Федеральное государственное бюджетное образовательное учреждение высшего образования

«Омский государственный технический университет»

На правах рукописи

Захаренкова Татьяна Романовна

АНАЛИТИКО-СТАТИСТИЧЕСКИЕ МЕТОДЫ РАСЧЕТА И ОПТИМИЗАЦИИ СИСТЕМ И СЕТЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ СО СТЕПЕННЫМИ ХВОСТАМИ РАСПРЕДЕЛЕНИЙ

05.13.18 — Математическое моделирование, численные методы и комплексы программ

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата технических наук

Научный руководитель
доктор технических наук, доцент
Задорожный Владимир Николаевич

Оглавление

Оглавление 2
Введение
Глава 1 Системы массового обслуживания со степенными распределениями и современные сети передачи данных
1.1 Современный сетевой трафик
1.2 Самоподобие, долговременная зависимость и распределения с тяжелыми
хвостами
1.2.1 Понятия фрактала и фрактальной случайной величины
1.2.2 Распределение Парето как частный случай распределений с тяжелыми хвостами
1.2.3 Самоподобие и долговременные зависимости
1.3 Системы массового обслуживания со степенными хвостами распределений и существующие методы исследования
Выводы
Глава 2 Разработка методов планирования и организации имитационных экспериментов при моделировании систем с РТХ
2.1 Особенности реализации РТХ в имитационных экспериментах
2.1.1 Генераторы случайных чисел
2.1.2 Реализация случайных величин с РТХ
2.1.3 Смещение моментов при реализации РТХ в имитационном моделировании
2.1.4 Реализация РТХ в имитационном моделировании
2.1.5 Метод ARAND
2.2 Организация последовательных и параллельных прогонов при моделировании СМО с РТХ
2.3 Доверительные интервалы при моделировании СМО со степенными хвостами распределений

2.3.1 Построение доверительных интервалов в зависимых испытаниях	54
2.3.2 Доверительные интервалы в расчетах классических очередей	56
2.3.3 Доверительные интервалы в расчете фрактальных очередей	59
Выводы	62
Глава 3 Методы расчета и уменьшения вероятности потерь заявок в сист	емах со
степенными хвостами	63
3.1 Расчет вероятностей потерь во фрактальных системах ускоренным м	
3.1.1 Ускоренный метод расчета вероятности потерь	
3.1.2 Проверка точности ускоренного метода	66
3.1.3 Расчет фрактальных систем М/Ра/n/m	68
3.1.4 Влияние параметра формы α	71
3.1.5 Другие фрактальные системы.	72
3.1.6 Предварительная оценка точности и коэффициента ускорения мет моделировании фрактальных систем.	•
3.2 Метод уменьшения вероятностей потерь пакетов в сетях с фракт трафиком	
3.2.1 Классические бесконечнолинейные СМО	76
3.2.2 Фрактальные бесконечнолинейные СМО	80
3.2.3 Фрактальные СеМО с многоканальными узлами	84
3.2.4 Задача и метод оптимального распределения каналов	86
3.3 Аппроксимация вероятности потерь P в системах $GI/GI/n/0$ вероятностояний p_k и хвостом $P(k \ge n)$ систем $GI/GI/\infty$	
3.3.1 Аппроксимация вероятности потерь P вероятностями состояний p	_k 89
3.3.2 Аппроксимация вероятности потерь P хвостом $P(k \ge n)$	92
3.3.3 Погрешность аппроксимации потерь P вероятностями состоян хвостом $P(k \ge n)$	
Выводы	96
Глава 4 Использование абсолютных приоритетов с дообслуживанием	98
4.1 Формирование приоритетных классов	99

