

АДАПТИВНАЯ СТРАТЕГИЯ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ В АВТОМАТИЗИРОВАННОЙ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКОЙ СИСТЕМЕ

И. В. Шевченко, И. Г. Оксанич, П. Д. Савушкин

Кременчугский национальный университет имени Михаила Остроградского
ул. Первомайская, 20, г. Кременчуг, 39600, Украина. E-mail: oksirena2017@gmail.com

Предложена адаптивная стратегия управления очередями при выборе маршрутов для множества разнородных бизнес-процессов и изменяющихся интенсивностей потоков заявок в рамках организационно-технической системы с учетом определенного критерия оптимальности. Стратегия усовершенствована за счет учета в качестве признаков степени изменения потока неоднородных заявок, степени компетенции исполнителей и степени загрузки исполнителей, что позволяет реализовать процесс структурной адаптации системы к изменяющейся оперативной обстановке. Разработан метод и алгоритм адаптации стратегии управления очередями, основанный на эволюционном подходе. При помощи эволюционного алгоритма подбираются оптимальные значения порогов отбора заявок по стоимости задержки и по уровню компетенции сотрудников. При изменении режима работы системы происходит адаптивное регулирование значений указанных порогов, что дает возможность при различных распределениях интенсивностей входных потоков выбрать наилучшие условия многофазного обслуживания и снизить потери ожидания. Каждая эпоха оптимизации режимов работы системы оставляет в качестве результата исходные данные сеанса моделирования и лучший набор значений рабочих параметров. Сохранение результатов в отдельной таблице в дальнейшем позволяет сразу находить значения рабочих параметров близкие к оптимальным.

Ключевые слова: организационная система, маршрутизация заявок, стратегия управления, адаптация.

АДАПТИВНА СТРАТЕГІЯ УПРАВЛІННЯ ЧЕРГАМИ У АВТОМАТИЗОВАНІЙ ОРГАНІЗАЦІЙНО-ТЕХНІЧНОЇ СИСТЕМІ

І. В. Шевченко, І. Г. Оксанич, П. Д. Савушкін

Кременчуцький національний університет імені Михайла Остроградського
вул. Першотравнева, 20, м. Кременчук, 39600, Україна. E-mail: oksirena2017@gmail.com

Метою роботи є розробити адаптивну стратегію управління чергами при виборі маршрутів для багатьох неоднорідних бізнес-процесів та змінах інтенсивностей потоків заявок в рамках організаційно-технічної системи. Це дає можливість реалізувати процес структурної адаптації системи до мінливого робочого середовища. При вирішенні задачі враховувались такі фактори: неоднорідність заявок за вектором параметрів через різноманітність процесів; вартість затримок операцій; невизначеність розподілу потоків різнорідних застосувань на вході в систему; неоднорідність приладів обслуговування через обмеження компетенцій працівників; той же пристрій обслуговує кілька типів додатків; режим обслуговування змінюється відповідно до довільного закону, від повної зупинки до перевантаження. З урахуванням цих факторів розроблена адаптивна стратегія зміни маршрутів бізнес-процесів шляхом динамічного вибору приладів обслуговування. Також розроблено метод вибору оптимальних граничних значень рівня компетентності при виборі виконавців та порогу поділу додатків на групи ризику відповідно до вартості затримок. Методи дослідження ґрунтуються на методах системного аналізу, методах керування чергами, теорії ймовірностей. Розроблено метод та алгоритм адаптації стратегії управління чергами на основі еволюційного підходу. Використовуючи еволюційний алгоритм, оптимальні значення граничних значень параметрів черг і заявок для вибору маршруту вибираються відповідно до вартості затримки та рівня компетентності працівників. При зміні режиму роботи відбувається адаптивне регулювання значень цих порогів. Це дає змогу вибрати найкращі умови для багатофазного обслуговування при різних розподілах інтенсивності вхідних потоків та зменшити втрати очікування. Кожна епоха оптимізації режимів роботи системи залишає в результаті вихідні дані сеансу моделювання та найкращий набір робочих параметрів. Збереження результатів в окремій таблиці в майбутньому дозволяє негайно знайти значення робочих параметрів, близьких до оптимальних.

Ключові слова: організаційна система, маршрутизація заявок, стратегія управління, адаптація.

АКТУАЛЬНОСТЬ РАБОТЫ. Организационно-техническая система (ОТС) представляет собой совокупность технических и человеческих ресурсов, предназначенную для выполнения определенного множества бизнес-процессов (БП). Каждый БП проходит ряд этапов, которые представляют собой отдельные бизнес-операции (БО), выполняемые на конкретных автоматизированных рабочих местах (АРМ). Любой БП создает поток заявок на выполнение БО. Распределение заявок между АРМ происходит, с одной стороны, согласно бизнес-логике БП, а с другой стороны – с учетом компетенций сотрудни-

ков ОТС и их текущей загрузки. Поскольку схожими компетенциями обычно обладают несколько сотрудников, возникает возможность варьировать маршруты БП через сеть АРМ. Таким образом, мы приходим к необходимости динамического распределения заявок или к динамическому формированию исполнительных структур при выполнении множества БП. Потребность в динамическом формировании исполнительных структур в рамках ОТС возникает из противоречия между требованием гарантированного выполнения множества бизнес-операций в установленные сроки и ограничениями

на технологические и организационные возможности ОТС.

В процессе функционирования ОТС с динамической маршрутизацией БО системе управления необходимо накапливать задания, ожидающие обработки, и формировать очереди заявок различных типов, к каждому обслуживающему узлу (АРМ). Кроме того, необходимо автоматически производить периодическое обновление очередей и уведомлять пользователя о наличии в очереди новых, еще не просмотренных заданий, заданий с высоким приоритетом или заданий с установленным предельным сроком выполнения [1]. Это станет возможным в полной мере, если будет разработана адаптивная стратегия маршрутизации бизнес-процессов через множество автоматизированных рабочих мест ОТС или, другими словами, адаптивная стратегия управления очередями.

