Моделирование глобальных процессов
Нелинейная динамика и гуманитарные науки

УДК 519.872 https://doi.org/10.18500/0869-6632-2019-27-5-53-71

Каталитическая модель массового обслуживания на примере циклической очереди

А. Т. Мустафин[1], *А. К. Кантарбаева*[2]

[1]Казахский национальный исследовательский технический университет им. К. И. Сатпаева
Казахстан, 050013 Алматы, ул. Сатпаева, 22
[2]Казахский национальный университет им. аль-Фараби
Казахстан, 050040 Алматы, просп. аль-Фараби, 71
E-mail: butsuri123@gmail.com, ratnakka@gmail.com
Автор для переписки Мустафин Алмаз Тлемисович, butsuri123@gmail.com
Поступила в редакцию 29.07.2019, принята к публикации 26.08.2019, опубликована 31.10.2019

Сформулирована и исследована детерминистическая («жидкостная») модель системы массового обслуживания с открытой на вход и выход циклической очередью с возможностью отказов от получения услуги. Модель представляет собой систему нелинейных обыкновенных дифференциальных уравнений, переменными которых являются клиенты, ожидающие обслуживания, клиенты, получившие услугу, и серверы (каналы обслуживания).

Цель работы – вывести минимальную математическую модель процесса оказания услуги как обобщённой псевдохимической реакции с катализатором. Особенность подхода состоит в представлении о провайдере услуги как о своего рода ферменте-катализаторе, который осуществляет трансформацию клиентов из категории ожидающих обслуживания в категорию получивших услугу, возвращаясь в исходное состояние после каждого акта конверсии. С междисциплинарной точки зрения, биохимическую реакцию и взаимодействие потребителя с провайдером объединяет образование короткоживущего комплекса – это соответственно фермент-субстрат и клиент-сервер. Построив базовую модель акта услуги, мы показываем её работоспособность на примере циклической системы массового обслуживания. Сформулированная модель циклической очереди исследуется **методами** качественной теории дифференциальных уравнений. Известный из практики факт, что среднее время обслуживания намного короче характерного времени ожидания, делает исходную систему уравнений сингулярно возмущённой. Методом многих масштабов система расщепляется на медленную и быструю подсистемы, описывающие соответственно динамику клиентов и серверов. **Результаты.** Показано, что в адиабатическом приближении количество занятых обслуживанием серверов мгновенно отслеживает текущую длину очереди (то есть спрос) в соответствии с известными соотношениями для квазистационарных концентраций ферментативной кинетики. Найдено физически допустимое стационарное решение медленной подсистемы и доказана его асимптотическая устойчивость. Проведён параметрический анализ стационарного состояния. Получен практический вывод о том, что стационарный спрос на услугу остаётся низким независимо от частоты отказов при входящем потоке клиентов, не превышающем некоторый критический уровень. Предложенный формализм также позволяет вывести в аналитическом виде клиринговую функцию – зависимость скорости обслуживания от текущего спроса. Показано, что формула клиринговой функции применима во всех случаях, когда время обслуживания меньше времени ожидания, и необязательно только в стационарном режиме работы системы.

Ключевые слова: очередь, отказ, повторное требование, время обслуживания, время ожидания, клиринговая функция.

Образец цитирования: Мустафин А. Т., Кантарбаева А. К. Каталитическая модель массового обслуживания на примере циклической очереди // Известия вузов. ПНД. 2019. Т. 27, № 5. С. 53–71. https://doi.org/10.18500/0869-6632-2019-27-5-53-71 https://doi.org/10.18500/0869-6632-2019-27-5-53-71

A catalytic model of service as applied to the case of a cyclic queue

A. Mustafin[1], *A. Kantarbayeva*[2]

[1]K. I. Satbayev Kazakh National Research Technical University
22 Satbayev St., Almaty, 050013, Kazakhstan
[2]al-Farabi Kazakh National University
71 al-Farabi Ave., Almaty, 050040, Kazakhstan
E-mail: butsuri123@gmail.com, ratnakka@gmail.com
Correspondence should be addressed to A. Mustafin, butsuri123@gmail.com
Received 29.07.2019, accepted for publication 26.08.2019, published 31.10.2019

The research is devoted to the development of a deterministic («fluid») model for the open cyclical service system with abandonment and re-entry. The model is the system of coupled nonlinear ordinary differential equations for the following variables: (i) potential customers awaiting the service in the queue, (ii) served customers, (iii) busy servers, and (iv) free servers.

Aim of the work is to derive a minimal mathematical model of the service process treated as a generalized pseudochemical reaction with catalyst. The key feature of our approach is the vision of service provider as a kind of enzyme that shifts customers from the category of «waiting in a queue» to the category of «served». The catalyst-facilitator does not get used up in the process and can continue to act repeatedly. From the interdisciplinary perspective, both the biochemical reaction, and the consumer-provider interaction share a common trait: formation of a short-lived intermediate complex (enzyme-substrate and client-server, respectively). Having constructed a basic model of a service act, we verify its workability by the example of a cyclic service system. The formulated model of a cyclic queue is investigated by **methods** of the qualitative theory of differential equations. The empirical fact that the average service time is much shorter than the characteristic waiting time makes the original system of equations singularly perturbed. Using the multiple-scale method the system is decomposed into slow and fast subsystems governing the respective dynamics of clients and servers. **Results.** Under the adiabatic approximation, the number of busy servers is shown to hastily instantly adapt to the momentary queue length (i.e. demand) in accordance with the well-known relationships for quasi-steady-state concentrations of enzyme kinetics. The physically feasible steady-state of the slow subsystem is found and proven to be asymptotically stable. A parametric analysis of the model's steady state is performed. A practical conclusion has been drawn, that that as long as the arrival rate of new customers remains below a certain threshold value, the steady-state demand will keep relatively low regardless of the abandonment rate. The proposed formalism also allows us to derive analytically the clearing function–functional response of the output (number of served customers in a unit of time) to the current demand, and to suggest the conditions of its applicability. It is shown that the clearing function formula remains valid in all cases when the service time is shorter than the waiting time, and not necessarily only in the steady-state mode of operation.

Key words: queue, abandonment, re-entry, service time, waiting time, clearing function.

Reference: Mustafin A., Kantarbayeva A. A catalytic model of service as applied to the case of a cyclic queue. *Izvestiya VUZ. Applied Nonlinear Dynamics*, 2019, vol. 27, no. 5, pp. 53–71. https://doi.org/10.18500/0869-6632-2019-27-5-53-71

Введение

Исторически сфера услуг долгое время относилась к периферийной экономической деятельности, не подпадавшей под категории сельского хозяйства и промышленности. Однако с 1980-х годов её доля в национальном хозяйстве большинства стран мира неуклонно расширялась. К настоящему времени вклад сферы услуг составляет две трети мирового валового внутреннего продукта, тогда как во многих развитых промышленных странах этот показатель превышает три четверти [1].

Быстрое развитие сферы услуг в современной экономике оживило интерес к моделям многоканальных систем массового обслуживания. Применение таких моделей особенно актуально в крупных системах обслуживания, занимающихся обработкой обращений и информированием по телефону (колл-центры), факсу, электронной и обычной почте, в интернет-чате (контакт-центры), а также в клиниках и других учреждениях здравоохранения [2–5]. Исследование операций, разделом которого является теория массового обслуживания, рассматривает оказание услуги и производство товара с единой кибернетической точки зрения – как операцию «вход–выход», выполняемую некоторым ящиком-преобразователем [6]:

$$\underset{\text{(Input)}}{\text{Вход}} \longrightarrow \boxed{\underset{\text{(Converter)}}{\text{Преобразователь}}} \longrightarrow \underset{\text{(Output)}}{\text{Выход}} \tag{1}$$

В случае промышленного производства вход – это затраты в виде сырья или сборочных деталей. На фабрике они превращаются рабочими и машинами в конечный выпускаемый продукт (выход).

В случае обслуживания вход и выход необязательно неодушевлённые и материальные объекты – это могут быть, например, физические лица или информация. При этом превращение затрат в выпуск может показаться менее очевидным. При оказании услуги трансформации подвергается не физическая сущность того, что поступает на вход, но его состояние. Согласно получившему широкое распространение определению, предложенному Т. Хиллом [7], услуга – это изменение состояния человека или имущества, принадлежащего некоторой экономической единице, которое осуществляется в результате деятельности другой экономической единицы по предварительному согласию ранее упомянутых человека или экономической единицы. В отличие от продукта, который сразу же после его выпуска приобретает обособленное положение по отношению как к производителю, так и к потребителю, услуга единосущна и с провайдером, и с клиентом, её невозможно хранить, она представляет собой не какой-то результат, а, скорее, акт или процесс [8].