4.2 Регулярная разметка оси трудоемкостей	. 101
4.3 Оптимизация шага регулярной разметки	. 104
4.3.1 Вывод расчетной формулы	. 105
4.3.2 Упрощение расчетной формулы с контролем погрешностей	. 106
4.3.3 Пример решения задачи оптимизации РР	. 108
4.4 Расчет $W(\Delta)$ при $\Delta \rightarrow 0$. 110
4.5 Введение приоритетов при $\alpha > 2$. 111
4.6 Экспоненциальная разметка	. 113
4.7 Эксперименты с вероятностями потерь при конечных размерах буферов.	. 114
4.7.1 Основная задача, решаемая введением абсолютных приоритетов	. 114
4.7.2 Эксперименты с системами M/Pa/1/ m при $1 < \alpha \le 2$. 115
4.7.3 Другие системы с РТХ	. 119
Выводы	. 120
Глава 5 Комплекс программ для исследования СМО и CeMO c PTX	. 122
5. 1 Реализация метода ARAND	. 122
5.2 Организация последовательных и параллельных прогонов	. 123
5.3 Программная реализация имитационных экспериментов для исследова	ания
вероятностей потерь	
Выводы	. 131
Заключение	. 132
Список сокращений и условных обозначений	. 135
Список использованных источников	. 136
ПРИЛОЖЕНИЕ А	. 148
ПРИЛОЖЕНИЕ Б	. 151
ПРИЛОЖЕНИЕ В	. 155
ПРИЛОЖЕНИЕ Г	157

Введение

Актуальность работы и степень разработанности темы исследования

Три десятилетия назад было обнаружено, что сетевой трафик — объём информации, передаваемой через компьютерную сеть за определённые периоды времени — обладает в современных сетях передачи данных свойствами «самоподобия» (масштабной инвариантности) т.е. является фрактальным [100]. Данный факт заставил ученых и специалистов в области сетей передачи данных поставить под сомнение пригодность широко используемых на тот момент моделей с пуассоновскими потоками и формул Эрланга, лежащих в основе проектирования сетей передачи данных. Фрактальный трафик характеризуется сильными пульсациями, т.е. является несглаживаемым на любых масштабах времени. Также не сглаживается он ни путем суммирования любого числа независимых потоков данных, ни путем случайного их просеивания. При таком трафике за короткие промежутки времени иногда приходит «катастрофически» большое количество пакетов, из-за чего при ограниченных размерах буферов в узлах сетей пакеты теряются с относительно высокой вероятностью [100].

Среди зарубежных ученных, занимающихся тематикой фрактального трафика, следует выделить У. Лелланда, У. Уиллингера, Д. Уилсона, М. Кровеллу, М. Такка, К. Парка, В. Паксона, С. Флойд и др., а среди отечественных ученых – О. И. Шелухина [44], В. И. Неймана, Б. С. Цыбакова [59].

Адекватными моделями сетевых устройств, функционирующих в условиях фрактального трафика, являются системы массового обслуживания (СМО) с асимптотически степенными хвостами распределений, задающих длительности интервалов поступления заявок и/или время их обслуживания. Распределения с асимптотически степенными хвостами, далее называемые степенными, относятся к распределениям с тяжелыми хвостами (РТХ). Степенные распределения — это единственный вид распределений, обладающих свойством масштабной инвариантности [40, 98].

Снижению вероятности потерь в сетях с фрактальным трафиком посвящено большое количество работ. Предлагаемые в них пути решения данной проблемы можно разделить на четыре основные направления, которые в терминах теории массового обслуживания определяются следующим образом:

- увеличение размеров буферных емкостей,
- повышение скорости (снижение коэффициентов загрузки) каналов,
- наращивание числа параллельно работающих каналов,
- введение дисциплин приоритетного обслуживания.

Увеличение размеров буферов является неэффективной стратегией [80] в силу того, что в условиях фрактального трафика вероятность потерь с ростом размера буфера уменьшается по степенному закону, а не по экспоненциальному, как в классической теории телетрафика [93, 101]. Кроме того, увеличение размеров буферов в узлах сетей, не решая проблемы потерь, приводит к другой проблеме – к чрезмерным задержкам пакетов и заторам.

Возможность снижения потерь пакетов путем уменьшения коэффициентов загрузки каналов рассматривается, например, в [79], но в итоге и этот подход признается неэффективным.

Более перспективными представляются подходы, связанные с наращиванием числа параллельно работающих каналов или с введением дисциплин приоритетного обслуживания [107]. Однако, как показал широкий обзор работ, опубликованных за последние 30 лет, систематические исследования этих двух направлений борьбы с потерями отсутствуют. Не найдены работы, в которых такие исследования завершались бы разработкой математических моделей и высокоэффективных методов, широко применяемых в инженерной практике проектирования сетей передачи данных.