Однако, как показывает практика, проблема оптимизации процесса динамического распределения нагрузки на АРМ при реализации БП по прежнему не решена. Поэтому, решение этой проблемы является весьма актуальной задачей.

Несмотря на то, что существует множество моделей, призванных описывать системы с динамической маршрутизацией, для всех них можно выделить ряд общих свойств. В основном, во всех моделях рассматривается пуассоновский поток входных заявок и предполагается, что время обслуживания распределено экспоненциально. Взятое при этом за основу предположение об однородности времени обслуживания, не всегда соответствует действительности, хотя и полезно, так как приводит к простым марковским системам, которые приближенно описываются системами дифференциально-разностных уравнений [2–3].

В работе [4] для систем обслуживания с отказами, произвольным распределением длительности обслуживания и двумя способами распределения заявок по неоднородным обслуживающим приборам (равнороятным распределением заявок по всем приборам (свободным и занятым) и равнороятным распределением заявок по свободным приборам) найдены стационарные распределения числа заявок и доказана их инвариантность относительно функций распределения длительности обслуживания. В работе [5] предложен метод определения характеристик системы обслуживания $M/G/1/m$ с функцией случайного отбрасывания заявок и распределением времени обслуживания, зависящим от длины очереди. В реальных ОТС ни случайное, ни детерминированное отбрасывание заявок не практикуется, бизнес-процессы прерываются только в особых случаях. В работах [6, 7] рассматривается процесс динамической маршрутизации заявок, поступающих в систему из N серверов в соответствии с распределением Пуассона с интенсивностью $N\lambda$ и имеющих Вейбулловское распределение времен обслуживания. Маршрутизация производится таким образом, что при поступлении в систему заявки в момент времени t , случайно выбираются K серверов из N , и заявка становится в очередь с минимальной дли-

ной. В реальных ОТС случайный выбор подмножества узлов невозможен из-за ограничений по компетенции исполнителей. Кроме того, выбор очереди с минимальной длиной не всегда оптимален, так как на состояние очереди влияет цена потерь задержки по каждой заявке.

В работах [8, 9] рассмотренные алгоритмы динамической маршрутизации распределяют поступающие в систему запросы, опираясь на текущие характеристики очередей таким образом, что при поступлении запроса в систему он направляется на обслуживающую станцию с очередью наименьшей длины или в очередь средней длины, если необходимо чтобы в системе всегда оставалось несколько достаточно свободных очередей, например, для быстрого обслуживания приоритетных запросов. Очевидно, что такие стратегии являются эмпирическими и не могут быть рекомендованы для реальных ОТС с неопределенным распределением потока разнородных заявок и изменяющимися интенсивностями процессов обслуживания и отсутствию отказов. Такие системы в кодировке Д. Кендалла имеют обозначение $G/G/m/n$, причем, параметры m , n изменяются во времени.

Таким образом, несмотря на то, что существует довольно развитый аналитический аппарат исследования многоканальных многофазных систем массового обслуживания, те ограничения, которые позволяют получить аналитические решения, являются неприемлемыми и соответствующие математические схемы становятся непригодны для описания реальных процессов с достаточной для практики точностью. Это говорит об актуальности проблемы разработки адаптивной стратегии управления очередями заявок в реальных ОТС.

Определим адаптивную стратегию, как такую стратегию, которая определенным образом меняется в процессе решения задачи на основе накопления данных о возможных результатах того или иного варианта решения.

Целью работы является разработка адаптивной стратегии управления очередями при выборе маршрутов для множества разнородных бизнес-процессов и изменяющихся потоков заявок в рамках организационно-технической системы с учетом определенного критерия оптимальности и некоторых ограничений. Для достижения поставленной цели в работе решаются следующие задачи:

- построение стратегии управления очередями при выборе маршрутов БП через множество АРМ, способной к адаптации в изменяющихся условиях работы ОТС;
- разработка метода адаптации стратегии управления очередями при изменениях в распределении потоков заявок разного типа;
- разработка имитационной модели для реализации метода оптимизации адаптивной стратегии.

Данная статья является продолжением работ, результаты которых опубликованы в [10, 11].

МАТЕРИАЛ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ. Рассматривая ОТС как СМО, нужно учесть следующие особенности системы:

1. Многокритериальность задачи оптимизации стратегии управления. В зависимости от ситуации может оказаться необходимым оптимизировать общие потери времени прохождения множества БП через ОТС, среднее число заявок в очереди, среднее время пребывания в очереди, какой либо другой критериальный параметр или совокупность параметров в рамках интегрального критерия.

2. Большое и неопределенное количество состояний системы из-за наличия многомерного операционного пространства на множестве БП, БО, АРМ.

3. Неоднородность заявок по вектору параметров из-за разнообразия БП и БО, стоимости задержки выполнения, приоритетам и т.д.

4. Неопределенность распределения потоков разнородных заявок на входе ОТС.

5. Неоднородность приборов обслуживания из-за ограничений в компетенциях сотрудников.

6. Один и тот же прибор (АРМ) обслуживает несколько типов заявок.

7. Режим обслуживания на любом АРМ изменяется по произвольному закону, от полной остановки до перегрузки.

Таким образом, ОТС является СМО конфигурации $GD/GD/m/n$, где аббревиатура GD означает, что входные потоки СМО являются последовательностями, члены которых могут быть независимы и, кроме того, они имеют различные распределения, вид которых заранее неизвестен.

Будем считать, что менеджер ОТС формирует задания на выполнение БП. При этом он использует какую либо нотацию, позволяющую осуществить декомпозицию БП на БО, относя каждую БО к заранее определенному классу, который, в свою очередь, соотношен с определенной компетенцией сотрудника.

