Информационные процессы, Том 6, № 4, 2006, стр. 396–444.

МАТЕМАТИЧЕСКИЕ МОДЕЛИ

ЭГАЛИТАРНОЕ РАЗДЕЛЕНИЕ ПРОЦЕССОРА[1]

С.Ф.Яшков*, **А.С.Яшкова****

**Институт проблем передачи информации им. А.А.Харкевича, Российская академия наук*
19, Большой Каретный переулок, 127994 Москва ГСП–4, Россия.
E-mail: yashkov@iitp.ru
***Муниципальный институт г.Жуковского,*
кафедра прикладной информатики и программирования,
15, улица Маяковского, 140180 г.Жуковский, Московской обл., Россия.
E-mail: annpost1@mail.ru

Поступила в редколлегию 14.11.2006, переработана 15.12.2006

Аннотация—Эта статья суммирует важнейшие результаты по точным решениям системы обслуживания M/GI/1 с эгалитарным разделением процессора (EPS — Egalitarian Processor Sharing). Материал почерпнут, в основном, из статей авторов, которые дополнены другими результатами. Главной целью является рассмотрение недавних достижений в точном анализе системы обслуживания M/GI/1—EPS с последующим упором на нестационарные (транзиентные) вероятностные распределения основных вероятностно–временных характеристик. В частности, обзор включает результаты по совместному нестационарному распределению времени пребывания в системе требования, поступающего в момент t с требуемой длительностью обслуживания (длиной) u, и числа требований в системе M/GI/1—EPS в момент $t-$, полученные в терминах многомерных преобразований. Демонстрируется также, как могут быть использованы нестационарные решения для получения известных и новых результатов, позволяющих предсказать поведение системы обслуживания и обнаружить ее новые неожиданные свойства.

ОГЛАВЛЕНИЕ

1 Введение 397
  1.1 Предварительные замечания 397
  1.2 Некоторые смежные направления 399
2 Теория системы обслуживания M/GI/1—EPS. Стационарный режим 401
  2.1 Вводные замечания 401
  2.2 Стационарное распределение числа требований 404
  2.3 Эскиз метода декомпозиции на элементы задержки 406
  2.4 Некоторые следствия. Интересные свойства системы M/GI/1—EPS. 410
  2.5 Некоторые частные случаи. 414
  2.6 Свойства монотонности. 415
  2.7 Асимптотика 415
  2.8 Связь с теорией страхового риска 417
  2.9 Другие асимптотические формулы. Предельные теоремы. 419
3 Теория системы обслуживания M/GI/1—EPS. Нестационарный режим. 422
  3.1 Нестационарное распределение числа требований 422

[1] Эта работа частично поддержана грантом по программе фундаментальных исследований Российской Академии наук (отделение информатики) “Новые физические и структурные решения в инфотелекоммуникациях"(Руководитель Н.А.Кузнецов).

3.2 Нестационарное совместное распределение числа требований и времени пребывания 429
3.3 Основные следствия. 430
4 Вместо заключения 434

1. ВВЕДЕНИЕ

1.1. Предварительные замечания

Одна из центральных проблем в прикладной теории вероятностей и теории очередей состояла в нахождении точного решения для стационарного времени пребывания требования в системе обслуживания M/GI/1 с дисциплиной эгалитарного (справедливого) разделения процессора (EPS — аббревиатура термина Egalitarian Processor Sharing)[2] Долгое время эта проблема оставалась открытой, впервые полное ее решение было получено Яшковым (Yashkov) с помощью предложенного им нового аналитического метода (первоначально для системы M/M/1 с дисциплиной кругового циклического доступа RR (RR — аббревиатура термина Round–Robin) [155] (1977)), адаптированного позднее для случая системы M/GI/1 с эгалитарным разделением процессора[3] (см., например, [78] (1978), [156] (1981), [158] (1983), а также [162], [167] для соответствующих ссылок). Этот метод, улучшенный немного позже, оказался весьма плодотворным для получения дальнейших результатов, например, для нестационарного распределения числа требований в рассматриваемой и родственной моделях [163] (1988), [164] (1989), [166] (1991), [167] (в западной литературе до конца прошлого столетия эти задачи считались неразрешимыми с аналитической точки зрения).

Отметим, что предположения, которые требуются для использования стационарных решений любой системы массового обслуживания, редко удовлетворяются в реальной жизни. Для реального применения теории очередей при проектировании и анализе технических систем использование только стационарных результатов теории во многих случаях явно недостаточно. Например, часто бывает необходимо исследовать поведение системы массового обслуживания, приближающееся к стационарному с ростом времени (если и когда стационарное состояние существует). При этом даже средняя длина очереди в момент t дает намного больше информации по сравнению со стационарным средним числом требований. Однако известно мало стохастических систем, для которых получены нестационарные (транзиентные) решения по вероятностным распределениям процессов обслуживания. Как правило, к таким системам относятся очереди M/GI/1 с классическим дисциплинами обслуживания (например, FCFS (первым пришел — первым обслужен), а также с относительными или абсолютными приоритетами, см. Джейсуол (Jaiswal) [65], Прабху (Prabhu) [112], Такач (Takács) [138, 139], а также Асмуссен (Asmussen) [10], Бочаров и Печинкин [25], Гнеденко и Коваленко [53], Коэн (Cohen) [36], Купер (Cooper) [38], Кокс и Смит (Cox and Smith) [40]). Кроме того, все нестационарные решения для очередей типа M/GI/1 получены в терминах двойных преобразований (по пространству и времени), из которых весьма затруднительно извлечь необходимую информацию относительно поведения системы обслуживания. (Подчеркнем дополнительно, что для получения нестационарных решений необходима более продвинутая математическая техника по сравнению с анализом системы в стационарном режиме.) Впрочем, некоторые исключения составляют вариации системы M/M/1, для которых известны транзиентные решения в явном виде. Как правило, точные результаты анализа в переходном режиме включают бесконечные суммы бесселевых функций (см., например, книгу Саати (Saaty) [119] или обзор Трипати (Tripathi) и Дуда (Duda) [140]). Впервые такое решение было получено Бейли (Baily) [15], который использовал производящие функции в дифференциальных уравнениях, описывающих

[2] Первоначальная идея EPS принадлежит Клейнроку (Kleinrock) [81] (1967).

[3] Система обслуживания EPS может рассматриваться как непрерывный предел очереди с дисциплиной RR (см. для деталей третий абзац раздела 2.1).

процесс числа требований. В более общих ситуациях получение точных решений в явном виде маловероятно. Поэтому были предложены различные численные и асимптотические методы, а также аппроксимации для вычисления вероятностей состояний и некоторых показателей производительности простых систем обслуживания. Однако мы не будем рассматривать такие аппроксимации в данной статье.

Прежде всего, эта статья посвящена точным решениям для вероятностных распределений основных характеристик системы M/GI/1—EPS. Полученные сначала для стационарного режима, эти решения в последнем десятилетии удалось распространить и на нестационарный случай. Иными словами, были найдены точные транзиентные решения в терминах многомерных преобразований. (В общем, такие решения могут рассматриваться как явные, по крайней мере в принципе, если мы считаем решение в терминах преобразований явным.) Эти решения развиты для получения совместного нестационарного распределения виртуального времени пребывания в системе требования, поступающего в момент t и имеющего в этот момент требуемую длительность обслуживания (длину) u, и числа требований в момент $t-$ в системе обслуживания M/GI/1—EPS (маргинальное нестационарное распределение времени пребывания требования является условным распределением по u).

Основные цели статьи состоит в более детальном по сравнению с [167] объяснении метода получения указанных результатов и обзоре точных аналитических решений, в особенности, полученных в последнем десятилетии. Асимптотическим и алгоритмическим решениям уделяется меньше внимания. Полностью оставлены в стороне и точные решения для системы M/GI/1 с другой популярной дисциплиной разделения процессора FBPS [4], несмотря на то, что эта модель исследована почти столь же глубоко как M/GI/1—EPS (в отличие от других вариантов дисциплин разделения процессора). (Относительно достижений в анализе модели FBPS см., например, [159] (1984), [110] (1986), [123] (1988), [164] (1989), [167] (1990), [25] (1995), [106] (2004), [175] (2004) и [9] (2005).[5]) Мы продемонстрируем также как большинство (если не все) известных или новых аналитических решений системы M/GI/1—EPS могут быть получены в качестве частных случаев, используя стандартные или нестандартные аргументы (например, с помощью абелевых/тауберовых (Abel's/Tauber's) теорем). В сущности, наши конструкции восходят к аналитическому методу из статей [155], [156], [158], детально отраженному и в монографическом виде [164] (1989). Основные результаты [164] доступны в виде обширного обзора [167] (1992) и для англоязычных читателей[6]. Статья продолжает, частично дополняет и углубляет обзоры [162] (отразивший достижения по 1987 г.), [167] (отразивший достижения по 1990 г.), но она отнюдь не претендует на полноту по сравнению, например, с обзорной статьей [167] [7]. За последнее десятилетие объем литературы по математической теории систем обслуживания с разделением процессора процессора настолько увеличился, что стало практически невозможно отразить в одной статье многие достижения, по которым, впрочем,

[4] FBPS — аббревиатура термина Foreground Background Processor Sharing. О различных вариантах дисциплин разделения процессора см. [162, 164, 167], здесь мы не останавливаемся на этом.

[5] Кроме того, в [1, 2, 3] (2004 — 2006) обнаружены новые факты относительно классов распределений, при которых дисциплина FBPS является неоптимальной.

[6] Однако представляется, что перевод на английский язык указанной обзорной статьи не вполне удовлетворителен с терминологической точки зрения, поскольку неизвестный переводчик, по–видимому, не является экспертом в теории очередей.

[7] Отметим также более ранние обзоры по этой проблематике и смежной тематике, принадлежащие Мак Кинни (McKinney) [96] (1969), Клейнроку (Kleinrock) (1970), (1972) (включенные позднее в его монографию [83] (1976)), Кобаяши и Конхейму (Kobayashi and Konheim) [86] (1977), Джейсуолу (Jaiswal) [66] (1982), Митрани (Mitrani) [98] (1986), Нейману [103] (1978) и Соловьеву [130] (1981). К этим работам примыкают, в частности, статья Варда и Витта (Ward and Whitt) [148] (2000) и диссертации Гришечкина [55] (1994) и Нунеса–Кьюеджа (Núñez–Queija) [105] (2000), изучивших новый аналитический метод, разработанный в [78, 158, 160, 164]. Некоторые частные вопросы теории систем с разделением процессора рассматриваются также в работах [17, 41, 44], [62, 73, 87], [89, 143], посвященных, в целом, другим проблемам.

появился ряд недавних обзорных статей. Некоторые из них будут упомянуты далее (без детального рассмотрения), что позволит нам сконцентрировать усилия только на сформулированных целях данной обзорной статьи. Конечно, авторы не стремились цитировать все работы по этой проблематике: их количество значительно превысило тысячу наименований (если учитывать более далекие от наших целей статьи явно прикладного характера), поэтому в обзоре указывались в первую очередь основные публикации, в которых были доказаны приведенные далее теоремы.

1.2. Некоторые смежные направления

Как уже отмечалось выше, мы обсуждаем только точные аналитические решения системы M/GI/1 с эгалитарным разделением процессора и некоторые специальные случаи этой системы обслуживания (важные проблемы аппроксимаций и асимптотических решений в стационарном режиме также почти не рассматриваются, поскольку известен ряд оригинальных и статей и обзоров по этой проблематике). В этом подразделе мы дадим лишь небольшое представление о некоторых смежных направлениях и их истокам.

Для обширной библиографии по различным типам аппроксимаций в теории очередей см., например, Бхат, Шелеби и Фишер (Bhat, Shalaby and Fisher) [21] (1979) (см. также [20]). Наиболее распространенным подходом к аппроксимации классических систем обслуживания является их анализ в условиях высокой загрузки (явно или неявно опирающийся на центральную предельную теорему), предложенный Кингманом (Kingman) [77](1961). Цель такого анализа состоит в выводе более простых выражений в явном виде для предельных распределений характеристик системы. Можно также отметить тесно связанную с этим подходом диффузионную аппроксимацию классических систем обслуживания, детально исследованную для многолинейных систем обслуживания Иглехартом и Виттом (Iglehart and Whitt) [61] (1970). Обзор других работ в этом направлении сделан, в частности, Глинном (Glynn) [52] (1990). Весьма популярна стала также флюидная, т.е. жидкостная аппроксимация (опирающуюся на усиленный закон больших чисел), предложенная Ньювеллом (Newell) в 1968 г. (см. [104, Ch. 6] (1971), [83, Ch. 2] (1976)).

С теоретической точки зрения случай большой загрузки интересен тем, что для классических систем обслуживания удается получить довольно простые асимптотические результаты в форме некоторых предельных теорем, которые зачастую можно распространить на более общие случаи (в частности, на случай рекуррентного входящего потока (типа GI) или даже типа G). В то же время точные результаты, если они достижимы в принципе, имеют очень сложный вид. Недавние достижения в этих направлениях для классических систем обслуживания хорошо изложены в монографиях [72, 152][8].

В последнее десятилетие весьма популярны стали аналогичные задачи, но уже для систем обслуживания с дисциплинами разделения процессора. В качестве примеров укажем на некоторые работы из многочисленных публикаций в этом направлении, см., например, [54, 55, 106, 125, 168, 184].

В общем, литература по асимптотическим решениям для стационарных характеристик систем с разделением процессора весьма (и даже чересчур) обширна. За последние годы в мировой литературе по этой проблематике появились десятки однообразных и дублирующих

[8] В сети Internet доступны даже дополнительные материалы к книге Витта [152] (суммарный объем этой книги с дополнением превышает 1000 стр.). Следует отметить также обзоры Стидхэма (Stidham) [132], Сайски (Syski) [137] и Джалалова (Dhalalow) [44] (1995), которые дают общее представление об основных достижениях классической теории очередей (и ее методах) за последние 50 лет. Тем не менее в статье [132], отражающей, впрочем, персональные взгляды автора, системы с разделением процессора упомянуты лишь мимоходом на стр. 199.

друг друга обзоров, информативность которых почти не увеличивается с быстрым ростом их количества (по–видимому, это может служить контрпримером для закона диалектики о переходе количества в качество). Поэтому мы сошлемся в этой связи только на небольшое количество оригинальных работ, как правило, стартующих от теорем из статей [78] (1978), [158] (1983), [122] (1984), [109] (1984), [160,162,167], [54], [125], [168]. В первую очередь к ним относятся работы [161] (1986), [185] (2000), [184] (2001) [69] (2002), [70] (2003), [55] (1994)(см. также [56] (2004) или статьи, цитированные в [70].В [161] выведены асимтотические оценки дисперсии стационарного времени пребывания в системе M/GI/1—EPS при малой и большой длине требования (в последнем случае эта оценка совпадает с оценкой Крамера (Cramér) [48] для риска разорения в обобщенном пуассоновском процессе). Это ветвь исследований была эффективно продолжена в [185] и [69, 70], в которых из формул работ [158, 162, 167] удалось извлечь дальнейшую информацию об асимптотическом поведении стационарного времени пребывания в системе M/GI/1—EPS для некоторых подклассов распределений длин требований с так называемым "тяжелым" хвостом "heavy tailed"[9]. Впрочем, основные результаты работ [70, 185] были предсказаны (но не доказаны) Клейнроком [83, p. 175] (1976).

Еще одно интересное достижение относится к явному виду асимптотического решения для числа требований в момент t (с нормировкой t) в системе M/GI/1—EPS, работающей в условиях перегрузки. Соответствующий результат был анонсирован в [167, теорема 6.4] (1990), впервые доказан (с довольно тяжеловесным доказательством) Гришечкиным [54] (1991), а позднее был распространен на случай рекуррентного входящего потока Жаном–Мари и Робером (Jean–Marie and Robert) [68](1994) (см.также [67]). Другие версии доказательств для случая пуассоновского входа содержат, в частности, работы [7] (в контексте дискриминаторного разделения процессора), [32, 181].

Относительно новым исследований является изучение свойств ветвящихся процессов, которые возникают при исследовании систем с разделением процессора. Классическим примером, восходящим к статьям Бореля (Borel) [26] (1942) и Кендалла (Kendall) [74] (1951), является получение распределения периода занятости в классической системе M/GI/1—FCFS (см. также [48]). В теории ветвящихся процессов (см. [88, гл. 13], [124], [142]) такие процессы изучены очень глубоко. Однако оказалось (см. [163], [164, стр. 95]), что исследование времени пребывания в системе M/GI/1—EPS тесно связано с более сложными и слабее изученными ветвящимися процессами Крампа– Мода–Ягерса [88, 57, 143] (такие процессы были введены при решении некоторых проблем биологии). Переформулировка метода и результатов анализа системы M/GI/1—EPS из работ [79, 156, 158, 160], где они излагались с позиций теории очередей, в терминах этих процессов была осуществлена в [54, 55]. Это позволило получить ряд интересных предельных теорем.

Тем не менее, системы обслуживания с разделением процессора — объект теории очередей, и в данном обзоре они рассматриваются с позиций именно этой теории с дополнительным привлечением некоторых нетривиальных вероятностных конструкций, в частности, случайной замены времени.

Раздел 2 содержит основные результаты теории системы обслуживания M/GI/1—EPS. Раздел состоит из трех подразделов, содержание которых более или менее соответствует основным этапам становления и развития этой теории. Принятые обозначения незначительно отличаются от обозначений, использованных в обзоре [167], поэтому не будем останавливаться на них. Некоторые дополнительные комментарии приведены в сносках. Для утверждений применяется двойная нумерация: первая цифра соответствует номеру раздела, а вторая цифра — номеру

[9] Теория таких вероятностных распределений впервые появилась в финансовой математике. Ее детальное изложение можно найти, в частности, в монографиях [46] (1997), [128] (1998). В этой связи упомянем также гл. 3 книги [141] (2005).

теоремы, формулы и т.п. в отдельных списках этих объектов по каждому разделу (например, теорема 2.2 указывает на 2–ю по счету теорему от начала раздела 2, а следствие 2.4 идентифицирует 4-е по счету следствие от начала раздела 2 из всего набора следствий этого раздела). Следуя примеру [128], литература (независимо от языка) дается единым списком в порядке, соответствующем латинскому алфавиту.

2. ТЕОРИЯ СИСТЕМЫ ОБСЛУЖИВАНИЯ M/GI/1—EPS. СТАЦИОНАРНЫЙ РЕЖИМ

2.1. Вводные замечания

В этом разделе мы представим наиболее существенные достижения (и ключевые особенности их вывода) в точном анализе системы M/GI/1 с эгалитарным разделением процессора. Понимание метода получения стационарных распределений основных вероятностно–временных характеристик позволит сократить объяснения для нестационарных решений. Основное внимание уделяется проблеме нахождения (условного) распределения времени пребывания требования в системе M/GI/1—EPS в терминах преобразований Лапласа–Стилтьеса (ПЛС), полученным решениям, методам решений и некоторым смежным результатам. По ходу изложения дается представление о наиболее важных достижениях, существовавших до наших работ и появившихся после них, о других подходах к исследованию, а также некоторая вспомогательная информация.

Рассмотрим систему M/GI/1 с дисциплиной эгалитарного разделения процессора (EPS)[10]. Эгалитарное разделение процессора обычно рассматривалось как предельная форма дисциплины RR при при бесконечно малой величине кванта обслуживания. Однако более перспективным оказалось описание различных разновидностей дисциплин разделения процессора как систем с переменными скоростями обслуживания требований, зависящими от состояния системы.

Дисциплина EPS введена Клейнроком [81] как предельный аналог дисциплины RR при устремлении величины кванта к нулю. Мы рассматриваем систему с дисциплиной EPS иначе — как систему, одновременно обслуживающую каждое из присутствующих требований с одинаковой скоростью, которая зависит от общего числа требований и равна $1/n$, если число требований в текущий момент равно n, $n = 1, 2, \ldots$. В моменты изменения числа обслуживаемых требований происходят скачки скорости обслуживания. Если называть остаточной длиной требования количество работы по его обслуживанию с единичной скоростью, измеряемое в единицах времени, то под скоростью обслуживания понимается предел отношения изменения остаточной длины требования за промежуток времени Δ к Δ при $\Delta \to 0$ (термин "длина" требования будет заменять неоднозначно трактуемый термин "длительность обслуживания"). Дисциплину EPS можно описать следующим образом. Поступившее в систему требование сразу начинает обслуживаться (очередь в традиционном ее понимании отсутствует) и обслуживается с переменной скоростью (понимаемой в кинематическом смысле) до тех пор, пока его остаточная длина не станет равной нулю. В моменты поступлений новых требований или ухода обслуженных происходят скачки скорости обслуживания. Скорость обслуживания

[10] Во всех классических дисциплинах обслуживания в однолинейных системах в любой момент времени обрабатывается не более одного требования, причем скорость обслуживания любого требования равна либо нулю при ожидании, либо единице при обслуживании. Другим классом дисциплин в таких системах обслуживания, включающим предельные формы дисциплин разделения времени, являются дисциплины разделения процессора. Наиболее важной из них является EPS. В частности, эта дисциплина наиболее широко используется в различных приложениях. Долгое время такая система оставалась изученной только на поверхностном уровне (удавалось получить формулы только для средних характеристик), поэтому проблема нахождения (стационарного) времени пребывания стала считаться аналитически неразрешимой. Краткий эскиз ее решения появился в [78] (1978), но полное доказательство дано впервые в [156] (1981) и усовершенствовано в [158] (1983).

флюктуирует во времени, а длительность пребывания отдельного требования в системе зависит не только от ранее поступивших, но и от последующих поступлений. Это на порядок усложняет анализ системы EPS по сравнению, например, с системой FCFS. В процессе обслуживания за бесконечно малый промежуток времени Δ происходит уменьшение остаточной длины каждого из n требований на $\Delta/n + o(\Delta)$. При этом ресурс процессора делится поровну (равномерно) между всеми обслуживаемыми требованиями. Такой способ разделения ресурса с точки зрения требований будет справедливым, что и подчеркивается в названии.

Практическое значение такой математической идеализации состоит в том, что она отражает наиболее существенную особенность систем с разделением времени (если под требованием понимать отдельное задание пользователя) или мультиплексных узлов коммутации пакетов в сети Internet (требованием является сообщение или его сегмент) — замедление обслуживания каждого требования, пропорциональное их общему числу в текущий момент. Пусть λ — интенсивность входящего пуассоновского потока, а длины требований — независимые и одинаково распределенные случайные величины (сл.в.) с произвольной функцией распределения (ф.р.) $B(x) = \mathrm{P}(B \le x),$, $B(0+) = 0$. Положим

$$\beta_1 = \int_0^\infty (1 - B(x))dx < \infty \quad \text{и} \quad \beta(s) = \int_0^\infty \mathrm{e}^{-sx}\, dB(x).$$

Необходимым и достаточным условием существования стационарного режима является

$$\rho = \lambda\beta_1 < 1. \tag{2.1}$$

Пусть $V(u)$ — стационарное время пребывания в системе требования, которое в момент поступления имеет длину u, т. е. реализацию сл.в. B, равную u (информация о величине u по каждому требованию недоступна для системы). Понятно, что $V(u)$ представляет собой также длительность фактического обслуживания требования длины u. Положим $v(s,u) \doteq \mathsf{E}[\exp(-\mathsf{sV(u)})]$ (здесь и далее u рассматривается как параметр). Отметим, что $v(s,u)$ есть ПЛС условного распределения $V(x|u) = \mathsf{P(V(u) \le x|B = u)}$.

Для системы M/M/1 с дисциплиной EPS (условное) распределение сл.в. $V(u)$ впервые получено Коффманом и др. [35] (1970); для системы M/G/1—EPS до недавнего времени был известен только первый момент ф.р. сл.в. $V(u)$, полученный впервые Сакатой и др. [120] (1969) До недавнего времени появлялись работы по различным способам вычисления $\mathsf{E[V(u)]}$ в этой системе (см., например, [12,51,83,107]). Стационарное распределение времени пребывания требования длины u в системе M/GI/1—EPS впервые найдено в [78,156,157] в терминах ПЛС с помощью обобщения и развития предложенного в [155] подхода к получению точного решения для $v(s,u)$ в системе M/M/1—RR (дополнительные подробности по различным методам и результатам анализа системы RR и ее вариантам можно найти [164, гл. 1]). В [158] усовершенствован метод вычисления распределения сл.в. $V(u)$ в системе M/GI/1—EPS и получен ряд новых вероятностных характеристик, на основе которых впервые выявлены характерные свойства системы, существенно дополняющие перечисленные в [83, гл.4]. Одна из основных тем данного раздела — описание метода и результатов вычисления $v(s,u)$ в системе M/GI/1—EPS. Следует отметить, что здесь предполагается существование стационарного эргодического распределения $V(x|u)$, поэтому во введенных обозначениях не указана зависимость от времени.

