УДК 156.6

Науменко Владимир Викторович, Копытов Владимир Вячеславович

ПРИМЕНЕНИЕ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ ДЛЯ МОДЕЛИРОВАНИЯ СОВМЕСТНО ВЫПОЛНЯЕМЫХ АДМИНИСТРАТИВНЫХ РЕГЛАМЕНТОВ

В статье рассматривается задача моделирования совместно выполняемых административных регламентов с использованием методов теории массового обслуживания, решение которой позволит выполнить проверку показателей предоставляемых государственных и муниципальных услуг на их соответствие нормативным срокам исполнения.

Ключевые слова: административный регламент, теория массового обслуживания, функциональная безопасность.

Naumenko Vladimir. V, Kopytov Vladimir V.

THE USE OF QUEUING THEORY FOR MODELING JOINTLY SATISFIABLE ADMINISTRATIVE REGULATIONS

The article gives the description of problem to modeling jointly satisfiable administrative regulations with methods queuing theory. It will check the parameters of administrative regulations to meet regulatory deadlines.

Key words: administrative regulations, queuing theory, functional safety.

Современная нормативно-правовая база [1, 2] требует от органов власти исполнения государственных функций и предоставления государственных услуг строго в соответствии с требованиями административных регламентов (АР), делая тем самым АР ключевым элементом системы государственного управления, от эффективности выполнения которого зависит эффективность всей системы государственного управления. В настоящее время информационная система государственного управления является критической социотехнической системой, неэффективная работа которой ухудшает социально-экономическое развитие и связана с невозможностью выполнения функций, закрепленных нормативно-правовыми актами и законами, что ухудшает социально-экономическое развитие. На функциональную безопасность информационной подсистемы системы государственного и муниципального управления напрямую влияет выполнимость АР, которая в свою очередь выражается в невыполнении требований АР. Следовательно существует актуальная научная задача по оценке выполнимости АР, которая связана с определением временных показателей предоставляемых государственных и муниципальных услуг.

Если рассматривать АР как алгоритмический процесс, реализующий государственную услугу, то его основу будут составлять процессы и операции, закрепленные в нормативно-правовом акте АР. Типовая структура АР R_m предполагает наличие описания административных процедур, каждая из которых достигается путем выполнения определенной последовательности задач z_n:

$$\forall R_m \to \exists (z_n)_{n=1}^{k}.$$

При этом каждая из задач связана с субъектом, который ее исполняет $\forall s_m \to \exists z_n$, и который может быть задействован в выполнении других задач (в том числе других АР) $\forall s_m \to \exists (z_k, ..., z_{k+p})$, образуя тем самым систему совместно выполняемых АР (см. рис. 1).

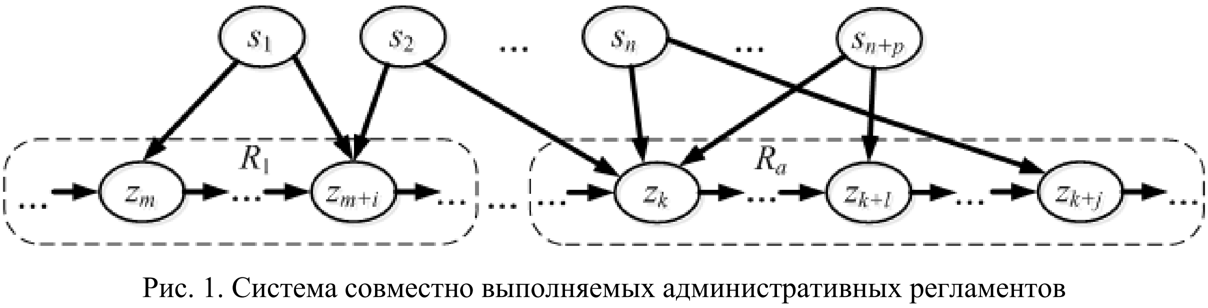

Рис. 1. Система совместно выполняемых административных регламентов

Процесс выполнения АР можно рассматривать как поток выполнения требований, встречающий ограниченные средства для их удовлетворения, поэтому множество совместно выполняемых АР можно рассматривать как систему массового обслуживания (СМО).

Согласно теории массового обслуживания (ТМО) АР представляет собой сеть узлов массового обслуживания. Учитывая, что интервал между заявками и длительность выполнения задач распределены по законам $A(\tau)$ и $B(\tau)$, которые необходимо определять отдельно для каждого случая (отдельно для каждого исполнителя и АР в целом) и которые также могут изменяться во времени, то следует воспользоваться общим решением для одноканальных систем массового обслуживания *G/G/1*, которое характеризует выполнение некоторого потока задачи z_k (или задач $z_k,\ldots,z_{k+p}$) исполнителем s_n.