В автоматизированной ОТС выполнение БП осуществляется многоагентной системой. В составе системы имеются программно реализованный агент-диспетчер, программно реализованные агенты-мониторы, программно реализованные агенты-исполнители (боты или b-агенты) и сотрудники – (h-агенты). Менеджер ОТС контактирует с агентом-диспетчером, который осуществляет динамическую маршрутизацию БП. Агент-диспетчер (в дальнейшем – диспетчер) контактирует с агентами-мониторами, каждый из которых контролирует некоторое подмножество АРМ. Состояние всех заявок (БО), всех БП и всех АРМ фиксируется в динамической таблице, модель которой в виде совокупности конечных автоматов описана в работе [10]. Фрагменты этой таблицы для отдельного АРМ формирует агент-монитор (далее – монитор).

Метод, модели и критерии динамической маршрутизации заявок описаны в работе [11]. Формально задача динамической маршрутизации записана следующим образом [12]. Необходимо определить оптимальную структуру исполнительской сети на период времени t

$$S_{opt}(t) = \{E^*(t), A^*(t), X^*(t), Y^*(t)\}, t \in T, \quad (1)$$

где $E^*(t) \in E(t)$ – допустимое множество элементов системы в период t ; $A^*(t) \in A(t)$ – допустимое множество взаимосвязей элементов системы в период t ; $X^*(t) \in X(t)$ – допустимое множество собственных функциональных характеристик элементов системы в период t ; $Y^*(t) \in Y(t)$ – допустимое множество системных функциональных характеристик в период t при выполнении ресурсных ограничений, в рамках которых осуществляется развитие и функционирование системы: $R[E(t), A(t), X(t), Y(t), t]$.

Установим связь, между вероятностью своевременного выполнения БО (компетенцией исполнителя) и затратами на выполнение БО. В работе [11] предложен показатель затратности выполнения роботизированной БО вида

$$C^{BO} = \frac{\alpha_H e_H TO_H + \alpha_B e_B TO_B + (\alpha_H + \alpha_B) R_M}{C_H^b}, \quad (2)$$

где $\alpha_H = 1 + (1 - P_{SH})$, $\alpha_B = 1 + (1 - P_{SB})$ – коэффициенты, учитывающие увеличение времени и расходов вследствие вероятности возникновения ошибки h-агента и b-агента, P_{SH} и P_{SB} – соответственно вероятности успешного (безошибочного) завершения операций h-агента и b-агента – эти величины будем в дальнейшем использовать как численные меры компетенции агентов; C_H^b – базовая величина себестоимости бизнес-операции, выполняемой исключительно человеком и без ошибок

$$C_H^b = e_H^b \cdot TO_H^b + R_M^b, \quad (3)$$

где e_H^b , TO_H^b , R_M^b – соответственно, базовая цена рабочего времени h-агента, базовое нормативное время, отведенное по результатам стоимостного анализа на выполнение данной операции человеком, базовые нормативные затраты на эксплуатацию и обслуживание технических средств в течение базового нормативного времени TO_H^b .

Обозначив через cd вероятную относительную стоимость задержки выполнения БО, показатель затратности выполнения БП, содержащего I БО, запишем в виде

$$C^{BPI} = \sum_{i=1}^I (C_i^{BO} + CD_i), i = \overline{1, I}. \quad (4)$$

Основные положения стратегии управления очередями. В работе [11], изложен метод динамической маршрутизации заявок в ОТС, на основании которого и будем строить адаптивную стратегию управления очередями.

Выполнение заявок в ОТС регламентировано режимом реального времени. Введем временные ограничения T_{max}^v на обслуживание заявки v -й категории η_v . Соответственно определим интенсивность обслуживания μ_v (величина обратная к средней номинальной длительности $\mu_v = 1/E(\eta_v)$, где $E(\eta_v)$ – математическое ожидание длительности обслуживания заявки v -й категории. Очевидно, что величина μ_v находится в функциональной зависимости от компетенции конкретного исполнителя Z_n , определяющей его производительность. С учетом временных ограничений можно записать $\mu_v = f(Z_n) \geq 1/T_{max}^v$.

Поскольку каждый исполнительный механизм обслуживает несколько категорий заявок, необходимо определить стратегию выбора маршрута БП, сообразуясь с целью минимизации длительности или стоимости задержки выполнения, а также соблюдения квазистационарного режима, при котором длина очередей ко всем АРМ не возрастает.

Цели стратегии:

1. Минимизация критерия (4) для каждого БП.
2. Обеспечение квазистационарного режима обслуживания заявок на всех АРМ.

Поскольку функция распределения длительности обслуживания для каждого АРМ заранее не определена, но известны временные ограничения T_{\max}^V , то адаптивную стратегию назначения приоритетов определим следующим образом.

1. Примем, что критерием для адаптации стратегии является CD_{kn} – стоимость задержки выполнения k -й заявки (БО) на n -м АРМ.

2. Множество заявок классифицируется по ограничению CD_{\max} на 2 группы срочности (ГС) (urgency group): UG_1 , UG_2 . На любой текущий момент существуют подмножества заявок в каждой ГС. k -я заявка включается в группу UG_1 при выполнении условия: $CD_k^{max} > CD_{\max}$. Интенсивность контроля в «старшей» группе выше, чем в «младшей», что дает экономию машинного времени. Кратность периодичности циклов контроля и переформирования очередей в группах можно регулировать.

3. Выбор подмножества АРМ для маршрутизации по критерию компетенции.

3.1 Для маршрутизации каждой заявки определяется подмножество АРМ по требуемой для данной БО компетенции. Согласно методу, изложенному в работе [11], отбирается рабочее подмножество АРМ, у которых компетенция h -агента больше заданного порога компетенции. Среди этих АРМ происходит выбор конкретного узла обслуживания по минимальному значению стоимости ожидания.

3.2 Порог требуемой компетенции для заявки зависит от относительной цены задержки cd_k . Значение порога определяется выражением

$$P_k = \frac{a}{1 + \exp(-\alpha \cdot cd_k + b)} + c, \quad (5)$$

где параметры a , b , c определяют границы изменения порога компетенции (0,3...0,9), а параметр α – крутизну изменения выбираемого порога.