Если получено $v(s,u)$, то ПЛС распределения безусловного времени пребывания V требования в системе находится как

$$v(s) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{sV}}] = \int_0^\infty \mathsf{v(s,u)\,dB(u)}. \tag{2.2}$$

Случайная величина времени ожидания требования длины u (точнее, потерянного времени сверх необходимого для обслуживания со скоростью единица) определяется равенством $W(u) = V(u) - u$. Отсюда, зная $v(s,u)$, легко получить $w(s,u) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{sW(u)}}]$.

Моменты распределения $V(x|u)$ вычисляются (с формальной точки зрения) как

$$v_j(u) = \lim_{s\downarrow 0}(-1)^j \partial v(s,u)/\partial s, \quad \mathsf{Var}[V(u)] = v_2(u) - v_1^2(u). \tag{2.3}$$

Кратко остановимся на распределении (стандартного) периода занятости П. Положим $\pi(s) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{s\Pi}}]$. Отсчет времени — с момента начала периода занятости. Все дисциплины разделения процессора принадлежат классу *консервативных* дисциплин. Свойство консервативности означает, что длины требований не зависят от дисциплины и отсутствуют искусственные простои прибора. Иными словами, при обслуживании не порождаются дополнительные и не уничтожаются имеющиеся требования (в частности, классу консервативных принадлежат дисциплины с дообслуживанием и без затрат времени на переключение прибора). Применительно к дисциплинам разделения процессора в однолинейных системах свойство консервативности означает, что в каждый момент времени на периоде занятости сумма мгновенных скоростей обслуживания всех имеющихся требований равна единице.

Теорема 2.1. *В системе M/GI/1 с любой дисциплиной разделения процессора ПЛС распределение (стандартного) периода занятости $\pi(s)$ удовлетворяет известному для дисциплин LCFS и FCFS функциональному уравнению Такача*

$$\pi(s) = \beta(s + \lambda - \lambda\pi(s)). \tag{2.4}$$

Уравнение (2.4) определяет единственную функцию $\pi(s)$, аналитическую в полуплоскости Re $s > 0$, *в которой $|\pi(s)| \le 1$. Функция $\pi(s)$ вполне монотонна. Если $\rho < 1$, то $\pi(0^+) =$* $\Pi(+\infty) = p^* = 1$, *в противном случае $\pi(0+) < 1$,* $\Pi(+\infty) = p^*$, *где p^* — единственный корень уравнения $p^* = \beta(\lambda - \lambda p^*)$, лежащий в (0,1).*

Комментарии к доказательству. Два варианта вывода (2.4) даны в [164, стр. 57, 62-63]. Один из них появился ранее в [78]. Доказательство в [164, §2.3] основано на редукции к ветвящимся процессам. Классическим примером служит вывод (2.4) для системы M/GI/1—FCFS (см., например, [48, стр. 542-544]). Для этого в теории очередей обычно рассматривались хорошо изученные ветвящиеся процессы, возникающие в системе M/GI/1 с дисциплиной LCFS (Last Come — First Served). Доказывалась справедливость (2.4) для системы M/G/1—LCFS и отмечалось, что ф.р. сл.в. П в системах LCFS и FCFS совпадают.

В [78, 156, 158] для этой цели были рассмотрены более сложно устроенные ветвящиеся процессы, возникающие в системе M/GI/1 с дисциплиной EPS. При этом, в отличие от [48], было введено новое правило порождения прямого потомства, основанное на равновероятном случайном выборе (подробности см. далее). После составления дифференциальных уравнений, описывающих время жизни $\Pi(x)$ соответствующего ветвящегося процесса, порожденного предком длины x, получено решение для

$$\pi(s,x) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{s\Pi(x)}}] = \mathrm{e}^{-\mathsf{x(s+\lambda-\lambda\pi(s))}}, \tag{2.5}$$

из которого после снятия условия по x вытекает (2.4), справедливое для системы M/GI/1—EPS.

Далее рассматривается процесс незаконченной работы $U(t)$ (сумма остаточных длин всех имеющихся требований в момент t). Поскольку сл.в. П есть интервал времени, на котором $U(t) > 0$, то отсюда следует, что ф.р. сл.в. П инвариантна относительно любой консервативной дисциплины в системе M/GI/1. Подклассом таких дисциплин являются все дисциплины разделения процессора. Они, в частности, удовлетворяют закону сохранения Клейнрока [83], и не используют информации о длинах требований. □

Замечание 2.1. Равенство (2.5) совпадает с ПЛС ф.р. стационарного времени пребывания $V(x)$ требования длины x в системе M/G/1 с дисциплиной LCFS–P (обслуживание в обратном порядке с прерываниями). При дисциплине LCFS–P каждому новому поступлению дается высший абсолютный приоритет, а в очереди, организованной по принципу стека, размещаются прерванные требования для дальнейшего дообслуживания.

2.2. Стационарное распределение числа требований

Ряд обозначений для модели M/GI/1—EPS введен выше. Долгое время для нее был известен только первый момент стационарного распределения времени пребывания требования длины u (см. (2.22)) и стационарное распределение числа требований (см. формулу (2.8)[11]). Введем в рассмотрение кусочно–линейный марковский процесс $X_0(t) = \{L(t), x_i(t), i = 1, \ldots, L(t); t \in \mathbf{R}_+\}$, характеризующий состояние системы в момент t. Здесь $L(t)$ — число требований в системе EPS в момент t, а дополнительные переменные $x_i(t)$ указывают на остаточную длину i–го требования в момент t (порядок нумерации требований не имеет значения в силу специфики дисциплины EPS). Пространство состояний процесса $X_0(t)$ есть $\{\mathbf{Z}_+ \times \{\emptyset \cup {\mathbf{R}_+}^1 \cup {\mathbf{R}_+}^2 \ldots\}\}$. Здесь $\emptyset$ обозначает состояние, когда $L(t) = 0$, $\mathbf{Z}_+ = \{0, 1, 2, \ldots\}$. Пусть также процесс $\{V_t(u),\ t \in \mathbf{R}_+\}$ описывает время пребывания требования, которое приходит в момент t и имеет в этот момент длину u. В силу следствия из теоремы Смита для регенерирующих процессов [131] (точками регенерации служат моменты окончания периодов занятости) справедлива

Теорема 2.2. *При выполнении условия (2.1) процессы $\{V_t(u)\}$ и $X_0(t)$ обладают единственными стационарными распределениями.*

Теорема 2.2 обеспечивает существование и единственность следующих пределов:

$$v(s, u) = \lim_{t\to\infty} \mathsf{E}[\mathrm{e}^{-\mathsf{s}\mathsf{V}_\mathsf{t}(\mathsf{u})}];$$

$$\mathsf{P}_n(x_1, \ldots, x_n) = \lim_{t\to\infty} \mathsf{P}_n(t; x_1, \ldots, x_n) = \lim_{t\to\infty} \mathsf{P}(L(t) = n; x_i(t) \in [x_i, x_i + dx_i), i = 1, 2, \ldots, n);$$
$$\mathsf{P}_n = \lim_{t\to\infty} \mathsf{P}(L(t) = n). \tag{2.6}$$

В (2.6) $\mathsf{P}_n(x_1, \ldots, x_n)$ означает стационарную плотность вероятности пребывания процесса $X_0(t)$ в состоянии $(n; x_1, \ldots, x_n)$. Здесь неявно предполагается, что финальное распределение процесса $X_0(t)$ при фиксированном числе требований $n \geq 1$ обладает плотностью. Это условие выполняется, если предположить для простоты, что ф.р. $B(x)$ имеет плотность $\beta(x)$.

Теорема 2.3. *В системе M/G/1— EPS при выполнении условия (2.1)*

$$\mathsf{P}_n(x_1, \ldots, x_n) = \mathsf{P}_0 \lambda^n \prod_{i=1}^{n} [1 - B(x_i)],\ n \geq 1,\ \text{где } \mathsf{P}_0 = \mathsf{P}\{\emptyset\}. \tag{2.7}$$

[11] Эскиз доказательства приведен впервые в [120]. Применён метод введения дополнительных переменных, в качестве которых использовались обслуженные части длин (достигнутое обслуживание) i-го требования, $i = 1, 2, \ldots, L(t)$. Поведение исследуемого процесса описывалось системой интегро–дифференциальных уравнений с некоторыми граничными условиями.

Эти результаты находятся после составления и решения дифференциальных уравнений Колмогорова для плотностей одномерных распределений компонент процесса $X_0(t)$. Те же результаты в общем случае можно получить при замене плотносей $\mathsf{P}_n(x_1,\ldots,x_n)$ дифференциалами соответствующих ф.р., используя обобщенные функции.

Плотность $\mathsf{P}_n(x_1,\ldots,x_n)$ симметрична относительно любой перестановки переменных x_i, $i = 1,2,\ldots,n$ в силу того, что предположение об упорядоченности дополнительных координат процесса $X_0(t)$ не делалось.

Отметим, что (2.7) было получено в [79] при рассмотрении процесса $A_0(t)=\{L(t),a_i(t),i=1,\ldots,L(t);t\in\mathbf{R}_+\}$ как вспомогательный результат, необходимый при доказательстве пуассоновского характера выходящего потока. Отличие процесса $A_0(t)$ от $X_0(t)$ заключается в том, что $a_i(t)$ имеет смысл не остаточной, а обслуженной части длины i–го требования. Плотности вероятностей состояний процесса $A_0(t)$ описываются равенством, аналогичным (2.7) при соответствующей замене переменных, т. е.

$$\mathsf{P}_n(a_1,\ldots,a_n)=\mathsf{P}_0\lambda^n\prod_{i=1}^{n}[1-B(a_i)],\ n\ge 1,\ \text{где } \mathsf{P}_0=\mathsf{P}\{\emptyset\}. \tag{2.8}$$

Попутно отметим, что от (2.8) нетрудно перейти к (2.7), однако найти простую связь (2.7) с (2.8) в обратную сторону не удается.

Следствие 2.1. *Число требований в системе M/G/1 с дисциплиной EPS имеет геометрическое распределение*

$$\mathsf{P}_n=(1-\rho)\rho^n,\ n=0,1,2,\ldots, \tag{2.9}$$

инвариантное относительно вида $B(x)$ при фиксированном среднем β_1.

Доказательство.

$$\mathsf{P}_n=\int_0^\infty\ldots\int_0^\infty \mathsf{P}_n(x_1,\ldots,x_n)\,dx_1\ldots dx_n=\mathsf{P}_0\rho^n,\ n\ge 1.$$

Из условия нормировки $\sum_{n=0}^{\infty}\mathsf{P}_n=1$ находим $\mathsf{P}_0=1-\rho$. □

Еще одно доказательство следствия 2.1 дано после следствия 3.3.

Свойства инвариантности такого же типа характерны и для многолинейных систем. Например, вероятность нахождения n требований в классической системе M/GI/∞ с пуассоновским входящим потоком интенсивности λ имеет вид $\mathsf{P}_n=\mathrm{e}^{-\rho}\,\rho^n/n!,\ n=0,1,2,\ldots$

Замечание 2.2. Отмеченная аналогия становится более понятной, если рассмотреть процесс обслуживания в системе $M/GI/1$—EPS в новом масштабе времени: интервалы времени, в течение которых в системе обслуживается $n>1$ требований, сжимаются в n раз [79] (эта идея предложена А.Д.Соловьевым в 1975). При таком изменении масштаба времени скорость обслуживания уже не является функцией от n, а постоянна и равна единице. Однако при этом усложняется структура входящего потока. Свойство отсутствия последействия сохраняется, но интенсивность становится переменной, равной $n\lambda$, если в системе находится n требований, $n\ge 1$, и равной λ, если $n=0$. Входящий поток на отрезке времени, когда в системе находится n требований, можно рассматривать как суперпозицию n независимых пуассоновских потоков интенсивности λ. Таким образом, в новом масштабе времени система $M/GI/1$—EPS по вероятностной структуре процесса обслуживания превращается в систему $\tilde{M}/GI/\infty$ с марковским входящим потоком переменной интенсивности $n\lambda$, в которой вновь поступающее требование с одинаковой вероятностью может занять любой свободный прибор. Содержательно описанное здесь преобразование случайных процессов, с помощью которого проводилось теоретическое обоснование наших новых аналитических методов анализа системы с дисциплиной EPS и другими дисциплинами разделения процессора [79, 160, 164, 166], является одним из примеров случайной замены времени [23,88] в теории случайных процессов. Случайная замена времени впервые введена Волконским в 1958 г. [147] в контексте абстрактной теории случайных процессов (см., например, [23]) и использовалась, в частности, в теории стохастического интегрирования [128] (обобщенная формула Ито для выпуклых функций от броуновского движения, преобразование любого непрерывного локального мартингала с квадратической вариацией, стремящейся к ∞, в броуновское движение, интегрирование по возрастающему процессу и т.п.). В теории очередей подобное преобразование, но для другого класса исследуемых случайных процессов — возникающих в системах с разделением процессора — впервые применено в [79,159,160], позволив, в частности, дать теоретическое обоснование введенным в [155,156,158] аналитическим методам.

Используя замечание 2.2 докажем еще одно следствие теоремы 2.3.

Следствие 2.2. *Стационарные вероятности состояний* $\mathsf{P}^\star_n(x_1, \ldots, x_n)$ *системы* $\tilde{M}/G/\infty$ *имеют вид*

$$\mathsf{P}^\star_n(x_1, \ldots, x_n) = \mathsf{P}^\star_0 \frac{\lambda^n}{n} \prod_{i=1}^{n} [1 - B(x_i)], \ \ n \geq 1, \tag{2.10}$$

где

$$\mathsf{P}^\star_n = \int_0^\infty \ldots \int_0^\infty \mathsf{P}^\star_n(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = \mathsf{P}^\star_0 \rho^n / n, \tag{2.11}$$

$$\mathsf{P}^\star_0 = \mathsf{P}\{\emptyset\} = (1 - \ln(1-\rho))^{-1}, \ \ \rho < 1.$$

Доказательство. Когда система M/GI/1—EPS содержит $n > 1$ требований, то "новое" время течет в ней в n раз быстрее реального времени (см. замечание 2.2). Поэтому отношение стационарных вероятностей пребывания в состояниях $(n; x_1, \ldots, x_n)$, $n \geq 1$ для систем M/GI/1—EPS и $\tilde{M}/GI/\infty$ равно n. В силу этого факта из формулы (2.7) вытекает (2.10). Равенство (2.10) приводит к (2.11), что позволяет получить с учетом условия нормировки и ряда Тейлора для натурального логарифма формулу для стационарной вероятности $\mathsf{P}^\star_0$ отсутствия требований в системе $\tilde{M}/GI/\infty$. □

2.3. Эскиз метода декомпозиции на элементы задержки

Дадим эскиз аналитического метода получения $v(s, u) = \mathsf{E}[\mathrm{e}^{-\mathsf{s}\mathsf{V}(\mathsf{u})}]$ в системе M/GI/1—EPS. Под сл.в. $V(u)$ будем понимать промежуток времени, за который остаточная длина требования уменьшится на u. Эквивалентно, это промежуток времени, за который требование (считаем, что оно поступает в систему в нулевой момент времени) получит приращение обслуженной части длины, равное u. Суть метода состоит в рассмотрении динамики прохождения по системе некоторого виртуального требования длины u (которое будем считать помеченным) при начальном условии, что в момент своего поступления оно застает систему в состоянии $(n; x_1, \ldots, x_n)$, и разложении сл.в. $V(u)$ на сумму "*элементов задержки*". Для этого с каждым из n требований, которых встречает помеченное при своем поступлении, и с собственно помеченным требованием (предками) связываются ветвящиеся процессы порождения потомков, т.е. некоторой части новых поступлений после прихода виртуального помеченного требования. Случайный механизм, относящий новые поступления к какому-либо потомству, основан на правиле: новое требование, застающее в системе n требований, считается прямым потомком любого (но только одного) из них с вероятностью $1/n$. Далее вводится сл.в. $\Phi(x, u)$ элемента задержки помеченного требования, которая представляет собой суммарную длительность обслуживания предка длины x и его прямых потомков за отрезок времени, в течение которого остаточная длина помеченного требования уменьшится на u. Случайную величину $\Phi(x,u)$ можно также трактовать как некоторый (марковский) функционал от соответствующего ветвящегося процесса (обрывающегося в момент ухода помеченного требования из системы), описывающий его время жизни, или, что проще, как длительность специфичного *обрывающегося (суб)периода занятости* (это одно из новых для теории очередей понятий, принципиально обобщающее хорошо известное понятие стандартного периода занятости), который открывается требованием длины x и обрывается в момент, когда приращение достигнутого обслуживания помеченного требования составляет величину u. В рассматриваемой системе EPS вероятностная структура компонент этого субпериода занятости напоминает структуру компонент обычного периода занятости с тем существенным отличием, что каждая последующая компонента зависит от момента обрыва и длины потомка, в связи с чем она "стохастически меньше" предыдущей (в смысле введенного Штойяном (Stoyan) [133, 134] отношения порядка $\overset{1}{\le}$ для функций распределения) [12]

Отметим, что при $u \to \infty$ сл.в. $\Phi(x,u)$ превращается в стандартный период занятости, у которого длина открывающего его требования фиксирована и равна x. (ПЛС ф.р. сл.в. $\Phi(x,\infty)$ дается правой частью (2.5). Иными словами, это время первого достижения уровня $x=0$ процессом незаконченной работы $U(t)$ при условии $U(0)=x$.) При $x \ge u$ сл.в. $\Phi(x,u)$ не зависит от x. В этом случае удобно ввести специальное обозначение $D(u) \overset{d}{=} \Phi(x,u)$ при $x \ge u$. Теперь можно показать, что при указанном выше начальном условии — в момент 0 система находится в состоянии $(n; x_1, \ldots, x_n)$ — справедлива декомпозиция

$$V(u) \overset{d}{=} \sum_{i=1}^{n} \Phi(x_i, u) + D(u), \tag{2.12}$$

где все составляющие — независимые сл. в. (Независимость этих сл.в. — нетривиальный факт, наиболее элегантно доказываемый с помощью случайной замены времени.) Соотношения типа (2.12) являются ключевыми для развиваемой далее теории. Подчеркнем, что компоненты правой части (2.12) — аналоги элементов задержки помеченного требования в системе RR [155,164], обобщенные на рассматриваемый случай: произвольное распределение длин $B(x)$ и одновременное обслуживание всех требований с переменной скоростью, при которой короткие по длине требования "обгоняют" по моменту ухода из системы более длинные требования, поступившие ранее. Положим $\varphi(s,x,u) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{s\Phi(x,u)}}]$ и $\delta(s,u) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{sD(u)}}]$. Рассматривая инфинитезимальные изменения сл. в. $\Phi(x,u)$, можно составить систему уравнений с начальными условиями, которым удовлетворяет $\varphi(s,x,u)$ и $\delta(s,u)$ в области $x<u$

$$\frac{\partial \varphi(s,x,u)}{\partial x} + \frac{\partial \varphi(s,x,u)}{\partial u} + \left[s + \lambda - \lambda \int_0^\infty \varphi(s,y,u)\, dB(y)\right] \varphi(s,x,u) = 0, \tag{2.13}$$

$$\frac{\partial \delta(s,u)}{\partial u} + \left[s + \lambda - \lambda \int_0^\infty \varphi(s,y,u)\, dB(y)\right] \delta(s,u) = 0, \tag{2.14}$$

$$\delta(s,0) = \varphi(s,0,u) = \varphi(s,x,0) = 1.$$

Для вывода этих уравнений удобно использовать следующие стохастические равенства, описывающие изменение элемента задержки виртуального требования:

[12] Приведем определения только двух часто используемых отношений порядка $\overset{(1)}{\le}$ и $\overset{(2)}{\le}$:

$B_i(x) \overset{(1)}{\le} B_j(x)$, если $1 - B_i(x) \le 1 - B_j(x)$ для $\forall x \ge 0$;

$B_i(x) \overset{(2)}{\le} B_j(x)$, если $\int_x^\infty (1 - B_i(y))\, dy \le \int_x^\infty (1 - B_j(y))\, dy$ для $\forall x \ge 0$.

$$\Phi(x+\Delta, u+\Delta) = \begin{cases} \Delta + \Phi(x,u), & \text{если не появились потомки,} \\ \Delta + \Phi(x,u) + \Phi(x,u), & \text{если появился потомок длины } x. \end{cases} \tag{2.15}$$

(Равенство (2.15) выполняется в области $x < u$.) Аналогично,

$$D(u+\Delta) = \begin{cases} \Delta + D(u), & \text{если не появились потомки,} \\ \Delta + D(u) + \Phi(x,u), & \text{если появился потомок длины } x. \end{cases} \tag{2.16}$$

Применяя к (2.15) and (2.16) формулу полного математического ожидания, получаем в результате предельного перехода при $\Delta \to 0$ систему уравнений (2.13) и (2.14).

После некоторых преобразований, сводящих систему уравнений (2.13)–(2.14) к дифференциальному уравнению с частными производными первого порядка и решения последнего методом характеристик, получаем

$$\varphi(s,x,u) = \begin{cases} \delta(s,u) & \text{при } x \geq u, \\ \delta(s,u)/\delta(s,u-x) & \text{при } x < u, \end{cases} \tag{2.17}$$

где

$$\delta(s,u) = \mathrm{e}^{-(s+\lambda)u}/\psi(s,u), \quad u \geq 0, \tag{2.18}$$

а функция ψ задается своим преобразованием Лапласа (ПЛ) по u (аргумент q) функции $\psi(s,u)$, являющейся ПЛ по t (аргумент s) функции $\Psi(t,u)$ двух переменных

$$\tilde{\psi}(s,q) \doteq \int_0^\infty e^{-qu}\psi(s,u)\,du = \frac{s+q+\lambda\beta(s+q+\lambda)}{(s+q+\lambda)[q+\lambda\beta(s+q+\lambda)]}, \quad Re\ s > 0,\ q > -\lambda\pi(s). \tag{2.19}$$

Применяя ПЛС к обеим сторонам равенства (2.12), приходим к теореме 2.4.

Теорема 2.4. *(1978) В системе M/GI/1—EPS*

$$\mathsf{E}[\mathrm{e}^{-\mathsf{s}\mathsf{V}(\mathsf{u})}\,|(\mathsf{n};\mathsf{x}_1,\ldots,\mathsf{x}_\mathsf{n})] = \delta(\mathsf{s},\mathsf{u})\prod_{\mathsf{i}=1}^{\mathsf{n}}\varphi(\mathsf{s},\mathsf{x}_\mathsf{i},\mathsf{u}), \quad \mathsf{Re}\ \mathsf{s} > 0, \tag{2.20}$$

где функции φ и δ задаются равенствами (2.17)– (2.19).

Комментарии к доказательству.[13] Впервые (краткий и требующий дополнительной аргументации) вывод (2.20) осуществлен в [78] (1978), два варианта полных доказательств даны в [156, 157] (1981), усовершенствованная версия доказательства содержится в [158] (1983). Основные результаты с дополнительным обсуждением введенных конструкций включены в книгу [164, гл.2] (1989). □

Замечание 2.3. При доказательстве теоремы 2.4 не использовалось равенство $\pi(0^+) = 1$, которое имеет место при выполнении условия (2.1), поэтому теорема 2.4 справедлива и при $\rho \geq 1$.

[13] Выше дано объяснение, в основном, к “вычислительной части” доказательства, лишь намекнув на одну из трудностей: доказательство независимости компонент декомпозиции (2.12). О другой трудности в рамках обзора говорить бессмысленно; впрочем, некоторое представление о ней и о способах ее преодоления с помощью нетривиальных конструкций из теории случайных процессов будет дано позже. По этим причинам доказательство неоднократно модифицировалось. Внешняя простота эскиза метода обманчива.

Замечание 2.4. В некоторых случаях могут быть полезны эквивалентные формы записи равенства (2.17). В качестве примера приведем одну из них

$$\varphi(s,x,u)=\mathrm{e}^{-(x\wedge u)(s+\lambda)+\lambda\int_0^{x\wedge u}\varphi_B(s,u-y)\,dy},\ x\in[0,\infty),$$

где

$$\varphi_B(s,t)\doteq\int_0^\infty\varphi(s,x,t)\,dB(x)=\int_0^t \mathrm{e}^{-\int_{t-x}^t(s+\lambda-\lambda\varphi_B(s,y))\,dy}\,dB(x)+(1-B(t))\,\mathrm{e}^{-\int_0^t(s+\lambda-\lambda\varphi_B(s,y))\,dy}.$$

Последняя формула представляет собой функциональное уравнение для ПЛС ф.р. обрывающегося в момент t периода занятости в системе M/GI/1—EPS. Решение этого уравнения получено нами в терминах функции ψ $(\psi(s,t)\doteq\exp(-\lambda\int_0^t\varphi_B(s,y)\,dy))$ (точнее, в терминах изображения по Лапласу функции ψ — см. формулу (2.19)).

