

КРАТКИЙ ОБЗОР РАБОТ В ОБЛАСТИ ИССЛЕДОВАНИЯ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С НЕНАДЕЖНЫМИ ОБСЛУЖИВАЮЩИМИ ПРИБОРАМИ

А. Н. Дудин, С. А. Дудин

*Белорусский государственный университет
Минск, Беларусь
e-mail: dudin@bsu.by*

Приводится краткая история вопроса и дается обзор недавно опубликованных работ по ненадежным системам массового обслуживания с коррелированными входными потоками.

Ключевые слова: системы массового обслуживания; ненадежные обслуживающие приборы.

BRIEF OVERVIEW OF RESULTS IN THE STUDY OF QUEUING SYSTEM WITH THE UNRELIABLE SERVERS

A. N. Dudin, S. A. Dudin

*Belarusian State University
Minsk, Belarus*

History of analysis of unreliable queueing models is described in brief and survey of recent publications devoted to analysis of such queues is presented.

Keywords: queueing systems; unreliable servers.

Общеизвестно, что теория массового обслуживания (ТМО) является хорошим математическим аппаратом для решения проблем планирования, оценивания производительности, оптимизации и оптимального управления многими реальными процессами в производственных, технических, экономических и других системах. Особенно большие приложения она нашла при исследовании и оптимизации телекоммуникационных сетей. Первые работы в области ТМО датского ученого А. К. Эрланга по многоканальным системам массового обслуживания (СМО) появились более ста лет назад и инспирированы именно приложениями к задачам планирования и дизайна телефонных систем связи.

В данном кратком обзоре не будем рассматривать однолинейные системы, в которых обслуживание запросов производится единственным прибором, а будем рассматривать только многолинейные системы. Поскольку соответствующая литература, включающая книги и статьи в математических, инженерных и телекоммуникационных журналах, практически необозрима, мы остановимся только на так называемых ненадежных многолинейных СМО. СМО называются ненадежными, если их приборы могут время от времени выходить из строя и

требовать восстановления (ремонта), только после которого они могут возобновить обслуживание запросов как новые.

Одной из первых работ, посвященных изучению ненадежных многолинейных СМО, можно назвать работу [1], в которой авторы рассмотрели систему обслуживания типа $M/M/N$, т. е. N -линейную систему с бесконечным буфером, в которую поступает стационарный пуассоновский поток запросов, а время обслуживания запроса в одном из идентичных приборов имеет показательное распределение. Время до выхода любого исправного прибора из строя, а также время его восстановления имеет показательное распределение. Поведение системы описано двумерной цепью Маркова, одна из компонент которой задает число запросов в системе, а другая – текущее число исправных приборов. Найдено условие эргодичности этой цепи и решена задача нахождения стационарного распределения этой цепи. Данная задача решена с помощью аппарата частичных производящих функций и использования свойства аналитичности этих функций в единичном круге комплексной плоскости.

В дальнейшем ввиду широкой применимости моделей ненадежных многолинейных СМО и их практической важности, исследования были расширены во многих направлениях. Например, популярным направлением является учет возможности того, что восстановление приборов происходит не независимо друг от друга, а имеется конечный пул ремонтников (число которых меньше числа приборов), которые осуществляют ремонт. При выходе прибора из строя, когда все ремонтники заняты, запросы на ремонт выстраиваются в специальную очередь. Как важную работу в этом направлении отметим статью [2], в которой авторы также рассмотрели систему обслуживания типа $M/M/N$. Данная задача решена с помощью матричного аппарата, впоследствии существенно развитого авторами для исследования многомерных цепей Маркова довольно общего вида. При этом решение получено в векторном виде в форме так называемого матрично-геометрического распределения. Получены алгоритмы для нахождения распределений времени ожидания в системе. Приведены результаты численных экспериментов, по результатам которых авторы предупреждают поклонников имитационного моделирования, что оно, в отличие от аналитических результатов, в применении к данной модели «не ловит» существующий эффект значительного долгосрочного увеличения очереди в системе после даже кратковременных периодов, когда число исправных приборов было малым.

Недостатком моделей, описанных в работах [1] и [2], как и подавляющего большинства более поздних работ по ненадежным многолинейным СМО, в приложении к реальным системам являются весьма жесткие предположения о показательном распределении всех времен, описывающих поведение системы (времен между моментами поступления запросов и поломок, времен обслуживания запросов и ремонта приборов). Благодаря достижениям в исследовании многомерных цепей Маркова, в первую очередь принадлежащих М. Ньютсу ([3, 4]), был создан математический фундамент для успешного исследования систем массового обслуживания с более сложными потоками и процессами обслуживания и ремонта, а именно с групповыми марковскими потоками *BMAP* (*Batch Markovian Arrival Process*) и распределениями фазового типа.