На основании решений для системы *G/G/1* можно определить, сколько времени будет выполняться задача при заданных параметрах системы (λ_k и t_{Zk}). Подсчитав для каждой задачи среднее время пребывания заявки в системе T_k (среднее время начиная от поступления заявки в очередь и заканчивая выполнением задачи), можно сравнить его с максимальным временем, указанном в АР $T_{\max}$, и таким образом определить, присутствует ли в системе нарушение временных нормативов, оценив таким образом выполнимость АР.

Основными параметрами, влияющими на выполнимость задачи АР (для системы *G/G/1*), являются:

- среднее фактическое время выполнения задачи t_{Zk};
- средняя интенсивность поступающих заявок λ_k;
- среднеквадратическое отклонение значения интервала между поступающими заявками σ_A;
- среднеквадратическое отклонение значения длительности выполнения задач σ_B.

Анализ многочисленных результатов (например результатов работ [3–5]) показывает, что наиболее удачное приближение для расчета среднего времени пребывания заявки для системы *G/G/1* дает формула

$$T_s \approx \frac{\rho t_s (v_{As}^2 + v_{Bs}^2)}{2(1-\rho)} f(v_{As}) + t_s, \tag{1}$$

где $\rho = \lambda_s t_s$, $f(v_{As}) = \begin{cases} \exp\left[-\dfrac{2(1-\rho)(1-v_{As}^2)^2}{3\rho(v_{As}^2+v_{Bs}^2)}\right], v_{As} < 1 \\ \exp\left[-(1-\rho)\dfrac{v_{As}^2-1}{v_{As}^2+4v_{Bs}^2}\right], v_{As} \geq 1 \end{cases}$.

Величины λ_s и t_s определяются следующим образом:

λ_s – суммарный поток заявок для исполнителя *s*, который равен $\lambda_s = \lambda_k + \ldots + \lambda_{k+p}$, где $\lambda_k,\ldots,\lambda_{k+p}$ – интенсивности поступления заявок на выполнение всех задач, выполняемых субъектом s_n;

t_s – среднее время выполнения задач субъектом s_n, которое определяется как

$$t_s = p_k t_k + \ldots + p_{k+p} t_{k+p}, \tag{2}$$

где $p_i = \dfrac{\lambda_i}{\sum_{i=k}^{k+p} \lambda_i}$.

Для определения коэффициента вариации интервала между поступающими заявками v_{As} воспользуемся уравнением средней дисперсии для мультиплексированного входящего потока системы *G/G/1*, приведенным в работе [6]:

$$\sigma_{As}^2 = \left(\frac{\lambda_k}{\sum_{i=k}^{k+p} \lambda_i}\right)^3 \sigma_{Azk}^2 + \ldots + \left(\frac{\lambda_{k+p}}{\sum_{i=k}^{k+p} \lambda_i}\right)^3 \sigma_{Az\ k+p}^2. \tag{3}$$

Коэффициент вариации интервала между поступающими заявками в свою очередь будет равен

$$v_{As} = \sigma_{As}(\lambda_k + \ldots + \lambda_{k+p}). \qquad (4)$$

Для определения коэффициента вариации длительности выполнения задач v_{Bs} необходимо учитывать, что субъект s_n выполняет задачи последовательно, поэтому нужно рассматривать множество значений длительности выполнения задач как совокупность с математическим ожиданием t_s, состоящей из групп с параметрами $t_k,\ldots,t_{k+p}$ и $\sigma_{Bk},\ldots,\sigma_{Bk+p}$. Учитывая независимость потоков заявок воспользуемся правилом сложения дисперсий для получения общей дисперсии длительности выполнения для s_n

$$\sigma_{Bs}^2 = \left(\frac{\lambda_k}{\sum\limits_{i=k}^{k+p} \lambda_i} \sigma_{Bzk}^2 + \ldots + \frac{\lambda_{k+p}}{\sum\limits_{i=k}^{k+p} \lambda_i} \sigma_{Bzk+p}^2 \right) + \left(\frac{\lambda_k}{\sum\limits_{i=k}^{k+p} \lambda_i} (t_k - t_s)^2 + \ldots + \frac{\lambda_{k+p}}{\sum\limits_{i=k}^{k+p} \lambda_i} (t_{k+p} - t_s)^2 \right). \qquad (5)$$