Вслед за Хиллом Ж. Гадре [9] предложил метафорическое понятие «треугольника услуги», в терминах которого он дал следующее определение: услуга есть действие, направленное на изменение состояния сущности C, принадлежащей потребителю B или находящейся у него в пользовании, которое осуществляется провайдером услуги A по просьбе B и, как правило, в сотрудничестве с последним, но не приводит к выпуску товара, способного обращаться в экономике независимо от среды C.

Рис. 1, построенный по работе Х. Фромма [6], иллюстрирует различные сценарии, возникающие в зависимости от участия физических лиц и их собственности в процессе предоставления услуги.

Чтобы предоставлять услуги, провайдер, вообще говоря, должен иметь как квалифицированный персонал (работников с необходимыми навыками, компетенциями и знаниями), так и материальные ресурсы (такие как помещения, оборудование, инструменты и расходные материалы). Схема на рис. 1 охватывает пять различных мыслимых ситуаций.

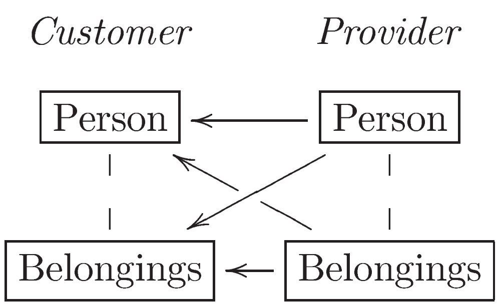

Рис. 1. К определению услуги по Т. Хиллу и его последователям. В общем случае потребителем (customer) услуги может быть физическое лицо (person) или имущество (belongings) этого лица. С другой стороны, провайдер (provider) услуги также может располагать специально обученным персоналом и оборудованием

Fig. 1. To the definition of service by T. Hill and his followers. In a general case, a customer, as the receiver of service, may be a person or the belongings of a person. On the other hand, a provider may also have people (i. e., persons with skills, competencies, and knowledge) and belongings (like facilities, tools, and materials) to deliver a service

1) Физическое лицо (со стороны провайдера) оказывает услугу другому лицу (со стороны потребителя) непосредственно. Имущества обеих сторон в процессе не участвуют. Напри-

мер, репетитор даёт частные уроки школьнику. При этом репетитор преобразует менее подготовленного ученика (вход) в более подготовленного (выход).

2) Физическое лицо, оснащённое оборудованием (со стороны провайдера), воздействует на другое физическое лицо (клиента). Например, оказывая лечебную услугу пациенту, врач использует специально оборудованные помещения (палату, операционную, койку), инструменты (стетоскоп, шприц, скальпель) и расходные материалы (медикаменты, перевязочный материал, дезинфицирующие средства). Тем самым клиника трансформирует больных (вход) в здоровых (выход).
3) Физическое лицо, оснащённое оборудованием (со стороны провайдера), воздействует на имущество потребителя услуги. Сам клиент в процессе не задействован. Такая ситуация имеет место при сдаче вещей в химчистку, техническом обслуживании и ремонте, пересылке почты и доставке грузов. Например, автомеханик трансформирует неисправный автомобиль (вход) в рабочий (выход).
4) Случай самообслуживания, когда услуга предоставляется провайдером без обслуживающего персонала. Провайдер устанавливает средства обслуживания и оборудование, которые клиент может использовать самостоятельно для изменения своего состояния. Примерами служат торговые автоматы, таксофоны, прокат автомобилей.
5) Имущество провайдера воздействует на имущество потребителя без участия людей с обеих сторон. Например, программное приложение, установленное в компьютере клиента, может самостоятельно запрашивать и получать регулярные обновления с удалённого сервера провайдера.

Занимаясь моделированием производства и систем массового обслуживания, исследование операций типично придерживается подхода «чёрного ящика», когда речь заходит о внутреннем устройстве «преобразователя», показанного на схеме (1). Подход чёрного ящика позволяет описывать акт трансформации, пользуясь информацией на интерфейсе, непосредственно доступной наблюдателю. В то же время «внутренняя кухня» преобразователя остаётся скрытой. Как следствие, любой оригинальный механизм трансформации можно заменить подходящей феноменологической передаточной функцией при условии, что она достаточно точно имитирует интерфейс. Подход чёрного ящика во многих случаях оправдан, так как робастная прикладная модель должна описывать производство или обслуживание, не будучи перегруженной детальной информацией о внутреннем устройстве. Самый известный пример чёрного ящика-преобразователя, используемого в исследовании операций, – это производственная функция.

Производственная функция представляет собой эконометрическое соотношение, устанавливающее связь между скоростью производства некоторого конечного либо промежуточного продукта Y (то есть выпуском) и набором факторов производства $\phi_1, \phi_2, \ldots, \phi_n$ (напр., [10]):

$$Y = F(\phi_1, \phi_2, \ldots, \phi_n). \tag{2}$$

В широком смысле фактор производства – это всё, что при увеличении его доступности способствует росту выпуска. Факторы производства могут иметь материальную, энергетическую, человеческую и финансовую природу. Выделяют два типа факторов: потребляемые и непотребляемые. Потребляемые факторы производства обычно называют вводимыми ресурсами или затратами. Под потреблением понимается необратимое превращение, физическое воплощение ресурса в материальный продукт. В процессе производства ресурс непременно потребляется, расходуется. Потребление ресурса означает тенденцию к уменьшению его доступности. Непотребляемые факторы производства, широко известные как первичные факторы производства, к которым классически относят землю, капитал и труд, часто объединяются под общим названием «фонды». Они не являются ресурсами в смысле данного выше определения. Фонды материально не преобразуются в продукт, который они производят. Они представляют собой инструменты, которые превращают вводимые ресурсы в продукт, но сами при этом физически не становятся частью продукта. Хотя фонды не расходуются, их количество может изменяться во времени по воле производителя, и они подвержены износу (амортизации). В отношении размерности выпуск Y в формуле (2), будучи количеством товара, произведённого за единицу времени, является потоковой переменной. Вводимые ресурсы чаще всего потоки, хотя в некоторых случаях могут быть и запасами. Фонды же всегда имеют размерность «запас».

Частной разновидностью производственной функции является клиринговая функция, понятие которой впервые появилось в управлении цепочками снабжения (напр., [11]). Она активно применяется в планировании производства. Клиринговая функция представляет собой детерминистическое соотношение между стационарным выпуском продукции Y и текущим объёмом ресурса (материалов, заготовок), вовлечённого в технологический процесс производства, W (который в англоязычной литературе носит название work-in-progress или WIP): $Y = f(W)$. Входная переменная W в этой формуле имеет размерность «запас».

Главная задача настоящей работы состоит в попытке дать правдоподобную реконструкцию событий, происходящих в чёрном ящике услуги, понимаемой как трансформация состояния клиента в результате действий провайдера. Иными словами, мы предлагаем минимальную нефеноменологическую модель процесса обслуживания, которая выводится не из регрессионного анализа, но скорее апеллирует к некоторым «первым принципам».

Мы преполагаем, что изменение состояния потребителя услуги в результате его взаимодействия с провайдером происходит примерно так же, как превращение молекулы субстрата в молекулу другого вещества в биохимической реакции, катализируемой ферментом. В самом деле, в живой клетке молекула субстрата S связывается с молекулой фермента E, образуя короткоживущий фермент-субстратный комплекс $[ES]$. Этот комплекс затем распадается на молекулу продукта P и исходную молекулу фермента, которая способна катализировать новую реакцию (напр., [12]):

$$S + E \rightleftarrows [ES] \rightarrow P + E. \tag{3}$$

На схеме (3), как принято в химической кинетике, стрелки показывают направление реакций. Двойная стрелка на первой стадии процесса превращения обозначает обратимую реакцию.

Трактуя услугу как разновидность процесса превращения, можно отождествить входящего (потенциального) клиента с субстратом, выходящего (обслуженного) клиента – с продуктом, а провайдера услуги, представленного каналом обслуживания (здесь и далее называемого сервером) – с ферментом. В приложении к процессу производства мы ассоциируем ресурс с субстратом, конечный товар с продуктом и фонды (работников и машины) – с ферментом. Сказанное выше сведено в таблицу.