3.3 Чем шире подмножество компетенций в ОТС, тем ниже устанавливается порог отбора по компетенциям в рабочее подмножество АРМ.

3.4 При повышении порога компетенции рабочее подмножество теоретически может оказаться пустым, поэтому необходим контроль мощности сформированного подмножества.

На рис. 1 показаны кривые для определения индивидуального порога требуемой компетенции по формуле (5) при значениях параметра $\alpha=14$ (кривая а) и $\alpha=8$ (кривая б).

Видно, что параметр α существенно влияет на выбор рабочего подмножества АРМ по компетенции исполнителей.

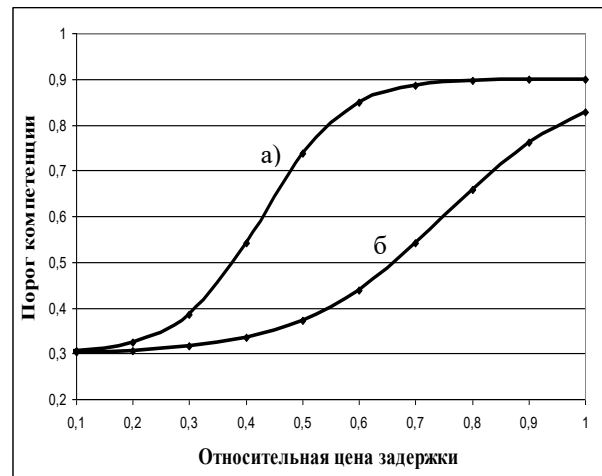


Рисунок 1 – Кривые выбора порога компетенции

4. Анализ очередей для выбора маршрута. После расчета суммарного времени выполнения впереди стоящих в каждой n -й очереди заявок TQ_n , с учетом цены задержки cd_k , можно вычислить стоимость задержки новой заявки $CD_{kn} = cd_k \cdot TQ_n$ для каждой n -й очереди из рабочего подмножества. Перебирая очереди в пределах рабочего подмножества, найдем минимальное значение CD_{kn}^* и поставим заявку k в очередь n^* .

5. Сортировка заявок в очередях по убыванию цены задержки cd_k .

6. Обеспечение квазистационарного режима по каждому АРМ.

Как известно, стационарный режим работы СМО должен быть независимым от режима работы системы (переходного или установившегося) [13]. Для обеспечения стационарного режима средние временные характеристики системы должны быть приведены в соответствие с временными ограничениями, которые определены исходными требованиями к системе. При динамической маршрутизации заявок можно обеспечить квазистационарный режим за счет обратной связи, когда диспетчер контролирует тенденции изменения длины очередей и принимает меры к тому, чтобы длина каждой очереди оставалась постоянной или уменьшалась. То есть, для любого канала обслуживания должно выполняться условие $\rho_n \leq \rho_T$, где ρ_T – пороговое значение загрузки, независимо от текущего распределения типов заявок.

6.1 Определение узких мест по компетенциям. Для каждого подмножества заявок типа v (по компетенции v) вычисляется текущее значение степени загрузки ρ_a средней за контрольный период скорости поступления заявок данного типа к средней скорости выхода обслуженных заявок данного типа. Если $\rho_a \geq \rho_{Tn}$, то по данной компетенции необходимо:

- ранжировать АРМ в рабочем подмножестве по уровню компетенции типа v ;
- с учетом установленного порога компетенции определить целевое АРМ с меньшим количеством заявок из группы риска UG_1 .

6.2. Если нагрузка по компетенции ниже нижнего порога загрузки ($\rho_a < \rho_{Tn}$), то есть, АРМ часто простаивает, осуществляется формирование очереди из заявок, стоящих в очередях на другие АРМ. Из хвостовых заявок в очередях к другим АРМ, образуется и ранжируется новая очередь в соответствии с компетенциями данного АРМ. Длина очереди выбирается по условию:

$$L_{\text{новой очереди}} \approx L_{\text{ср}} \text{ длины имеющихся очередей.}$$

6.3 Ранжировать очереди по убыванию относительной цены задержки.

Сформулированная выше стратегия была экспериментально апробирована на имитационной модели и показала свою работоспособность.

Таким образом получила дальнейшее развитие адаптивная стратегия управления очередями в ОТС, за счет учета в качестве признаков степени изменения потока неоднородных заявок, степени компетенции исполнителей, степени загрузки агентов-исполнителей, что позволяет реализовать процесс структурной адаптации ОТС к изменяющейся оперативной обстановке.

Разработка метода и алгоритма адаптации стратегии управления очередями. Для выбора параметров стратегии используется имитационная модель, описанная ниже. Рабочими переменными в процессе адаптации являются: порог $X1$ разделения заявок на группы риска по стоимости CD_k ; $X2$ – параметр α выражения (5) для определения порога компетенции.

Метод оптимизации реализуется в виде следующих этапов:

1. Для имитационного моделирования процесса многофазного обслуживания и формирования маршрутов заявок задаются вероятностные распределения потока поступления БП, значения компетенций, временные ограничения на выполнение заявок разных типов, соответствующих типам БО, относительные цены задержек БО, характерные для данной ОТС.

2. При помощи эволюционного алгоритма, изложенного ниже, генерируется популяция допустимых решений в виде двоек $(X1_r, X2_r)$, $r=1 \dots NP$, где NP – число членов популяции.

3. Каждая двойка (особь) $(X1_r, X2_r)$ используется в процессе имитационного моделирования, для вычисления значения критерия оптимальности.