Замечание 2.5. Более полная информация о различных формах интерпретации уравнений (2.13), (2.14) и способах их составления содержится в [164]. В [157] продемонстрирована возможность использования еще одного приема, слегка облегчающего составление таких уравнений: приема введения дополнительного события (окраски требований), интенсивно развивавшегося, в частности, в работах [84], [97], [100] по очередям с приоритетами и системам с другой и более простой для исследования разновидностью дисциплины разделения процессора [14].

При пуассоновском входящем потоке поступление требования создает момент случайного наблюдения за системой (свойство PASTA), следовательно вероятности состояний в момент поступления совпадают с вероятностями состояний в произвольный момент времени. Поэтому при $\rho<1$ мы вправе снять в (2.20) условие по $(n;x_1,\ldots,x_n)$ с помощью теоремы 2.3, описывающей мультипликативную форму стационарной плотности совместной ф.р. числа $L=n$ требований с остаточными длинами, принадлежащими бесконечно малым окрестностям точек $x_1,\ldots,x_n$. Это дает

$$v(s,u)\doteq\mathsf{E}[\mathrm{e}^{-\mathsf{sV(u)}}]=\frac{(1-\rho)\delta(\mathsf{s},\mathsf{u})}{1-\lambda\int_0^\infty\varphi(\mathsf{s},\mathsf{x},\mathsf{u})(1-\mathsf{B}(\mathsf{x}))\,\mathsf{dx}},$$

что с учетом (2.17) — (2.19) приводит к

Теорема 2.5. *(1978) В системе M/GI/1—EPS при выполнении условия (2.1)*

$$v(s,u)=\frac{(1-\rho)\,\mathrm{e}^{-u(s+\lambda)}}{\psi(s,u)-\lambda\int_0^u\mathrm{e}^{-x(s+\lambda)}\,\psi(s,u-x)(1-B(x))\,dx-\lambda\,\mathrm{e}^{-u(s+\lambda)}\int_u^\infty(1-B(x))\,dx},\quad(2.21)$$

$$Re\ s>0.$$

[14] Такая дисциплина (точнее, класс дисциплин) назван в [84] “пакетным обслуживанием в режиме разделения процессора”; в сущности, это обслуживание по правилу EPS только некоторой группы требований, формируемой по принципу: отсечь влияние большинства будущих поступлений на время пребывания в системе любого требования из обслуживаемой группы. В момент окончания обслуживания группы требований в прибор поступают все накопившиеся в очереди требования, а последующим поступлениям запрещается вход в прибор до следующего момента освобождения прибора от обслуживания всей текущей группы. Тем самым создаются предпосылки к применению классических методов теории очередей, не приспособленных для нахождения $\mathsf{E}[\mathrm{e}^{-\mathsf{sV(u)}}]$ в системе M/GI/1— EPS в ее “чистом” виде. В западной литературе такие модели называются “*gated processor–sharing*”. Они введены Клейнроком в 1970 г. в контексте значительно более широкого класса “эгоистических” (*selfish*) дисциплин разделения процессора (см. ссылку [49] в [83, гл.4], а также [167, стр.13–14]), средние характеристики таких систем найдены в [108]; стационарные распределения характеристик получены в [84].

Заметим, что (2.21) — преобразование безгранично делимого распределения на $[0,\infty)$. Иными словами, положительный корень m-й степени $v^{1/m}(s,u)$, $m=1,2,\dots$, есть ПЛС некоторого распределения вероятностей[15].

Без особых трудностей теорема 2.5 обобщается на на случай нескольких классов требований [164, стр. 70]. Чтобы дать представление о таком обобщении, положим $\lambda=\lambda_1+\lambda_2$ (нижние индексы соответствуют классу требований $m=1,2$), $\beta_m(s)$ будет обозначать ПЛС распределения длин требований $B_m(x)$ класса m, имеющего первый момент $\beta_{1m}<\infty$. Пусть также $\rho_m=\lambda_m\beta_{1m}$ и $\rho=\rho_1+\rho_2<1$. Тогда при $\beta(s)=\lambda^{-1}(\beta_1(s)+\beta_2(s))$ (в рассматриваемом случае это ПЛС ф.р. $B(x)=\lambda^{-1}(B_1(x)+B_2(x))$) теорема 2.5 дает ПЛС ф.р. стационарного времени пребывания требования длины u, принадлежащего классу m (необходимости в присутствии индекса m в левой части (2.21) не возникает в силу того, что сл.в. $V_m(u)$ одинаково распределена для каждого m, поскольку скорость обслуживания требований каждого класса определяется только общим числом $n=i+j$ присутствующих требований независимо от числа требований класса 1 (i) и класса 2 (j)). При таком обобщении формула (2.2) для ПЛС ф.р. безусловного времени пребывания в системе требования класса m, очевидно, принимает вид:

$$v_m(s)=\int_0^\infty v(s,u)\,dB_m(u).$$

Приведем также аналог следствия 2.1 для системы EPS с 2-мя классами требований:

$$\mathsf{P}_{ij}=\left(1-\sum_{m=1}^{2}\rho_m\right)\binom{i+j}{i}\rho_1^i\rho_2^j,\quad i,j=0,1,2,\dots,$$

где $\rho_1=\lambda_1\beta_{11}$, $\quad\rho_2=\lambda_2\beta_{12}$.

2.4. Некоторые следствия. Интересные свойства системы M/GI/1—EPS.

В качестве следствий теоремы 2.5 получены, в частности, математическое ожидание и дисперсия ф.р. сл. в. $V(u)$

Следствие 2.3.

$$\mathsf{E}[\mathsf{V}(\mathsf{u})]=\mathsf{u}/(1-\rho). \tag{2.22}$$

Следствие 2.4.

$$\mathsf{Var}[V(u)]=u^2(1-\rho)^{-2}-2(1-\rho)^{-1}\int_0^u\delta_1(y)\,dy, \tag{2.23}$$

где

$$\delta_1(y)=\mathsf{E}[\mathsf{D}(\mathsf{u})]=\mathsf{u}+\int_0^{\mathsf{u}}(\mathsf{u}-\mathsf{x})\sum_{\mathsf{n}=1}^{\infty}\rho^{\mathsf{n}}\mathsf{f}^{\mathsf{n}*}(\mathsf{x})\,\mathsf{dx}. \tag{2.24}$$

Здесь $f^{n*}(x)$ — n-кратная свертка плотности распределения остаточной длины требования $f(x)=(1-B(x))/\beta_1$, $f^{0*}(x)\equiv\delta(x)$ (δ-функция Дирака).

[15] Для доказательства этого утверждения удобнее установить сначала факт безграничной делимости ф.р. с ПЛС $\delta(s,u)$ (это достигается с помощью некоторых преобразований уравнения (2.29) из [164], но можно воспользоваться формулой (2.18) и теоремами из [48, гл.13, §7]), то же утверждение оказывается справедливым для ф.р. с ПЛС $\varphi(s,x,u)$; затем можно применить декомпозицию (2.12) с учетом факта безграничной делимости сложного геометрического распределения или использовать аналог теоремы де Финетти (de Finetti) из теории характеристических функций.

Формула (2.22) — хорошо известный (с 1969) результат, упоминавшийся выше [16], равенство (2.23) анонсировано в [78] и доказано в [79, 178] с помощью решения системы уравнений, которым удовлетворяют $\mathsf{E}[\Phi(x,u)^j]$ и $\mathsf{E}[D(u)^j]$, $j = 1, 2$ (эти уравнения являются более простым вариантом уравнений (2.13) и (2.14))[17]. Явные выражения для дисперсии при распределениях длин типа M, E_2 и H_2, полученные из (2.23), (2.24), приводятся в [79, 156, 178] и [164, стр. 81–82].

Приведем также другую полезную форму записи дисперсии, к которой можно прийти после простых преобразований (2.23) и (2.24) [156, 164]

$$\mathsf{Var}[V(u)] = 2(1-\rho)^{-2} \int_0^u (u-x)(1-W(x))\,dx, \tag{2.25}$$

где

$$W(x) = (1-\rho) \sum_{n=0}^{\infty} \rho^n F^{n*}(x), \tag{2.26}$$

$$F^{n*}(x) = \int_0^x F^{(n-1)*}(x-y)\,dF(y), \tag{2.27}$$

$$F(x) = F^{1*}(x) = \beta_1^{-1} \int_0^x (1-B(y))\,dy. \tag{2.28}$$

Здесь $F^{0*}(x) = \mathbf{1}(x)$ (функция Хевисайда).

В (2.26) $W(x) = \lim_{t\to\infty} \mathsf{P}\{U(t) \le x\}$ — предельное распределение процесса незаконченной работы, совпадающее со стационарным распределением времени ожидания в системе M/GI/1—FCFS[18].

Рекуррентные формулы для вычисления любого момента ф.р. сл.в. $V(u)$ получены в [173, 174] (1999), используя прием, предложенный О.Хевисайдом (Heaviside) [60] еще в конце XIX века. Доказательство заняло всего несколько строк (см. следствие 3.9), в отличие от более громоздкого вывода в [185](2000).

Замечание 2.6. Тем не менее, предыдущий способ вычисления моментов [164, §2.7] до сих пор представляет самостоятельный интерес, поскольку при этом проливается дополнительный свет на структуру полученных формул. В частности (см. [164, формула (2.66)]),

$$\mathsf{E}[V(u)] = u/(1-\rho) = \delta_1(u) + \lambda(1-\rho)^{-1} \int_0^\infty \varphi_1(x,u)(1-B(x))\,dx,$$

[16] Известно не менее десятка различных способов вывода (2.22).

[17] Подчеркнем, что ввиду большой сложности ПЛС ф.р. сл.в. $V(u)$ (формула (2.21)) стандартный способ вычисления моментов этой ф.р. с помощью (2.3) срабатывает фактически только при нахождении математического ожидания (формула (2.22)). При получении $\mathsf{Var}[V(u)]$ необходимо дважды продифференцировать $v(s,u)$ по s и перейти далее к пределу при $s \to 0$. Эти вычисления настолько громоздки, что их успешное завершение нереально. Именно по этой причине формула (2.23) была выведена с помощью решения системы уравнений, которым удовлетворяют $\mathsf{E}[\Phi(x,u)^j]$ и $\mathsf{E}[D(u)^j]$, $j = 1, 2$ [79, 164, 178]. В частности, (2.24) — результат обращения преобразования Лапласа $s^{-2}[1-\lambda(1-\beta(s))/s]^{-1}$. (Подробности см. в [164, с.78-81].) Аналогичным способом удалось вывести и третий момент ф.р. сл.в.$V(u)$. Несколько позже (см. далее следствие 3.9) в [173, 174, 185] были найдены рекуррентные соотношения для моментов любого порядка.

[18] Побочным результатом нашего вычисления $\mathsf{Var}[V(u)]$ в системе M/GI/1—EPS является знаменитая формула Поллачека–Хинчина для ф.р. стационарного времени ожидания в системе M/GI/1-FCFS. Известно не менее нескольких десятков различных способов ее получения. Проведенный анализ дисциплины EPS дает, помимо искомых результатов, и новый вывод формулы Поллачека–Хинчина (2.26) (это же равенство в терминах ПЛС приведено в (3.35)). Подчеркнем, что (2.26) (или (3.35)) является краеугольным камнем многих исследований в теории очередей. Например, такое направление как теория приоритетных очередей возникло из обобщения системы M/GI/1—FCFS на случай двух классов требований с различными фиксированными приоритетами.

где $\delta_1(u)$ дается (2.24), а $\varphi_1(x,u) \doteq \mathsf{E}[\Phi(\mathsf{x},\mathsf{u})] = \delta_1(\mathsf{u}) - \delta_1(\mathsf{u}-\mathsf{x})$ при $x < u$ и $\varphi_1(x,u) = \delta_1(u)$ при $x \geq u$. Это явилось одним из факторов, позволивших доказать нетривиальную предельную теорему 2.7, а также выяснить причины возникновения так называемого “парадокса”, связанного с крайне простым видом формулы (2.22). В сущности, приведенное выше разложение (частный случай декомпозиции (2.12) после снятия условий) заранее дает ответ на поставленный в статье [50, p. 674] (1989) вопрос о “парадоксальности” вида формулы (2.22). В [50] дано объяснение этому “парадоксу”, которое можно рассматривать как дополнительное к нашему более раннему объяснению.

Замечание 2.7. Можно снять в (2.20) условие только по $(x_1,\ldots,x_n)$, получив в итоге некоторый аналог теоремы из [155], т.е. вычислив $\mathsf{E}[\mathrm{e}^{-\mathsf{sV(u)}}|\mathsf{n}]$ для системы M/GI/1—EPS. Ограничимся здесь лишь формулой для среднего (условного) времени пребывания требования в системе

$$(a) \qquad \mathsf{E}[\mathsf{V}|(\mathsf{u},\mathsf{n})] \equiv \mathsf{E}[\mathsf{V}(\mathsf{u})|\mathsf{n}] = \delta_1(\mathsf{u}) + \mathsf{n}\beta_1^{-1}\int_0^\infty \varphi_1(\mathsf{x},\mathsf{u})(1-\mathsf{B}(\mathsf{x}))\,\mathsf{dx}.$$

Из (2.24) и (2.26) вытекает, что

$$(b) \qquad \delta_1(u) = (1-\rho)^{-1}\int_0^u W(x)\,dx.$$

Тогда с учетом формул из замечания 2.6, равенство (a) сводится к

$$(c) \qquad \mathsf{E}[\mathsf{V}(\mathsf{u})|\mathsf{n}] = \mathsf{u}/(1-\rho) + (\mathsf{n}-\rho/(1-\rho))\rho^{-1}\int_0^{\mathsf{u}}(1-\mathsf{W}(\mathsf{x}))\,\mathsf{dx}.$$

Приведем далее эквивалентную форма равенства (2.26), которая может оказаться полезной для более глубокого изучения вероятностных законов, управляющих случайными флюктуациями времени ожидания в системе M/GI/1—FCFS при некоторых специальных классах распределений $B(x)$ с тяжелым хвостом.

Замечание 2.8. В эквивалентной форме равенство (2.26) можно представить как [19]

$$W(x) = \mathrm{e}^{-\lambda}\sum_{n=0}^\infty \lambda^n H^{n*}(x)/n!, \quad \text{где } H(x) = H^{1*}(x) = -(\ln(1-\rho))^{-1}\sum_{n=1}^\infty \rho^n F^{n*}(x)/n\,.$$

Отметим также взаимосвязь констант перед знаком суммы в этой формуле для $H(x)$, $\mathsf{P}_0^\star$ в следствии 2.2 и константы в правой части равенства, приведенного в [175, p. 183] между формулой (3.6) и следствием 3.4.

Выражение (2.25) позволило установить ряд других новых свойств системы M/GI/1—EPS [156,158,164]. Здесь отметим только два из них. Формула (2.26) совпадает с обратным преобразованием ПЛС ф.р. стационарного времени ожидания в системе M/GI/1—FCFS, выполненным впервые Бенешем (Beneš) [18]. Важно подчеркнуть, что геометрические весовые коэффициенты в (2.26) равны стационарным вероятностям системы M/M/1—FCFS, которые не совпадают с вероятностями состояний системы M/GI/1—FCFS (но совпадают с вероятностями состояний системы M/GI/1—EPS в силу следствия 2.1). Попытки дать объяснение этого простого, но интригующего факта, предпринятые Клейнроком [82, стр.221] и Купером [37, p.219], оказались безуспешными. Однако это сделал Прабху [112], который интерпретировал формулу

[19] Мы не утверждаем, что такое представление в виде сложного распределения Пуассона хорошо известной ф.р. $W(x)$ является новым, но, по крайней мере, оно нетипично в теории очередей.

(2.26) в терминах лестничных процессов, а также Купер и Чен [39], объяснившие эту формулу с помощью результатов анализа системы M/GI/1—LCFS–P. Другой способ аналогичной интерпретации формулы (2.26) с помощью сравнения систем FCFS и EPS обсуждается в [164, стр. 86] (идея состоит в совпадении распределений процессов незаконченной работы в системе M/GI/1 с дисциплинами EPS и FCFS). Формула (2.23) позволила установить еще одно новое свойство системы M/GI/1—EPS, которое неоднократно уточнялось в течение четверти века (см. [156, 158], [164], [170, предложение 3.1], [180], [34]). После недавнего уточнения оно формулируется так:

Следствие 2.5. *Если у ф.р. $B(x)$ существует производящая функция моментов $\beta(-s) < \infty$, $s > 0$, то распределение случайной величины $V(u)$ принадлежит $\mathcal{L}$–классу стареющих распределений*[20].

Иными словами, распределение сл.в.$V(u)$ имеет распределение $\mathsf{P(V(u) \le x|B = u)}$, для которого $\int_0^\infty e^{-rx}(1 - \mathsf{P(V(u) \le x|B = u))dx} \ge \mathsf{E[V(u)]}/(1 + \mathsf{rE[V(u)]})$ (в сущности, этим и дается определение класса $\mathcal{L}$). Сказанное позволяет найти точные границы хвоста распределения сл.в. $V(u)$, различные оценки моментов и т.п.

Принципиально новое доказательство теоремы 2.5 дано Шассбергером (Schassberger) [122] (1984). Его результаты (вывод формулы (2.21) в другом виде) получены при анализе некоторой модификации дисциплины RR. Модификация состоит в том, что вновь поступающее требование немедленно получает свой квант обслуживания (это возможно, поскольку соответствующий вариант дисциплины RR рассматривается в дискретном, а не в непрерывном времени) и только после этого становится в конец очереди, если не закончит обслуживание за отведенный квант. Тем самым удается обойти основную трудность, связанную с тем, что порядок обслуживания требований в последовательных раундах при стандартной дисциплине RR не сохраняется — именно это обстоятельство не позволяет получить точное решение для $v(s, u)$ в системе M/GI/1 со стандартной дисциплиной RR (см. окончание раздела 2.4 в [170] или [164, §1.5]). В качестве дискретной единицы времени используется постоянный размер кванта, пуассоновский входящий поток аппроксимируется потоком Бернулли, а $B(x)$ — произвольным арифметическим распределением с тем же временны́м дискретом. Одним из основных результатов [122] является производящая функция для дискретного распределения условного времени пребывания требования в такой системе. Далее доказывается, что последовательность стационарных длительностей пребывания в модифицированной системе RR при стремящемся к нулю размере кванта сходится по распределению к сл.в. $V(u)$ в системе M/GI/1—EPS, и вид соответствующего предела (в терминах ПЛС) определяется формулой, которую можно редуцировать к (2.21).

Статьи Отта [109] (1984) и Гришечкина [54] (1991) содержат обобщения теоремы 2.5 непринципиального характера (получение $\mathsf{E}[e^{-\mathsf{sV(u)}}\mathsf{z^L}]$ и $\mathsf{E}[e^{-\mathsf{sV(u)}}]$ для той же самой системы EPS и системы EPS с групповыми поступлениями, соответственно); доказательства, в сущности, повторяют аргументацию [78, 156, 158] в переформулированном виде (например, в терминах теории ветвящихся процессов плюс балласт непривычной для теории очередей терминологии в [54][21]). Впрочем, [54] содержит ряд новых интересных предельных теорем. Это резко контрастирует с [109], являющейся, в сущности, производной от статей [78, 156, 158], а также от неопубликованных результатов Яшкова (см. 2–ую ссылку в литературе [109]).

Отметим другую статью Риджа и Сенгупты [115] (1994), в которой раскрыты дополнительные возможности изложенного выше метода, примененного к системе с дисциплиной дискриминаторного разделения процессора.

[20] О классах ф.р. в теории надежности см., например, [134]. Относительно $\mathcal{L}$– класса распределений см., в частности, [80], а также цитированные в этой статье работы. К примеру, $\mathcal{L}$–класс ф.р. шире класса HNBUE.

[21] Более простое решение той же задачи содержит статья Риджа и Сенгупты [114], опирающаяся на метод [158].

Близко к теореме 2.5 могли подойти также Эйсер (Asare) и Фостер (Foster) [12] (1983). Они использовали подобные нашим функционалы от ветвящихся процессов, но смогли вычислить для них только математические ожидания. Кроме того, разложение (2.12) не вводилось. В итоге главным результатом [12] оказалось лишь равенство (c) из замечания 2.7. Тем не менее, Фостер — один из первых, кто применил понятия ветвящихся процессов (на поверхностном уровне) для анализа системы M/M/1—EPS в статье [51] (1973). Выводу (c) из замечания 2.7 в частном случае $B(x)$ типа M посвящен §6 этой статьи, где указана и "an important unsolved problem" — обобщение на случай $B(x)$ типа GI, сделанное позднее в [12]. Напомним, что (c) является элементарным следствием теоремы 2.4.

Ван ден Берг [19] (1990) дал еще одно доказательство теоремы 2.5 с помощью рассмотрения некоторого обобщения системы M/M/1 с одним из вариантов бернуллиевской обратной связи. Некоторые комментарии к [19] (и [116], дополняющей ее) содержатся в [167, стр.27–28]. Использованный метод анализа восходит к методу Шассбергера [122], искусственно модифицированному с помощью некоторых довольно громоздких конструкций из теории очередей с обратной связью (об очередях с обратной связью см., например, книгу [42]).

Описанный выше аналитический метод декомпозиции на элементы задержки был распространен на систему M/M/1—EPS с ненадежным прибором, а также на дисциплину EPS в случайной среде [105]. Дальнейшее развитие разработанного метода позволило впервые исследовать нестационарные процессы в системе M/GI/1—EPS и найти их распределения [164,167,172], [182,175] в терминах многомерных преобразований. Эти результаты до сих пор остаются, по–видимому, недостижимыми с помощью других упомянутых нами методов.

2.5. Некоторые частные случаи.

Рассмотрим случай $B(x) = 1 - \mathrm{e}^{-\mu x}$, тогда $\beta(s) = \mu/(s+\mu)$ и $\beta_1 = 1/\mu$. Прямым следствием теоремы 2.5 является

Следствие 2.6. *В системе M/M/1—EPS при* $\rho < 1$

$$v(s,u) = \frac{(1-\rho)(1-\rho r^2)\exp(-u(s+\lambda-\lambda r))}{(1-\rho r)^2 - \rho(1-r)^2\exp(-\mu u(1-\rho r^2)/r)}, \tag{2.29}$$

где $r = \pi(s)$ *— решение уравнения (2.4) при* $\beta(s) = \mu/(s+\mu)$.

Доказательство. См. [164, с.73-75], [179]. Основные этапы: уравнение (2.4) сводится к квадратному относительно r; уравнение (2.14) упрощается, позволяя в итоге преобразовать равенство (2.21) к виду

$$v(s,u) = (1-\rho)\,\mathrm{e}^{-u(s+\lambda)}\left[\psi(s,u) + \mu^{-1}\partial\psi(s,u)/\partial u\right]^{-1};$$

формула (2.19) сводится к $\tilde{\psi}(s,q) = (q+s+\mu)/[(q-q_1)(q-q_2)]$, где $q_1 = -\lambda r$, $q_2 = -\mu/r$ — два простых полюса, и находится ее обратное преобразование (по аргументу q) с помощью классической теоремы Коши о вычетах. □

Формула (2.29) впервые получена в [35] (1970). Метод доказательства в [35] основан на использовании процессов рождения и гибели для описания процесса обслуживания; технически он довольно громоздок (приходится, например, решать дифференциальные уравнения с частными производными первого порядка) и непригоден для анализа более общих случаев. Отметим, что (2.29) можно также вывести из теоремы статьи [155] после выполнения предельного перехода, отражающего устремление к нулю размера кванта. При этом приходится решать только два обыкновенных дифференциальных уравнения.

Рассмотрим теперь случай детерминированного (вырожденного) распределения длин требований с ПЛС $\beta(s) = \mathrm{e}^{-su}$; при этом $\beta_i = u^i$, $i = 1, 2, \ldots$; $v(s,u) = v(s)$. Справедливо

Следствие 2.7. *В системе M/D/1—EPS при* $\rho = \lambda u < 1$

$$v(s) = \frac{(1-\rho)(s+\lambda)^2 \, \mathrm{e}^{-u(s+\lambda)}}{s^2 + \lambda[s + (s+\lambda)(1-\rho)] \, \mathrm{e}^{-u(s+\lambda)}}, \quad Re\ s > 0. \tag{2.30}$$

Доказательство. См. [164, с.73], [179]. Стартовая точка: формула перед (2.21). После преобразований, дело сводится к решению уравнения Бернулли

$$\partial\delta(s,t)/\partial u + (s+\lambda)\delta(s,t) - \lambda\delta^2(s,t) = 0$$

с начальным условием $\delta(s,0) = 1$. Решение имеет вид

$\delta(s,t) = (s+\lambda)/\left(\lambda + s\,\mathrm{e}^{t(s+\lambda)}\right), \quad t \le u$. В итоге приходим к (2.30). □

Решение для $v(s)$, представленное равенством (2.30), получено также в [19] и [109]. Еще одним из следствий является второй момент сл.в. V в системе M/D/1—EPS:

Следствие 2.8.

$$\mathsf{Var}[V] = u^2(1-\rho)^{-2} - 2u^2\rho^{-2}(1-\rho)^{-1}(\mathrm{e}^{\rho} - 1 - \rho).$$

2.6. Свойства монотонности.