Таблица. Биохимический катализ в сравнении с производством и обслуживанием
Table. Biochemical catalysis vs. production and service

	Агенты со сходными функциями Agents with similar functions		
Преобразовательный процесс Transformation process	Вход Input	Выход Output	Преобразователь Converter
Ферментативная реакция Enzymatic reaction	Субстрат Substrate	Продукт Product	Фермент Enzyme
Производство Manufacturing	Ресурс Resource	Готовое изделие Finished goods	Фонды Funds
Обслуживание Service	Потенциальный клиент Potential customer	Обслуженный клиент Served customer	Провайдер Provider

Впервые мысль о том, что такой фактор производства, как труд, подобен химическому катализатору, высказал Н. Георгеску-Рёген: «Предполагается, что труд, как обязательный фактор при производстве любого товара, может неограниченно заменяться другими факторами. Иными словами, в любой отрасли хозяйства выпуск произвольно большого объёма можно получить с использованием сколь угодно малого количества труда при условии, что другие ресурсы будут доступны в неограниченном количестве. Поскольку тут просматривается явная аналогия с катализатором в химической реакции, то можно говорить, что труд выступает катализатором в отрасли G_k, если он обладает указанным свойством в этой отрасли» (перевод наш. – *Авторы*) [13, p. 319].

Более детально аналогия между биосинтезом и промышленным производством была проработана И. А. Полетаевым, по мысли которого «типичным прообразом процессов типа производства–потребления является химическая реакция на катализаторе и процесс промышленного производства. В процессах этих типов участвуют три рода компонентов: входные компоненты (субстрат реакции, сырьё), фондовые компоненты (катализатор, оборудование) и выходные компоненты (продукт реакции, изделия)» [14, с. 72].

Практически в те же годы, но независимо, Д. С. Чернавский указывал: «По духу и методам исследования к излагаемому нами предмету близко примыкает моделирование экономических, производственных процессов. Нет ничего удивительного в том, что биологические системы с их основными переменными – концентрациями веществ – похожи на экономические, где в качестве переменных выступают количества тех или иных продуктов или предметов, а роль концентрации ферментов играет число станков в цехе или автоматической линии. В этом смысле и кинетические модели биофизики и биохимии, и модели экономические являются частями одной общей отрасли кибернетики, так называемой теории сложных систем» [15, с. 134].

К сожалению, предложенная перечисленными исследователями перспективная аналогия осталась надолго невостребованной в теории исследования операций. Положение дел начало меняться после признания эволюционной экономики [16] самостоятельным научным направлением и после появления концепции «промышленного метаболизма» [17]. С тех пор развитие моделирования в логистике сопровождалось растущим осознанием того, что мультиферментная цепь биохимического синтеза в живой клетке подобна производственной линии на фабрике, где продукт, производимый одной машиной, используется другой машиной как ресурс для её собственного выпуска [18]. В последние годы инструментальные методы исследования операций всё чаще находят себе применение в изучении внутриклеточных метаболических сетей [19, 20]. Данное направление получило название «биологистики» [21].

Однако до сих пор делались лишь отдельные попытки реконструкции чёрного ящика обслуживания с помощью междисциплинарного подхода. В немногочисленных же работах, затрагивающих этот вопрос, не уделяется внимания одной существенной общей характеристике фермента и машины (провайдера услуги – применительно к обслуживанию): они взаимодействуют с трансформируемым объектом, образуя короткоживущий промежуточный комплекс, который после дезинтеграции выдаёт конечный результат. Так, например, Дж. Нийирора и Дж. Чжуан [4] предложили жидкостную модель очереди для клиники скорой помощи, в которой хотя и вводится переменная, представляющая промежуточный комплекс «больничная койка–пациент», но не сравниваются масштабы временнóй динамики койко-мест и пациентов. Тем не менее, как это хорошо известно из химической кинетики, наличие в системе двух существенно разноскоростных процессов может иметь важные следствия. Этим обстоятельством мотивирована наша попытка применить ферментативную модель к системе обслуживания в более последовательной манере. Настоящее исследование частично основывается на результатах предшествующей работы авторов по производственным цепям [22].

Мы выверяем работоспособность ферментативной модели оказания услуги на примере циклической системы массового обслуживания. Среди непрерывно увеличивающегося разнообразия моделей центров массового обслуживания наша система относится к разновидности открытых (на вход и выход) циклических очередей из клиентов одного и того же класса с множеством одинаковых параллельных серверов, буфером неограниченного размера, возможностью отказа от обслуживания и повторными требованиями. В настоящее время аппарат теории вероятностей служит основным математическим инструментом теории массового обслуживания для описания процессов, происходящих в очередях (напр., [23]). Вместе с тем, в широком круге задач оказываются полезными так называемые «жидкостные» детерминистические модели [24,25]. Они относятся к категории приближённых и первоначально предназначались для анализа устойчивости исходных дискретных стохастических моделей [26], но к настоящему времени показали свою эффективность в изучении важных динамических характеристик сетей массового обслуживания. Когда регулярные вариации параметров модели превалируют над случайными флуктуациями, то учётом последних можно пренебречь. Кроме того, если количество клиентов в системе изменяется в широких пределах, то можно также проигнорировать их дискретную природу, совершить, используя терминологию статистической физики, предельный переход от системы с очень большим числом частиц к сплошной среде, к гидродинамическому описанию. Дискретный поток требований заменяется непрерывным потоком жидкости со скалярным полем плотности и векторным полем скоростей, течение которой следует по тому же маршруту. Получающаяся в результате система обыкновенных дифференциальных уравнений (ОДУ) баланса массы и есть жидкостная модель. Жидкостные модели часто легче поддаются исследованию и решению по сравнению с исходными стохастическими моделями теории массового обслуживания.

Модель системы массового обслуживания, предлагаемая в настоящей работе, уподобляет серверы молекулам фермента, которые переводят клиентов из популяции «потенциальных» в популяцию «обслуженных». Это позволяет записать модель в жидкостной форме в виде системы ОДУ химической кинетики. Предположение, что в расчёте на один сервер время обслуживания много короче времени ожидания, даёт возможность применить квазистационарное приближение к популяции занятых серверов. Эта переменная, как показано, быстро подстраивается к текущему спросу на услугу, следуя знаменитой гиперболической кривой Михаэлиса–Ментен ферментативной кинетики. В результате мы легко получаем клиринговую функцию системы, которая, как выясняется, тождественна функции Кармаркара [27], известной в производственной логистике. Предложенная нами модель позволяет выписать явное стационарное решение и провести анализ его устойчивости. Мы также исследуем влияние скорости появления новых клиентов на стационарный спрос.

1. Модель

Рассмотрим открытую систему обслуживания с группой параллельных одинаковых серверов, обрабатывающих входящие требования (обслуживающих клиентов). Постоянный поток потенциальных потребителей услуги поступает в систему извне, и если все имеющиеся серверы заняты, то прибывший становится в очередь. Потенциальный клиент может проявить нетерпение и через некоторое время покинуть очередь совсем. Такие клиенты считаются потерянными для системы – они могут направиться в другую компанию того же профиля.

Клиенты, ожидающие обслуживания, собираются в общем буфере, снабжающем все серверы. Размер буфера предполагается неограниченным. Каждый сервер может принять к обработке любое из поступивших требований, причём в состоянии обслуживать одновременно только одного клиента. Закончив обслуживание, сервер принимает следующего клиента из накопителя.

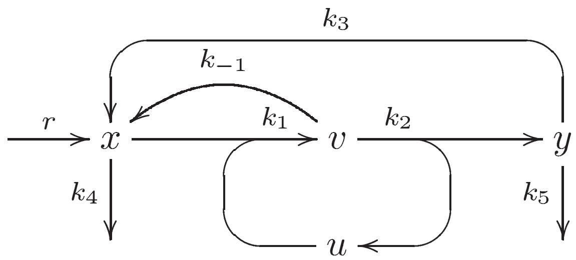

Рис. 2. Открытая циклическая система массового обслуживания. Здесь x – популяция потенциальных клиентов (иными словами, длина очереди на обслуживание или спрос), u – популяция свободных серверов, v – популяция серверов, занятых обслуживанием, y – популяция потребителей, получивших услугу (бывших клиентов). Кинетические константы r, k_{-1}, $k_1, \ldots, k_5$ характеризуют скорости соответствующих процессов

Fig. 2. The open cyclic service system. Here x is the population of potential customers (in other words, queue length or demand), u is the population of free servers, v is the population of busy (operative) servers, and y is the population of successfully served (past) customers. The constants $r, k_{-1}, k_1, \ldots, k_5$ depict the various rates with which these processes proceed

В нашей модели такие производственные термины как «запасы» (inventory), «работа, находящаяся в процессе выполнения» (work-in-progress, WIP) и «буферный запас» (buffer stock) считаются синонимами, и все означают текущую длину очереди из потенциальных клиентов. Фактически это мгновенный спрос на услугу. Популяция потенциальных потребителей не включает в себя клиентов, обслуживаемых в данный момент.