4. При помощи эволюционных процедур, описанных ниже, отбирается наилучшее решение $(X1^*, X2^*)$ по критерию минимизации суммарной стоимости затрат и задержек (6)

$$Q = \sum_{j=1}^J \sum_{i=1}^{I_j} (C_{ij}^{BO} + CD_{ij}) \rightarrow \min, i = \overline{1, I}, j = \overline{1, J} \quad (6)$$

где J – количество запущенных в работу БП, I_j – количество БО в j -м БП.

Разработка эволюционного алгоритма для решения задачи оптимизации. Предварительные исследования показали, что критерий (5) является мультимодальной

функцией, которая к тому же изменяется для разных ОТС и бизнес-процессов. С учетом этого факта для решения поставленной задачи целесообразно использовать адаптивные методы случайного поиска, развитием которых являются эволюционные алгоритмы [14]. Основная продуктивная идея эволюционных алгоритмов состоит в том, что поиск ведется по всей области допустимых значений и в процессе его выполнения изменяются условия генерации вектора рабочих параметров. Обычно на начальных этапах распределение точек является равномерным, а потом количество точек увеличивается в тех областях, где предположительно находятся локальные экстремумы целевой функции. Уточнение значений локальных экстремумов можно производить с помощью случайного поиска в узкой области, которая предположительно содержит локальный экстремум. Вероятность быстрого схождения алгоритма при этом значительно возрастает и, тем самым, ускоряется весь процесс поиска оптимального решения. Таким образом, эволюционные алгоритмы используют информацию о рельефе поверхности функции цели. С учетом изложенного сформулируем основные правила, которые должен использовать алгоритм поиска оптимального решения:

1. Обеспечение равномерного поиска во всей области определения функции цели;

2. Обеспечение сужения границ случайного поиска в отдельных областях, которые предположительно содержат локальные экстремумы;

3. Узкие границы локального поиска адаптивно смещаются по мере уточнения значения локального экстремума;

4. Выбор для локального анализа не только элитарных точек, но и нескольких случайных точек;

5. С учетом того, что анализ популяции ведется при помощи имитационной модели, необходимо принять меры по сокращению объема вычислений при анализе популяции:

- ограничение количества членов популяции;
- дискретизация уровней рабочих переменных в диапазоне $[0,3 \dots 0,9]$ с шагом 0,1;
- проверка абсолютной идентичности новой особи с экземплярами, имеющимися в глобальной таблице особей.

Ниже приведены шаги алгоритма для реализации разработанного метода:

Шаг 1. Генерация исходной популяции. Синтезируется NP точек \overline{X}_k ($k = 1, \dots, NP$), в которых определяется значение целевой функции (6) (при помощи имитационной модели и усреднения за NI реализаций).

Шаг 2. Сортировка наборов \overline{X}_k (особей) по возрастанию значений (поскольку в решаемой задаче ищется минимум целевой функции).

Шаг 3. Из всех P точек отбирается η первых точек – претендентов на лидерство.

Шаг 4. Вычисляются границы поиска уточненно-го значения экстремума для каждой из η выбранных точек:

$$a' = x^* - A(b-a)/2, \quad b' = x^* + A(b-a)/2, \quad (7)$$

где a', b' – суженные границы поиска; a, b – начальные границы поиска; x^* – координата выбранного решения; A – размер границы для уточненного поиска экстремума.

Шаг 6. Генерируется M точек в границах a', b' .

Шаг 7. Для каждой точки $X_l, l \in M$ вычисляется значение целевой функции. Если новое значение меньше предшествующего, соответственно изменяются значения x^* .

Как результат в массиве популяции для выбранных точек фиксируются новые значения координат и локальных экстремумов.

Шаг 8. Координаты и значения целевой функции η лидеров заносятся в массив лидеров.

Шаг 9. Если количество заданных итераций не исчерпано, то для всех особей генерируются новые случайные значения координат и происходит переход к шагу 3. Иначе переход к шагу 10.

Шаг 10. Массив лидеров сортируется по возрастанию целевой функции. Первый элемент в списке особей является найденным оптимальным решением. Конец.

Алгоритм имеет следующие параметры настройки:

- 1) NP – количество членов популяции.
- 2) NI – количество реализаций в имитационной модели.
- 3) η – количество претендентов в элиту.
- 4) A – коэффициент уменьшения границ для уточненного поиска экстремума.
- 5) M – количество пробных точек в узких границах уточненного поиска.

Каждая эпоха оптимизации режимов работы СМО оставляет в качестве результата исходные данные сеанса моделирования и лучший набор значений рабочих параметров. Сохранение результатов в отдельной таблице в дальнейшем позволяет сразу находить значения параметров XI, XI близкие к оптимальным.

Таким образом, разработан метод и алгоритм адаптации стратегии управления очередями в ОТС. При изменении режима работы ОТС (изменения относительного числа разнородных БП и БО) появляется возможность адаптивного регулирования значений порогов XI, XI , что дает возможность при изменении распределения входных потоков выбрать наилучшие условия многофазного обслуживания в процессе случайного поступления требований на выполнение бизнес-процессов известных классов.

Имитационная модель. Поскольку аналитически решить сформулированную выше задачу не представляется возможным, был выбран путь имитационного моделирования. Использование имитацион-

ной модели позволяет провести интересные в теоретическом и практическом отношении исследования СМО. В первую очередь это задачи анализа системы. К ним относятся определение показателей эффективности и других свойств системы по известным ее параметрам: интенсивности потока требований, количеству каналов и их характеристикам, времени обслуживания и т. д. Большой интерес представляет исследование влияния вариаций параметров системы на показатели, характеризующие ее основные свойства. В частности, имитационная модель необходима для исследования эффективности различных вариантов стратегий оперативного управления режимами работы ОТС.

Структура имитационной модели показана на рис. 2. Все операции программы диспетчирует блок управления ходом моделирования. Пользователь через блок интерфейса и с помощью блока формирования исходных данных фиксирует желаемые вероятностные распределения потока поступления заявок, режимы генерации потоков событий, матрицу значений компетенций h -агентов, устанавливает множество типов заявок и нормы времени выполнения в зависимости от компетенции. Заполняется также массив цен задержек выполнения БО. Зависимость времени выполнения от уровня компетенции h -агента определяется коэффициентом увеличения времени $\alpha_H = 1 + (1 - P_{SH})$, который учтен в формуле (2).