В работе [5], опубликованной ровно через 30 лет после работы [2], результаты были распространены на модель, в которой времена обслуживания имеют распределение фазового типа. В последние годы большую активность в исследовании многолинейных ненадежных СМО проявляет исследовательская группа из Тайваня под руководством профессора Ке.

Как наиболее общую модель, в которой предполагается наличие *ВМАР*-потока запросов, *МАР*-потока поломок и распределений фазового типа времен обслуживания и восстановления, следует указать модель, исследованную в [6], в которой дополнительно предположено, что запросы, поступающие в моменты, когда все приборы заняты, совершают повторные попытки попасть на обслуживание.

Следует отметить близость моделей ненадежных СМО и моделей СМО с отрицательными запросами, например, в статье [7] рассмотрена система *ВМАР/РН/Н* с *МАР*-потоком отрицательных запросов. Прибытие отрицательного запроса имеет похожий эффект с наступлением поломки, заключающийся в прерывании обслуживания запроса. Разница в основном в том, что после наступления поломки прибор должен ремонтироваться и временно недоступен для обслуживания запросов. Поступление отрицательного запроса вызывает удаление запроса с обслуживания, но не вызывает поломку прибора. Интересной особенностью модели, изученной в [7], является то, что некоторые фазы в *РН* распределенном времени обслуживания являются «защищенными». Если отрицательная заявка пытается удалить с обслуживания запрос, находящийся на «защищенной» фазе обслуживания, эта попытка игнорируется. Отрицательная заявка уходит из системы, не вызвав удаления из системы обычной заявки.

Близким к понятиям поломок и отрицательных запросов является понятие прерывания обслуживания (service interruption). Большая обзорная статья по таким системам [8] опубликована в 2014 г. Из последних статей по системам с прерыванием обслуживания отметим статью [9], в некотором смысле обобщающую результаты из [7].

Интересным вариантом отрицательных запросов являются так называемые катастрофы (disasters). В отличие от обычного отрицательного запроса, удаляющего из системы один запрос, прибытие катастрофы вызывает удаление из системы всех находящихся в ней запросов (с поломкой всех приборов или без нее). Из работ по многолинейным системам такого вида отметим статью [10].

Следует также отметить близость ненадежных СМО и моделей СМО с отдыхами приборов (server vacations), например в книге [11]. Похожесть моделей заключается в том, что и в тех и других прибор (или группа приборов) временно становится недоступным для обслуживания запросов. Разница заключается в том, что в моделях ненадежных СМО в основном предполагается, что поломки возникают в случайные моменты времени, не связанные с процессами обслуживания и поступления запросов, а в моделях СМО с отдыхами приборов предполагается временное отключение приборов, предпринимаемое намеренно. Заметим, что в последние годы появилось много работ, в основном китайских и индийских исследователей, по так называемым системам с рабочими отдыхами (working vacations), во время которых допускается обслуживание запросов (возможно, в ограниченном объеме или с уменьшенной скоростью).

Более общим видом СМО, чем ненадежные СМО, являются так называемые СМО, функционирующие в случайной среде. У таких систем предполагается, что параметры системы зависят от состояний некоторого марковского случайного процесса, называемого случайной средой. При изменении случайной среды параметры системы мгновенно изменяются. Для информации об исследовании систем обслуживания, функционирующих в случайной среде, – работы [12–14].

Кратко упомянем о наличии работ [15–18] по ненадежным многолинейным СМО, в которых с целью повышения вероятности успешного обслуживания запроса,

при наличии нескольких свободных приборов в момент поступления запроса он копируется и параллельно обслуживается несколькими приборами.

В завершение данного краткого обзора отметим еще две разновидности работ по ненадежным многолинейным СМО. Первый из них связан с использованием ТМО для анализа так называемых систем когнитивного радио. В математическом плане работа таких систем упрощенно может быть описана следующим образом. В многолинейную СМО поступают два потока запросов. Один поток – основной, имеющий абсолютный приоритет. Запросы второго типа (когнитивные запросы) обслуживаются в системе только при наличии свободных приборов. Приход запроса первого типа, не нашедшего свободного прибора, приводит к удалению с обслуживания запроса второго типа. Более подробно актуальность исследования таких моделей и состояние исследуемого вопроса описаны в недавних работах [19, 20].