Контакт клиента с сервером не обязательно заканчивается получением желаемой услуги. Иногда клиент может засомневаться и отложить обслуживание. Провайдер также может отклонить требование, если, например, заявка составлена не по форме. В обоих случаях клиент возвращается обратно в очередь.

После успешного обслуживания потребитель пополняет ряды бывших клиентов. Спустя некоторое время он может либо совсем утратить интерес к данной услуге, либо опять присоединиться к очереди на обслуживание. Последнее имеет место, если изменение состояния клиента носит временный характер. Например, приобретённый покупателем товар может быть скоропортящимся, или подлежащим замене (ремонту) вследствие постоянной эксплуатации, или подверженным моральному износу, или входить в категорию модных товаров, и тому подобное. Таким образом, очередь в буфере всегда состоит из смеси новых и бывших клиентов.

Систематическое исследование циклических очередей было инициировано в работах Дж. Тейлора и Р. Джексона [28] и Э. Кёнигсберга [29, 30]. Перечень наиболее примечательных результатов, полученных к настоящему времени в этой области вероятностными методами, даётся в обзоре Дж. Шортла с соавторами [31, ch. 5, pp. 213–254]. Наш подход, однако, основан на жидкостном формализме.

Перечисленную выше совокупность процессов, происходящих в изучаемой циклической очереди, можно представить в виде схемы псевдохимических реакций, как показано на рис. 2. Схема кодирует как последовательность превращений, так и скорости их протекания. Она может быть формализована в виде системы кинетических уравнений для скоростей изменения количеств участвующих агентов. Для записи соответствующих уравнений воспользуемся известным в химии законом действующих масс (напр., [32, с. 350–354]), согласно которому скорость образования продукта в реакции взаимодействия пропорциональна количеству каждого из участвующих реагентов. Коэффициент пропорциональности в законе действующих масс – это константа скорости соответствующего акта превращения, которая предполагается известной. Константы каждого процесса обозначены на рис. 2 метками у стрелок. Они имеют следующий смысл:

– r – скорость притока клиентов в систему извне;
– k_1 – частота приёма требований индивидуальным сервером;
– k_{-1} – частота отзывов (или отклонений) требований на обслуживание после установления контакта между клиентом и сервером;
– k_2 – число оборотов сервера, то есть максимальное количество удовлетворённых им заявок за единицу времени. Таким образом, $(k_{-1} + k_2)^{-1}$ представляет собой среднее время занятости сервера;

– k_3 – частота возвращения клиентов в очередь на повторное обслуживание. Тем самым k_3^{-1} характеризует время, в течение которого сохраняется ценность полученной услуги, или время жизни изменённого состояния клиента;
– k_4 – частота, с которой нетерпеливые потребители покидают очередь;
– k_5 – частота, с которой провайдер услуги по разным причинам утрачивает лояльность бывших клиентов.

Записывая балансовое соотношение для каждого из участников схемы на рис. 2, мы получаем систему ОДУ

$$\dot{x} = r + k_{-1}v + k_3 y - k_1 ux - k_4 x, \tag{4a}$$

$$\dot{y} = k_2 v - (k_3 + k_5)y, \tag{4b}$$

$$\dot{u} = (k_{-1} + k_2)v - k_1 ux, \tag{4c}$$

$$\dot{v} = k_1 ux - (k_{-1} + k_2)v, \tag{4d}$$

где точки сверху означают дифференцирование по времени t. Все параметры модели неотрицательны. (Заметим, что модель, рассмотренная в работе [4], является частным случаем системы (4) при $k_{-1} = k_3 = k_5 = 0$.)

Складывая уравнения (4c) и (4d), выявляем первый интеграл системы, смысл которого состоит в сохранении общего числа имеющихся серверов u_0

$$u + v = u_0. \tag{5}$$

С помощью соотношения (5) систему (4) можно упростить, исключая u либо v. Произвольно выбирая для исключения u, мы следующим шагом приводим три оставшихся уравнения к безразмерному виду путём перехода к новым переменным

$$\begin{gathered} \tau = t/T, \quad \xi = x/K, \quad \eta = y/K, \quad \zeta = v/u_0, \\ \text{где} \quad T = \frac{1}{k_1 u_0} \quad \text{и} \quad K = \frac{k_{-1} + k_2}{k_1}. \end{gathered} \tag{6}$$

Согласно формулам (6) безразмерное время τ отсчитывается в единицах T. Величина T имеет смысл характерного времени, проводимого потенциальным клиентом в очереди перед началом обслуживания. Это время обратно пропорционально общему количеству серверов u_0. Обоснование нормировки x и y на величину K будет дано в следующем разделе. Что касается переменной v, то её нормировка на u_0 представляется очевидной.

Безразмерные жидкостные уравнения принимают вид

$$\dot{\xi} = \rho + \alpha\eta + (1 - \mu)\zeta + \zeta\xi - (1 + \beta)\xi, \tag{7a}$$

$$\dot{\eta} = \mu\zeta - (\alpha + \gamma)\eta, \tag{7b}$$

$$\varepsilon\dot{\zeta} = \xi - \zeta\xi - \zeta, \tag{7c}$$

где точки сверху теперь подразумевают дифференцирование по τ. Величины

$$\begin{gathered} \alpha = \frac{k_3}{k_1 u_0}, \qquad \beta = \frac{k_4}{k_1 u_0}, \qquad \gamma = \frac{k_5}{k_1 u_0}, \\ \varepsilon = \frac{k_1 u_0}{k_{-1} + k_2}, \quad \mu = \frac{k_2}{k_{-1} + k_2}, \quad \rho = \frac{r}{(k_{-1} + k_2)u_0} \end{gathered} \tag{8}$$

– новые безразмерные параметры. Следует заметить, что $\mu < 1$. Согласно определению (8), ε – это отношение времени обслуживания $(k_{-1} + k_2)^{-1}$ к времени ожидания $(k_1 u_0)^{-1}$.

Из физических соображений имеют смысл только неотрицательные решения уравнений (7).

2. Результаты

При выводе уравнений (7) мы не делали никаких специальных предположений относительно параметров. Однако известно [5], что время ожидания часто длиннее времени обслуживания, поэтому следует считать параметр ε малым. При $\varepsilon \ll 1$ система (7) становится сингулярно-возмущённой. «Медленные» переменные в ней – это ξ и η, соответствующие числу ожидающих обслуживания и числу получивших услугу. «Быстрая» переменная – количество занятых серверов ζ. Стандартной практикой редукции таких систем служит метод многих масштабов, посредством которого быстрая переменная адиабатически исключается. Законность процедуры адиабатического исключения устанавливается в каждом конкретном случае. В случае системы (7) теорема Фенихеля–Тихонова требует, помимо всего прочего, чтобы 1) неподвижная точка $\bar{\zeta}(\xi, \eta)$ быстрого уравнения (7c) была изолированным корнем алгебраического уравнения $\dot{\zeta} = 0$ и сохраняла устойчивость при всех допустимых значениях медленных переменных ξ и η, и 2) начальное условие ζ_0 попадало в область влияния неподвижной точки $\bar{\zeta}$ при всех начальных условиях ξ_0 и η_0 [33, ch. 3, p. 53–70]. Заметный вклад в обоснование применимости квазистационарного приближения к ферментативным реакциям внесли В. Клоновский [34] и Л. Зегель [35].

Для расщепления системы (7) на быструю и медленную части введём быстрое время $\vartheta = \tau/\varepsilon$. Совершая подстановку $\tau \to \vartheta\varepsilon$ и полагая $\varepsilon = 0$, получаем

$$\begin{aligned} &\xi' = 0, \quad \eta' = 0, \\ &\zeta' = \xi - \zeta\xi - \zeta, \end{aligned} \tag{9}$$

где «штрихи» означают дифференцирование по ϑ. Это быстрая подсистема, где ξ и η заменены их начальными значениями и рассматриваются как параметры. Она порождает внутреннее решение, пригодное для $\tau = \mathcal{O}(\varepsilon)$.

Полагая в системе (7) $\varepsilon = 0$, получаем медленную подсистему

$$\dot{\xi} = \rho + \alpha\eta + (1-\mu)\zeta + \zeta\xi - (1+\beta)\xi, \tag{10a}$$

$$\dot{\eta} = \mu\zeta - (\alpha+\gamma)\eta, \tag{10b}$$

$$0 = \xi - \zeta\xi - \zeta, \tag{10c}$$

которая даёт внешнее решение, справедливое на масштабах времени $\tau = \mathcal{O}(1)$. В сингулярном пределе $\varepsilon \to 0$ медленная подсистема определяет поток на медленном многообразии, задаваемом алгебраическим уравнением (10c). Внешнее решение правомерно для всех значений ξ и η, при которых квазистационарное решение быстрой подсистемы (9) устойчиво.

