода расчета моментов времени пребывания заявки в системе $M/H_2/n – H_2$ и его верификация на ИМ.
4) Разработка ИМ СеМО с ограничением времени жизни заявок.
5) Верификация с ее помощью численного метода расчета сети, представленного в статье.

2. Потокоэквивалентная декомпозиция сети

Потокоэквивалентная декомпозиции является одним из наиболее перспективных методов, позволяющих повысить вычислительную эффективность и получить достаточно точные оценки характеристик моделей СеМО, которые принципиально не могут быть получены точными методами. В общем случае расчет сети проходит следующие этапы.

1) Решение уравнений баланса потоков для узлов сети и нахождение интенсивностей входящих в них потоков.
2) Расчет узлов сети как изолированных СМО и нахождение моментов времени пребывания заявок в них.
3) Расчет моментов случайного времени пребывания заявки в сети в целом.

Поскольку при этом игнорируются корреляции распределений количества заявок в узлах, данный метод является приближенным.

2.1. Предположения и шаги алгоритма

Расчет рассматриваемой сети предполагает следующие допущения:

- циклические маршруты отсутствуют;
- узлы представляют собой многоканальные СМО с фазовыми распределениями времен обслуживания и пребывания;
- потоки между узлами – простейшие;
- ограничение накладывается на полное время пребывания заявки в сети.

Пусть сеть обслуживания состоит из *рабочих узлов*, занумерованных от 0 до *M*, *источника* (узел «0») и *стока* (узел «M+1»). Новая заявка рождается в источнике, с попаданием заявки в сток формируется окончание ее пребывания в сети. Маршрут заявки в сети случаен и определяется неразложимой матрицей передач $R=\{r_{i,j}\}$, $i,\ j=\overline{0, M+1}$, образованной вероятностями перехода из i-го в j-й узел. Эти вероятности не зависят от маршрута, уже пройденного заявкой. Неразложимость сети предполагает невозможность ее разделения на несвязанные допустимыми переходами подсети.

Узлы исследуемой сети предварительно перенумеровываются в соответствии с отношением предшествования – алгоритм Форда [7]. Расчет проходит следующие этапы.

1) Решаем систему уравнений баланса межузловых потоков без учета потерь заявок

$$\lambda_j=\Lambda r_{0,j}+\sum_{i=1}^{j-1}\lambda_i r_{i,j}, \quad j=\overline{1,M}.$$

Здесь Λ – интенсивность внешнего входящего потока.

2) Для узлов $j=\overline{1,M}$: если $j=1$, то берется исходное распределение времени жизни заявок с моментами $\{\overline{g}_k\}$. В противном случае:

a) просматривая узлы-предшественники ($i<j$), накапливаем интенсивности входящего в j-й узел потока

$$\lambda_j=\sum_{i<j}\lambda_i\pi_i r_{i,j}$$

и суммы

$$\sum_{i<j}\lambda_i g_{i,k},$$

где $g_{i,k}$ – k-й момент распределения оставшегося времени жизни у заявок, благополучно завершивших обслуживание в i-м узле, π_i – вероятность такого завершения;

b) рассчитываем средневзвешенные моменты распределения оставшегося времени жизни для входящих в j-й узел заявок:

$$\overline{g}_k=\sum_{i<j}\lambda_i g_{i,k}\Big/\lambda_j.$$

По этим моментам подбираем параметры H_2-распределения;

c) каждый узел $j=\overline{1,M}$ рассчитывается как система с ограничением времени пребывания заявками. При этом определяются:

- доля π_j заявок, успешно прошедших обслуживание;
- начальные моменты распределений времени ожидания $\{w_{j,k}\}$ и времени пребывания $\{\upsilon_{j,k}\}$ в узле, $k=\overline{1,3}$;
- «частные» моменты $\{v_{j,k}\}$ распределения времени пребывания заявки в сети (на момент завершения j-обслуживания) – как свертка средневзвешенных аналогичных моментов на входе и собственно в j-м узле;
- если $j<M$, считаются моменты распределения остаточного времени жизни заявок (подробности обсуждаются ниже).

3) Если $\max_j\{\pi_j-\pi'_j\}>\varepsilon$, заменить $\{\pi_j\}$ на $\{\pi'_j\}$ и перейти к этапу 1.