Другая разновидность работ по ненадежным многолинейным СМО предполагает распределенное во времени групповое поступление поломок (поступление поломок в составе так называемых атак). Атаки поступают в соответствии с *МАР*-поток. Поступление атаки влечет выход из строя одного из занятых приборов (с потерей обслуживаемого запроса) и запуск механизма дальнейшего поступления разнородных поломок в составе данной атаки. Разнородные поломки поступают в составе атаки в соответствии с *МАР*-поток с поглощающим состоянием и время восстановления прибора после поломки имеет показательное распределение с параметром, зависящим от типа поломки. В работе [21] поставлена и решена задача исследования стандартных вероятностно-временных характеристик системы (длина очереди, время ожидания и т. д.) и относительно новой характеристики системы, которая по-русски может быть названа жизнеспособностью (существующее английское название этой характеристики – *survivability* – дословно – способность выживать).

БИБЛИОГРАФИЧЕСКИЕ ССЫЛКИ

1. Mitrani I. L., Avi-Itzhak B. A many-server queue with server interruptions // *Operations Research*. 1968. Vol. 16, № 3. P. 628–638.
2. Neuts M. F., Lucantoni D. A Markovian queue with N servers subject to breakdowns and repair // *Management Science*. 1979. Vol. 25, № 9. P. 849–861.
3. Neuts M. F. *Matrix-geometric solutions in stochastic models*. The Johns Hopkins University Press : Baltimore, 1981.
4. Neuts M. F. *Structured stochastic matrices of $M/G/1$ type and their applications*. Marcel Dekker : N. Y., 1989.
5. Yang X., Alfa A. F. A class of multi-server queueing systems with server failures // *Computers and Industrial Engineering*. 2009. Vol. 56. P. 33–43.
6. Klimenok V. I., Orlovsky D. S., Kim C. S. The BMAP/PH/N retrial queue with Markovian flow of breakdowns // *European J. of Operational Research*. 2008. Vol. 189, № 3. P. 1057–1072.
7. Klimenok V. I., Dudin A. N. A BMAP/PH/N Queue With Negative Customers and Partial Protection of Service // *Communications in Statistics – Simulation and Computation*. 2012. Vol. 41, № 7. P. 1062–1082.
8. Krishnamoorthy A., Pramod P. K., Chakravarthy S. R. Queues with interruptions: a survey // *Top*. 2014. Vol. 22. P. 290–320.
9. Dudin A., Jacob V., Krishnamoorthy A. A Multi-Server Queueing System with Service Interruption, Partial Protection and Repetition of Service // *Annals of Operations Research*. 2015. Vol. 233, № 1. P. 101–121.

10. Семенова О. В., Дудин А. Н. Система массового обслуживания М/М/Н с управляемым режимом обслуживания и катастрофическими сбоями // Автоматика и вычислит. техника. 2007. № 6. С. 72–80.
11. Tian N., Zhang Z. G. Vacation Queueing Models - Theory and Applications. Springer : Heidelberg. 2006.
12. Erlang loss queueing system with batch arrivals operating in a random environment / Ch. Kim [et al.] // Computers & Operations Research. 2009. Vol. 36, № 3. P. 674–697.
13. The BAMP/PH/N retrial queueing system operating in Markovian random environment / Ch. Kim [et al.] // Computers & Operations Research. 2010. Vol. 37, № 7. P. 1228–1237.
14. Analysis of MMAP/PH₁,PH₂/N/∞ queueing system operating in a random environment / Ch. Kim [et al.] // International J. Applied Mathematics and Computer Science. 2014. Vol. 24, № 3. P. 485–502.
15. Kim Y., Richter R., Wolff R. Job replication on multiserver systems // Advances in Applied Probability. 2009. Vol. 41. P. 546–575.
16. Дудин А. Н., Сунь Б. Многолинейная система MAP/PH/N с управляемым ширококвещательным обслуживанием ненадежными приборами // Автоматика и вычислит. техника. 2009. Т. 43, № 5. С. 32–43.
17. Дудин А. Н., Сунь Б. Многолинейная ненадежная система с управляемым ширококвещательным обслуживанием // Автоматика и телемеханика. 2009. Т. 70, № 12. С. 147–160.
18. Дудин А. Н., Сунь Б. Исследование многолинейной системы MAP/PH/N с ненадежными приборами и различными дисциплинами обслуживания // Информатика. 2010. № 2. С. 82–93.
19. Analysis of multiserver queueing system with opportunistic occupation and reservation of servers / Sun B. [et al.] // Mathematical Problems in Engineering. 2014. ID 178108. P. 1–13.
20. MAP+MAP/M₂/N/∞ Queueing System with Absolute Priority and Reservation of Servers servers / Sun B. [et al.] // Mathematical Problems in Engineering. 2014.
21. Generalised Survivability Analysis of Systems with Propagated Failures / K. Al-Begain [et al.] // Computers and Mathematics with Applications. 2012. Vol. 64, № 12. P. 3777–3791.