Квазистационарное состояние быстрой подсистемы (9) суть $\zeta = \xi/(1+\xi)$, и оно асимптотически устойчиво для любого положительного ξ. Напомним, что ζ – это доля занятых серверов, v/u_0. Покуда нормированный спрос (длина очереди) мал, то есть $\xi \ll 1$, эта доля тоже адекватно мала, $\zeta \simeq \xi$, означая низкую загруженность серверов. Однако при высоких уровнях спроса, когда $\xi \gg 1$, практически все серверы заняты обслуживанием: $\zeta \to 1$. Когда $\xi = 1$, доля занятых серверов составляет половину от их общего числа в системе. В этом состоит причина выбора величины $K = (k_{-1} + k_2)/k_1$ в качестве масштабной единицы для размеров популяций потенциальных и бывших клиентов. В биохимии K носит название константы Михаэлиса.

Из уравнений (10) следует, что на временах порядка $\tau = \mathcal{O}(1)$ динамика системы массового обслуживания (7) управляется двумя уравнениями

$$\dot{\xi} = \rho + \alpha\eta - \frac{\mu\xi}{1+\xi} - \beta\xi, \tag{11a}$$

$$\dot{\eta} = \frac{\mu\xi}{1+\xi} - (\alpha+\gamma)\eta. \tag{11b}$$

Первый член в правой части уравнения (11b) представляет собой выпуск рассматриваемой системы, то есть количество клиентов, получающих услугу в единицу времени

$$\begin{aligned} \tilde{Y} &= \frac{\mu\xi}{1+\xi} \quad \text{(безразмерная форма)}, \\ Y &= \frac{k_2 u_0 x}{K+x} \quad \text{(размерная форма)}. \end{aligned} \tag{12}$$

Это есть не что иное как клиринговая функция, поскольку она зависит от числа потенциальных потребителей, стоящих в очереди, то есть от WIP – по терминологии логистики. Точнее, это клиринговая функция в форме, первоначально введённой У. Кармаркаром [27]. Однако в отличие от формулы, эмпирически полученной Кармаркаром для логистических цепей, аргумент ξ в нашей формуле совсем не обязан быть стационарной величиной. При выводе клиринговой функции мы не предполагали, что система обслуживания функционирует в стационарном режиме. Тем не менее, количество занятых серверов ζ, будучи быстрой переменной, после кратковременного переходного процесса продолжительностью порядка $\mathcal{O}(\varepsilon)$ будет находиться в квазистационарном состоянии по отношению к текущему спросу, то есть к ξ. В уравнении (11a) приток потенциальных клиентов извне, задаваемый параметром ρ, необязательно постоянен, и если он варьирует намного медленнее характерного времени обслуживания, то формула (12) для клиринговой функции остаётся справедливой.

В формуле (12) можно узнать знаменитое уравнение Михаэлиса–Ментен из ферментативной кинетики [12, гл. 2]. Гиперболическая форма отклика вход–выход типа (12) нередко встречается в естественных науках: она описывает соотношение между количеством активных центров адсорбции на поверхности адсорбента и парциальным давлением газа-адсорбата (уравнение Ленгмюра), сигмоидную кривую связывания кислорода гемоглобином или долю макромолекул, насыщенных лигандом, как функцию от концентрации лиганда (уравнение Хилла), скорость роста микроорганизмов в питательном растворе (уравнение Моно), отклик численности хищника на плотность популяции жертвы (уравнение Холлинга II-го типа) и им подобные.

Отличительной чертой уравнения (12) является насыщение выпуска в ответ на увеличение спроса. При низком спросе выпуск приблизительно пропорционален спросу. Однако при высоком спросе поток получивших услугу стремится к постоянной величине.

Система (11) имеет два стационарных состояния. Из них лишь одно всегда лежит в первом квадранте фазовой плоскости $\xi\eta$, то есть физически осмысленно:

$$\begin{aligned} \bar{\xi} &= \frac{\sqrt{B^2+4\beta(\alpha+\gamma)^2\rho}-B}{2\beta(\alpha+\gamma)}, \\ \bar{\eta} &= \frac{B+2(\alpha+\gamma)\rho-\sqrt{B^2+4\beta(\alpha+\gamma)^2\rho}}{2\gamma(\alpha+\gamma)}, \end{aligned} \tag{13}$$

где ради краткости обозначено $B=(\alpha+\gamma)(\beta-\rho)+\gamma\mu$. Эта неподвижная точка всегда устойчива по типу узла, поскольку одновременно выполняются условия $\operatorname{Tr}\mathbf{J}=-\alpha-\beta-\gamma-\mu/(1+\bar{\xi})^2<0$, $\det\mathbf{J}=\beta(\alpha+\gamma)+\gamma\mu/(1+\bar{\xi})^2>0$ и $(\operatorname{Tr}\mathbf{J})^2-4\det\mathbf{J}>0$, где $\mathbf{J}$ – матрица Якоби системы (11), вычисленная в равновесии (13). Из вида уравнений (11) можно заключить, что фазовые траектории входят в первый квадрант слева и снизу, делая тем самым узловую точку (13) глобально устойчивой для всех мыслимых начальных условий.

3. Обсуждение

В нашей системе массового обслуживания сосуществуют две взаимозависимые, но непосредственно не взаимодействующие популяции клиентов. Одна, с нормированной численностью ξ, представляет собой очередь на получение услуги и состоит как из вновь прибывших, так и из тех, кто уже прежде обслуживался. Другая, с нормированной численностью η, состоит из потребителей, требования которых на данный момент удовлетворены. В стационарном состоянии, как это следует из уравнений (11), отношение размеров популяций выражается формулой $\bar{\xi}/\bar{\eta} = (\alpha + \gamma)(1 + \bar{\xi})/\mu$. Нетрудно видеть, что при низком стационарном спросе, $\bar{\xi} < 1$, это отношение принимает минимальное значение $(\alpha+\gamma)/\mu = (k_1 u_0)^{-1}(1 + k_{-1}/k_2)(k_3 + k_5)$. Множитель $1 + k_{-1}/k_2$ скорее всего имеет порядок $\mathcal{O}(1)$ при реалистическом предположении, что $k_{-1} < k_2$ (вероятность отзыва заявки меньше вероятности её выполнения). Тем самым, выражение для $\min(\bar{\xi}/\bar{\eta})$ может быть интерпретировано просто как отношение времени ожидания $(k_1 u_0)^{-1}$ к времени пребывания клиента в категории бывших $(k_3 + k_5)^{-1}$. Чем меньше времени потребитель проводит в очереди на обслуживание, чем дольше длится «последействие» услуги и чем дольше сохраняется поддержка потребителем данного бренда или продукта, тем меньше отношение «текущий спрос–лояльный контингент».

Потребовав, чтобы стационарный нормированный спрос, даваемый первой формулой из (13), не превышал единицы, находим, при каком соотношении параметров модели очередь на обслуживание остаётся относительно короткой:

$$\bar{\xi} < 1 \quad \forall(\rho \leqslant \rho^* \wedge \beta > 0) \vee (\rho > \rho^* \wedge \beta > \rho - \rho^*), \tag{14}$$

где $\rho^* = \gamma\mu/[2(\alpha+\gamma)]$ – критическая скорость притока клиентов в систему. Примечательно, что при докритической скорости притока новых клиентов, такой что $\rho \leqslant \rho^*$, стационарный спрос остаётся низким независимо от частоты уходов из очереди (отказов), как показано на рис. 3.

В первом приближении для стационарного спроса в окрестности параметров $\rho = \rho^*$ и $\beta = 0$ получаем: $\bar{\xi} = 1 + 2(\rho - \rho^*)/\rho^* - 2\beta/\rho^*$. Как и следовало ожидать, увеличение входящего потока требований удлиняет очередь, а повышение частоты отказов – сокращает.