Для $\mathsf{Var}[V(u)]$ доказано следующее свойство монотонности [79, 156, 164]. Пусть индексы M, E_m, H_m обозначают тип ф.р. $B(x)$ с одинаковым первым моментом.

Следствие 2.9. *В системе M/GI/1—EPS при* $\rho < 1$

$$\mathsf{Var}_{E_m}[V(u)] \le \mathsf{Var}_M[V(u)] \le \mathsf{Var}_{H_m}[V(u)]. \tag{2.31}$$

Идея доказательства. Вывод неравенства (2.31) основан на установлении связи между распределениями случайных величин $V(u)$ в системах EPS (с распределениями длин из классов E_m, M, H_m) в терминах отношений стохастического порядка типа $\overset{(2)}{\le}$ (см. сноску 12).

При этом основную роль играет формула (2.25). Подробности см. в [156, 164]. □

Приведем еще одно довольно простое свойство монотонности ф.р. сл.в. $V(u)$ в системе EPS.

Следствие 2.10. *Вероятность* $\mathsf{P}(V(u) > t)$ *не убывает по* $u \ge 0$. *Отсюда вытекает, что каждый момент* $v_j(u) = \mathsf{E}[\mathsf{V(u)^j}]$, $j \in \mathsf{N}$ *не убывает по* u.

Доказательство. На всех реализациях управляющих последовательностей в системе M/G/1—EPS точка выхода требования длины $u_2 > u_1$ находится правее точки выхода требования длины $u_1 \quad (u_1, u_2 \in \mathbf{R}_+)$. □

Интересный пример свойства монотонности, обнаруженного с помощью формулы (2.25), содержит [150].

2.7. Асимптотика

Асимптотические оценки дисперсии. Рассмотрим способ упрощения вычисления $\mathsf{Var}[V(u)]$ в системе M/GI/1—EPS, состоящий в получении асимптотических оценок дисперсии при малой и большой длине требования. Используя разложение правой части (2.25) в ряд Тейлора в точке $u = 0$ и ограничиваясь первыми четырьмя членами разложения, приходим к утверждению

Следствие 2.11. *В системе $M/GI/1$—EPS при выполнении условия (2.1)*

$$\mathsf{Var}[V(u)] \sim \rho(1-\rho)^{-2}u^2 - \lambda(3(1-\rho))^{-1}u^3, \quad u \to 0. \tag{2.32}$$

Предположим, что:
(a) существуют первые три момента распределения $B(x)$ (β_i, $i = 1,2,3$). Интерпретируем $W(x)$ в (2.26) как распределение времени жизни обрывающегося процесса восстановления с распределением (несобственным) промежутков времени между последовательными моментами восстановления $G(x) = \rho F(x)$ ($G(0) = 0$, $G(+\infty) = \rho < 1$) и вероятностью обрыва $(1-\rho)$. Здесь $F(x)$ приведено в (2.28). В дополнение к предположению (a): $\beta_i < \infty$, $i = 1,2,3$, будем считать, что:
(b) распределение $G(x)$ обладает тем свойством, что для некоторой константы $\Theta > 0$

$$\int_0^\infty \mathrm{e}^{\Theta x}\, dG(x) = \lambda \int_0^\infty \mathrm{e}^{\Theta x}(1-B(x))\, dx = 1 \text{ и } C = \lambda \int_0^\infty \mathrm{e}^{\Theta x}\, x(1-B(x))\, dx < \infty.$$

Введем обозначения $r_1 = \int_0^\infty (1-W(y))\, dy = \rho(1-\rho)^{-1} f_1$, $f_j = \beta_{j+1}((j+1)\beta_1)^{-1}$, $j = 1,2,\ldots$,

$$r_2 = \int_0^\infty y(1-W(y))\, dy = 2(1-\rho)^{-1}\left(f_2 + 2\rho(1-\rho)^{-1} f_1^2\right).$$

Следствие 2.12. *Для системы $M/G/1$—EPS при $\rho < 1$ и предположениях (a), (b), сделанных выше,*

$$\mathsf{Var}[V(u)] \sim 2(1-\rho)^{-2}\left(r_1 u - r_2 + (1-\rho)\,\mathrm{e}^{-\Theta u}/[C\Theta^3]\right), \quad u \to \infty. \tag{2.33}$$

Доказательство. (Впервые — в [156], уточнения — в [161], см. также [164, с.82-85].) В предположении (a) перепишем (2.25) в терминах r_1 и r_2

$$\mathsf{Var}[V(u)] = 2(1-\rho)^{-2}\left[r_1 u - r_2 + \int_u^\infty (x-u)(1-W(x))\, dx\right], \tag{2.34}$$

где r_1 и r_2 приведены выше. (Равенство (2.34) — третья эквивалентная форма записи $\mathsf{Var}[V(u)]$; первые две даются равенствами (2.23) и (2.25).) Дополнительно к (a) введем теперь предположение (b). В силу следствия из теоремы 2 [48, гл.11, §6, с.443] для $1-W(x)$ при $x \to \infty$ справедлива оценка $1-W(x) \sim (1-\rho)\,\mathrm{e}^{-\Theta x}/[C\Theta]$, которая совпадает с известной асимптотической оценкой Крамéра для риска разорения в обобщенном пуассоновском процессе. Тогда последний член в правой части равенства (2.34) при $u \to \infty$ асимптотически равен $(1-\rho)\,\mathrm{e}^{-\Theta u}/[C\Theta^3]$, что приводит к (2.33). □

Отметим, что при $B(x) = 1 - \mathrm{e}^{-\mu x}$ формула (2.33) становится точной для всех $u \geq 0$ и может быть преобразована к виду

$$\mathsf{Var}[V(u)] = 2\rho u/[\mu(1-\rho)^3] - 2\rho[1-\mathrm{e}^{-\mu(1-\rho)u}]/[\mu^2(1-\rho)^4].$$

Более грубая по сравнению с (2.33) асимптотическая оценка $\mathsf{Var}[V(u)] \sim 2(r_1 u - r_2)/(1-\rho)^2$ при $u \to \infty$ получена ранее в [79, 157]. Десять лет спустя такой же результат получил ван ден Берг [19]. Отметим, что точность этой формулы, а также асимптотики (2.32) достаточно высока (см. [156, 164] и [19] для численных примеров).

Замечание 2.9. Во многих приложениях вполне можно пользоваться даже еще более грубой по сравнению с приведенными выше асимптотической оценкой дисперсии: $\mathsf{Var}[V(u)] \sim 2r_1u/(1-\rho)^2$ при $u \to \infty$ в предположении, что $\beta_2 < \infty$. Эта формула эквивалентна следующему асимптотическому разложению второго момента ф.р. сл.в. $V(u)$ в системе M/GI/1—EPS

$$v_2(u) = v_1^2(u) + \lambda\beta_2(1-\rho)^{-3}u + o(u), \quad u \to \infty.$$

Главный член асимптотики совпадает с точными формулами при всех $u \geq 0$ для дисперсии сл.в. $\Pi(u)$ длительности стандартного периода занятости, открываемого требованием длины u в системе M/G/1 (ПЛС ф.р. сл.в. $\Pi(u)$ приведено в (2.5)), и для $\mathsf{Var}[V(u)]$ в системе M/GI/1—LCFS–P (ср. с замечанием 2.1).

Анализ остаточного члена асимптотического разложения в замечании 2.9 проведен в следствии 2.12, которое дает, в частности, и скорость сходимости $\int_u^\infty (x-u)(1-W(x))\,dx$ в (2.34) к нулю при $u \to \infty$.

2.8. Связь с теорией страхового риска

Условия (a), (b), введенные перед следствием 2.12, известны как условия Крамера–Лундберга в теории страхового риска. В более общем случае условие (a) задается как

$$f(-\Theta) = \int_0^\infty \mathrm{e}^{\Theta x}\, dF(x) = \rho^{-1} < \infty, \quad \Theta > 0. \tag{2.35}$$

Это эквивалентно предположению о том, что существует производящая функция моментов (экспоненциальный момент) распределения $F(x)$. Другими словами, ПЛС $f(s) = (1-\beta(s))/(s\beta_1)$ существует в непустой окрестности нуля (у $f(s)$ абсцисса абсолютной сходимости $s_b < 0$), откуда вытекает, что хвост $1-F(x)$ (а следовательно, и $1-B(x)$) экспоненциально ограничен: существуют $K < \infty$, $\epsilon > 0$ и $x_0 \geq 0$, такие, что

$$1 - B(x) \leq K\,\mathrm{e}^{-\epsilon x}, \quad x \geq x_0. \tag{2.36}$$

Таким образом, условие (2.35) выделяет во множестве ф.р. $B(x)$ класс распределений с так называемым легким хвостом (примеры: гамма-распределения, ф.р. Вейбулла $B(x) = 1 - \exp(-cx^b)$ при $c > 0$ и $b \geq 1$ и любая ф.р. с ограниченным носителем). Условие (2.35) не выполняется, когда $f(s)$ имеет существенную особенность в нуле ($f(-\epsilon) = \infty$ при $\forall \epsilon > 0$), и это выделяет класс $\mathcal{K}$ распределений на $(0,\infty)$ с так называемым тяжелым хвостом (heavy-tailed) (примеры: ф.р. Вейбулла при $c > 0$ и $b \in (0,1)$, логнормальное распределение, семейство ф.р. Парето и т.д.). Распределения из класса $\mathcal{K}$ имеют ряд патологических свойств, основные из которых: отсутствие некоторых (или всех) старших моментов, отсутствие характеризации в терминах ПЛС, более медленное, чем экспоненциальное, убывание хвоста. Известны различные подклассы таких распределений и о некоторых из них мы напомним. Самым узким является класс ф.р. $B(x)$ ($B(x) < 1$ при $\forall x \geq 0$) с правильно меняющимся (regularly varying) хвостом, обозначаемый далее через $RV(-a)$:

$$B(x) \in RV(-a),\ a \geq 0,\ \text{если}\ \lim_{x\to\infty} \frac{1-B(tx)}{1-B(x)} = t^{-a},\ \forall t > 0. \tag{2.37}$$

Из (2.37) вытекает, что $1 - B(x) \sim x^{-a}\ell(x)$, где $\ell(x)$ — медленно меняющаяся (на бесконечности) функция, т.е. положительная измеримая функция, для которой $\ell(tx)/\ell(x) \to 1$ при $x \to \infty$ и $\forall t > 0$ (примеры $\ell(x)$: функции, сходящиеся к положительной константе, логарифмы, $\ln(\ln(\mathrm{e}+x))$ и т.д.). Отметим, что для $\forall \epsilon > 0\ \ \exists\, x_0$, такое, что

$$x^{-\epsilon} \leq \ell(x) \leq x^{\epsilon}, \quad x > x_0.$$

Другой подкласс $\mathcal{K}$ — класс $\mathcal{S}$ субэкспоненциальных ф.р. ($1 - B(x) > 0$ при $x \geq 0$), введен В.П.Чистяковым (1964) в контексте ветвящихся процессов, рассмотрен в [13] (1972) и изучен в книге А.А. Боровкова (1972) в контексте процессов в теории очередей. Класс $\mathcal{S}$ определяется так:

$$B(x) \in \mathcal{S}, \text{ если } \lim_{x\to\infty} \frac{1 - B^{n*}(x)}{1 - B(x)} = n, \quad n \geq 2. \tag{2.38}$$

Если условие в (2.38) выполняется для некоторого $n \geq 2$, то оно справедливо для всех $n \geq 2$.

В эквивалентной форме (2.38) можно переписать как

$$\mathsf{P}(B_1 + \ldots + B_n > x) \sim \mathsf{P}(\max\{B_1, \ldots, B_n\} > x) \sim n(1 - B(x)),$$

где $B_1, \ldots, B_n$ — независимые копии сл.в. B с ф.р. $B(x)$.

Если $B(x) \in \mathcal{S}$, то $\lim_{x\to\infty} \mathrm{e}^{\epsilon x}(1 - B(x)) = \infty$, $\epsilon > 0$, т.е. (2.36) не выполняется. Полезным свойством $\mathcal{S}$ является: для $\forall\epsilon > 0$ существует положительная константа $K(\epsilon) < \infty$, такая, что при $n \geq 1$ и $x \geq 0$ имеет место

$$\frac{1 - B^{n*}(x)}{1 - B(x)} \leq K(\epsilon)(1 + \epsilon)^n.$$

Некоторые другие свойства $\mathcal{S}$: если $B(x) \in \mathcal{S}$ и $1 - B(x) \sim 1 - G(x)$, то $G(x) \in \mathcal{S}$;
если $B(x) \in \mathcal{S}$ и $1 - G(x) = o(1 - B(x))$ при $x \to \infty$, то $B(x) * G(x) \in \mathcal{S}$, причем $1 - B(x) * G(x) \sim 1 - B(x)$;
если $B(x) \in \mathcal{S}$ и $1 - G(x) = c(1 - B(x))$ при $x \to \infty$ и $c > 0$, то $G(x) \in \mathcal{S}$, $B(x) * G(x) \in \mathcal{S}$, причем $1 - B(x) * G(x) \sim (1 + c)(1 - B(x))$.
Отметим еще, что класс $\mathcal{S}$ не замкнут относительно операции свертки.

Обозначим через $\mathcal{L}_\circ$ класс ф.р. $B(x)$ с длинным хвостом (long-tailed), который определяется асимптотическим соотношением

$$1 - B(x - y) \sim 1 - B(x) \text{ при } x \to \infty \text{ и всех } y > 0.$$

Из (2.37) следует, что $B(x) \in \mathcal{L}_\circ$ тогда и только тогда, когда $1 - B(\log x)) \in RV(0)$, где $RV(0)$ — класс медленно меняющихся функций. Основное свойство $\mathcal{L}_\circ$: если сл.в. $B \in \mathcal{L}_\circ$ и Y — любая неотрицательная сл.в., не зависящая от B, то $\mathsf{P}(B - Y > x) \sim \mathsf{P}(B > x)$.

Между введенными классами вероятностных распределений выполняются следующие отношения включения $RV(-a) \subset \mathcal{S} \subset \mathcal{L}_\circ \subset \mathcal{K}$. Более подробные сведения по классам ф.р. с тяжелым хвостом и их свойствам можно найти в [22], [46] и частично в [48]. Важным является то обстоятельство, что, в частности, неизвестно, как распознать ф.р. $B(x) \in \mathcal{S}$ по виду ее ПЛС $\beta(s)$ (не существует характеризации класса $\mathcal{S}$ в терминах ПЛС, класс $\mathcal{S}$ характеризуется в других терминах, см., например, [22, p.430][22]). Однако исходные данные и результаты решения задач теории очередей, как правило, формулируются в терминах ПЛС. Поэтому при изучении очередей с ф.р. из классов $RV(-a)$ [23] или $\mathcal{S}$ необходимо дополнительно применять нетрадиционные инструменты исследований: различные варианты теоремы Карамата

[22] В частности, асимптотическая эквивалентность $1 - W(x)$ и $1 - F(x)$ в (2.40) — одна из характеризаций класса $\mathcal{S}$: $W(x) \in \mathcal{S} \Leftrightarrow F(x) \in \mathcal{S}$.

[23] Для класса $RV(-a)$ известна, в частности, характеризация асимптотического поведения $1 - B(x)$ при $x \to \infty$ в терминах остаточного члена разложения ПЛС $\beta(s)$ в ряд Тейлора при $s \downarrow 0$ (см. теорему 8.1.6 в [22, pp.333-334], называемую иногда тауберовой теоремой Бингхема и Дони (Bingham and Doney) (1974)). В сущности, упомянутая теорема является одним из вариантов формального обоснования для класса $RV(-a)$ первого принципа операционного исчисления, установленного Хевисайдом [60] в более общем случае; см. по этому поводу [5].

О накладываемых на $B(x)$ условиях, при которых достигается обоснование принципов Хевисайда, см. также [43, p.254].

при $B(x) \in RV(-a)$, тесно связанной с тауберовой теоремой Харди–Литтлвуда, другие тауберовы и абелевы теоремы, нетривиальные асимптотические разложения, разработанные в асимптотической теории экстремальных порядковых статистик, теорию больших уклонений, а также ряд новых результатов из продвинутой в последние десятилетия теории правильно меняющихся функций, называемой теорией де Хаана [24], и т.д.

Это — область очень активных исследований. Причина состоит в том, что в теории очередей привыкли иметь дело с вероятностными распределениями, имеющими легкий хвост. Но пространственно-временна́я динамика потоков трафика в современных компьютерных сетях описывается распределениями со степенными хвостами и, как правило, бесконечной дисперсией (так называемый *the Noah effect* в информатике), кроме того, убывание автокорреляционных функций процессов трафика имеет гиперболический, а не экспоненциальный характер (феномен "самоподобности" (*self-similar*) или "фрактальности" (*fractal*), другой термин — *the Joseph effect*). Такие неожиданные эффекты впервые обнаружены только в 1993 г. в сети ETHERNET.

Подобные случаи изучались в теории страхового риска [46, 139], однако переформулировка результатов этой теории на системы обслуживания часто оказывается нетривиальной. Относительно просто получаются результаты, когда $B(x) \in RV(-a)$ и $a \notin \mathbf{N}$. (Случаи, когда $a = 1, 2, \ldots$, сложнее и изучаются с помощью теории де Хаана.) Часто (но не всегда) те же результаты оказываются справедливыми и в более общих ситуациях, когда, например, $B(x) \in \mathcal{S}$. Отметим, что практическая важность ф.р. с тяжелым хвостом в теории страхового риска была обнаружена на несколько десятилетий раньше, чем в информатике и теории очередей — именно при таком классе ф.р. возникает иной механизм разорения, и поэтому необходимы более точные оценки вероятности разорения по сравнению с возможностями классических схем. Отметим, что в контексте теории страхового риска вероятность разорения за бесконечный промежуток времени в случае положительных страховых сумм и премий есть $1 - W(x)$, где ф.р. $W(x)$ дается (2.26) [25]. Даже простая формула (3.35) до сих пор позволяет находить некоторые новые асимптотические свойства системы FCFS для ф.р. $B(x)$ из класса $\mathcal{S}$. Аналогом формулы для вероятности разорения $1-W(x,t)$ на конечном промежутке времени (в терминах двойных преобразований) служит решение уравнения Такача в теории очередей (см. равенство (3.30)).

2.9. Другие асимптотические формулы. Предельные теоремы.

Приведем сначала результат (анонсированный в [184, 185] и доказанный в [170]), позволяющий изучать случай, когда условие Крамера–Лундберга не выполняется, в частности, не справедливо (2.35) и $\beta_2 = \infty$.

Следствие 2.13. *Пусть $B(x) \in RV(-a)$ (см. определение (2.37)) при $a \in (1,2)$. Тогда в системе $M/GI/1{-}EPS$ при $\rho < 1$*

$$\mathsf{Var}[V(u)] \sim \frac{2\Gamma(2-a)}{\Gamma(4-a)} \frac{\lambda}{(1-\rho)^3(a-1)} u^{3-a}\ell(u), \quad u \to \infty, \tag{2.39}$$

где $\ell(u)$ — медленно меняющаяся (на бесконечности) функция и гамма-функция $\Gamma:\ (0,\infty) \to \mathbf{R}_+$ определяется как $\Gamma(t) = \int_0^\infty x^{t-1}\,\mathrm{e}^{-x}\,dx, \quad t > 0$.

---

[24] В теории правильно меняющихся функций (теории Карамата (J.Karamata), который дал характеризацию таких функций) изучаются асимптотические соотношения вида $h(tx)/h(x) \to g(t) = O(t^{-a}), \quad x \to \infty$ при $\forall t > 0$. В теории де Хаана (L.de Haan) изучаются более общие соотношения вида $(h(tx)-h(x))/h(x) \to k(t) \in \mathbf{R}, \quad x \to \infty$ при $\forall t > 0$.

[25] Х.Крамер в 1930 г. нашел преобразование Фурье ф.р. $W(x)$, а ПЛС этой ф.р. в контексте теории очередей (формула (3.35) нашли Поллачек (F.Pollaczek) (1930) и А.Я.Хинчин (1932). Еще одна ее интерпретация — распределение максимума процесса случайного блуждания с отрицательным сносом (см. также замечание 2.8).

Доказательство. При $1 - B(x) \sim x^{-a}\ell(x),\ a \in (1,2)$ из утверждения (d) предложения A3.8 в [46] вытекает, что $\beta_2 = \infty$, следовательно, теорема 2.12 и замечание 2.9 не применимы в этом случае. В силу леммы 1 из [48, с.190] (или леммы 8.1.5 в [22]) существует момент $\beta_{1+\varepsilon} < \infty$ порядка $1+\varepsilon,\ \varepsilon \in (0,1)$, который позволяет получить искомое асимптотическое приближение из формулы (2.25) с помощью двух теорем. По теореме Коэна–Пэйкса [22, Th.8.10.3], эквивалентной варианту II теоремы Крамéра–Лундберга в теории страхового риска [46, Th.1.3.8]

$$1 - W(x) \sim \rho(1-\rho)^{-1}(1 - F(x)), \quad x \to \infty, \tag{2.40}$$

причем ф.р. $W(x)$ и $F(x)$ принадлежат классу $\mathcal{S}$. [26]. Так как $RV(-a) \subset \mathcal{S}$, то после использования в (2.40) формулы (2.28) и применения одного из вариантов теоремы Карамата ([22, Prop.1.5.8], [48, c.341])[27], правая часть асимптотического равенства (2.40) редуцируется к виду $\lambda(1-\rho)^{-1}(a-1)^{-1}x^{1-a}\ell(x)$. Теперь предел при $u \to \infty$ интеграла в (2.25) после подстановки $x = ut$ ($\ell(ut) \sim \ell(t)$ при $u \to \infty$ и $\forall t$) сводится к бета-интегралу (интегралу Эйлера первого рода) $B(p,q) = \int_0^1 (1-t)^{p-1}t^{q-1}\,dt$ при $p = 2$ и $q = 2 - a$. Множитель перед бета-интегралом стал асимптотически равен $2\lambda(1-\rho)^{-3}(a-1)^{-1}u^{3-a}\ell(u)$. Осталось вспомнить, что бета-интеграл выражается через гамма-функции как $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$, что приводит к (2.39). $\square$

Изучение асимптотического поведения систем обслуживания при некоторых дополнительных условиях можно проводить в форме предельных теорем. Многие из них получены в условиях высокой загрузки.

Изучение систем обслуживания в условиях большой загрузки проводится с помощью двух подходов. Первый подход связан с асимптотическим анализом явных формул или уравнений, описывающих стационарные распределения числа требований или времени пребывания в системе. Первые предельные теоремы в теории очередей о виде ф.р. надлежащим образом нормированных сл.в. числа требований или времени пребывания в условиях высокой загрузки были получены именно таким образом. Подмеченные закономерности оказались справедливыми и в более общих условиях, когда явные формулы уже отсутствовали. Соответствующие результаты получали с помощью второго подхода, связанного с изучением на основе теории слабой сходимости предельного поведения случайных процессов, возникающих в рассматриваемых системах. Ю. В. Прохоровым было установлено, что в основе явлений, возникающих в нагруженных системах, лежит принцип инвариантности (ныне называемый принципом Донскера–Прохорова), позволяющий с помощью винеровского процесса аппроксимировать довольно сложные процессы, порожденные суммами случайных величин [27, 36, 88, 152]. В этих исследованиях в качестве предельного фигурирует винеровский процесс или его многомерные аналоги.

Применение общих методов исследования поведения систем с разделением процессора в условиях большой загрузки сталкивается с серьезными трудностями из-за более сложной структуры возникающих процессов по сравнению с классическими системами обслуживания.

[26] Отметим, что (2.40) принципиально отличается от асимптотической формулы Крамéра (см. формулу для $1 - W(x)$ в доказательстве следствия 2.12) в том смысле, что оценку Крамéра нельзя получить из (2.40) с помощью формального распространения (2.40) на класс ф.р. с легким хвостом.