Таким образом, проблема снижения вероятностей потерь в сетях с фрактальным трафиком является актуальной проблемой, эффективное решение которой позволило бы значительно повысить качество информационного обслуживания пользователей сетей и существенно снизить затраты на борьбу с потерями.

Основным инструментом исследования СМО и сетей массового обслуживания (СеМО), описываемых посредством РТХ, в том числе посредством степенных (со степенными хвостами) распределений, в силу трудности получения аналитических решений является имитационное моделирование (ИМ). Параметрам фрактального трафика современных сетей передачи данных, как правило, соответствуют такие степенные распределения, которые имеют конечное математическое ожидание и бесконечную дисперсию [100]. Применение ИМ для исследования подобных СМО и СеМО тоже связано с рядом трудностей, к которым относятся:

- смещение моментов РТХ при их реализации в ИМ [105];
- низкая скорость сходимости оценок, обусловленная бесконечной дисперсией элементов выборок и/или долговременными зависимостями (ДВЗ) между элементами выборки, получаемой в одном прогоне имитационной модели;
 - длительные переходные процессы, также обусловленные ДВЗ [76];
- необходимость оценивания и сравнения малых вероятностей при разработке методов их снижения [47, 48];
- проблематичность вычисления в ИМ градиентов при решении задач оптимизации [48, 17].

Методам решения перечисленных проблем применения ИМ для исследования СМО и СеМО с РТХ посвящено очень мало работ, хотя такое применение ИМ играет большую роль в инженерных приложениях [101]. В этой связи разработка методов, позволяющих решать перечисленные проблемы и корректно применять ИМ для исследования СМО и СеМО с РТХ, также является актуальной задачей, имеющей самостоятельное значение. В данном диссертационном исследовании эту задачу необходимо решить прежде, чем ИМ будет использовано для разработки методов снижения вероятностей потерь в системах и сетях с РТХ, содержащих конечные буферы.

Цель и задачи

Целью работы является решение проблемы больших вероятностей потерь в современных сетях передачи данных с фрактальным трафиком.

Системы и сети массового обслуживания с длительностями интервалов поступления и/или обслуживания заявок, описываемыми степенными распределениями с конечным математическим ожиданием и бесконечной дисперсией наиболее точно отражают особенности функционирования сетевых устройств в условиях фрактального трафика и могут широко применяться в практике проектирования сетей передачи данных.

Для достижения поставленной цели в диссертации сформулированы и решены следующие задачи.

- 1. Исследование проблем, возникающих при использовании существующих генераторах стандартных псевдослучайных чисел (ГСПЧ) для ИМ СМО и СеМО со степенными распределениями: проблемы смещения моментов РТХ, проблемы длины периода ГСПЧ, недостаточной для исследования случайных процессов с долговременными зависимостями, и разработка специальных ГСПЧ, удовлетворяющих повышенным требованиям к ним при моделировании таких СМО и СеМО.
- 2. Вывод выражений для определения доверительных интервалов по выборкам с долговременными зависимостями.
- 3. Разработка ускоренного регенеративного метода для расчета зависимостей вероятности потерь заявок от размера буфера в системах со степенными распределениями и бесконечной дисперсией времени обслуживания.
- 4. Разработка ускоренного аналитико-имитационного метода для снижения вероятности потерь заявок в CeMO со степенными распределениями за счет оптимального распределения каналов по узлам сети.
- 5. Исследование влияния приоритетных дисциплин на характеристики очередей в системах со степенными распределениями и разработка метода снижения вероятностей потерь, основанного на данном исследовании.
- 6. Разработка программного комплекса, обеспечивающего возможность корректного высокоточного моделирования систем с очередями и реализующего разработанные в диссертации методы.