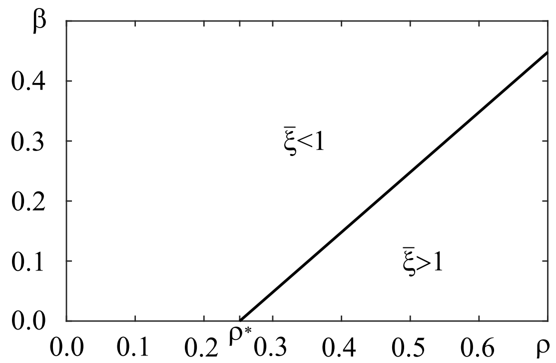

Рис. 3. Проекция линии уровня $\bar{\xi} = 1$ стационарного решения (13) системы (11) на плоскость параметров $\rho\,\beta$ («поток требований–частота отказов»). Прямая $\beta = \rho - \rho^*$, где $\rho^* = \gamma\mu/[2(\alpha+\gamma)]$ – критический поток входящих клиентов, делит плоскость на две области: выше прямой стационарный спрос низкий, ниже – высокий. Параметры расчёта: $\alpha = 1.1$, $\gamma = 1.4$, $\mu = 0.9$, $\rho^* = 0.252$

Fig. 3. Contour line $\bar{\xi} = 1$ of the steady-state solution (13) to the system (11) projected onto the plane $\rho\,\beta$ («customers arrival rate–frequency of abandonment»). Line $\beta = \rho - \rho^*$, where $\rho^\star = \gamma\mu/[2(\alpha+\gamma)]$ is the critical arrival rate, partitions the positive quadrant into two domains: above the line the steady-state demand is low, and below is high. The parameters chosen for simulation: $\alpha = 1.1$, $\gamma = 1.4$, $\mu = 0.9$, and $\rho^* = 0.252$

В размерном виде условие (14) выглядит так:

$$\bar{x} < K \quad \forall (0 < r \leqslant r^* \wedge k_4 > 0) \vee (r > r^* \wedge k_4 > (r - r^*)/K),$$

где $K = (k_{-1} + k_2)/k_1$ и $r^* = k_2 u_0 k_5 / [2(k_3 + k_5)]$.

При ограниченной вместимости буфера снизить угрозу его переполнения можно прежде всего путём повышения критической скорости притока потенциальных клиентов. В выражении для r^* множитель $k_5/(k_3 + k_5)$ заключён в пределах от 0 до 1: нижнее значение соответствует случаю, когда среднее время сохранения лояльности бывшего клиента к провайдеру k_5^{-1} намного продолжительнее времени жизни результата обслуживания k_3^{-1}; верхнее – случаю, когда услуга долговременная, а бывшие клиенты быстро утрачивают интерес или потребность в получении повторной услуги у данного провайдера. Поэтому более действенным средством отодвигания критического порога может послужить увеличение фактора $k_2 u_0$ – максимально возможной скорости обслуживания клиентов всеми серверами, имеющимися в системе. Это достигается укомплектованием системы обслуживания дополнительным персоналом или оборудованием либо повышением числа оборотов k_2 каждого канала обслуживания.

На ходу, однако, такую реорганизацию провести затруднительно. Более практичной мерой представляется увеличение частоты отказов k_4 сразу по достижении притоком клиентов критической скорости r^*. Учёт отказов от обслуживания является важной особенностью рассматриваемой модели. И не только потому что реальной очереди без отказов не бывает – таково поведение потребителя. Без отказов стационарный режим в модели становится невозможным при превышении входящего трафика клиентов над максимальной скоростью обслуживания, в результате чего длина очереди неограниченно растёт. Включение в модель отказов делает её более робастной и тем самым – более реалистичной. При необходимости частотой отказов в системе массового обслуживания намного проще манипулировать в динамике по сравнению с другими параметрами. Можно, например, постоянно следить за скоростью прибытия клиентов, и в случае достижения ею критической величины побуждать потенциальных потребителей отложить получение услуги на более поздний срок путём объявлений о задержке обслуживания.

Обсуждение свойств открытой системы массового обслуживания было бы неполным без упоминания особенностей её замкнутого варианта. К формулировке замкнутой модели можно прийти, если в уравнениях (4) положить $r = k_4 = k_5 = 0$. Получившаяся в результате система кинетических уравнений содержит две сохраняющиеся величины: общее количество серверов u_0, даваемое уравнением (5), и общую численность потребителей x_0 – потенциальных, обслуживаемых в данный момент и бывших:

$$x + v + y = x_0.$$

Таким образом, исходная система из четырёх уравнений упрощается до двух, которые после перехода к безразмерным переменным

$$\tau = t k_1 x_0, \quad \xi = x/x_0, \quad \eta = y/x_0$$

принимают следующий вид:

$$\begin{aligned} \dot{\xi} &= c\eta - (\xi + a)(\xi + \eta - 1) - \varepsilon\xi, \\ \dot{\eta} &= -b(\xi + \eta - 1) - c\eta, \end{aligned} \tag{15}$$

где

$$a = \frac{k_{-1}}{k_1 x_0}, \quad b = \frac{k_2}{k_1 x_0}, \quad c = \frac{k_3}{k_1 x_0}, \quad \varepsilon = \frac{u_0}{x_0},$$

и точки сверху означают дифференцирование по новому времени τ.

При реалистическом предположении $\varepsilon \ll 1$ система (15) имеет лишь одно стационарное решение, обладающее физическим смыслом. С точностью до $\mathcal{O}(\varepsilon)$ получаем:

$$\bar{\xi} = 1 - \frac{\varepsilon(b+c)}{c(a+b+1)}, \quad \bar{\eta} = \frac{\varepsilon b}{c(a+b+1)}. \tag{16}$$

Это устойчивый узел на фазовой плоскости $\xi\eta$ для всех допустимых значений параметров модели, поскольку $\det \mathbf{J} = c(a+b+1) + \mathcal{O}(\varepsilon) > 0$, $\operatorname{Tr} \mathbf{J} = -a - b - c - 1 + \mathcal{O}(\varepsilon) < 0$ и $(\operatorname{Tr} \mathbf{J})^2 - 4 \det \mathbf{J} = (a+b-c+1)^2 + \mathcal{O}(\varepsilon) > 0$, где $\mathbf{J}$ – матрица Якоби системы (15), вычисленная в неподвижной точке (16). Как видно, стационарная популяция бывших клиентов всегда относительно немногочисленна в меру малости параметра ε. Количество клиентов, получивших услугу, находится в прямой зависимости от числа оборотов сервера и в обратной зависимости от всех остальных параметров модели. Дополнительно заметим, что замкнутая система такого типа рассматривалась В. П. Миловановым [36, гл. 7, с. 196–203]. Однако автор цитируемой работы, не выписав в обозримой форме стационарное решение и не получив окончательное выражение для дискриминанта характеристического уравнения, сделал поспешный вывод о возможности устойчивого фокуса.

В качестве побочного результата настоящего исследования мы дали вывод клиринговой функции Кармаркара – формулы, используемой сегодня в планировании производства. С чисто технической стороны, это известный результат для стационарного режима. Однако мы распространили понятие клиринговой функции на нестационарный случай и выявили условия применимости этой формулы. Такие вопросы как «Насколько хорошо приближение клиринговой функции?» и «При каких условиях она работает, а при каких – нет?» актуальны в современной теории очередей, как следует из обзора Д. Армбрустера [11]. Итак, в рассмотренной нами модели системы обслуживания присутствуют два сильно различающихся масштаба времени: более продолжительный, соответствующий динамике популяций клиентов, и более короткий, соответствующий динамике популяции занятых серверов. Такая временнáя иерархия делает возможным квазистационарное приближение и, следовательно, гиперболическую зависимость выхода от входа в виде клиринговой функции Кармаркара, идентичной формуле Михаэлиса–Ментен. В терминах нашей «ферментативной» модели с двумя шкалами времени оказывается возможным указать более точные рамки применимости клиринговой функции: текущее количество занятых серверов будет пребывать в квазистационарном состоянии по отношению к спросу при условии, что время обслуживания значительно короче времени ожидания. Следовательно, условие $\varepsilon \ll 1$ можно полагать достаточным для пригодности приближения клиринговой функции в нестационарных условиях.

Заключение

Нами построена жидкостная модель системы массового обслуживания, представляющей собой открытую циклическую очередь. Мы обратились к биологической аналогии провайдера услуги с ферментом-катализатором, составив дифференциальные уравнения для зависящих от времени средних количеств клиентов и серверов, подобные уравнениям химической кинетики для концентраций молекул. Фактически эти уравнения представляют собой непрерывное детерминистическое приближение первого порядка дискретного марковского процесса. Выгодным отличием жидкостных моделей от дискретных стохастических является то, что их сложность не зависит от числа обрабатываемых требований. Это позволяет исследовать с их помощью поведение очереди при большом числе потребителей, а также учитывать возможные резкие всплески числа клиентов, связанные с массовым приходом или уходом.

Анализ модели позволил сделать ряд содержательных выводов о свойствах и поведении системы в стационарном режиме и в динамике, которые затруднительно получить в рамках вероятностной постановки.

Показано, что модель допускает только одно физически осмысленное стационарное состояние, которое всегда асимптотически устойчиво по типу узла. Тем самым исключается возможность появления затухающих колебаний спроса в случае вывода системы из равновесия.

В результате параметрического анализа стационарного решения уравнений модели вычислен критический поток входящих клиентов, ниже которого длина очереди остаётся малой при любых допустимых значениях параметров. При закритическом потоке потенциальных потребителей наиболее практичной оперативной мерой регулирования длины очереди может быть увеличение частоты отказов.