[27] В сущности, теорема Карамата утверждает, что интегралы от правильно меняющихся функций снова правильно меняются и тем же свойством обладают урезанные моменты ф.р. из класса RV. В частности, при $x_0 > 0$ таком, что $\ell(x)$ локально ограничена на $[x_0, \infty)$), имеют место:

$$(a) \qquad \int_{x_0}^{x} t^a \ell(t)\,dt \sim \ell(x)\int_{x_0}^{x} t^a\,dt \sim (a+1)^{-1}x^{a+1}\ell(x) \quad \text{при } x \to \infty \quad \text{и } a > -1,$$

$$(b) \qquad \int_{x}^{\infty} t^a \ell(t)\,dt \sim -(a+1)^{-1}x^{a+1}\ell(x) \quad \text{при } x \to \infty \quad \text{и } a < -1.$$

Основываясь на приведенных выше результатах,новые предельные теоремы для систем с разделением процессора можно получить значительно проще с помощью первого подхода. Приведем некоторые из них.

Теорема 2.6. *В системе M/GI/1—EPS*

$$\lim_{\rho\uparrow 1} \mathsf{P}(\mathsf{L}(1-\rho)/\rho < \mathsf{x}) = 1 - \exp(-\mathsf{x}),\ \mathsf{x} > 0. \tag{2.41}$$

Доказательство. См. [165, 54]. □

То же самое утверждение справедливо и для системы M/GI/1—FCFS, но при дополнительном предположении, что $\beta_2 < \infty$. Это хорошо известный результат, восходящий к классическим статьям Кингмана, см., например, [36, Ch.7, §2]. Случай $\beta_2 = \infty$ (например, когда $B(x) \in RV(-a),\ 1 < a < 2$) значительно сложнее и только с конца XX века наметился прогресс в доказательствах соответствующих предельных теорем для системы FCFS. Приведенный результат для дисциплины EPS верен и в случае $\beta_2 = \infty$.

Следующее утверждение доказано в [54, 125, 168] с помощью различных приемов. Доказательство в [168],повидимому, наиболее простое наиболее простое; оно основано на разложении в ряд Тейлора формулы (2.21) для $\mathsf{E}[\mathrm{e}^{-\mathsf{sV(u)}}]$ в системе EPS.

Теорема 2.7. *В системе M/GI/1—EPS при любом фиксированном* $u \in [0, \infty)$

$$\lim_{\rho\uparrow 1} \mathsf{P}(\mathsf{V(u)}(1-\rho)/\mathsf{u} < \mathsf{x}) = 1 - \exp(-\mathsf{x}),\ \mathsf{x} \ge 0. \tag{2.42}$$

Следствие 2.14.

$$\lim_{\rho\uparrow 1} v(s(1-\rho)) = \int_0^\infty (1+su)^{-1}\, dB(u). \tag{2.43}$$

Доказательство. Из определения $v(s,u)$ в §2.1, теоремы 2.7 и теоремы о мажорируемой сходимости. См. [168]. □

Замечание 2.10. Формула (2.43) дает ПЛС ф.р. произведения двух независимых сл.в.: B и экспоненциальной с единичным средним.

Приведем результат для предельного поведения числа требований $L(t)$ в момент t в системе M/GI/1—EPS при $\rho > 1$

Теорема 2.8. *В системе M/GI/1—EPS при любом* $\rho = \lambda\beta_1 > 1$ *выполняется следующий вариант усиленного закона больших чисел*

$$\lim_{t\to\infty} \frac{L(t)}{t} = \gamma \quad \text{с вероятностью } 1, \tag{2.44}$$

где $\gamma = \lambda(1-\pi(0)) > 0$ *— решение уравнения*

$$\lambda \int\limits_0^\infty \mathrm{e}^{-\gamma x}(1-B(x))dx = 1. \tag{2.45}$$

Доказательство. Впервые доказано в [54], однако доказательство сложное и опирается на экзотические теоремы из теории ветвящихся процессов. Утверждение вытекает в качестве простого следствия теоремы из [164] о ф.р. нестационарного процесса числа требований, доказательство см. в [181]. □

Рассмотрим континуум систем M/GI/1—EPS и для каждой из них будем наблюдать стационарное время пребывания требования длины u. Эти наблюдения дают процесс $\{V(u) : u \geq 0\}$. Из результатов [158, 50] вытекает, что этот процесс имеет стационарные и независимые приращения (но не одинаково распределенные). Пусть $\mathcal{N}(x)$ — ф.р. нормального закона со средним 0 и дисперсией 1.

Теорема 2.9. *Если $\rho < 1$ и $\mathsf{Var}[V(u)] < \infty$, то в системе M/GI/1—EPS*

$$\lim_{u\to\infty} \mathsf{P}\left(\frac{\mathsf{V(u)} - \mathsf{u}(1-\rho)^{-1}}{(\lambda\beta_2\mathsf{u})^{1/2}(1-\rho)^{-3/2}} \leq \mathsf{x}\right) = \mathcal{N}(\mathsf{x}). \tag{2.46}$$

Теорема 2.10. *Если $\rho = 1$ и $\mathsf{Var}[V(u)] < \infty$, то в системе M/G/1—EPS*

$$\lim_{u\to\infty} \mathsf{P}\left(\lambda\beta_2\mathsf{V(u)}/\mathsf{u}^2 \leq \mathsf{x}\right) = \mathsf{G}_{1/2}(\mathsf{x}), \tag{2.47}$$

где $G_{1/2}(x) = 2[1 - \mathcal{N}(x^{-1/2})]$, $x > 0$ устойчивое распределение с показателем 1/2 и ПЛС $\mathrm{e}^{-\sqrt{2s}}$.

Две последние теоремы восходят к Прабху [112, гл.3], который доказал их аналоги в контексте теории запасов и для другой системы, интерпретируемой как система M/GI/1—LCFS–P в теории очередей. Для обсуждения других предельных теорем см. [54, 165, 184].

Приведем еще один асимптотический результат. Точное решение для ПЛС безусловного времени пребывания в системе M/GI/1—EPS дается (2.2) и (2.21).

Теорема 2.11. *Для системы M/GI/1—EPS при $B(x) \in RV(-a)$, $a > 1$, $a \notin \mathbf{N}$ и $\rho < 1$ следующие утверждения эквивалентны:*

a) $1 - B(x) \sim x^{-a}\ell(x)$ *при* $x \to \infty$;

b) $\mathsf{P(V > x)} \sim (1-\rho)^{-\mathsf{a}}\mathsf{x}^{-\mathsf{a}}\ell(\mathsf{x})$ *при* $\mathsf{x} \to \infty$.

Аналитическое доказательство, опирающееся на результаты [158] (см. §2.3 и §2.4 обзора), дано впервые в [185]. Для различных обобщений теоремы и других способов доказательств см. [69, 70, 170] и цитированные там работы. В вероятностных доказательствах обычно используется неравенство Маркова (следствием неравенства Маркова является неравенство Чебышева).

3. ТЕОРИЯ СИСТЕМЫ ОБСЛУЖИВАНИЯ M/GI/1—EPS. НЕСТАЦИОНАРНЫЙ РЕЖИМ.

3.1. Нестационарное распределение числа требований

Обсудим ряд результатов исследования системы EPS, обобщающих и усиливающих теоремы 2.4 и 2.5, дающих обоснование метода анализа и приводящих, кроме того, к получению нестационарного распределения процесса числа требований $\{L(t),\ t \in \mathbf{R}_+\}$. Пусть в момент $t = 0^+$ начинается первый период занятости системы и в этот момент она находится в состоянии $(n; x_1, \ldots, x_n)$, т.е. содержит $L(0^+) = n$ требований, $n > 0$, с остаточными длинами $x_i \in [x_i,\ x_i + dx_i)$, $i = 1, 2, \ldots$ Положим $\zeta = \inf\{t > 0 :\ L(t) = 0\}$. Рассмотрим процесс $\{L(t),\ t \in [0, \zeta)\}$ на первом периоде занятости (считаем $L(t) = 0$ при $t > \zeta$). Введем непрерывный аддитивный функционал, построенный на траекториях процесса $\{L(t)\}$ (точнее, на траекториях процесса $X_0(t)$, описанного перед теоремой 2.2):

$$X(t) = \int_0^t \frac{du}{L(u) \vee 1}. \tag{3.1}$$

Будем предполагать, что $X(0) = 0$, $X(t) = X(\zeta^-)$ при $t > \zeta$ и $\lim_{t\to\infty} X(t) = \infty$. Заметим, что $X(t) < \infty$ при $t < \infty$.

Случайная замена времени, связанная с процессом $\{X(t)\}$, определяется как

$$\tau(t) = \inf\{u > 0 : \ X(u) \geq t\}, \quad t \in [0, \infty). \tag{3.2}$$

Это возрастающая кусочно гладкая функция. Семейство таких функций образует процесс $\{\tau(t)\}$ со стационарными приращениями. Пусть $\tau(0) = 0$ и $\tau(t) = \infty$, если $\{u > 0 : \ X(u) \geq t\} = \emptyset$, что доопределяет $\{\tau(t)\}$ для всех ω. Отметим, что при каждом фиксированнном t сл.в. $\tau(t)$ есть марковский момент процесса $X_0(t)$ относительно возрастающего семейства σ-алгебр. Положим

$$M(t) = L(\tau(t)) \tag{3.3}$$

и будем считать $M(t) = 0$ при $\tau(t) = \infty$. Равенство (3.3) означает, что процесс $\{M(t)\}$, определенный почти наверное, получен из процесса $\{L(t)\}$ случайной заменой времени (3.2). (О случайной замене времени см., например, [23], [136] или [149].) Процесс $\{M(t)\}$ непрерывен справа; рассматриваемый вместе со своими дополнительными координатами $x_1(\tau(t)), \ldots, x_{L(\tau(t))}(\tau(t))$, этот процесс обладает строго марковским свойством. Другие подробности о строении преобразованных процессов в этой и родственных системах можно найти в [159,160,163,164].

Пусть $\zeta_m = \inf\{t : \ M(t) = 0\}$. Отметим, что $X(\zeta) = \zeta_m$. Кроме того,

$$\tau(X(t)) = t, \quad X(\tau(t)) = t \quad \text{для каждого } t. \tag{3.4}$$

Теорема 3.1. *Между траекториями процессов $\{L(t)\}$ и $\{M(t)\}$ с соответствующими дополнительными координатами существует взаимно-однозначное соответствие, при котором для каждого фиксированного t*

$$\tau(t) = \int_0^t M(u)\, du. \tag{3.5}$$

Доказательство. (Подробности см. в [164, с.93-94].) Равенство (3.5) доказывается с помощью формулы Лебега, играющей центральную роль в случайной замене времени [23]: если $\{X(t), t \geq 0\}$ — неотрицательный процесс с монотонно возрастающими траекториями и $\{\tau(t),\ t \geq 0\}$ — процесс, определяемый (3.2) (т.е. с помощью обратной к $X(t)$ функции), то для любого процесса $\{Y(t),\ t \geq 0\}$ с измеримыми по Борелю траекториями выполняется равенство

$$\int_0^\infty Y(t)\, X(dt) = \int_0^{X(\infty)} Y(\tau(t))\, dt.$$

Теперь применим эту формулу при фиксированном t к процессу $\{X(t),\ t \geq 0\}$, определяемому равенством (3.1), полагая $Y(t) = \mathbf{1}_{(0,\tau(b)]}(t) L(t)$. Учитывая (3.4), имеем $\mathbf{1}_{(0,\tau(b)]}(\tau(t)) = \mathbf{1}_{(0,\tau(b)]}$. Тогда правая часть формулы Лебега в силу (3.3) сводится к $\int_0^b M(t)\, dt$. Левая часть с учетом (3.1) принимает вид

$$\int_0^\infty \mathbf{1}_{(0,\tau(b)]}(t)(L(t) \vee 1)\frac{1}{L(t) \vee 1}\, dt = \tau(b).$$

Это дает (3.5). □

Из теоремы следует, что (строго возрастающие) траектории процесса $\{\tau(t)\}$ абсолютно непрерывны. Для каждого фиксированного t сл.в. $\tau(t)$ в (3.5) интерпретируется так: $\tau(t)$ на множестве $\{\zeta_m > t\}$ представляет собой сумму обслуженных к моменту t длин требований.

Процесс $\{M(t)\}$ интерпретируется как процесс числа требований в системе $\tilde{M}/G/\infty$ с марковским входящим потоком переменной интенсивности $\lambda L(t)$ (см. также замечание 2.2). Еще одна интерпретация процесса $\{M(t)\}$ состоит в рассмотрении каждого требования как частицы со временем жизни, распределенным согласно $B(x)$ (ср. с эскизом метода). Каждая частица за время своей жизни порождает другие с интенсивностью λ. Тогда $\{M(t)\}$ представляет собой сумму n параллельно протекающих независимых ветвящихся процессов порождения новых частиц (потомков), каждый из которых образуется единственной частицей (предком) из $L(0) = n$ с оставшимся временем жизни x и обрывается в момент t. Сказанное позволяет представить $\{M(t)\}$ при начальном условии $(n; x_1, \ldots, x_n)$ как

$$M(t) = \sum_{i=1}^{n} N_{x_i}(t), \tag{3.6}$$

где $N_x(t)$ — число частиц в момент t в отдельном ветвящемся процессе, порожденном единственным предком со временем жизни x. Траектории каждого процесса $\{N_x(t)\}$ непрерывны справа.

Время жизни такого ветвящегося процесса, обрывающегося в момент t, определяется как

$$\Phi(x,t) = \int_0^t N_x(u)\,du. \tag{3.7}$$

(В иной терминологии $\Phi(x,t)$ было введено как "элемент задержки".) В (3.6) и (3.7) время измеряется по новой шкале.

Независимость составляющих декомпозиции (2.12) становится теперь очевидной.

Введем обозначение $\varphi(z,s,x,t) \doteq \mathsf{E}[\mathrm{e}^{-\mathsf{s}\Phi(\mathsf{x},\mathsf{t})}\,\mathsf{z}^{\mathsf{N}_\mathsf{x}(\mathsf{t})}]$ для преобразования совместного распределения времени жизни $\Phi(x,t)$ и числа частиц $N_x(t)$ в ветвящемся процессе с единственным предком длины x.

Теорема 3.2. *(1986) На первом периоде занятости системы, в которую преобразована система $M/GI/1{-}EPS$ посредством случайной замены времени (3.2), выполняется равенство*

$$\mathsf{E}[\mathrm{e}^{-\mathsf{s}\tau(\mathsf{t})}\,\mathsf{z}^{\mathsf{M}(\mathsf{t})}\mathbf{1}_{(\zeta_\mathsf{m}>\mathsf{t})}|(\mathsf{n};\mathsf{x}_1,\ldots,\mathsf{x}_\mathsf{n})] = \prod_{\mathsf{i}=1}^{\mathsf{n}}\varphi(\mathsf{z},\mathsf{s},\mathsf{x}_\mathsf{i},\mathsf{t}), \tag{3.8}$$

где

$$\varphi(z,s,x,t) = \begin{cases} \delta(z,s,t) & \text{при } x \ge t, \\ \delta(z,s,t)/\delta(z,s,t-x) & \text{при } x < t, \end{cases} \tag{3.9}$$

$$\delta(z,s,t) = z\,\mathrm{e}^{-(s+\lambda)t}/\psi(z,s,t), \quad t \ge 0. \tag{3.10}$$

Здесь функция ψ задается своим преобразованием Лапласа по t (аргумент q) функции $\psi(z,s,t)$

$$\tilde{\psi}(z,s,q) = \int_0^\infty \mathrm{e}^{-qt}\,\psi(z,s,t)\,dt = \frac{q+s+\lambda-\lambda z+\lambda z\beta(q+s+\lambda)}{(q+s+\lambda)(q+\lambda\beta(q+s+\lambda))}, \tag{3.11}$$

$$Re\ s > 0,\ q > -\lambda\pi(s),\ |z| \le 1,\ \pi(s) = \mathsf{E}[\mathrm{e}^{-\mathsf{s}\zeta}] \text{ – решение (2.4).}$$

Комментарии к доказательству. Подробности см. в [164]. Формулы (3.9) – (3.11) являются решением системы уравнений, которые аналогичны по типу уравнениям (2.13), (2.14) и отличаются от них только наличием еще одного переменного z в неизвестных функциях $\delta(z,s,t)$, $\varphi(z,s,x,t)$ и начальными условиями $\varphi(z,s,0,t) = 1$, $\varphi(z,s,x,0) = \delta(z,s,0) = z$. Эти условия отражают тот факт, что $\Phi(0,t)=\Phi(x,0)=0$, $N_0(t)=0$ и $N_x(0)=1$. Уравнения выводятся при рассмотрении инфинитезимальных изменений функций φ и $\delta=\varphi$ при $x\geq t$. Составление уравнений аналогично выводу уравнений (2.14) и (2.13) для $\delta(s,u)=\delta(1,s,u)$ и $\varphi(s,x,u)=\varphi(1,s,x,u)$. Техническая часть доказательства теоремы 3.2 заключается в формальном распространении доказательства теоремы 2.4 на случай учета текущего числа потомков. Отметим лишь, что при этом функция ψ определяется равенством (ср. с замечанием 2.4)

$$\psi(z,s,t)\doteq\exp\left[-\lambda\int_0^t\varphi_B(z,s,u)\,du\right],\ \text{где } \varphi_B(z,s,t)=\int_0^\infty\varphi(z,s,x,t)\,dB(x),$$

откуда следует

$$\int_0^t\varphi_B(z,s,y)\,dy=-\lambda^{-1}\ln\psi(z,s,t), \tag{3.12}$$

а также $\varphi_B(z,s,t)=-\lambda^{-1}\frac{\partial}{\partial t}\psi(z,s,t)/\psi(z,s,t)$ и $\pi(s)=-\lambda^{-1}\lim_{t\to\infty}\ln\psi(1,s,t)$. □

Замечание 2.3 справедливо и для теоремы 3.2. Пусть в момент $t=0$ система M/GI/1—EPS находится в состоянии $(n;x_1,\dots,x_n)$. Введем обозначения

$$p_{n,\mathbf{x}}(z,s)=\int_0^\infty \mathrm{e}^{-st}\,\mathsf{E}[\mathsf{z}^{\mathsf{L(t)}}\mathbf{1}_{(\zeta>\mathsf{t})}|(\mathsf{n};\mathsf{x}_1,\dots,\mathsf{x}_\mathsf{n})]\,\mathsf{dt},\ \ \mathsf{n}\geq1,\ \ \mathsf{Re\ s}>0,\ \ |\mathsf{z}|\leq1, \tag{3.13}$$

$$p_n(z,s)=\int_0^\infty \mathrm{e}^{-st}\,\mathsf{E}[\mathsf{z}^{\mathsf{L(t)}}\mathbf{1}_{(\zeta>\mathsf{t})}|\mathsf{L}(0)=\mathsf{n}]\,\mathsf{dt}.$$

Используя теорему Фубини, можно переписать (3.13) в следующей эквивалентной форме

$$p_{n,\mathbf{x}}(z,s)=\mathsf{E}\left[\int_0^\zeta \mathrm{e}^{-\mathsf{st}}\,\mathsf{z}^{\mathsf{L(t)}}|(\mathsf{n};\mathsf{x}_1,\dots,\mathsf{x}_\mathsf{n})\right]\,\mathsf{dt}. \tag{3.14}$$

Заменяя в (3.14) переменную интегрирования на $\tau(t)$ (мы интегрируем по возрастающему процессу) и снова применяя теорему Фубини, можно получить с учетом (3.3) и (3.5)

$$p_{n,\mathbf{x}}(z,s)=z\int_0^\infty\frac{\partial}{\partial z}\left(\mathsf{E}[\mathrm{e}^{-\mathsf{s}\tau(\mathsf{t})}\,\mathsf{z}^{\mathsf{M(t)}}\mathbf{1}_{(\zeta_\mathsf{m}>\mathsf{t})}|(\mathsf{n};\mathsf{x}_1,\dots,\mathsf{x}_\mathsf{n})]\right)\,dt. \tag{3.15}$$

После снятия условия по остаточным длинам требований, находящихся в системе в момент $t=0$, приходим с учетом (3.12) к утверждению [163, 164, 165]

Теорема 3.3. *При любом конечном ρ справедливо следующее интегральное представление для преобразования Лапласа производящей функции числа требований в момент t на первом периоде занятости системы M/G/1—EPS, который начинается при наличии n требований*

$$p_n(z,s)=\frac{z}{\lambda^n}\int_0^\infty\frac{\partial}{\partial z}\left[-\frac{\partial}{\partial t}\ln\psi(z,s,t)\right]^n dt,\ \ Re\ s>0,\ \ |z|\leq1, \tag{3.16}$$

где преобразование Лапласа функции ψ задается равенством (3.11).

Ограничиваясь случаем $n=1$ (первый период занятости открывается единственным требованием), можно получить $p_1(z,s)$ в явном виде.

Теорема 3.4. *В системе M/G/1—EPS при любом конечном ρ*

$$p_1(z,s)=\frac{z(1-\pi(s))}{s+\lambda(1-z)(1-\pi(s))},\quad Re\ s>0,\ \ |z|\leq1. \tag{3.17}$$

Схема доказательства. (Полное доказательство дано впервые в [164, с.97-98], см. также [165, 166,176].) Простые алгебраические преобразования позволяют представить $\tilde{\psi}(z,s,q)$ (см. (3.11)) в виде

$$\tilde{\psi}(z,s,q)=\tilde{\kappa}_1(s,q)-z\tilde{\kappa}_2(s,q).$$

В этом равенстве

$$\tilde{\kappa}_1(s,q)=[q+\lambda\beta(q+s+\lambda)]^{-1},\ \tilde{\kappa}_2(s,q)=(\lambda-\lambda\beta(q+s+\lambda))\tilde{\kappa}_1(s,q)(q+s+\lambda)^{-1}. \qquad (3.18)$$

Из (3.18) следует, что неизвестные функции $\kappa_i(s,t)=\mathcal{L}^{-1}(\tilde{\kappa}_i(s,q))$, $i=1,2$ являются плотностями функций восстановления обрывающихся процессов восстановления. Здесь $\mathcal{L}^{-1}$ — оператор обращения (двумерного) преобразования Лапласа (контурный интеграл Бромвича); аргументу q соответствует переменная t. (Нет необходимости применять $\mathcal{L}^{-1}$ к одномерному преобразованию Лапласа $\kappa_i(s,t)$.)

Методом преобразования несобственного уравнения восстановления к собственному [48, гл.11, §6], опирающимся на предельную теорему для обрывающихся процессов восстановления, находятся асимптотики $\kappa_i(s,t)\,\mathrm{e}^{\lambda\pi(s)t}$ при $t\to\infty$. Подчеркнем, что для этого следует сначала проверить некоторые условия, гарантирующие возможность преобразований несобственных уравнений восстановления, и далее ввести новую вероятностную меру (в теории страхового риска подобное преобразование случайной меры называется преобразованием Эсшера (Esscher transformation)[47], в итоге из-за него появляется дополнительный экспоненциальный множитель в левой части приводимых далее формул). Асимптотические решения двух полученных собственных уравнений восстановления находятся с помощью узловой теоремы восстановления [131]и имеют вид

$$\lim_{t\to\infty}\kappa_1(s,t)\,\mathrm{e}^{\lambda\pi(s)t}=[C_1(s)(\lambda-\lambda\pi(s))]^{-1},\ \lim_{t\to\infty}\kappa_2(s,t)\,\mathrm{e}^{\lambda\pi(s)t}=[C_2(s)(s+\lambda-\lambda\pi(s))]^{-1},$$

где $C_1(s)<\infty$ и $C_2(s)<\infty$ — некоторые зависящие от s константы. Из упомянутых выше условий вытекает, что $C_1(s)=C_2(s)$ при $s>0$. Тогда из приведенных равенств следует

$$\lim_{t\to\infty}\kappa_2(s,t)/\kappa_1(s,t)=\lambda(1-\pi(s))/(s+\lambda-\pi(s)). \qquad (3.19)$$

При $n=1$ в подынтегральном выражении в правой части равенства (3.16) можно поменять порядок дифференцирования по z и по t, получая после интегрирования

$$p_1(z,s)=\frac{z}{\lambda}\lim_{t\to\infty}\frac{\kappa_2(s,t)}{\kappa_1(s,t)-z\kappa_2(s,t)}.$$

Формула (3.19) позволяет привести последнее равенство к виду (3.17). □

Распределение процесса $\{L(t)\}$ на всей положительной полуоси $[0,\infty)$ можно теперь найти с помощью более стандартных результатов теории восстановления [65], рассматривая два процесса восстановления, образованные моментами начала и окончания последовательных периодов занятости. Предположим для простоты, что в момент $t=0$ система свободна от требований [28]. Введем обозначение $g_0(z,s)\doteq\int_0^\infty \mathrm{e}^{-st}\,\mathsf{E}[\mathsf{z}^{\mathsf{L(t)}}|\mathsf{L}(0)=0]\,\mathsf{dt}$. Нетрудно показать [164, с.99], что функция $g_0(z,s)$ удовлетворяет уравнению $g_0(z,s)=\eta(s)/\lambda+\eta(s)p_1(z,s)$, в котором $\eta(s)=\lambda/(s+\lambda-\lambda\pi(s))$ есть преобразование Лапласа плотности восстановления в процессе восстановления, образованном моментами начала очередного периода занятости. Следовательно,

$$g_0(z,s)=[1+\lambda p_1(z,s)](s+\lambda-\lambda\pi(s))^{-1},\ Re\ s>0,\ |z|\le 1. \qquad (3.20)$$

[28] Переход от нулевых к произвольным начальным условиям не вызывает принципиальных затруднений, но приводит к более громоздким результатам.