Объект и предмет диссертационного исследования

Объектом диссертационного исследования являются реальные системы с очередями, функционирующие в информационных системах, финансовых учреждениях, службах преодоления чрезвычайных ситуаций и т.д., ярким примером которых являются современные сети передачи данных, обозначенные Международным союзом электросвязи (ITU), как сети следующего поколения (NGN), подробное определение которых приведено в рекомендациях серии Y.2000 «Глобальная информационная структура, аспекты протокола Интернет и сети следующего поколения». Проблемы, возникающие в сетях с фрактальным трафиком, структурируются вокруг очередей, поэтому формальное описание NGN сетей логично формулировать в терминах теории массового обслуживания. Соответственно, предметом диссертационного исследования являются системы и сети массового обслуживания (СеМО) со степенными хвостами распределений, имеющих конечное математическое ожидание и бесконечную дисперсию.

Научная новизна результатов, представленных в диссертации, состоит в следующем:

- 1. Получены выражения доверительных интервалов для математических ожиданий, оцениваемых по выборкам с долговременными зависимостями элементов, позволяющие определять точность результатов моделирования СМО и СеМО со степенными распределениями и отличающиеся от известных выражений отсутствием необходимости предварительного разбиения выборки на большие слабо зависимые группы элементов. Установлено, что с ростом объема таких выборок доверительные интервалы сокращаются со степенной скоростью с показателем степени, абсолютная величина которого лежит в интервале (0, 1/2) и может быть при тяжелых хвостах распределений сколь угодно близка к нулю.
- 2. Разработан ускоренный метод регенеративного моделирования СМО со степенными распределениями, отличающийся возможностью за один прогон модели получать зависимость вероятности потерь от размера буфера на произвольной длины отрезке возможных размеров буфера, и с помощью этого метода найдены виды зависимостей вероятности потерь от размера буфера для различных классов СМО со степенными распределениями.

- 3. Разработан оригинальный ускоренный аналитико-имитационный метод оптимального распределения каналов по узлам CeMO со степенными распределениями, обеспечивающий возможность кардинального снижения вероятности потерь в сетях с фрактальным трафиком за счет эффективного использования ограниченной аппаратной избыточности.
- 4. Предложен и исследован оригинальный метод кардинального снижения вероятностей потерь в СМО с бесконечной дисперсией времени обслуживания за счет введения абсолютных приоритетов с бесконечным числом приоритетных классов, определяемых специальными разметками диапазона возможных значений времени обслуживания. Метод отличается тем, что при бесконечном буфере в таких системах, обусловливающем бесконечное стационарное среднее время ожидания, применение предлагаемого способа назначения приоритетов делает среднее время ожидания конечным.
- 5. Разработан эффективный численный метод оптимизации специальных разметок (регулярной и экспоненциальной) по критерию минимума среднего времени ожидания, обеспечивающий возможность широкого применения метода абсолютных приоритетов со специальными бесконечными разметками в проектировании современных компьютерных сетей.

Методы исследования

В диссертационной работе используются методы теории вероятностей, теории массового обслуживания, математического программирования, вычислительной математики, методы имитационного и аналитико-имитационного моделирования.

Теоретическая и практическая значимость

Теоретическая значимость результатов диссертационной работы заключается в том, что они вносят вклад в развитие актуальных методов и моделей теории массового обслуживания, применяемых в проектировании современных компьютерных сетей. Модель СМО с абсолютными приоритетами обобщена на случай применения распределений времени обслуживания с бесконечной дисперсией и использования бесконечного числа приоритетных классов. Метод ведения абсолютных приоритетов с бесконечным числом приоритетных классов, определяемых специальными раз-

метками, исключительно за счет эффективного изменения порядка обслуживания заявок позволяет в системах с бесконечной дисперсией времени обслуживания, в случае бесконечного буфера уменьшить стационарное среднее время ожидания с бесконечного до конечного, а в случае конечного буфера – увеличить скорость убывания вероятности потерь с ростом размера буфера со степенной до экспоненциальной. Разработанный численный метод оптимизации специальных разметок (применяемых для определения приоритетных классов) по критерию минимума среднего времени ожидания позволяет минимизировать как задержки заявок в очередях сетей, так и вероятности потерь заявок. Методы аналитико-имитационного моделирования, включая метод ускоренного регенеративного моделирования СМО с потерями заявок при бесконечной дисперсии времени обслуживания и метод оптимизации распределения каналов по узлам сети с экспоненциальными распределениями, позволяют выполнять исследование и оптимизацию широкого класса СМО и СеМО, применяемых при решении задач проектирования компьютерных сетей и сетевых устройств. Выражения, полученные для построения доверительных интервалов по выборкам с долговременными зависимостями, а также разработанные специальные генераторы псевдослучайных чисел позволяют осуществлять высокоточное имитационное моделирование СМО и СеМО рассматриваемого класса на соответствующих этапах применения разработанных аналитико-имитационных методов.