Найдена гиперболическая функциональная зависимость скорости обслуживания от спроса, идентичная уравнению Михаэлиса–Ментен ферментативной кинетики и клиринговой функции Кармаркара для логистики. Показано, что такой вид клиринговой функции обусловлен квазистационарностью числа серверов по отношению к более медленно изменяющемуся числу клиентов.

Важно подчеркнуть, что мы использовали язык химических реакций лишь для удобства описания процессов взаимодействия экономических агентов, когда переход агента из одного состояния в другое осуществляется через образование и дезинтеграцию короткоживущего промежуточного комплекса. Этот формализм может с успехом быть применён и для моделирования других явлений, где имеют место акты трансформации вход–выход.

Библиографический список

1. The World Bank. World development indicators. Table 4.2: Structure of output. Washington, DC: The World Bank Group, 2019. Access mode: http://wdi.worldbank.org/table/4.2.
2. *Armony M., Shimkin N., Whitt W.* The impact of delay announcements in many-server queues with abandonment // Operations Research. 2009. Vol. 57, no. 1. P. 66–81.
3. *Yom-Tov G.B., Mandelbaum A.* Erlang-R: A time-varying queue with reentrant customers, in support of healthcare staffing // Manufacturing & Service Operations Management. 2014. Vol. 16, no. 2. P. 283–299.
4. *Niyirora J., Zhuang J.* Fluid approximations and control of queues in emergency departments // European Journal of Operational Research. 2017. Vol. 261, no. 3. P. 1110–1124.
5. *Whitt W.* Time-varying queues // Queueing Models and Service Management. 2018. Vol. 1, no. 2. P. 79–164.
6. *Fromm H., Cardoso J.* Foundations // Fundamentals of Service Systems / Ed. by J. Cardoso, H. Fromm, S. Nickel et al. New York, NY: Springer, 2015. Service Science: Research and Innovations in the Service Economy. P. 1–32.
7. *Hill T.P.* On goods and services // Review of Income and Wealth. 1977. Vol. 23, no. 4. P. 315–338.
8. *Gallouj F.* Innovation in services and the attendant old and new myths // The Journal of Socio-Economics. 2002. Vol. 31, no. 2. P. 137–154.
9. *Gadrey J.* The characterization of goods and services: An alternative approach // Review of Income and Wealth. 2000. Vol. 46, no. 3. P. 369–387.
10. *Shephard R.W.* Theory of Cost and Production Functions. Princeton, NJ: Princeton University Press, 2016.
11. *Armbruster D.* The production planning problem: Clearing functions, variable lead times, delay

equations and partial differential equations // Decision Policies for Production Networks / Ed. by D. Armbruster, K. G. Kempf. London: Springer, 2012. P. 289–302.

12. *Корниш-Боуден Э.* Основы ферментативной кинетики. М.: Мир, 1979. 280 с.
13. *Georgescu-Roegen N.* Some properties of a generalized Leontief model // Analytical Economics: Issues and Problems. Cambridge, MA: Harvard University Press, 1966. P. 316–337.
14. *Колесова Г.И., Полетаев И.А.* Некоторые вопросы исследования систем с лимитирующими факторами // Управляемые системы: Сборник трудов. Вып. 3 / под ред. В. А. Евстегнеева. Новосибирск: Институт математики СО АН СССР, 1969. С. 71–80.
15. *Романовский Ю.М., Степанова Н.М., Чернавский Д.С.* Что такое математическая биофизика: Кинетические модели в биофизике. М.: Просвещение, 1971. 136 с.
16. *Нельсон Р., Уинтер С.* Эволюционная теория экономических изменений. М.: Финстатинформ, 2000. 472 с.
17. Industrial Metabolism: Restructuring for Sustainable Development / Ed. by R.U. Ayres, U.E. Simonis. Tokyo: United Nations University Press, 1994. 376 p.
18. Networks of Interacting Machines: Production Organization in Complex Industrial Systems and Biological Cells / Ed. by D. Armbruster, K. Kaneko, A.S. Mikhailov. Singapore: World Scientific, 2005. Vol. 3 of World Scientific Lecture Notes in Complex Systems. 267 p.
19. *Levine E., Hwa T.* Stochastic fluctuations in metabolic pathways // Proceedings of the National Academy of Sciences. 2007. Vol. 104, no. 22. P. 9224–9229.
20. *Hochendoner P., Ogle C., Mather W.H.* A queueing approach to multi-site enzyme kinetics // Interface Focus. 2014. Vol. 4. P. 1–11.
21. *Helbing D., Armbruster D., Mikhailov A.S., Lefeber E.* Information and material flows in complex networks // Physica A: Statistical Mechanics and its Applications. 2006. Vol. 363, no. 1. P. xi–xvi.
22. *Mustafin A., Kantarbayeva A.* Opening the Leontief's black box // Heliyon. 2018. Vol. 4, no. 5. P. e00626.
23. *Hopp W. J., Spearman M.L.* Factory Physics: Foundations of Manufacturing Management. 3rd edition. New York, NY: McGraw-Hill, 2008. 720 p.
24. *Armbruster D., Marthaler D., Ringhofer C.* Kinetic and fluid model hierarchies for supply chains // Multiscale Modeling & Simulation. 2003. Vol. 2, no. 1. P. 43–61.
25. *Gamarnik D.* Fluid models of queueing networks // Wiley Encyclopedia of Operations Research and Management Science / Ed. by J.J. Cochran, L.A. Cox, P. Keskinocak et al. Hoboken, NJ: Wiley, 2011.
26. *Bramson M.* Stability of Queueing Networks. Lecture Notes in Mathematics no. 1950. Berlin; Heidelberg: Springer, 2008. 208 p.
27. *Karmarkar U.S.* Manufacturing lead times, order release and capacity loading // Logistics of Production and Inventory / Ed. by S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin. Amsterdam: North Holland, 1993. Vol. 4 of Handbook in Operations Research and Management Science. P. 287–329.
28. *Taylor J., Jackson R.R.P.* An application of the birth and death process to the provision of spare machines // Operational Research Quarterly. 1954. Vol. 5, no. 4. P. 95–108.
29. *Koenigsberg E.* Cyclic queues // Operational Research Quarterly. 1958. Vol. 9, no. 1. P. 22–35.
30. *Koenigsberg E.* Twenty five years of cyclic queues and closed queue networks: A review // Journal of the Operational Research Society. 1982. Vol. 33, no. 7. P. 605–619.

31. *Shortle J.F., Thompson J.M., Gross D., Harris C.M.* Fundamentals of Queueing Theory. Wiley Series in Probability and Statistics. 5th edition. Hoboken, NJ: Wiley, 2018.
32. *Ландау Л.Д., Лифшиц Е.М.* Статистическая физика. Часть 1. М.: Физматлит, 2010. 616 с.
33. *Kuehn C.* Multiple Time Scale Dynamics. New York, NY: Springer, 2014. Vol. 191 of Applied Mathematical Sciences. 814 p.
34. *Klonowski W.* Simplifying principles for chemical and enzyme reaction kinetics // Biophysical Chemistry. 1983. Vol. 18, no. 2. P. 73–87.
35. *Segel L.A., Slemrod M.* The quasi-steady-state assumption: A case study in perturbation // SIAM Review. 1989. Vol. 31, no. 3. P. 446–477.
36. *Милованов В.П.* Неравновесные социально-экономические системы: Синергетика и самоорганизация. М.: Эдиториал УРСС, 2001. 264 с.