Замечание 3.1. Формула (3.20), связывающая характеристики процесса числа требований на $[0, \infty)$ при начальном условии $\mathsf{P}(\mathsf{L}(0) = 0) = 1$ с характеристиками этого процесса на периоде занятости, справедлива для любой консервативной дисциплины. Дисциплиной определяется только вид функции $p_1(z, s)$, что является наиболее трудным этапом анализа.

Равенства (3.17) и (3.20) приводят к утверждению [164, 165, 166]

Теорема 3.5. *Функция*

$$g_0(z, s) = [s + \lambda(1 - z)(1 - \pi(s))]^{-1} \tag{3.21}$$

задает преобразование Лапласа производящей функции распределения числа требований в момент t в системе M/GI/1—EPS при начальном условии $\mathsf{P}(\mathsf{L}(0) = 0) = 1$.

Одним из следствий теоремы 3.5 является вид преобразования Лапласа вероятности отсутствия требования в системе EPS в момент t

Следствие 3.1.

$$g_0(0, s) \equiv \tilde{p}_{00}(s) = \int_0^\infty \mathrm{e}^{-st}\, \mathsf{P}_{00}(\mathsf{t})\, \mathsf{dt} = [\mathsf{s} + \lambda - \lambda\pi(\mathsf{s})]^{-1}. \tag{3.22}$$

Здесь и далее дополнительный нижний индекс 0 будет указывать на начальное состояние системы в момент $t = 0$.

Формула (3.22) показывает, что $\mathsf{P}_{00}(\mathsf{t})$ принадлежит классу стандартных p-функций, введенному Кингманом при изучении феноменов регенерации[29]. Следовательно, многие свойства $\mathsf{P}_{00}(\mathsf{t})$ вытекают из свойств p-функций. В частности, $|\mathsf{P}_{00}(\mathsf{t}_2) - \mathsf{P}_{00}(\mathsf{t}_1)| < \rho|\mathsf{t}_2 - \mathsf{t}_1|$ для всех положительных t_1 и t_2, поскольку $\mathsf{P}'_{00}(0) = -\rho$. Отсюда следует, что $\mathsf{P}_{00}(\mathsf{t})$ абсолютно непрерывна относительно меры Лебега. Это означает, что если ввести функцию $A(t) = (1 - \mathsf{P}_{00}(\mathsf{t}))/\rho$ $(\mathsf{P}_{00}(0) = 1,\ \mathsf{P}_{00}(\infty) = 1 - \rho)$, то она имеет плотность $\alpha(t)$ и $0 \le \alpha(t) \le 1$. (Однако $A(t)$ может не быть дифференцируемой при всех t, примером служит система M/D/1.)

Другая форма следствия 3.1 (а заодно и теоремы 2.1):

Следствие 3.2. *В системе M/GI/1—EPS преобразование Лапласа $\tilde{p}_{00}(s)$ удовлетворяет функциональному уравнению*

$$\tilde{p}_{00}(s) = [s + \lambda - \lambda\beta(1/\tilde{p}_{00}(s))]^{-1}.$$

Доказательство. Положим $\xi = \xi(s) = s + \lambda - \lambda\pi(s)$, $\mathrm{Re}\ s > 0$ в формуле (2.4). Утверждение вытекает из (3.22) и (2.4). □

Ответ на вопрос о среднем числе требований в системе M/GI/1—EPS в момент t дает

Следствие 3.3. [30]

$$\int_0^\infty \mathrm{e}^{-st}\, \mathsf{E}[\mathsf{L}(\mathsf{t})]\, \mathsf{dt} = \lambda(1 - \pi(\mathsf{s}))/\mathsf{s}^2\ . \tag{3.23}$$

[29] В общем, $\mathsf{P}_{00}(\mathsf{t})$ — давно изученная рядом авторов характеристика системы M/GI/1—FCFS; следствие 3.1 формально устанавливает, что в точности той же характеристикой обладает и система M/GI/1—EPS, и в этом нет неожиданного. Следствие 3.1 гарантирует, что все известные утверждения, связанные с этой характеристикой системы M/GI/1—FCFS (например, аналоги теоремы 3.6), автоматически распространяются и на систему M/GI/1—EPS. Тем не менее, некоторые пробелы в изучении $\mathsf{P}_{00}(\mathsf{t})$ следовало заполнить.

[30] В классической системе M/G/1—FCFS $\mathsf{E}[\mathsf{L}(\mathsf{t})]$ в изображении по Лапласу имеет отличающийся от (3.23) вид, зависящий и от $\beta(s)$.

Доказательство. Элементарное вычисление предела $\partial g_0(z,s)/\partial z$ при $z \to 1$. □

Дадим теперь другое доказательство следствия 2.1.
При $\rho < 1$ равенство $\mathsf{P_n} = (1-\rho)\rho^{\mathsf{n}}$, $\mathsf{n} = 0, 1, 2, \ldots$, вытекает из (3.21) после применения классической тауберовой теоремы $\lim_{s\downarrow 0} g_0(z,s) = \lim_{t\to\infty} \mathsf{E}[\mathsf{z}^{\mathsf{L(t)}}] = (1-\rho)/(1-\rho\mathsf{z})$ и последующего обращения полученной производящей функции. □

Из теоремы 3.5 нетрудно извлечь много других новых результатов по характеризации переходного поведения системы M/GI/1—EPS, но далее мы ограничимся кратким обсуждением времени релаксации.

Можно выписать в явном виде нестационарные вероятности состояний системы M/M/1—EPS в терминах модифицированных функций Бесселя первого рода (они совпадают с вероятностями состояний системы M/M/1—FCFS [36],поскольку инфинитезимальный оператор (generator) процесса числа требований в обеих системах один и тот же при $B(x) = 1 - \mathrm{e}^{-\mu x}$, и привести асимптотику для среднего числа требований в момент t при $t \to \infty$ ($\rho < 1$) (см., например, формулу (2.55) для $\mathsf{E}[\mathsf{L(t)}|\mathsf{L(0)} = 0]$ в [167]), из которой следует, что $\mathsf{E}[\mathsf{L(t)}|\mathsf{L(0)} = 0]$ стремится к $\rho/(1-\rho)$ экспоненциальным образом. Константа $T_c = [\mu(1-\sqrt{\rho})^2]^{-1}$ при $\rho = \lambda/\mu < 1$ характеризует скорость сходимости $\mathsf{E}[\mathsf{L(t)}]$ к стационарному значению при $t \to \infty$ и называется *временем релаксации* (ср. с [36], а также со статьями Блана (J.P.Blanc) и Кейлсона (J.Keilson).

Более удобное определение времени релаксации[31] дается равенством

$$T_c = T_c(\mathsf{P_{00}(t)}) = \inf\{\mathsf{T} \geq 0: \ \mathsf{P_{00}(t)} - \mathsf{P_{00}}(\infty) = \mathsf{O}(\mathrm{e}^{-\mathsf{t/T}})\}.$$

Это определение пригодно для любых однолинейных систем обслуживания при соблюдении условий: $\rho < 1$, ф.р. интервалов между поступлениями и длин требований имеют легкий хвост. Поскольку нас интересует система M/GI/1, то достаточно предположить, что $\beta(s) \uparrow \infty$ при $s \downarrow s_b$, где $s_b < 0$ — абсцисса абсолютной сходимости ПЛС $\beta(s)$. В этом случае $\beta(s)$ имеет аналитическое продолжение на часть левой полуплоскости, следовательно, $B(x) \notin \mathcal{K}$. (Варьируя определение T_c и используя вместо $\mathsf{P_{00}(t)}$ другие нестационарные характеристики системы, можно находить для них оценки скорости сходимости, отличающиеся от приведенной далее, которая получена для простейшей характеристики $\mathsf{P_{00}(t)} - (1-\rho)$.)

Теорема 3.6. *При сделанном предположении относительно $\beta(s)$ время релаксации T_c для системы M/G/1—EPS находится из равенства $T_c = -s_0^{-1}$, где s_0 — особая точка с наибольшей действительной частью (исключая полюс в точке $s = 0$) функции $\tilde{p}_{00}(s)$, определяемой равенством (3.22).*

Доказательство. [32] Из известных принципов операционного исчисления, восходящих к Хевисайду [60] (и позднее обоснованных в теории преобразований Лапласа для рассматриваемого класса ф.р.), вытекает, что асимптотическое поведение функции $\mathsf{P_{00}(t)} - \mathsf{P_{00}}(\infty)$ при $t \to \infty$ зависит от положения той особой точки s_0 преобразования Лапласа по t этой функции, которая является ближайшей к мнимой оси из особых точек, находящихся в левой полуплоскости $\mathrm{Re}\, s \leq 0$. Преобразование Лапласа $\tilde{p}_{00}(s)$ — аналитическая функция при $\mathrm{Re}\, s \leq 0$ за исключением конечного числа изолированных особых точек; при $\mathsf{P_{00}}(\infty) \neq 0$ в точке $s = 0$ появляется дополнительный полюс, не играющий роли при вычислении T_c. Это приводит к утверждению теоремы.

[31] Понятие времени релаксации ввел впервые Морз (P.Morse) (1955), но его определение отличается от нашего. См. также по этому поводу [83, §2.8].

[32] Различные варианты доказательств аналогичной теоремы для системы M/GI/1—FCFS давались Кейлсоном и Серви (1987), Бланом и ван Дорном (1986) , Коэном [36] (1982) и др. Нами рассматривался случай M/GI/1—EPS. Из первых строк доказательства станет ясно, что особой оригинальности ни в одном из доказательств не было, поскольку они опирались на идеи Хевисайда.

Более формально теорема доказывается так [167, с.37]. Из (3.22) вытекает, что $\xi = \xi(s) = s + \lambda - \lambda\pi(s)$, $\mathrm{Re}\, s > 0$ есть единственное решение уравнения $\beta(\xi) = (s + \lambda - \xi)/\lambda, \mathrm{Re}\, \xi > 0$. В силу предположения о виде $\beta(s)$, единственным решением уравнения $\beta'(\xi) = -1/\lambda$ является ξ_0 на $(s_b, 0)$. Поэтому в точке $s_0 = \xi_0 + \lambda(\beta(\xi_0) - 1)$ функция $\xi(s)$ имеет точку ветвления. Аналитическое продолжение функции $\xi(s)$ на область $\mathrm{Re}\, s > s_0$ имеет единственный нуль при $s = 0$, являющийся полюсом функции $\tilde{p}_{00}(s)$. Следовательно, s_0 — особая точка с наибольшей действительной частью функции $\tilde{p}_{00}(s)$, за исключением полюса в точке $s = 0$. □

Теорема 3.6 позволяет вычислять значения T_c для системы M/G/1—EPS при различных видах ф.р. $B(x)$ и параметрах системы. Если $\beta(s)$ имеет дробно–рациональный вид со знаменателем степени 1 или 2, то такие вычисления просты. При этом легко убедиться в том, что T_c — возрастающая функция ρ, s_0 и коэффициента вариации.

Более интересный результат — теорема 2.8 об асимптотическом поведении $L(t)/t$ в системе M/GI/1—EPS при $t \to \infty$ и $\rho > 1$. Константа, к которой сходится п.н. это отношение, имеет совершенно другой вид по сравнению с более или менее известной константой $(\rho - 1)/\beta_1$ для системы M/GI/1—FCFS.

Замечание 3.2. Теоремы 3.3 и 3.4 (наряду с теоремами 2.4 и 2.5) — наиболее существенные из наших результатов, рассмотренных в статье. Они содержат не только решения проблем, считавшихся неразрешимыми. Даже более важным представляется новый аналитический метод декомпозиции на элементы задержки, специально построенный для решения этих и смежных задач. При этом, с помощью конструкций из теории случайных процессов мы переложили на математический язык и разработали смутное предвидение Клейнрока, отметившего по поводу (всего лишь) дисциплины FCFS, что “*... изучение периода занятости в действительности является изучением неустановившихся явлений и это одна из причин того, что развитие теории натолкнулось на трудности и затормозилось*” [33].

.

3.2. Нестационарное совместное распределение числа требований и времени пребывания

Метод разложения на элементы задержки оказался настолько мощным, что его дальнейшее усовершенствование позволило получить точное решение задачи, указанной в заголовке. Поскольку рассматривается нестационарный режим, необходимо некоторое начальное условие; для простоты будем предполагать, что рассматриваемая система обслуживания свободна от требований в момент $t = 0$. Приведем результаты точного решения проблемы вычисления совместного распределения сл.в. $V(t, u)$ (времени пребывания помеченного требования длины u, поступившего в момент t) и $L(t)$ (числа требований в момент t) в системе M/GI/1—EPS [172, 173], [174, 169], [170, 171].

Пусть $\zeta = \inf(t > 0 : \ L(t) = 0)$ and $\pi(s) = \mathsf{E}[\mathrm{e}^{-s\zeta}]$ — ПЛС распределения стандартного периода занятости (см. теорему 2.1), т.е. единственное положительное решение (с наименьшим абсолютным значением) функционального уравнения (2.4). Введем определение

$$\tilde{v}_0(z, r, s, u) \doteq \int_0^\infty \mathrm{e}^{-st} \mathsf{E}\left[\mathrm{e}^{-rV(t,u)} z^{L(t)} \,|\, L(0) = 0\right] dt \quad (\mathrm{Re}\, s, r > 0,\ |z| \leq 1). \tag{3.24}$$

Справедлива

[33] В оригинальном издании: “*... our study the busy period has really been the study of a transient phenomenon and this is one of the reasons that the development bogged down*”. (цитируется по книге [82, p. 226]).

Теорема 3.7. *(1997) Для любой загрузки $\rho = \lambda\beta_1$, точное решение для функции $\tilde{v}_0$ имеет вид:*

$$\tilde{v}_0(z,r,s,u) = \tilde{p}_{00}(s)\frac{\delta(r,u)}{1-\tilde{a}(z,r,s,u)/\psi(r,u)} \tag{3.25}$$

где $\tilde{p}_{00}(s) = [s+\lambda-\lambda\pi(s)]^{-1}$, $\delta(r,u) = \mathrm{e}^{-u(r+\lambda)}/\psi(r,u)$, $\psi(r,u)$ дается своим преобразованием Лапласа по u (аргумент q):

$$\tilde{\psi}(r,q) = \frac{q+r+\lambda\beta(q+r+\lambda)}{(q+r+\lambda)(q+\lambda\beta(q+r+\lambda))} \quad (r \geq 0,\, q > -\lambda\pi(r)), \tag{3.26}$$

$$\begin{aligned}\tilde{a}(z,r,s,u) = z\lambda\Bigg\{&\psi(r,u) * \left[\mathrm{e}^{-u(r+\lambda\pi(s)-s)}\int_u^\infty \mathrm{e}^{-y(s+\lambda-\lambda\pi(s))}\,dB(y)\right]\\ &+ \mathrm{e}^{-u(r+\lambda\pi(s)-s)}\int_u^\infty \mathrm{e}^{-y(s+\lambda-\lambda\pi(s))}(1-B(y))\,dy\Bigg\}.\end{aligned} \tag{3.27}$$

Здесь $*$ — символ стилтьесовской свертки and $\pi(s)$ — минимальное решение функционального уравнения (2.4).

Комментарии к доказательству. Доказательство основывается на распространении наших аргументов из [158, 164] (см. §2.3 и теоремы 2.4, 2.5 и 3.3) на рассматриваемый случай, используя при этом такую технику как обрывающийся период занятости и случайная замена времени, а также некоторые конструкции из теории восстановления и теории ветвящихся процессов. Частные случаи $z = 1 \quad (t \to \infty, \rho < 1)$ и $r = 0 \quad (0 < \rho < \infty)$ представлены выше. В сущности, мы вывели выражение для $\tilde{v}_0$ с помощью представления условного времени пребывания $V(t,u)$ и числа требований $L(t)$ в момент t в виде некоторого обобщенного функционала от (нетривиального) ветвящегося процесса типа Крампа–Мода–Ягерса на первом периоде занятости системы. Используя структуру этого ветвящегося процесса, мы составили и решили в терминах многомерных преобразований систему дифференциальных уравнений с частными производными первого порядка, определяющую $\tilde{v}_0$ на первом периоде занятости. Вид $\tilde{v}_0$ на всей положительной полуоси времени был найден с помощью стандартных приемов теории восстановления. Решение содержит контурные интегралы Бромвича, т.е. операторы обращения преобразований Лапласа $\mathcal{L}^{-1}$ (например, $\psi(r,u) = \mathcal{L}^{-1}(\tilde{\psi}(r,q))$ в (3.27)).

Замечание 3.3. Теорема 3.7 может быть представлена в нескольких эквивалентных формах. Здесь выбран вид представления результата, напоминающий формулу (5.2) из [162].

Замечание 3.4. Теорема 3.7 справедлива не только для привычного условия устойчивости $\rho < 1$, но и для $\rho \geq 1$.

Замечание 3.5. Теорема 3.7 содержит в качестве частных случаев все утверждения предыдущих разделов. Из нее можно также извлечь много новых результатов.

3.3. Основные следствия.

Формула (3.25) очень сложна и представляет скорее теоретический интерес. Однако в частных случаях, необходимых в прикладных исследованиях, она значительно упрощается. Покажем, как можно получить несколько важных следствий теоремы 3.7.

Следствие 3.4. *Для любого ρ, преобразование Лапласа по t (аргумент s) вероятностной производящей функции (аргумент z) ф.р. числа требований $L(t)$ в момент t, имеет вид:*

$$\tilde{v}_0(z,0,s,u) = g_0(z,s) = [s+\lambda(1-z)(1-\pi(s))]^{-1}. \tag{3.28}$$

Доказательство. Полагая $r = 0$, (3.27) редуцируется к $\tilde{a}(z, 0, s, u) = z\tilde{p}_{00}(s)(\lambda - \lambda\pi(s))e^{-\lambda u}$. Теперь утверждение следует из (3.25). □

Этот результат доказан в [164, формула (2.108)] (1989) (см. также [163, теорема 2], [166, теорема 3.2]) в качестве самостоятельной теоремы, которая отражена в обзоре как теорема 3.5. Теперь мы получили новое доказательство формулы (3.21) исходя из теоремы 3.7.

Следствие 3.5. *При любом ρ преобразование Лапласа по t (аргумент s) от ПЛС (аргумент r) ф.р. сл.в. $V(t, u)$ —- виртуального времени пребывания в системе требования длины u, поступающего в момент t, дается формулой (3.25), где $\tilde{a}$ получается из (3.27) как $\tilde{a}(1, r, s, u)$. Формула для $\tilde{a}$ опускается, поскольку она имеет почти тот же самый вид как (3.27).*

Доказательство. [174, 173] Положим $z = 1$ в (3.25) и (3.27). □

Этот результат после снятия условия по u дает характеристику системы M/GI/1—EPS подобную решению уравнения Такача для системы M/GI/1–FCFS [138, p.51-56](конечно, наше решение принципиально отличается от решения Такача как методом получения, так и объектом применения этого метода: дисциплины FCFS и EPS — два полярных предельных варианта дисциплины RR при размере кванта $\Theta \to \infty$ и $\Theta \to 0$ соответственно; кроме того, Такач получил решение для ф.р. виртуального времени ожидания, а не виртуального времени пребывания в системе).

Более подробный комментарий. Утверждения следствия 3.5 после усреднения по ф.р. $B(\cdot)$ дают точное решение значительно болеее трудной проблемы нахождения ф.р. (безусловного) виртуального времени пребывания требования, поступающего в момент t в систему M/GI/1—EPS, в терминах трехмерных преобразований (соответствующая сл.в. обозначается через $V(t)$):

$$\int_0^\infty e^{-st}\,\mathsf{E}\left[e^{-\mathsf{r}\mathsf{V}(\mathsf{t})}\,|\mathsf{L}(0) = 0\right]\,\mathsf{dt} = \int_0^\infty \frac{\delta(\mathsf{r}, \mathsf{u})}{1 - \tilde{\mathsf{a}}(1, \mathsf{r}, \mathsf{s}, \mathsf{u})/\psi(\mathsf{r}, \mathsf{u})}\,\mathsf{dB}(\mathsf{u}), \tag{3.29}$$

где $\psi(r, u) = \mathcal{L}^{-1}(\tilde{\psi}(r, q))(r, u)$ и $\tilde{\psi}(r, q)$ дается равенством (3.26).

Метод получения этого решения является дальнейшим развитием метода разложения на элементы задержки, описывающего принципы составления уравнений для процесса достигнутого требованием обслуживания к моменту t. Скорость течения времени в этом процессе определенным образом увеличивается, а процесс удается декомпозировать так, чтобы вычислялись распределения функционалов от его компонент в новой шкале времени (после решения некоторых уравнений). Обратный переход к исходной скорости течения времени приводит к искомому результату.

Упомянутый результат Такача, который, пожалуй, является одним из наиболее блестящих достижений в теории очередей, состоит в точном решении в терминах двойных преобразований классической и намного более простой проблемы нахождения ф.р. виртуального времени ожидания требования, поступающего в момент t в систему M/GI/1—FCFS (соответствующая сл.в. обозначается здесь через $W(t)$):

$$\int_0^\infty e^{-st}\,\mathsf{E}\left[e^{-\mathsf{r}\mathsf{W}(\mathsf{t})}\,|\mathsf{L}(0) = 0\right]\,\mathsf{dt} = \frac{\mathsf{s} + \lambda - \lambda\pi(\mathsf{s}) - \mathsf{r}}{(\mathsf{s} + \lambda - \lambda\beta(\mathsf{r}) - \mathsf{r})[\mathsf{s} + \lambda - \lambda\pi(\mathsf{s})]}. \tag{3.30}$$

Результат Такача (1955) получен им методом составления и решения некоторого уравнения для процесса незаконченной работы.

Замечание 3.6. Одно из принципиальных различий между дисциплинами FCFS и EPS заключается в том, что при дисциплине FCFS очень длинное требование, поступившее на обслуживание, полностью монополизирует прибор, создавая весьма большую задержку каждому из ожидающих требований. При дисциплине EPS такое же требование никогда не сможет полностью монополизировать ресурс прибора; конечно, оно создает определенные задержки будущим поступлениям, но поскольку все требования обслуживаются одновременно, то достаточно короткие из них получат полностью свое обслуживание и выйдут из системы, обогнав длинное требование. Приведенные теоремы позволили количественно охарактеризовать эти качественные различия, причем как точно, так и асимптотически.

Следствие 3.6. $\mathsf{E}[V(t,u)]$ *задается своим преобразованием Лапласа по* t *(аргумент* s*)*

$$v_1(s,u) = \frac{\delta_1(u)}{s^2\tilde{p}_{00}(s)} - \frac{\lambda\delta_1(u)}{s^2\tilde{p}_{00}(s)} * \left[\mathrm{e}^{u/\tilde{p}_{00}(s)} \int_u^\infty \mathrm{e}^{-y/\tilde{p}_{00}(s)}\, dB(y)\right]$$

где $\delta_1(u) = u + \int_0^u \sum_{n=1}^\infty \rho^n F^{n*}(x)\, dx$ *при* $\rho < 1$ *(две другие эквивалентные формы* $\delta_1(u)$ *даны в (2.24) и в равенстве (b) из замечания 2.7). Здесь* $\tilde{p}_{00}(s)$ *приведено в теореме 3.7 (см. также следствия 3.1, 3.2) и* $F^{n*}(x)$ *определяется равенством (2.27*[34]*).*

Доказательство. Результат следует из (3.25) при $z = 1$, используя

$$v_1(s,u) = -\lim_{r\downarrow 0} \partial\tilde{v}_0(1,r,s,u)/\partial r$$

. □

Следствие 3.7. *Распределение стационарного времени пребывания* $V(u)$ *требования длины* u *при* $\rho < 1$ *дается равенством*

$$v(r,u) \doteq \mathsf{E}[\mathrm{e}^{-rV(u)}] = \frac{(1-\rho)\delta(r,u)}{1-\tilde{a}(1,r,0,u)/\psi(r,u)}\,, \tag{3.31}$$

где

$$\tilde{a}(1,r,0,u) = \lambda\psi(r,u) * \left[\mathrm{e}^{-u(r+\lambda)}(1-B(u))\right] + \lambda\mathrm{e}^{-u(r+\lambda)}\int_u^\infty (1-B(x))dx. \tag{3.32}$$

Доказательство. Известно, что $\pi(0) = 1$ при $\rho < 1$. Следовательно, полагая $z = 1$ и применяя классическую тауберову теорему, мы имеем

$$\lim_{s\downarrow 0} s\tilde{v}_0(1,r,s,u) = \lim_{t\to\infty} \frac{1}{t}\int_0^t \mathsf{E}[\mathrm{e}^{-rV(t,u)}]dt = v(r,u).$$

Отсюда и из (3.25), (3.27) после простых преобразований с применением правила Лопиталя вытекает (3.31). □

Это следствие есть не что иное, как теорема 2.5 (с учетом небольшой разницы в обозначениях).