После фиксации исходных данных блок управления запускает генератор потока событий и блок маршрутизации, который использует текущие значения пороговых параметров стратегии управления очередями. Блок формирования результатов производит мониторинг процесса по параметрам загрузки узлов сети обслуживания, вычисляет и фиксирует текущие и результирующие значения таких параметров: среднее за контрольный период число заявок в очередях; среднее время ожидания; дисперсия времени ожидания; потери ожидания; степень загрузки APM .

Блок эволюционных процедур реализует алгоритм отбора лучших вариантов сочетаний значений пороговых параметров и формирует глобальную таблицу особей для предотвращения повторения особей.

Блок генерации и вывода отчетов формирует отчет по результатам моделирования для пользователя.

Пользователь может запустить цикл автоматической смены значений основных исходных данных (тип распределения для входного потока, шаговые изменения интенсивности поступления и обслуживания заявок, распределение типов заявок).

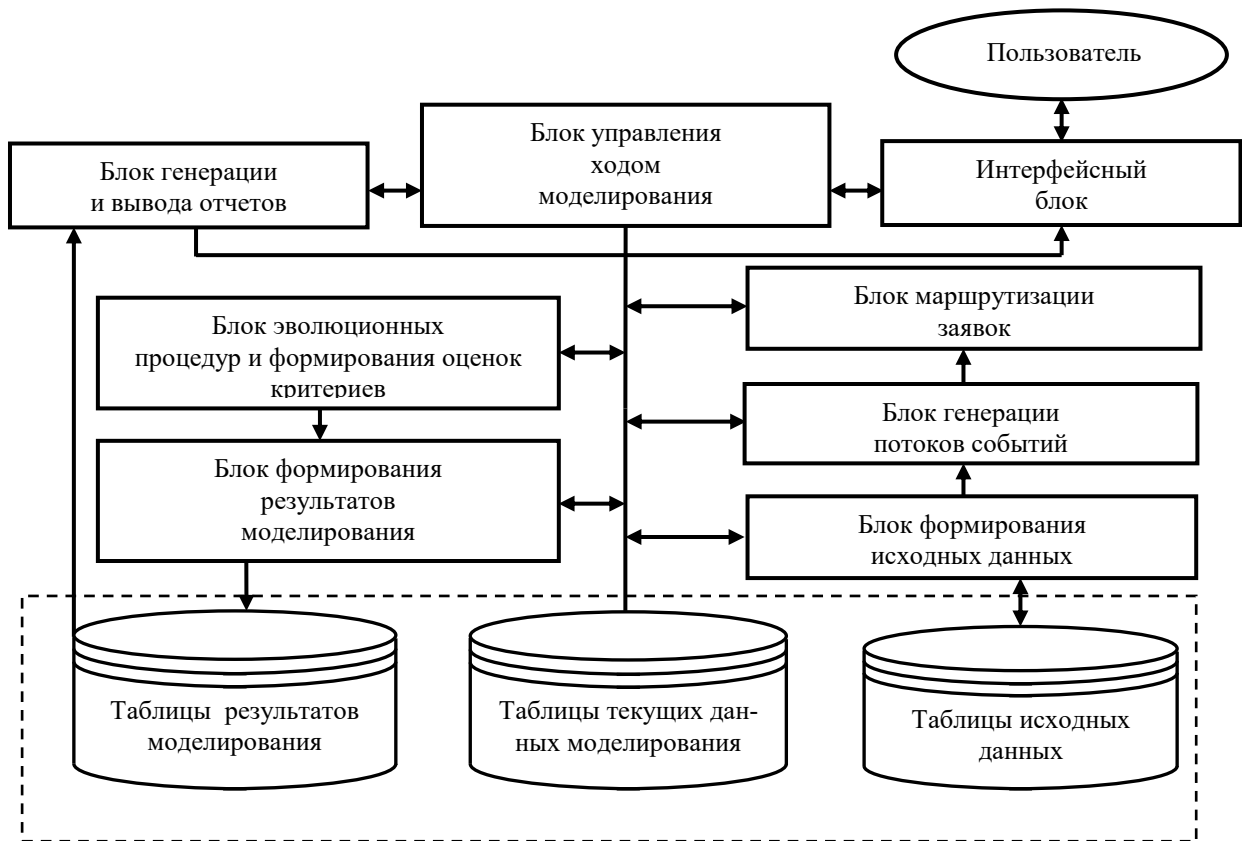
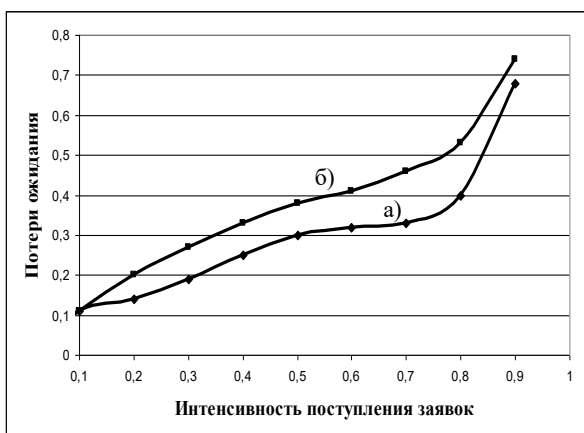