Отсюда коэффициент вариации по определению равен

$$v_{Bs} = \frac{\sigma_{Bs}}{t_s}.$$

В случае если одну задачу (один поток) выполняет несколько субъектов, то необходимо воспользоваться решением для параллельных каналов с различным временем облуживания [7], то есть поступающее требование направляется на один из каналов (к одному из исполнителей) с вероятностью:

$$p_{ki} = 1 / (t_{iZk} \sum_{i=n}^{n+p} \frac{1}{t_{iZk}}). \qquad (6)$$

Поэтому поток заявок для каждого из исполнителей задачи z_k будет равен

$$\lambda_{ki} = p_{ki} \lambda_k. \qquad (7)$$

Таким образом, выражение (4) позволяет оценить среднее время пребывания заявки у исполнителя s_n. Однако использовать данное значение для оценки выполнимости АР не возможно, так оно является общим для всех задач, выполняемых исполнителем s_n. Поэтому необходимо определить среднее время обслуживание для каждой задачи $z_k,\ldots, z_{k+p}$, выполняемой s_n – $T_k,\ldots,T_{k+p}$. Для этого рассмотрим следующую модель (рис. 2).

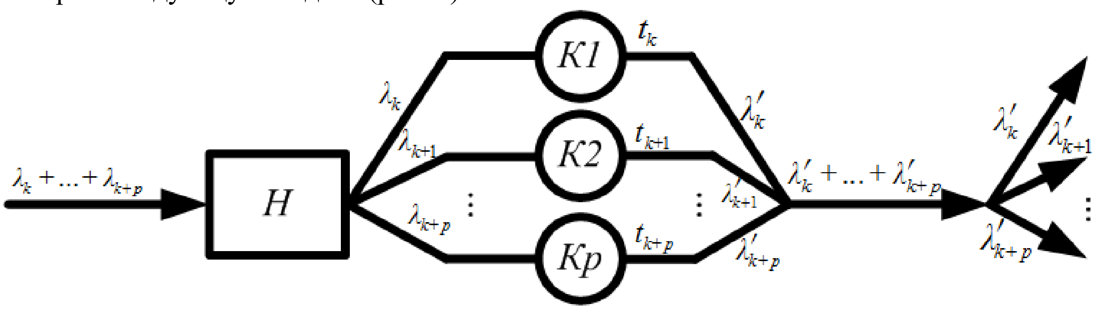

Рис. 2. Модель системы обслуживания потока различных задач АР

Рассмотрим узел СМО (отдельного исполнителя) как многоканальную систему (см. рис. 1), где каждая заявка перед обслуживанием попадает в бесконечный накопитель H, а каждый канал K_N связан с выполнением определенной задачи из множества $z_k,\ldots, z_{k+p}$.

Учитывая, что накопитель H является общим для всех потоков $\lambda_k,\ldots, \lambda_{k+p}$, то среднее время ожидания в очереди будет одинаковым для всех заявок $z_k,\ldots, z_{k+p}$. Поэтому с учетом выражения 1 время пребывания заявки в системе для каждой из задач, выполняемых исполнителем s_n будет определяться по формуле

$$T_i \approx \frac{\rho t_s (v_{As}^2 + v_{Bs}^2)}{2(1-\rho)} f(v_{As}) + t_{Zi}. \qquad (8)$$

Просчитав для каждой задачи среднее время выполнения T_i и сравнив их с максимальным временем $T_{i\,\max}$, определенным нормативными регуляторами АР, можно сказать, что если $T_i > T_{i\,\max}$, то задача не выполняется с заданными условиями, то есть существует нарушение требований АР.

В том случае, если задачу z_i выполняет несколько исполнителей $s_n,\ldots,s_{n+l}$, то для сравнения выбирается максимальное значение времени пребывания заявки в системе:

$$T_i = \max\left(T_{i\ n},\ldots,T_{i\ n+l}\right). \tag{9}$$

Необходимо отметить, что законы распределении интервала времени между заявками и длительности выполнения задач $A(\tau)$ и $B(\tau)$ трудно выявить в процессе исполнения государственных и муниципальных услуг. При этом с течением времени могут изменяться основные параметры (λ_k, σ_A, t_{Zk}, σ_B), используемые для расчета среднего времени выполнения, а также в связи с нестабильностью процессов в социотехнической системе могут поменяться и законы распределения данных величин. Следовательно, необходимо разработать подход, позволяющий оперативно определять основные параметры, влияющие на выполнимость АР.