Как и в теореме 3.7, решение содержит контурные интегралы.

Следствие 3.8. *(1999) Эквивалентной формой (3.31) (без контурных интегралов) является*

$$\frac{1}{v(r,u)} = \sum_{n=0}^\infty \frac{r^n}{n!}\xi_n(u), \tag{3.33}$$

[34] Сл.в. с ф.р. $F(x)$ имеет около десятка названий: случайная модификация времени обслуживания, эксцесс, интегрированный хвост (integrated tail) времени обслуживания, остаточное время обслуживания, прямое время возвращения и т.д.

где

$$\xi_0(u) = 1, \quad \xi_n(u) = \frac{n}{(1-\rho)^n} u^{n-1} * W^{(n-1)*}(u), \quad n = 1, 2, \ldots \tag{3.34}$$

Здесь $W^{(n-1)*}(u)$ is $(n-1)$-кратная свертка ф.р. стационарного времени ожидания $W(u)$ в системе $M/GI/1$—FCFS ($W^{0*}(u) = \mathbf{1}(u), W^{1*}(u) = W(u)$), ПЛС ф.р.$W(u)$ задается формулой Поллачека–Хинчина

$$w(q) = (1-\rho)/(1-\rho f(q)), \tag{3.35}$$

где $f(q) = (1-\beta(q))/(q\beta_1)$ есть ПЛС ф.р. $F(x)$.

Доказательство. [174, 173] (1999). Перепишем (3.31) в виде формулы :

$$v(s,u) = (1-\rho)\delta(r,u) \left[1 - \rho\delta(r,u) \left(\int_0^u \delta^{-1}(r, u-x)\, dF(x) + 1 - F(u)\right)\right]^{-1}.$$

Для этого используем преобразование Лапласа по u от $1/\delta(r,u)$ (аргумент q), которое дается равенством (3.26) как $\tilde{\psi}(r, q-r-\lambda), \quad r \geq 0,\ q > r + \lambda - \lambda\pi(r)$. После простых преобразований получаем

$$\int_0^\infty \mathrm{e}^{-qu} \frac{1}{v(r,u)} du = \frac{1}{q}\left[1 + \frac{1}{1-\rho}\frac{r}{q}\frac{1}{1 - \frac{1}{1-\rho}\frac{r}{q}w(q)}\right] = \frac{1}{q}\left[1 + \sum_{n=1}^\infty \left(\frac{1}{1-\rho}\frac{r}{q}\right)^n w(q)^{n-1}\right] \tag{3.36}$$

где $w(q)$ дается равенством (3.35). Отметим, что $\left|\frac{rw(q)}{(1-\rho)q}\right| < 1$ при $q > r + \lambda - \lambda\pi(r),\ \rho < 1$. Теперь нетрудно найти обратное преобразование (по аргументу q) каждого члена разложения в степенной ряд по r преобразования Лапласа по u функции $1/v(r,u)$) (см. равенство (3.36). Результат представлен равенством (3.34), откуда вытекает (3.33), правая часть которого — степенной ряд по r с коэффициентами $\xi_n(u)/n!$. □

Идея такого разложения восходит к Хевисайду [60, Ch.5]. Аналогичный результат получен также в [185] (2000) как теорема 3.1. Однако доказательство в [185] более громоздко, поскольку оно стартует от менее удобного представления нашего результата, а именно — с формулы (5.6) при $z = 1$ из [162] и использует информацию о виде $\mathsf{Var}[V(u)]$ (см. (2.25). В нашем доказательстве не требуется предварительного знания о $\mathsf{Var}[V(u)]$. В действительности, вид $\mathsf{Var}[V(u)]$ наводит на догадку о возможности такого разложения.

Следствие 3.9. *(1999) Пусть $v_n(u) = \mathsf{E}[V(u)^n]$, $n = 1, 2, \ldots$. Тогда справедлива следующая рекуррентная формула*

$$v_n(u) = \sum_{i=1}^n \binom{n}{i} v_{n-i}(u)\xi_i(u)(-1)^{i+1} \tag{3.37}$$

Доказательство. [174, 173, 169]. Разложим в ряд Тейлора преобразование $v(r,u)$ при малых $r > 0$

$$v(r,u) = 1 - \frac{r}{1!}v_1(u) + \frac{r^2}{2!}v_2(u) - \frac{r^3}{3!}v_3(u) + \ldots. \tag{3.38}$$

Произведение рядов (3.38) и (3.33)дает

$$\begin{aligned} -\frac{r}{1!}[v_1(u) - \xi_1(u)] &+ \frac{r^2}{2!}[v_2(u) - 2v_1(u)\xi_1(u) + \xi_2(u)] \\ &- \frac{r^3}{3!}[v_3(u) - 3v_2(u)\xi_1(u) - 3v_1(u)\xi_2(u) - \xi_3(u)] + \ldots = 0 \end{aligned}$$

и это приводит к (3.37) после n-кратного дифференцирования по r и перехода к пределу при $r \to 0$. $\square$

Формула (3.37) была получена в [79, 156, 158] при $n = 2$ и в [174, 173, 185] при произвольном $n \in \mathbf{N}$. Она может быть полезной при получении асимптотических оценок $v_n(u)$ при $u \to \infty$ и $u \to 0$ в духе таких оценок для $\mathsf{Var}[V(u)]$ в [161] (см. (2.32) и (2.33)).

Очередным шагом в исследованиях являлось дальнейшее распространение указанных точных транзиентных решений на случай системы M/GI/1—EPS, в которой постоянно присутствует K, $K = 0, 1, 2 \ldots$ перманентных требований бесконечной длительности [182]. Задачи такого рода рассматриваются также в [19, 28, 151] для систем разделения процессора в стационарном режиме.

Интересные транзиентные решения по распределениям числа требований при некоторых модификациях дисциплины EPS содержат статьи [58, 92], но в этом обзоре они не будут рассматриваться.

Упомянем еще о следующих результатах. В [127] описан алгоритм вычисления ф.р. сл.в. $V(u)$ в системе M/PH/1—EPS, использующий некоторые формулы §2.3 и процедуру униформизации. Появились некоторые новые результаты для средних характеристик системы EPS с более общим входящим потоком (рекуррентным или типа G) [30, 154]. Обсуждение возможностей численного обращения ПЛС, возникающих при анализе систем с разделением процессора, содержится в [148]. (Хорошо известны обзоры Витта по соответствующим проблемам, теперь эти обзоры дополняет недавняя статья [63].) В дополнение к асимптотическим решениям типа теоремы 2.11, найдена асимптотика хвоста с.в. $V(u)$ в случае легких хвостов $B(x)$ для системы M/GI/1—EPS [180] (и в ее частном случае при детерминированном распределении длин [45]). В обзор не включены эти результаты.

Отметим, что до сих пор не удалось найти обратное преобразование для ПЛ ф.р.сл.в. V в системе EPS с ф.р. длин, отличающейся от экспоненциального распределения (следует напомнить, что для системы M/M/1—EPS это было сделано Моррисоном [101](1985)).

4. ВМЕСТО ЗАКЛЮЧЕНИЯ

Теория систем с разделением процессора — новое направление теории очередей, возникшее в 1967 г. после появления статей Клейнрока. Первые нетривиальные результаты в ней принадлежат Сакате и др. [120] (1969). Дальнейшие попытки более глубокого изучения введенных Клейнроком моделей оказались, в сущности, почти безуспешными. Ситуация стала качественно изменяться в лучшую сторону только после появления статей [155, 78, 79, 156, 158, 122]. За последние три десятилетия удалось добиться существенного прогресса, в основном, в получении точных и асимптотических решений задач вычисления стационарных (а затем и нестационарных) распределений вероятностно–временны́х характеристик. В обзоре дано достаточно глубокое (хотя и не полное) представление об этих и других достижениях построенной математической теории системы M/GI/1 с разделением процессора. Почти полностью оставлены в стороне системы обслуживания с другими разновидностями дисциплин разделения процессора в силу естественных ограничений по объему статьи.

Модели разделения процессора начали интенсивно применяться при анализе производительности систем с разделением времени, но особую актуальность они приобрели в настоящее время в связи с созданием специальных протоколов управления передачей данных в компьютерных сетях и развитием асинхронных режимов передачи и обработки информации. Они широко используются также при создании сетей сотовой подвижной связи, при проектировании узлов компьютерных сетей, в системах спутниковой связи и т.п. Более того, системы с разделением процессора начали применяться не только в информатике, но также в экономических и социальных науках (см., например, [59, 102]).

Кроме того, такие модели являются источником новых и нетривиальных математических задач, стимулирующих прогресс в развитии вероятностных методов анализа.

СПИСОК ЛИТЕРАТУРЫ

1. S.Aalto, U.Ayesta. On the non–optimality of the FB discipline within the service time distribution class IMRL. Report. CWI. Amsterdam, Sept. 2005, PNA–E0509, 16 pages.
2. S.Aalto, U.Ayesta. Mean delay analysis of a multi level processor sharing disciplines. *Proc. INFOCOM'06*, 2006. 11 pages.
3. S.Aalto, U.Ayesta and E.Nyberg–Oksanen. Two level processor–sharing scheduling disciplines: mean delay analysys. *ACM SIGMETRICS Performance Evalution Review*, 2004, vol. 32, no. 1, pp. 97–105.
4. J.Abate and W.Whitt. Transient behavior of the M/G/1 workload process. *Operations Research*, 1994, vol. 42, no. 4, pp. 751–764.
5. J.Abate, G.L.Choudhuri and W.Whitt. Waiting–time probabilities in queues with long-tail service–time distributions. *Queueing Systems*, 1994, vol. 16, no. 3–4, pp. 311–333.
6. E.Altman, K.Avrachenkov and U.Ayesta. A survey on discriminatory processor sharing. *Queueing Systems*, 2006, vol. 53, no. 1–2, pp. 53–63.
7. E.Altman, T.Jiménez and D.Kofman. DPS queues with stationary ergodic service times and the performance of TCP in overload. *Proc. IEEE INFOCOM'04*, paper 0-783-8356-7/04, 2004. 9 pages.
8. Г.Т.Артамонов, О.М.Брехов. *Аналитические вероятностные модели функционирования ЭВМ.* М.: Энергия, 1978. 368 стр.
9. C.D'Apice and A.V.Pechinkin. Non–stationary characteristics in MAP/G/1/∞ queue with the foreground–background processor–sharing discipline. *Proc. of the 17th IMACS World Congress 2005 on Sci. Computation, Appl. Math. and Simulation* (Paris, France, July 2005). 2005. 6 pages.
10. S.Asmussen. *Applied Probability and Queues.* Heidelberg: Springer, 2003 (2nd edition). 448 pages. ISBN 0-387-00211-1.
11. S.Asmussen, C.Klüppelberg and K.Sigman. Sampling at subexponential times, with queueing applications. *Stochastic processes and their Applications*, 1999, vol. 79, pp. 265–286.
12. B.K.Asare and F.G.Foster. Conditional response times in the M/G/1 processor–sharing system. *J. Appl. Probab.*, 1983, vol. 20, no. 4, pp. 910–915.
13. K.B.Athreya and P.E.Ney. *Branching Processes.* New York: Springer, 1972.
14. K.Avrachenkov, U.Ayesta and P.Brown. Batch arrival processor–sharing with application to multi–level processor–sharing scheduling. *Queueing Systems*, 2005, vol. 50, no. 4, pp. 459–480.
15. N.T.J.Bailey. A continuous time treatment of a simple queue using generating functions. *J. of Roy. Statist. Soc. ser.B*, 1954, v. 16, pp. 288-291.
16. N.Bansal. Analysis of the M/G/1 processor sharing queue with bulk arrivals. *Oper. Res. Lett.*, 2003, vol. 31, pp. 401–405.
17. Г.П.Башарин, А.Л.Толмачев. Теория сетей массового обслуживания и ее приложения к анализу информационно–вычислительных систем. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика.* 1983, том 21, стр. 3–119. Пер. на англ. яз.: G.P.Bashrin and A.L.Tolmachev. Theory of queueing networks anf its application to the analysis of information–computing systems. *J. of Soviet Mathematics*, 1985, vol. 99, no. 1 (Apr.), pp. 951–1050.
18. V.E.Beneš. On queues with Poisson arrrivals. *Annals of Math. Statist.*, 1957, vol. 28, pp. 670-677.
19. J.L. van den Berg. *Sojourn Times in Feedback and Processor-Sharing Queues.* PhD thesis. Utrecht: Rijksuniversiteit, 1990. 126 pages.

20. U.N.Bhat. *Introduction to Queueing Theory.* (an unpublished lecture course). 2005. Available at http://www.faculty.smu.edu/nbhat.

21. U.N.Bhat, M.Shalaby and M.J.Fisher. Approximation techniques in the solution of queueing problems. *Naval Research Logist. Quart.*, 1979, vol. 26 , pp. 311–326.

22. N.H.Bingham, C.M.Goldie and J.L.Teugels. *Regular Variation.* (Encycl. of Math. and its Appl., Vol. 27). Cambridge: Cambridge Univ. Press, 1987.

23. R.Blumenthal and R.Getoor. *Markov Processes and Potential Theory.* New York: Academic Press, 1968.

24. П.П.Бочаров, В.М.Вишневский. G–сети: развитие теории мультипликативных сетей. *Автоматика и телемеханика,* 2003, № 5, стр. 46–74..

25. П.П.Бочаров, А.В.Печинкин. *Теория массового обслуживания.* М.: РУДН, 1995. 529 стр. ISBN 5-209-00796-0.

26. E.Borel. Sur l'emploi du théorème de Bernoulli pour faciliter le calcul d'une infinité de coefficients. Application au problème de l'attente à un guichet. *Comptes Rendus Hebd. des Séanc. de l'Académie des Sciences*, 1942, vol. 214, pp. 452–456.

27. А.А.Боровков. *Асимптотические методы в теории массового обслуживания.* М.: Наука, 1980.

28. A.Brandt, M.Brandt. On the sojourn times for many–queue head–of–the–line processor–sharing systems with permanent customers. *Math. Methods of Operations Research*, 1998, vol. 47, pp. 181–220.

29. A.Brandt, M.Brandt. A sample path relation for the sojourn times in G/G/1–PS systems and its applications. *Queueing Systems*, 2006, vol. 52, no. 4, pp. 281–286. Preliminary version: ZIB–Report 03–18. Berlin, June 2003, pp. 1–14 (preprint Konrad–Zuse–Zentrum für Informationstechnik Berlin, available at http://www.zib.de/Publications/Reports/ZR-03-18.pdf).

30. A.Brandt, M.Brandt. On the stability of the multi–queue multi–server processor sharing with limited service. ZIB–Report 06–25. Berlin, May 2006, pp. 1–16 (preprint Konrad–Zuse–Zentrum für Informationstechnik Berlin, available at http://www.zib.de/Publications/Reports/ZR-06-25.pdf).

31. A.Brandt, M.Brandt. Waiting times for M/M systems under generalized processor sharing. ZIB–Report 06–46. Berlin, Nov. 2006, pp. 1–25 (preprint Konrad–Zuse–Zentrum für Informationstechnik Berlin, available at http://www.zib.de/Publications/Reports/ZR-06-46.pdf).

32. H.Chen, O.Kella and G.Weiss. Fluid approximation for a processor–sharing queue. *Queueing Systems*, 1997, vol. 27, no. 1–2, pp. 99–125..

33. S.–K.Cheung, J.L. van den Berg and R.J.Boucherie. Decomposing the queue length distribution of processor–sharing models into queue lengths of permanent customer queues. *Proc. of IFIP's Performance'2005* (Juan–les–Pins, France, Oct. 3–7, 2005), 2005..

34. S.–K.Cheung, H. van den Berg and R.J.Boucherie. Insensitive bounds for the moments of the sojourn time distribution for the M/G/1 processor–sharing queue. *Queueing Systems*, 2006, vol. 53, no. 1–2, pp. 7–18.

35. E.G.Coffman, R.Muntz and H.Trotter. Waiting time distributions for processor–sharing systems. *J. Assoc. Comput. Mach.*, 1970, vol. 17, no. 1, pp. 123–130.

36. J.W.Cohen. *The Single Server Queues.* Amsterdam: North–Holland, 1982 (2nd edition). 692 pages.

37. R.B.Cooper. *Introduction to Queueing Theory.* New York: North-Holland, 1981 (2nd edition). ISBN 0-444-00379-7. 347 pages.

38. R.B.Cooper. Queueing theory. Ch. 10 in: *Handbooks in Oper. Res. and Manag. Sci.* Vol. 2: Stochastic Models. Eds. D.P.Heyman and M.J.Sobel. New York: Elsevier, 1990, pp. 469–518.

39. R.B.Cooper and Shun–Chen Niu. Beneš's formula for M/G/1—FIFO "explained" by preemptive resume LIFO. *J. Appl. Probab.*, 1986, vol. 23, no. 2, pp. 550–554.

40. D.R.Cox and W.L.Smith. *Queues.* London: Methuen, 1961. Пер. на рус. яз.: Д.Р.Кокс, У.Л.Смит. *Теория очередей.* М.: Мир, 1966. 218 стр.

41. H.Daduna. *Queueing networks with discrete time scale: explicit expressions for the steady state behavior of discrete time stochastic queueing networks* (Lecture Notes on Computer Science, Vol. 2046), Heidelberg: Springer, 2001.

42. R.L.Disney and P.Kiessler. *Traffic Processes in Queueing Networks. A Markov Renewal Approach.* Baltimore: The John Hopkins Univ. Press, 1987. 252 pages.

43. G.Doetsch. *Introduction to the Theory and Applications of the Laplace Transformation.* New York: Springer, 1974.

44. J.H.Dshalalow. An antology of classical queueing methods. *Advances in Queueing: Theory, Methods, and Open Problems.* Ed. J.H.Dhalalow (Probability and Stochastic Series), Boca Raton: CRC Press, 1995, pp. 1–42.

45. R.Egorova, B.Zwart and O.Boxma. Sojourn time tails in the M/D/1 processor sharing queue. *Probab. in the Engrg. and Informational Sci.*, 2006, vol. 20, no. 3 (July), pp. 429–446.

46. P.Embrechts, C.Klüppelberg and T.Mikosch. *Modelling Extremal Events for Insurance and Finance.* Heidelberg: Springer, 1997. 617 pages.

47. F.Esscher. On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift.*, 1932, vol. 15, pp. 175–195.

48. W.Feller. *Introduction to Probability Theory and its Applications, Vol. 2.* New York: Wiley, 1966. Пер. на рус. яз. Ю.В.Прохорова: В.Феллер. *Введение в теорию вероятностей и ее приложения. Том 2.* М.: Мир, 1967. 752 стр.

49. L.Flatto. The waiting time distribution for the random order service M/M/1 queue. *Ann. Appl. Probab.*, 1997, vol. 7, no. 2, pp. 382–409.

50. R.D.Foley and G.–A.Klutke. Stationary increments in the accumulated work process in processor–sharing queues. *J. of Appl. Probab.*, 1989, vol. 26, no. 3, pp. 671–677.

51. F.G.Foster. Stochastic processes. *Operational Research'72. Proc. 6th IFORS Int. Conf.* (Dublin, Aug. 21–25, 1972). Ed. M.Ross. Amsterdam: North–Holland, 1973, pp. 223–239.

52. P.W.Glynn. Diffusion approximation. Ch. 4 in: *Handbooks in Oper. Res. and Manag. Sci.* Vol. 2: Stochastic Models. Eds. D.P.Heyman and M.J.Sobel. New York: Elsevier, 1990, pp. 145–198.

53. Б.В.Гнеденко, И.Н.Коваленко. *Введение в теорию массового обслуживания.* М.: Наука, 1987 (2-ое издание). 336 стр. Пер. на англ. яз.: B.V.Gnedenko and I.N.Kovalenko. *Introduction to Queueing Theory.* Boston: Birkhäuser, 1991.

54. С.А.Гришечкин. Ветвящиеся процессы Крампа-Мода-Ягерса как метод исследования системы M/G/1 с разделением процессора. *Теория вероятн. и ее примен.* 1991, том 36, № 1, стр. 16–33. Пер. на англ. яз.: S.A.Grishechkin. The Crump–Mode–Jagers branching processes as a method for studying the processor–sharing system M/G/1. *Theory of Probab. and its Appl.*, 1991, vol. 36, no. 1, pp. 19–35.

55. С.А.Гришечкин. Ветвящиеся процессы в теории массового обслуживания. Автореферат докторской диссертации. М.: мехмат МГУ, 1994. 30 стр.

56. H.C.Gromoll. Diffusion approximation for a processor sharing queue in heavy traffic. *Annals of Appl. Probab.*, 2004, vol. 14, no. 2, pp. 555–611.

57. P.Haccou, P.Jagers and V.A.Vatutin. *Branching Processes: Variation, Growth, and Extinction of Populations.* (Ser.: Cambridge Studies in Adaptive Dynamics), Cambridge: Cambridge Univ. Press, 2005.

58. R.C.Hampshire, M.Harchol–Balter and W.A.Massey. Fluid and diffusion limits for transient sojourn times of processor sharing queues with time varying rates. *Queueng Systems*, 2006, vol. 53, no. 1–2, pp. 19–30.

59. R.Hassin and M.Haviv. *To Queue or not to Queue: Equilibrium Behavior in Queueing Systems.* Boston: Kluwer, 2002.

60. O.Heaviside. *Electromagnetic Theory.* London: The Electrician Co. Vol. 1, 1893; Vol. 2, 1899; Vol. 3, 1912.

61. D.L.Iglehart and W.Whitt. Multiple channel queues in heavy traffic. I, II. *Adv. in Appl. Probab.*, 1970, vol. 2, pp. 150–177, pp. 355–369.

62. В.А.Ивницкий. *Теория сетей массового обслуживания.* М.: Физматлит, 2004.

63. P. den Iseger. Numerical transform inversion using Gaussian quadrature. *Probab. in the Engrg. and Informational Sci.*, 2006, vol. 20, no. 1 (Jan.), pp. 1–44.

64. P.Jagers. *Branching Processes with Biological Applications.* London: Wiley, 1975.

65. N.K.Jaiswal. *Priority Queues.* New York: Academic Press, 1968. Пер. на рус. яз.: Н.Джейсуол. *Очереди с приоритетами.* М.: Мир, 1973. 280 стр.

66. N.K.Jaiswal. Performance evaluation studies for time–sharing computer systems. *Performance Evaluation,* 1982, vol. 2, no. 4, pp. 223–236.

67. A.Jean–Marie. On overloaded queue. *30–th Int. Conf. on the Mathematics of Operations Research* (Lunteren, The Netherlands, Jan. 18, 2005), 2005.

68. A.Jean–Marie and Ph.Robert. On transient behavior of the processor–sharing queue. *Queueing Systems,* 1994, vol. 17, pp. 129–136.

69. P.Jelenković and P.Momčilović. Resource sharing with subexponential distributions. *Proc. INFOCOM'02* (June 2002), New York, 2002, pp. 179–190.

70. P.Jelenković and P.Momčilović. Large deviation analysis of subexponential waiting times in a processor sharing queue. *Mathematics of Operations Research*, 2003, vol. 28, no. 3, pp. 587–608.

71. J.F.C.M. de Jongh. *Share Scheduling in Distributed Systems.* PhD thesis. Delft: Techn. Univ., 2002.

72. F.I.Karpelevich and A.Ya.Kreinin. *Heavy Traffic Limits for Multiphase Queues.* (Translations of Math. Monographs, Vol. 137), Providence: AMS, 1994.

73. М.Я.Кельберт, Ю.М.Сухов. Математические вопросы теории сетей с очередями. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика.* 1988, том 26, стр. 3–96. Пер. на англ. яз.: M.Ya.Kel'bert and Yu.M.Sukhov. Mathematical theory of queueing networks. *J. of Soviet Mathematics*, 1990, vol. 50, no. 3 (June), pp. 1527–1600.