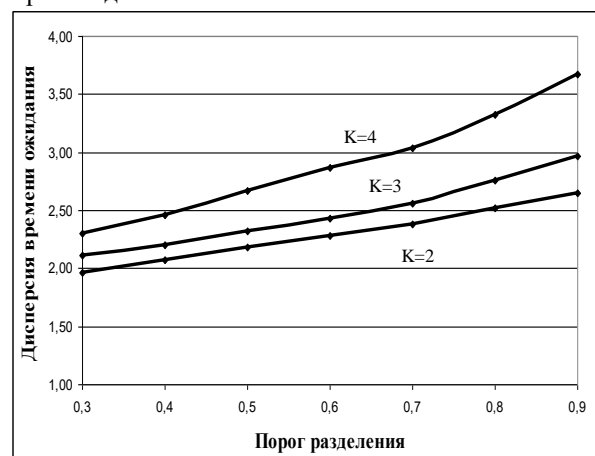
Рисунок 2 – Структура имитационной модели

Результаты экспериментов. Главным результатом экспериментов на имитационной модели – подтверждение гипотезы о том, что предложенная стратегия управления очередями работоспособна и может быть адаптирована к изменяющимся условиям работы ОТС, а именно – к изменениям интенсивности потоков разнородных заявок, степени компетенции исполнителей, степени загрузки АРМ.

На рис. 3 показана зависимость дисперсии времени ожидания от значений порога разделения заявок на группы по стоимости задержек. При увеличении кратности периодичности циклов контроля и перформирования очередей дисперсия увеличивается.

Рисунок 3 – Зависимость дисперсии времени ожидания от порога разделения заявок на группы и кратности K периодичности циклов контроля очередей в группах

На рис. 4 показана зависимость критерия потерь от средней интенсивности поступления заявок в систему. Кривая а) отражает эту зависимость при случайных значениях параметров $X1$ и $X2$. Кривая б) показывает ту же зависимость при оптимальных значениях параметров $X1$ и $X2$. На кривой б) виден участок стабилизации потерь за счет оперативного изменения маршрутов заявок. По мере увеличения интенсивности потока заявок стабилизация становится практически невозможной и потери возрастают. Вместе с тем видно, что адаптивная стратегия является эффективным инструментом снижения потерь ожидания.

Рисунок 4 – Зависимость потерь ожидания от интенсивности поступления заявок. Кривая а) – при случайных значениях параметров $X1$ и $X2$; кривая б) – при оптимальных значениях этих параметров

Среднее нормированное время пребывания в очереди для заявок из группы UG1 составляло 0,63 при случайных значениях параметров настройки и 0,45 при оптимальных значениях этих параметров. Для группы UG2 эти значения соответственно составили 0,68 и 0,56.

ВЫВОДЫ. Получила дальнейшее развитие адаптивная стратегия управления очередями в ОТС, за счет учета в качестве признаков степени изменения потока неоднородных заявок, степени компетенции исполнителей, степени загрузки агентов-исполнителей, что позволяет реализовать процесс структурной адаптации ОТС к изменяющейся оперативной обстановке.

Разработан метод и алгоритм адаптации стратегии управления очередями в ОТС, основанный на эволюционном подходе к подбору оптимальных значений пороговых параметров отбора заявок по стоимости задержки и подбора компетенции сотрудников. При изменении режима работы ОТС (изменения относительного числа разнородных бизнес-процессов и бизнес-операций) происходит адаптивное регулирование значений указанных порогов, что дает возможность при различных распределениях интенсивностей входных потоков выбрать наилучшие условия многофазного обслуживания и снизить потери ожидания.

ЛИТЕРАТУРА

1. Аксенов К. А., Гончарова Н. В. Динамическое моделирование мультиагентных процессов преобразования ресурсов. Екатеринбург: ГОУ ВПО УГТУ-УПИ, 2006. 311 с.
2. Введенская Н. Д., Добрушин Р. Л., Карпелевич Ф. И. Система обслуживания с выбором наименьшей из двух очередей – асимптотический подход. *Проблемы передачи информации*. 1996. Т. 32. № 1. С. 20–34.
3. Suhov Yu. M., Vvedenskaya N. D. Fast Jackson Networks with Dynamic Routing. *Problems of Information Transmission*. 2002. Vol. 38. No. 2. PP. 136–153.
4. ЖерновЫй Ю. В. Многолинейные системы с отказами и равновероятным распределением заявок. *Информационные процессы*. 2013. Том 13. № 1. С. 19–34.
5. ЖерновЫй Ю. В., ЖерновЫй К. Ю. Метод потенциалов для системы M/G/1/m с активным управлением очередью. *Информационные процессы*. 2015. Том 15. № 1. С. 66–77.
6. Аленичев А. В. Система массового обслуживания с динамической маршрутизацией и распределением Вейбулла времени обслуживания. *Информационные процессы*. 2005. Том 5. № 5. С. 432–442.
7. Аленичев А. В., Лиханов Н. Б. Динамическая маршрутизация в системе с заявками, имеющими степенной закон распределения времени обслуживания. *Информационные процессы*. 2005. Т. 5. № 3. С. 213–226.
8. Корячко В. П., Перепелкин Д. А. Математическая модель адаптивной маршрутизации в корпоративных сетях. *Информационные технологии в образовании*. 2012. С. 47–55.
9. Поздняк И. С. Формирование множества допустимых маршрутов с использованием алгоритма адаптивной маршрутизации. *Тез. докл. XIV российской научной конференции ПГАТИ* (г. Самара, Россия, 2007). С. 61–62.
10. Оксанич И. Г., Шевченко И. В., Краснопольская Ю. А. Отображение описания бизнес-процесса в операционное пространство организационно-технической системы. *Радиоэлектроника та інформатика*. 2019. № 2. С. 54–60.
11. Шевченко И. В., Оксанич И. Г., Конох И. С. Модель и метод динамического формирования исполнительных структур в роботизированных организационно-технических системах. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. 2019. Вип. 5 (118). С. 103–109.
12. Габалин А. В., Разбегин В. П. Анализ и синтез структуры Workflow-систем. *Международная конференция «Системы проектирования, технологической подготовки производства и управления этапами жизненного цикла промышленного продукта CAD/CAM/PDM»* (г. Москва, Россия, 26-28 октября 2011 г.) С. 34.
13. Алиев Т. И. Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.
14. Курейчик В. М. Генетические алгоритмы и их применение. Таганрог: ТРТУ, 2002. 242 с.