4) Расчет итоговых характеристик:

- доли заявок, успешно прошедших всю сеть:

$$e=\sum_{j=1}^{M}\lambda_j\pi_j r_{j,M+1}\Big/\Lambda.$$

Напомним, что первоначальные интенсивности потоков $\{\lambda_j\}$ на входе в узлы $j=\overline{2,M}$ пересчитываются с учетом ухода нетерпеливых заявок из их предшественников;

- начальных моментов распределения времени пребывания прошедших всю сеть заявок

$$\upsilon_k = \sum_{j=1}^{M} \lambda_j \pi_j r_{j,M+1} v_{j,k} \Bigg/ \sum_{j=1}^{M} \lambda_j \pi_j r_{j,M+1}, \quad k = \overline{1,3}.$$

5) Конец алгоритма.

По моментам $\{\upsilon_k\}$ можно построить дополнительную функцию распределения (ДФР) времени пребывания заявки в сети.

2.2. Остаточное время жизни заявок

Пусть для j-го узла:

- $\upsilon(t)$ – плотность распределения длительности пребывания заявки в сети до успешного выхода из данного узла;
- $g(t)$ и $\bar{G}(t)$ – соответственно плотность и ДФР стартового (на входе в узел) распределения времени жизни.

Тогда условная плотность распределения вероятности «пожить» еще t единиц времени

$$\gamma(t,\tau) = \frac{g(t+\tau)}{\bar{G}(\tau)} \upsilon(\tau),$$

а безусловная

$$g(t) = \int_0^\infty \gamma(t,\tau)\, d\tau . \tag{1}$$

Моменты этого распределения

$$g_k = \int_0^\infty t^k g(t)\, dt . \tag{2}$$

Вычислить несобственные интегралы в правых частях (1) и (2) можно по квадратурной формуле Чебышева-Лагерра

$$\int_0^\infty e^{-t} t^k \varphi(t)\, dt = \sum_{i=1}^{L} A_i \varphi(t_i), \tag{3}$$

обширные таблицы узлов $\{A_i\}$ и весов $\{t_i\}$ для которой приводятся в [8].

Строго говоря, для расчета моментов различного порядка следует пользоваться различными наборами узлов и весов (соответствующие таблицы в [8] имеются). Однако численные эксперименты показали приемлемость использования в формуле (3) их общего набора, соответствующего k=0.

Приведем реализующий эти вычисления фрагмент псевдокода программы на Фортране 90:

```
s=0   ! Обнуление сумм для трех моментов
do i=1,L    ! Узлы для работы с формулой (2)
     g=0
     do j=1,L     ! Считаем интегральную сумму для (1)
             g=g+A(j)*γ(t(i),t(j))*exp(t(j))
     end do
     c=A(i)*t(i)*exp(t(i))
     do k=1,3     ! Порядок моментов
             s(k)=s(k)+c*g
             c=c*t(i)
     end do
end do
```

Отметим также необходимость включения в этот фрагмент программы умножения интегрируемой функции на дополнительную экспоненту – для приведения задачи к виду (3).

3. Заключение

Представлен метод расчета моментов случайного времени пребывания успешно прошедших СеМО заявок при ограничении на полное время пребывания в ней. Предложена методика нахождения остаточного времени жизни заявок для расчета узлов сети. В настоящее время ведется разработка численного метода расчета узла, запрограммирована и верифицирована его имитационная модель.

Исследуемая сеть обслуживания имеет широкую область потенциальных приложений, в том числе при оценивании оперативности систем управления, информация в которых со временем теряет актуальность.

Список литературы

1. Mahdipour E., Rahmani A., Setayeshi S. Importance Sampling for a Markov Modulated Queuing Network with Customer Impatience until the End of Service // Informatica Economica. 2009. Vol. 13, No. 3. P. 106-118.
2. Вишневский В.М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003. – 512 с.
3. Рыжиков Ю.И. Теория очередей и управление запасами. СПб.: Питер, 2001. 384 с.
4. Roubos A., Jouini O. Call Centers with Hyperexponential Patience Modeling // International Journal of Production Economics. 2013. Vol. 141. P. 307-315.
5. Дудин С.А., Дудина О.С. Модель функционирования колл-центра как система $MAP/PH/N/R\!-\!N$ с нетерпеливыми запросами // Проблемы передачи информации. 2011. № 47. С. 68-83.
6. Рыжиков Ю.И., Уланов А.В. Опыт расчета сложных систем массового обслуживания // Информационно-управляющие системы. 2009. № 2. С. 56-62.
7. Городецкий В.И, Рыжиков Ю.И. Математическое программирование и массовое обслуживание. Л.: ВИКИ им. А.Ф. Можайского, 1975. 182 с.
8. Крылов В.И., Шульгина Л.Т. Справочная книга по численному интегрированию. М.: Наука, Физматгиз, 1966. 372 с.