74. D.G.Kendall. Some problems in the theory of queues, *J. of Roy. Statist. Soc. Ser.B*, 1951, vol. 13, pp. 151–185.

75. G. van Kessel. *Limiting Regimes and Approximations for Discriminatory Processor Sharing.* Master's thesis. Eindhoven: Univ. of Technology, 2005. 68 pages.

76. J.Kim and B.Kim. Sojourn time distribution in the M/M/1 queue with discriminatory processor sharing. *Performance Evaluation*, 2004, vol.58, no.4, pp. 341–365.

77. J.F.C.Kingman. The single server queue in heavy traffic. *Proc. Cambridge Philos. Soc.*, 1961, vol. 57, pp. 902–904.

78. М.Ю.Китаев, С.Ф.Яшков. Распределение условного времени пребывания в системе с разделением времени обслуживания. *Изв. АН СССР. Техн. кибернетика*, 1978, № 4, стр. 211–215. Пер. на англ. яз.: M.Yu.Kitayev and S.F.Yashkov. Distribution of the conditional sojourn time in a system with division of time of servicing. *Engrg. Cybernetics*, 1978, vol. 16, no.4, pp. 162–167.

79. М.Ю.Китаев, С.Ф.Яшков. Анализ одноканальной системы обслуживания с дисциплиной равномерного разделения прибора. *Изв. АН СССР. Техн. кибернетика*, 1979, № 6, стр. 64–71. Пер. на англ. яз.: M.Yu.Kitayev and S.F.Yashkov. Analysis of a single–channel queueing systems with the discipline of uniform sharing of a device. *Engrg. Cybernetics*, 1979, vol. 17, no. 6, pp. 42–49.

80. B.Klar. A note on the $\mathcal{L}$–class of life distributions. *J. Appl. Probab.*, 2002, vol. 39, no. 1, pp. 11–19.

81. L.Kleinrock. Time–shared systems: a theoretical treatment. *J. Assoc. Comput. Mach.*, 1967, vol. 14, no. 2, pp. 242–251.

82. L.Kleinrock. *Queueing systems, Vol. 1: Theory.* New-York: Wiley, 1975. ISBN 0-471-49110-1. 417 pages. Пер. на рус. яз.: Л.Клейнрок. *Теория массового обслуживания.* М.: Машиностроение, 1979. 432 стр.

83. L.Kleinrock. *Queueing Systems, Vol. 2: Computer Applications.* New–York: Wiley, 1976. 549 pages. ISBN 0-471-49111-X. Пер. на рус. яз.: Л.Клейнрок. Вычислительные системы с очередями. М.: Мир, 1979. 600 стр.

84. Г.П.Климов, А.К.Ляху, В.Ф.Матвеев. *Математические модели систем с разделением времени.* Кишинев: Штиинца, 1983. 112 с.

85. Ch.Knessl. A diffusion model for two parallel queues with processor sharing: transient behavior and asymptotics. *J. of Appl. Math. and Stochastic Analysis*, 1999, vol. 12, no. 4, 311–338.

86. H.Kobayashi and A.Konheim. Queueing models of computer communications system analysis. *IEEE Trans. Commun.*, 1977, vol. 25, no. 1, pp. 2–28.

87. I.N.Kovalenko. Rare events in queueing systems — A survey. *Queueing Systems*, 1994, vol. 16, pp. 1–49.

88. И.Н.Коваленко, Н.Ю.Кузнецов, В.М.Шуренков. *Случайные процессы. Справочник.* Под ред. А.В.Скорохода. Киев: Наукова думка, 1983. 368 стр. Пер. на англ. яз.: I.N.Kovalenko, N.Yu.Kuznetsov and V.M.Shurenkov. *Models of Random Processes. A Handbook.* Ed. by A.V.Skorokhod. Boca Raton: CRC Press, 1996. 448 pages.

89. Д.Кёниг, В.В.Рыков, Ф.Шмидт. Стохастические системы массового обслуживания с зависимостями. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика.* 1981, том 18, стр. 95–186. Пер. на англ. яз.: D.König, V.V.Rykov and V.Schmidt. Stationary queuing systems with dependencies. *J. of Soviet Mathematics*, 1983, vol. 21, no. 6 (Apr.), pp. 938–994.

90. B.Krishna Kumar and D.Arivudainambi. Transient solution of an M/M/1 queue with catastrophes. *Computers and Mathematics with Applications*, 2000, vol. 40, pp. 1233–1240.

91. V.G.Kulkarni. Fluid models for single buffer systems. *Frontiers in Queueing: Models and Applications in Science and Engineering.* Ed. J.H.Dshalalow. Boca Raton: CRC Press, 1997, pp. 321–338.

92. Q.-L. Li, Z.Lian and L.Liu. A RG–factorization approach for a BMAP/M/1 generalized processor–sharing queue. *Stochastic Models*, 2005, vol. 21, no. 2–3, pp. 507–530.

93. Q.-L. Li, C.Lin. The M/G/1 procesor-sharing queue with disasters. *Computers & Mathematics with applications*, 2006, vol. 51, pp.987–998.

94. V.Limic. On the behavior of LIFO preemptive resume queues in heavy traffic. *Elec. Comm. in Probab.*, 1999, vol. 4, pp. 13–27.

95. H.Masuyama and T.Takine. Sojourn time distribution in a MAP/M/1 processor–sharing queue. *Oper. Res. Lett.*, 2003, vol. 31, pp. 406–412.

96. J.McKinney. A survey of analytical time–sharing models. *Computing Surveys*, 1969, vol 1, no. 2, pp. 105–116.

97. В.Ф.Матвеев, В.Г.Ушаков. *Системы массового обслуживания. Учебное пособие для вузов.* М.: МГУ, 1984. 240 стр.

98. I.Mitrani. Response time problems in communication networks. *J. of the Royal Statist. Soc., ser. B*, 1986, vol. 47, no. 3, pp. 396–406.

99. I.Mitrani. *Probabilistic Modeling.* Cambridge: Cambridge Univ. Press, 1997.

100. Г.К.Мишкой. *Вероятности состояния приоритетных систем в нестационарном режиме.* Кишинев: Штиинца, 1979.

101. J.Morrison. Response time for a processor–sharing system. *SIAM J. Appl. Math.*, 1985, vol. 45, no. 1, pp. 152–167.

102. H.Moulin. Minimizing the worst slowdown: off–line and on–line. *Working paper*, 2005.

103. В.И.Нейман. Сети связи электронных вычислительных машин. *Итоги науки и техники. Сер. Электросвязь.* 1978, том 9, стр. 5–119.

104. G.F.Newel. *Applications of Queueing Theory.* London: Chapman and Hall, 1971.

105. R.Núñez–Queija. *Processor Sharing Models for Integrated Services Networks.* PhD thesis, Eindhoven: Univ. of Technology, 2000. 193 pages.

106. M.Nuyens. *The Foreground–Background Queue.* PhD thesis. Amsterdam: Amsterdam Univ., 2004. 160 pages.

107. T.M.O'Donovan. Direct solutions of M/G/1 processor sharing models. *Operations Research*, 1974, vol. 22, no. 6, pp. 1232–1235.

108. T.M.O'Donovan. The queue M/G/1 when jobs are scheduled within generations. *Operations Research*, 1975, vol. 23, no. 4, pp. 821–824.

109. T.Ott. The sojourn–time distribution in the M/G/1 queue with processor sharing. *J. Appl. Probab.*, 1984, vol. 21, no. 2, pp. 360–378.

110. А.В.Печинкин. Анализ однолинейных систем массового обслуживания с различными дисциплинами обслуживания. Автореферат докторской диссертации. М.: мехмат МГУ, 1986. 34 стр.

111. F.Pollaczek. La loi de l'attente des appels téléphoniques. *Comptes Rendus Hebd. des Séanc. de l'Académie des Sciences*, 1946, vol. 222, pp. 352–355.

112. N.U.Prabhu. *Stochastic Storage Processes.* New York: Springer, 1980. Пер. на рус. яз.: Н.Прабху. *Стохастические процессы теории запасов.* М. Мир, 1984. 185 стр.

113. V.Rakočević. *Dynamic Bandwidth Allocation in Multi–Class IP Networks Using Utility Functions.* PhD thesis. London: Dept. of Electr. Engrg. Queen Mary, Univ. of London., 2002. 164 pages.

114. K.M.Rege and B.Sengupta. The M/G/1 processor–sharing queue with bulk arrivals. *Modelling and Performance Evaluation of ATM Tecnology.* Amsterdam: Elsevier, 1993, pp. 417–432.

115. K.M.Rege and B.Sengupta. A decomposition theorem and related results for the discriminatory processor sharing queue. *Queueing Systems*, 1994, vol. 18, no. 3–4, pp. 333–351.

116. J.A.C.Resing, G.Hooghiemstra and M.S.Keane. The M/G/1 processor sharing queue as the almost sure limit of feedback queues. *J. Appl. Probab.*, 1990, vol. 0, no. 0, pp. 0–0.

117. A.Riedl. *Routing Optimization and Capacity Assignment in Multi–Service IP Networks.* Dr.–Ing. Dissertation. München: Techn. Univ., 2003.

118. T.Rolski. *Twisting in Applied Probability* (Draft of the lection course). Edinburgh: Heriot–Watt Univ., Dept. of Acturial Math. and Statistics, 2004. 90 pages.

119. T.L.Saaty. *Elements of Queueing Theory with Applications.* New York: Mc Graw Hill Book Co., 1961. Пер. на рус. яз.: Т.Саати. *Элементы теории массового обслуживания и ее приложения.* М.: Сов. радио, 1971. 520 стр..

120. M.Sakata, S.Noguchi and J.Oizumi. Analysis of a processor shared model for time sharing systems. *Proc. 2nd Hawaii Int. Conf. on System Sci.*, Honolulu: Univ. of Hawaii, 1969, pp. 625–628.

121. R.Schassberger. *Warteschlangen.* Wien: Springer, 1973.

122. R.Schassberger. A new approach to the M/G/1 processor–sharing queue. *Adv. Appl. Probab.*, 1984, vol. 16, no. 1, 202–213.

123. R.Schassberger. The steady state distribution of spent service present in the M/G/1 foreground–background processor–sharing queue. *J. Appl. Probab.*, 1988, vol. 25, no. 1, pp. 194–203.

124. Б.А.Севастьянов. Теория ветвящихся процессов. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика. 1967*, М.: ВИНИТИ, 1968, стр. 5–46.

125. B.Sengupta. An approximation for the sojourn–time distribution for the GI/G/1 processor–sharing queue. *Stochastic Models*, 1992, vol. 8, no. 1, pp. 35–57.

126. B.Sengupta and D.L.Jagerman. A conditional response time of the M/M/1 processor–sharing queue. *AT & T Technical J.*, 1985, vol. 64, pp. 409–421.

127. B.Sericola, F.Guillemin and J.Boyer, Sojourn times in the M/PH/1 processor sharing queue, *Queueing systems*, 2005, vol.50, no. 1, pp. 109–130.

128. А.Н.Ширяев. *Основы стохастической финансовой математики. Том 1: Факты. Модели; Том 2: Теория.* М.: Фазис, 1998, 1017 стр. ISBN 5-7036-0043-X (Т. 1); 5-7036-0044-8 (Т. 2). Пер. на англ. яз.: A.N.Shiryaev. *Essentials of Stochastic Finance: Facts, Models, Theory.* Singapore: World Scientific, 1999.

129. K.Sigman. A primer on heavy tailed distributions. *Queueing Systems*, 1999, vol. 33, pp. 261–275.

130. А.Д.Соловьев. Анализ системы M/G/1 для различных дисциплин обслуживания. *Теория массового обслуживания. Труды Всес. шк.-совещания* (Баку, сент. 1978). Под ред. Б.В. Гнеденко, В.В. Калашникова. М.: ВНИИСИ, 1981, 172-178.

131. W.Smith. Regenerative stochastic processes. *Proc. of the Royal Soc., Ser. A*, 1955, vol. 232, pp. 6–31.

132. S.Stidham. Analysis, design, and control of queueing systems. *Operations Research*, 2002, vol 50, no. 1, pp. 197–216.

133. D.Stoyan. *Qualitative Eigenschaften und Abschätzungen Stochastischer Modelle.* Berlin: Akademie, 1977. Пер. на рус. яз.: Д.Штойян.*Качественные свойства и оценки стохастических моделей.* М.: Мир, 1979. 270 стр.

134. D.Stoyan. *Comparison Methods for Queues and Other Stochastic Models.* New York: Wiley, 1983.

135. B.W.Stuck and E.Arthurs. *A Computer & Communications Network Performance Analysis Primer.* Englewood Cliffs: Prentice–Hall, 1985. Available at http://www.signallake.com/publications/Primer.

136. R.Syski. Markov functionals in teletraffic theory. *Teletraffic Analysis and Computer Performance Evaluation.* Eds. O.J.Boxma and H.C.Tijms. Amsterdam: North–Holland, 1986, pp. 303–317.

137. R.Syski. A personal view of queueing theory. *Frontiers in Queueing: Models and Applications in Science and Engineering.* Ed. J.Dshalalow. Boca Raton: CRC Press, 1997, pp. 3–18.

138. L.Takács. *Introduction to the Theory of Queues.* New York: Oxford Univ. Press, 1962.

139. L.Takác. *Combinatorial Methods in the Theory of Stochastic Processes*, New York: Wiley, 1967. Пер. на рус. яз.: Л.Такач.*Комбинаторные методы в теории случайных процессов.* М.: Мир, 1971. 264 стр.

140. S.K.Tripathi and A.Duda. Time–dependent analysis of queueing systems. *Info. Systems and Operational Researh (INFOR).* 1986, vol. 24, no. 3, pp. 199–219.

141. Г.Ш.Цициашвили, М.А.Осипова. *Кооперативные эффекты в стохастических моделях.* М.: Наука, 2005. 198 стр. ISBN 5-02-033708-0.

142. В.А.Ватутин, А.М.Зубков. Ветвящиеся процессы. I. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика.*, 1985, том. 23, стр. 3–67. Пер. на англ. яз.: V.A.Vatutin and A.M.Zubkov. Branching processes. I. *J. of Soviet Mathematics*, 1987, vol. 39, no. 1 (Oct.), pp. 2431–2475.

143. V.A.Vatutin and A.M.Zubkov. Branching processes. II. *J. of Soviet Mathematics.* 1993, vol. 67, no. 6 (Dec.), pp. 3407–3485.

144. В.М.Вишневский. *Теоретические основы проектирования компьютерных сетей.* М.: Техносфера, 2003. 512 стр. ISBN 5-94836-011-3.

145. В.М.Вишневский, О.Семенова. Математические методы исследования систем поллинга. *Автоматика и телемеханика,* 2006, № 2, стр. 3–56.

146. D.Villela, P.Pradhan and D.Rubenstein. Provisioning servers in the application tier for e–commerce systems. *Proc. 12th IEEE Int. Workshop on Quality of Service* (Montreal, June, 2004) 2004, pp. 57–66.

147. В.А.Волконский. Случайная замена времени в строго марковских процессах. *Теория вероятн. и ее примен.*, 1958, том 3, no. 3, стр. 332–350.

148. A.Ward and W.Whitt. Predicting response times in processor–sharing queues. *Proc. of the Fields Institute Conf. on Commun. Networks* (Toronto, 1998). Eds. P.Glynn, D.MacDonald and S.Turner. AMS, 2000, pp. 1–29 (available at http://www.columbia.edu/∼ww2040/response.pdf).

149. H.von Weizsäcker and G.Winkler. *Stochastic Integrals.* Braunschweig: Fridr. Vieweg, 1990.

150. A.Wierman and M.Harchol–Balter. Classifying scheduling policies with respect to higher moments of conditional response time. *SIGMETRICS'05. Proc. ACM 2005 Sigmetrics Int. Conf. on Measurement and Modeling of Computer Systems* (Bannf, Alberta, Canada, June 6–10, 2005), 2005, pp. 229–240.

151. W.Whitt. The M/G/1 processor–sharing queue with long and short jobs. *Unpublished manuscript*, Sept. 1998 (available at http://www.research.att.com/resourses/trs/TRs/ 98/98.37/98.37.1/body.ps).

152. W.Whitt. *Stochastic–Process Limits: An introduction to stochastic–process limits and their applications to queues*, New York: Springer, 2002. 727 pages.

153. R.W.Wolff. *Stochastic Modeling and the Theory of Queues.* Englewood Cliffs: Prentice–Hall, 1989.

154. D.-A.Wu and H.Takagi. Processor–sharing and random–service queues with semi–Markovian arrivals. *J. Appl. Probab.*, 2005, vol. 42, no. 2, pp. 478–490.

155. С.Ф.Яшков. Распределение условного времени ожидания в системе с разделением времени. *Изв. АН СССР. Техн. кибернетика*, 1977, № 5, стр. 88–94. Пер. на англ. яз.: S.F.Yashkov. Distribution of the conditional waiting time in a system with division of time. *Engrg. Cybernetics*, 1977, vol. 15, no. 5, pp.44–52.

156. С.Ф.Яшков. О некоторых результатах анализа вероятностной модели систем телеобработки. *Автоматика и вычисл.техника*, 1981, № 4, стр. 3–11. Пер. на англ. яз.: S.F.Yashkov. Some results of analyzing a probabilistic model of remote processing systems. *Autom. Control and Computer Sci.*, 1981, vol. 15, no. 4, pp.1–8.

157. С.Ф.Яшков. Система с разделением процессора и ее применение в моделях вычислительных сетей. *VIII Всес. конф. по теории кодирования и передаче информации.* М.–Куйбышев: Научн. Сов. по киберн. АН СССР, 1981, часть 3, стр. 67–72.

158. S.F.Yashkov. A derivation of response time distribution for an M/G/1 processor–sharing Queue. *Problems of Control and Info. Theory*, 1983, vol. 12, no. 2, pp. 133–148.

159. С.Ф.Яшков. Анализ системы с приоритетным разделением процессора. *Автоматика и вычисл. техника*, 1984, № 3, стр. 29–38. Пер. на англ. яз.: S.F.Yashkov. Analysis of a system with priority–based processor–sharing. *Autom. Control and Computer Sci.*, 1984, vol. 18, no. 3, pp. 27–36.

160. S.F.Yashkov. New application of random time change to analysis of processor–sharing queues. *4–th Int. Vilnius Conf. on Prob. Theory and Math. Statistics. Abstracts* (June 24–29,1985), Vilnius: Inst. of Math. and Cybern., 1985, vol. 4, pp. 343–345.

161. S.F.Yashkov. A note on asymptotic estimates of the sojourn time variance in the M/G/1 queue with processor-sharing. *Systems Analysis. Modelling. Simulation*, 1986, vol. 3, no. 3, pp. 267–269.

162. S.F.Yashkov. Processor–sharing queues: some progress in analysis (invited paper). *Queueing Systems*, 1987, vol. 2, no. 1, pp. 1–17.

163. S.F.Yashkov. The non–stationary distribution of numbers of calls in the M/G/1 processor-sharing queue. *Systems Analysis and Simulation 1988. Proc. 3rd Int. Symp.* (Berlin, Sept. 12–16, 1988). Eds. A.Sydow et al. Berlin: Akademie, 1988, vol. 2 (Ser. Math. Research, Bd. 47), pp. 158–162. Reprinted in: *Advances in Simulation.* Eds. P.A.Lukar and B.Schmidt. Berlin: Springer, 1988, vol. 2, pp. 158–162.

164. С.Ф.Яшков. *Анализ очередей в ЭВМ.* М.: Радио и связь, 1989. 216 стр. ISBN 5-256-00304-6. Рез. и оглавл. на англ. яз. (Рецензия в *Math. and Comput. in Simul.*, 1991, vol. 33, no. 2 (Aug.), pp.177–178).

165. С.Ф.Яшков. О некоторых характеристиках неклассических моделей очередей с разделением процессора. *Модели и методы информационых сетей.* М.: Наука, 1990, стр. 14–19.

166. S.F.Yashkov. Time–dependent analysis of processor–sharing queue. *Queueing, Performance and Control in ATM. Proc. ITC–13* (Copenhagen, June 19–26, 1991). Eds. J.W.Cohen and C.D.Pack. Amsterdam: Elsevier, pp. 199–204.

167. С.Ф.Яшков. Математические вопросы теории систем обслуживания с разделением процессора. *Итоги науки и техники. Сер. Теория вероятностей. Матем. статистика. Теоретич. кибернетика.*, 1990, том. 29, стр. 3–82. Пер. на англ. яз.: S.F.Yashkov. Mathematical problems in the theory of shared–processor systems. *J. of Soviet Mathematics*, 1992, vol. 58, no. 2 (Jan.), pp. 101–147.

168. S.F.Yashkov. On a heavy traffic limit theorem for the M/G/1 processor–sharing queue. *Stochastic Models*, 1993, vol. 9, no. 3, pp. 467–471.

169. S.F.Yashkov. Modelling processor sharing queue. *Proc. of the 16th IMACS World Congress 2000 on Sci. Computation, Appl. Math. and Simulation* (Lausanne, EPFL, Aug. 21–25, 2000). Eds. M.Deville and R.Owens. EPFL and Rutgers Univ., 2000 (4 pages in pdf format on CD–ROM, Track 799–4, ISBN 3-9522075-1-9).

170. С.Ф.Яшков. *Математические модели систем с разделением времени.* М.: ИППИ РАН, 2002. 85 стр.

171. S.F.Yashkov. On sojourn time problem in processor sharing queue. *Int. Conf. "Kolmogorov and Contemporary Mathematics". Abstracts* (Moscow, June 16–21, 2003). Moscow: Moscow State Univ., 2003, pp. 594–595.

172. S.F.Yashkov and A.S.Yashkova. The M/G/1 processor–sharing system: the transient solutions. *Distributed Computer Commun. Networks (DCCN'97). Proc. 2nd Int. Conf.* (Tel–Aviv, Nov. 4–8, 1997). Moscow: Inst. for Info. Transm. Problems, 1997, pp. 261–272.

173. S.F.Yashkov and A.S.Yashkova. Processor sharing queue: additional results. *Distributed Comput. Commun. Networks (DCCN'99). Proc. 3rd Int. Conf.* (Tel–Aviv, Nov. 9–13, 1999). Moscow: Inst. for Info. Transm. Problems, 1999, pp. 216–221.

174. S.F.Yashkov and A.S.Yashkova. Processor sharing queue: transient solutions. *Computer Sci. and Info. Technologies. Proc. 2nd Int. Conf.* (Yerevan, Aug. 17–22, 1999), Yerevan: National Acad. of Sci. of Armenia, 1999, pp. 99–103.

175. S.F.Yashkov and A.S.Yashkova. The time–dependent solution of the M/G/1—FBPS queue. *Information Processes*, 2004, vol. 4, no. 2, pp. 175–187 (available at http://www.jip.ru/).

176. S.F.Yashkov and A.S.Yashkova. A note on asymptotics associated with limit theorem for terminating renewal process in processor sharing queue. *Information Processes*, 2004, vol. 4, no. 3, pp. 256–260 (available at http://www.jip.ru/).

177. S.F.Yashkov and A.S.Yashkova. Some insight to the time–dependent proreties of the queue length process in the M/G/1—EPS and LCFS–P queues. *Information Processes*, 2005, vol. 5, no. 2, pp. 102–105 (available at http://www.jip.ru/).

178. S.F.Yashkov and A.S.Yashkova. The moments of the sojourn time in the M/G/1 processor sharing system. *Information Processes*, 2006, vol. 6, no. 3, pp. 237–249 (available at http://www.jip.ru/).

179. S.F.Yashkov and A.S.Yashkova. Two special cases of the M/G/1—EPS queue. *Information Processes*, 2006, vol. 6, no. 3, pp. 250–255 (available at http://www.jip.ru/).

180. С.Ф.Яшков, А.С.Яшкова. Асимптотика времени пребывания в системе M/GI/1 с разделением процессора в случае легких хвостов. *Обозрение прикл. и промышл. математики*, 2006, том 13, № 4.

181. A.S.Yashkova. A note on limit theorem for overloaded processor sharing queue. *Information Processes*, 2003, vol. 3, no. 2, pp. 151–153 (available at http://www.jip.ru/).

182. A.S.Yashkova and S.F.Yashkov. Distribution of the virtual sojourn time in the M/G/1 processor sharing queue. *Information Processes*, 2003, vol. 3, no. 2, pp. 128–137 (available at http://www.jip.ru/).

183. R.D.Yates. *High Speed Round Robin Queueing Networks.* PhD thesis. Massachusetts Inst. of Technology, May 1990.

184. A.P.Zwart. *Queueing Systems with Heavy Tails.* PhD thesis. Eindhoven: Univ. of Technology, 2001. 225 pages.

185. A.P.Zwart and O.J.Boxma. Sojourn time asymptotics in the M/G/1 processor sharing queue. *Queueing systems*, 2000, vol. 35, pp. 141–166.

Статью представил к публикации член редколлегии В.И.Венец