ADAPTIVE QUEUE MANAGEMENT STRATEGY IN THE AUTOMATED ORGANIZATIONAL-TECHNICAL SYSTEM

I. Shevchenko, I. Oksanych, P. Savushkin

Kremenchuk Mykhailo Ostrohradskyi National University

vul. Pershotravneva, 20, Kremenchuk, 39600, Ukraine. E-mail: oksirena2017@gmail.com

Purpose. The goal is to develop an adaptive queue management strategy when choosing routes for many heterogeneous business processes and changing application flows within the organizational and technical system, taking into account a certain criterion of optimality and some restrictions. **Methodology.** Research methods are based on methods of system analysis, methods of queuing systems, probability theory. When solving the problem, the following factors were taken into account: heterogeneity of applications by the parameter vector due to the diversity of processes; cost of delayed operations; the uncertainty of the distribution of flows of heterogeneous applications at the input of the system; heterogeneity of service devices due to limitations in the competencies of employees; the same device serves several types of applications; the service mode changes according to an arbitrary law, from a complete stop to an overload. Based on these factors, an adaptive strategy for changing application routes has been developed. A method has also been developed for selecting the optimal thresholds for the level of competency when selecting performers and the threshold

for dividing applications into groups according to the cost of delays. **Originality.** The adaptive queue management strategy was further developed by taking into account, as signs of the degree of change in the flow of heterogeneous applications, the degree of competence of the executors, the degree of loading of the executors, which makes it possible to implement the process of structural adaptation of the system to a changing operational environment. A method and algorithm for adapting a queue management strategy based on an evolutionary approach has been developed. Using the evolutionary algorithm, the optimal values of the thresholds for selecting applications are selected according to the cost of delay and the level of competence of employees. **Practical value.** When changing the operating mode, adaptive regulation of the values of these thresholds occurs. This makes it possible to select the best conditions for multiphase maintenance with various distributions of the intensities of the input flows and reduce the loss of waiting. Each epoch of optimization of the system operating modes leaves as the result the initial data of the simulation session and the best set of operating parameters. Saving the results in a separate table in the future allows you to immediately find the values of the operating parameters close to optimal. References 14, figure 4.

Key words: organizational system, application routing, management strategy, adaptation.

REFERENCES

1. Aksenov, K.A., Goncharova, N.V. (2006), *Dynamic modeling of multi-agent resource conversion processes* [Dinamicheskoe modelirovanie multiagentnykh protsessov preobrazovaniia resursov], GOU VPO UGTU-UPI, Ekaterinburg, Russia, 311 p.
2. Vvedenskaya, N.D., Dobrushin, R.L., Karpelevich, F.I. (1996), "Service system with the choice of the smallest of the two queues - asymptotic", *Problems of Information Transmission*, vol. 32, no. 1, pp. 20-34.
3. Suhov, Yu.M., Vvedenskaya, N.D. (2002), "Fast Jackson Networks with Dynamic Routing", *Problems of Information Transmission*, vol. 38, no. 2, pp. 136-153.
4. Zhernovyi, Yu.V. (2013), "Multiserver systems with failures and equiprobable distribution of customers", *Information Processes*, vol. 13, no. 1, pp. 19-34.
5. Zhernovyi, Yu.V., Zhernovyi, K.Yu. (2015), "Method of potentials for the system M/G/1/m with active queue management", *Information Processes*, vol. 15, no. 1, pp. 66-77.
6. Alenichev, A.V. (2005), "Queuing system with dynamic routing and Weibull distribution of service time", *Information Processes*, vol. 5, no. 5, pp. 432-442.
7. Alenichev, A.B., Likhanov, N.B. (2005), "Dynamic routing in a system with applications having a power law distribution of service time", *Information Processes*, vol. 5, no. 3, pp. 213-226.
8. Koriachko, V.P., Perepelkin, D.A. (2012), "The mathematical model of adaptive routing in corporate networks", *Information Technology in Education*, pp. 47-55.
9. Pozdniak, I.S. (2007), "Formation of the set of valid routes using the adaptive routing algorithm", *Tez.dokl. XIV rossiiskoi nauchnoi konferentsii PGATI*, Samara, Russia, pp. 61-62.
10. Oksanych, I.G., Shevchenko, I.V., Krasnopol'skaja, J.A. (2019), "Displaying a description of a business process in the operational space of an organizational-technical system", *Radioelectronics & Informatics*, no. 2, pp. 54-60.
11. Shevchenko, I., Oksanych, I., Konokh, I. (2019), "Model and method of dynamic formation of executive structures in robotized organizational-technical systems", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, iss. 5 (118), pp. 103-109.
12. Gabalin, A.V., Razbegin, V.P. (2011), "Analysis and synthesis of the structure of Workflow systems" ["Analiz i sintez struktury Workflow-sistem"], *International Conference "CAD / CAM / PDM Systems for Design, Technological Preparation of Production and Life Cycle Management of an Industrial Product [Mezhdunarodnaia konferentsiia «Sistemy proektirovaniia, tekhnologicheskoi podgotovki proizvodstva i upravleniia etapami zhiznennogo tsikla promyshlennogo produkta CAD/CAM/PDM»]*, Moscow, Russia, 26-28 October, pp. 34.
13. Aliev, T.I. (2009), *Discrete Systems Modeling Basics* [Osnovy modelirovaniia diskretnykh sistem], SpbGU ITMO, Sankt Peterburg, Russia, 363 p.
14. Kureichik, V.M. (2002), *Genetic Algorithms and Their Application* [Geneticheskie algoritmy i ikh primeneniie], TRTU, Taganrog, Russia, 242 p.

Стаття надійшла 10.12.2019.