References

1. The World Bank. World development indicators. Table 4.2: Structure of output. Washington, DC: The World Bank Group, 2019. Access mode: http://wdi.worldbank.org/table/4.2.
2. Armony M., Shimkin N., Whitt W. The impact of delay announcements in many-server queues with abandonment. *Operations Research*, 2009, vol. 57, no. 1, pp. 66–81.
3. Yom-Tov G.B., Mandelbaum A. Erlang-R: A time-varying queue with reentrant customers, in support of healthcare staffing. *Manufacturing & Service Operations Management*, 2014, vol. 16, no. 2, pp. 283–299.
4. Niyirora J., Zhuang J. Fluid approximations and control of queues in emergency departments. *European Journal of Operational Research*, 2017, vol. 261, no. 3, pp. 1110–1124.
5. Whitt W. Time-varying queues. *Queueing Models and Service Management*, 2018, vol. 1, no. 2, pp. 79–164.
6. Fromm H., Cardoso J. Foundations // Fundamentals of Service Systems / Ed. by J. Cardoso, H. Fromm, S. Nickel et al. New York, NY: Springer, 2015. Service Science: Research and Innovations in the Service Economy. P. 1–32.
7. Hill T. P. On goods and services. *Review of Income and Wealth*, 1977, vol. 23, no. 4, pp. 315–338.
8. Gallouj F. Innovation in services and the attendant old and new myths. *The Journal of Socio-Economics*, 2002, vol. 31, no. 2, pp. 137–154.
9. Gadrey J. The characterization of goods and services: An alternative approach. Review of Income and Wealth, 2000, vol. 46, no. 3, pp. 369–387.
10. Shephard R. W. Theory of Cost and Production Functions. Princeton, NJ: Princeton University Press, 2016.
11. Armbruster D. The production planning problem: Clearing functions, variable lead times, delay equations and partial differential equations // Decision Policies for Production Networks / Ed. by D. Armbruster, K. G. Kempf. London: Springer, 2012. P. 289–302.
12. Cornish-Bowden A. Fundamentals of Enzyme Kinetics. 4th edition. Weinheim: Wiley-Blackwell, 2012. 498 p.
13. Georgescu-Roegen N. Some properties of a generalized Leontief model. In: Analytical Economics: Issues and Problems. Cambridge, MA: Harvard University Press, 1966. P. 316–337.
14. Kolesova G.I., Poletaev I.A. Nekotorye voprosy issledovaniia sistem s limitiruiushchimi faktorami. Vypusk 3 [Selected problems in research of the systems with limiting factors. Issue 3].

In: Upravliaemye sistemy [Controllable systems]. Novosibirsk: Institute of Mathematics, Siberian Branch of the USSR Academy of Sciences, 1969. P. 71–80 (In Russian).

15. Romanovskii Yu.M., Stepanova N.M., Chernavskii D.S. Chto takoe matematicheskaia biofizika: Kineticheskie modeli v biofizike [What is mathematical biophysics: Kinetic models in biophysics]. Moscow: Prosveshchenie, 1971. 136 p. (In Russian).
16. Nelson R.R., Winter S.G. An Evolutionary Theory of Economic Change. Cambridge, MA: Belknap Press of Harvard University Press, 1982. 437 p.
17. Industrial Metabolism: Restructuring for Sustainable Development. Ed. by R.U. Ayres, U.E. Simonis. Tokyo: United Nations University Press, 1994. 376 p.
18. Networks of Interacting Machines: Production Organization in Complex Industrial Systems and Biological Cells. Ed. by D. Armbruster, K. Kaneko, A. S. Mikhailov. Singapore: World Scientific, 2005. Vol. 3 of World Scientific Lecture Notes in Complex Systems. 267 p.
19. Levine E., Hwa T. Stochastic fluctuations in metabolic pathways. *Proceedings of the National Academy of Sciences*, 2007, vol. 104, no. 22, pp. 9224–9229.
20. Hochendoner P., Ogle C., Mather W.H. A queueing approach to multi-site enzyme kinetics. *Interface Focus*, 2014, vol. 4, pp. 1–11.
21. Helbing D., Armbruster D., Mikhailov A.S., Lefeber E. Information and material flows in complex networks. *Physica A: Statistical Mechanics and its Applications*, 2006, vol. 363, no. 1, pp. xi–xvi.
22. Mustafin A., Kantarbayeva A. Opening the Leontief's black box. *Heliyon*, 2018, vol. 4, no. 5, p. e00626.
23. Hopp W.J., Spearman M.L. Factory Physics: Foundations of Manufacturing Management. 3rd edition. New York, NY: McGraw-Hill, 2008. 720 p.
24. Armbruster D., Marthaler D., Ringhofer C. Kinetic and fluid model hierarchies for supply chains. *Multiscale Modeling & Simulation*, 2003, vol. 2, no. 1, pp. 43–61.
25. Gamarnik D. Fluid models of queueing networks. In: Wiley Encyclopedia of Operations Research and Management Science, ed. by J. J. Cochran, L. A. Cox, P. Keskinocak et al. Hoboken, NJ: Wiley, 2011.
26. Bramson M. Stability of Queueing Networks. Lecture Notes in Mathematics no. 1950. Berlin; Heidelberg: Springer, 2008. 208 p.
27. Karmarkar U.S. Manufacturing lead times, order release and capacity loading. In: Logistics of Production and Inventory, ed. by S.C. Graves, A.H.G. Rinnooy Kan, P.H. Zipkin. Amsterdam: North Holland, 1993. Vol. 4 of Handbook in Operations Research and Management Science. P. 287–329.
28. Taylor J., Jackson R. R.P. An application of the birth and death process to the provision of spare machines. *Operational Research Quarterly*, 1954, vol. 5, no. 4, pp. 95–108.
29. Koenigsberg E. Cyclic queues. *Operational Research Quarterly*, 1958, vol. 9, no. 1, pp. 22–35.
30. Koenigsberg E. Twenty five years of cyclic queues and closed queue networks: A review. *Journal of the Operational Research Society*, 1982, vol. 33, no. 7, pp. 605–619.
31. Shortle J.F., Thompson J.M., Gross D., Harris C.M. Fundamentals of Queueing Theory. Wiley Series in Probability and Statistics. 5th edition. Hoboken, NJ: Wiley, 2018.
32. Landau L.D., Lifshitz E.M. Statistical Physics. Part 1. 3rd edition. Burlington, MA: Elsevier Butterworth-Heinemann, 2013. Vol. 5 of Course of Theoretical Physics. 544 p.
33. Kuehn C. Multiple Time Scale Dynamics. New York, NY: Springer, 2014. Vol. 191 of Applied Mathematical Sciences. 814 p.

34. Klonowski W. Simplifying principles for chemical and enzyme reaction kinetics. *Biophysical Chemistry*, 1983, vol. 18, no. 2, pp. 73–87.

35. Segel L. A., Slemrod M. The quasi-steady-state assumption: A case study in perturbation. *SIAM Review*, 1989, vol. 31, no. 3, pp. 446–477.

36. Milovanov V.P. Neravnovesnye sotsial'no-ekonomicheskie sistemy: sinergetika i samoorganizatsiia [Nonequilibrium social-economic systems: synergetics and self-organization]. Moscow: Editorial URSS, 2001. 264 p. (in Russian).

Мустафин Алмаз Тлемисович – родился в Караганде, Казахстан (1955), окончил Московский физико-технический институт (1978) и аспирантуру МГУ (1983). Доктор технических наук по специальности 05.13.16 «Применение вычислительной техники, математического моделирования и математических методов в научных исследованиях» (Институт математики НАН Казахстана, 2001). Опыт работы: стажёр-исследователь Физического института АН СССР (1978–1980); научный сотрудник Института математики и механики АН КазССР (1984–1991) и Института космических исследований НАН Казахстана (1991–2001); преподаватель университета им. Дж. Вашингтона (США, 2001–2002) и Мэрилендского университета (США, 2002–2007); с сентября 2007 года – профессор Казахского национального исследовательского технического университета им. К. И. Сатпаева. Область профессиональных интересов -– математическое моделирование нелинейных явлений в биофизике, квантовой электронике, социально-экономических системах. Опубликовал более пятидесяти научных работ по вышеуказанным направлениям исследований. Член Американского физического общества (APS) и Общества индустриальной и прикладной математики США (SIAM).

Казахстан, 050013 Алматы, ул. Сатпаева, 22
Казахский национальный исследовательский технический университет им. К. И. Сатпаева
E-mail: butsuri123@gmail.com

Кантарбаева Алия Кажбековна – родилась в Алматы, Казахстан (1946), окончила Ивановский химико-технологический институт (1970) и аспирантуру Совета по изучению производительных сил при Госплане СССР (1978). Кандидат экономических наук по специальности «Экономика и управление народным хозяйством» (Институт экономики НАН Республики Казахстан), доктор экономических наук по институционально-эволюционной теории предпринимательства (университет Киото, Япония, 1999). Работала в Институте экономических исследований при Госплане КазССР, в министерствах промышленности и экономики Казахстана, в Казахском национальном исследовательском техническом университете. С сентября 2018 года – профессор Высшей школы экономики и бизнеса Казахского национального университета им. аль-Фараби. Автор четырёх монографий и около 90 научных статей по экономике промышленности, предпринимательству и малому бизнесу, институционально-эволюционной экономике, менеджменту и маркетингу, управлению производственным процессом фирмы. Член Американской экономической ассоциации (AEA) и Японской ассоциации эволюционной экономики (JAFEE). Научный редактор экономического журнала «АльПари» в 1997–2009 годах.

Казахстан, 050040 Алматы, просп. аль-Фараби, 71
Казахский национальный университет им. аль-Фараби
E-mail: ratnakka@gmail.com