Для отслеживания изменения законов распределения $A(\tau)$ и $B(\tau)$ авторами предлагается использовать понятие однородности двух независимых выборок, которое указывает на одинаковый закон их распределения. Для этого временной ряд интервалов времени между заявками или длительности выполнения задач аппроксимируется интервалы, включающие равное количество N значений (см. рис. 3).

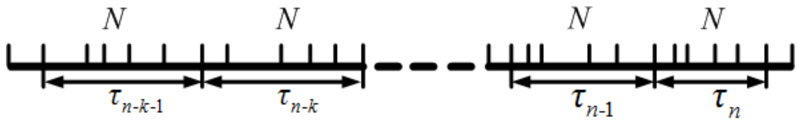

Рис. 3. Временной ряд интервалов времени между заявками или длительности выполнения задач

Затем проверяется гипотеза равенства распределения τ на последнем интервале τ_n (включающего N значений) и интервале $\tau' = \tau_{n-1}+\ldots+\tau_{n-k}$ (причем каждая из выборок на любом из интервалов $\tau_{n-1},\ldots,\tau_{n-k}$ однородна по отношению к соседним), который включает в себя M значений H_0: $A(\tau) = A(\tau')$ (или $B(\tau) = B(\tau')$), причем $M >> N$. В том случае если гипотеза H_0 верна, то значения λ_k, σ_A (t_{Zk}, σ_B) определяются на интервале $\tau_n + \tau'$, в противном случае значения определяются на интервале τ_n.

В качестве метода определения однородности двух выборок в настоящей работе применен критерий Вилкоксона [8] (для независимых выборок). Выбор данного метода обусловлен тем, что он подходит для проверки однородности любых независимых выборок с единственным требованием непрерывности случайной величин. Также необходимо отметить простоту вычислений для проверки гипотезы равенства законов распределения, а также оптимальную эффективность при использовании выборок малых объемов, что является наиболее важным для решения поставленной задачи.

Данный метод сводится к объединению выборок в один вариационный ряд в возрастающем порядке и нахождению критерия $W_{\text{набл}}$ – суммы порядковых номеров значений первой выборки

$$W_n = i_{N1} + i_{N2} + \ldots + i_{Nk}\ .$$

Далее на основе табличных значений критических точек критерия Вилкоксона определяется нижняя критическая точка $\omega_{\text{нижн.кр}}$ и на основе нее определяется верхняя критическая точка по формуле:

$$\omega_{\text{нижн. кр}} = (N + M + 1) - \omega_{\text{нижн. кр}}\ .$$

В случае если $W_{\text{набл.}} < \omega_{\text{нижн.кр.}}$ или $W_{\text{набл.}} > \omega_{\text{верхн.кр.}}$, гипотеза H_0 об однородности выборок отвергается, а если $\omega_{\text{нижн.кр.}} < W_{\text{набл.}} < \omega_{\text{верхн.кр.}}$, то выборки N и M однородны.

Для решения данной задачи критерий Вилкоксона был применен со значениями $N = 6$ и $M = 24$ для первой и второй выборок соответственно. Размер выборок обусловлен примерным диапазоном значений промежутков времени между поступающими заявками и длительности выполнения задач, которые колеблются в пределах от нескольких минут до нескольких дней.

Предлагаемый способ расчета основных параметров, влияющих на выполнимость АР реализована виде циклического алгоритма (см. рис. 4), определяющего значения λ_k, σ_A (t_{Zk}, σ_B) через каждые N интервалов.

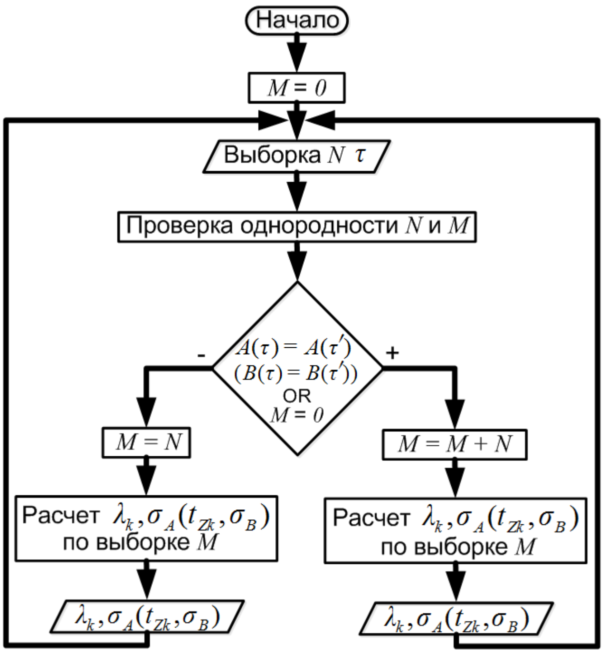

Рис. 4. Алгоритм расчета основных параметров, влияющих на выполнимость административных регламентов

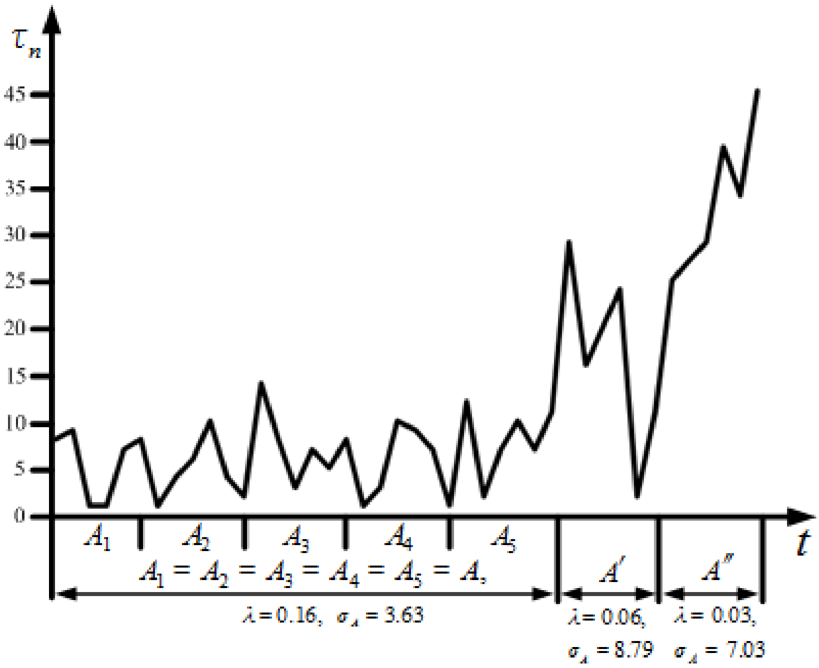

Рис. 5. Пример расчета интенсивности потока заявок

Пример расчета интенсивности потока заявок на промежутке времени, составляющем семь интервалов между соседними заявками приведен на рис. 5.

Таким образом, последовательно рассчитав для каждой из задач АР, входящих в СМО, значения интенсивности поступающих заявок, времени их выполнения, а также значения среднеквадратичных отклонений данных величин, можно выявить нарушение нормативных сроков как для отдельной задачи (сравнив T_i с $T_{i\ \max}$), так и для АР в целом, сравнив сумму значений среднего времени выполнения всех задач АР с максимальным временем исполнения АР, и полностью решить поставленную задачу по оценке выполнимости совместно выполняемых АР.

Литература

1. Федеральный закон «Об организации предоставления государственных и муниципальных услуг» № 210-ФЗ от 27 июля 2010 года. URL: http://www.rg.ru/2010/07/30/gosusl-dok.html (дата обращения: 01.03.2013).

2. Постановление Правительства РФ от 16.05.2011 № 373 «О разработке и утверждении административных регламентов исполнения государственных функций и административных регламентов предоставления государственных услуг». URL: http://www.rg.ru/2011/05/31/gosuslugi-site-dok.html (дата обращения: 01.03.2013).

3. Шнепс М. А. Системы распределения информации. Методы расчета. М.: Связь, 1979. 344 с.

4. Алиев Т. И. Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.

5. Нгуен Дык Тай. Методы и средства исследования распределенных сетей передачи данных с неоднородным трафиком на основе неэкспоненциальных моделей: дис. … канд. техн. наук: 05.13.13. М., 2009. 145 с.

6. Бахарева Н. Ф. Аппроксимативные методы и модели массового обслуживания для исследования компьютерных сетей: дис. … д-ра техн. наук: 05.13.15. Пенза, 2011. 335 с.

7. Саати Т. Л. Элементы теории массового обслуживания и её приложения. М.: Советское радио, 1971. 520 с.

8. Гмурман В. Е. Теория вероятностей и математическая статистика. М.: ВШ, 2003. 478 с.