Результаты диссертационной работы могут быть использованы непосредственно в практике проектирования современных компьютерных сетей и сетевых устройств. Предложенные в диссертации методы и модели позволяют повысить качество информационного обслуживания в таких сетях путем существенного снижения вероятностей потерь передаваемых данных, и, одновременно, задержек передаваемых данных в очередях. Разработанный программный комплекс позволяет эффективно решать задачи исследования и оптимизации сетей и сетевых устройств, функционирующих в условиях фрактального трафика, с использованием доступных систем моделирования и математических пакетов.

Разработанные методы и средства могут применяться и для решения задач в других областях, в которых адекватные модели случайных величин требуют применения распределений с тяжелыми хвостами, в том числе в финансовой математике и в теории катастроф.

Достоверность полученных результатов подтверждается математически корректными выводами и доказательствами положений, представленных в работе, согласованностью полученных аналитических результатов с результатами, ранее полученными для соответствующих частных случаев, вычислительными экспериментами с использованием программных реализаций предложенных моделей и методов.

Положения, выносимые на защиту

- 1. Выражения для определения доверительных интервалов по выборкам с долговременными зависимостями, получаемым при моделировании СМО и СеМО со степенными распределениями.
- 2. Ускоренный регенеративный метод расчета зависимости вероятности потерь от размера буфера для СМО со степенными распределениями и бесконечной дисперсией времени обслуживания.
- 3. Теорема о снижении бесконечного стационарного среднего времени ожидания в системе M/Pa/1 с бесконечной дисперсией времени обслуживания до конечной величины при введении абсолютных приоритетов, определяемых регулярной разметкой с конечным шагом.
- 4. Методы кардинального снижения вероятностей потерь в системах и сетях с бесконечными дисперсиями времени обслуживания:
- ускоренный аналитико-имитационный метод оптимального распределения каналов по узлам CeMO по критерию минимума вероятности потерь;
- метод введения абсолютных приоритетов с бесконечным числом приоритетных классов, определяемых специальными бесконечными разметками диапазона возможных значений времени обслуживания.
- 5. Численный метод оптимизации специальных бесконечных разметок, используемых при введении абсолютных приоритетов с бесконечным числом приоритетных классов, по критерию минимума среднего времени ожидания.

Личное участие автора в получении результатов, изложенных в диссертации. Постановка задач моделирования компьютерных сетей и сетевых устройств, функционирующих в условиях фрактального трафика, принадлежит научному руководителю, д. т. н., доценту В. Н. Задорожному, который принимал участие и в создании комплекса программ и алгоритмов, реализующих полученные

в диссертации модели и методы. Автор лично участвовала в получении этих моделей и методов, а именно, в разработке специальных генераторов псевдослучайных чисел, удовлетворяющих повышенным к ним требованиям при моделировании СМО и степенными распределениями, в выводе формул для доверительных интервалов по выборкам с долговременными зависимостями, в CMO разработке **ускоренного** метода регенеративного моделирования рассматриваемых классов, в разработке метода оптимального по критерию минимума вероятности потерь распределения каналов по узлам СеМО, в разработке метода абсолютных приоритетов со специальными бесконечными разметками, в частности, разработке расширенной модели СМО с абсолютными приоритетами, выводе описывающих ее аналитических соотношений, получении соответствующих эффективных расчетных формул, формулировке и доказательстве теоремы о преобразовании очередей с бесконечным средним временем ожидания в очереди с конечным средним временем ожидания за счет применения специальных бесконечных разработке разметок. Α также В программного реализующего все полученные в диссертации модели и методы. В совместных публикациях д. т. н., доценту В. Н. Задорожному, принадлежат постановки задач и формулировка основных направлений исследований. Вклад других соавторов состоит в обсуждении практических аспектов распространения и внедрения рассматриваемых методов и в участии в разработке программного комплекса для анализа и оптимизации СМО и СеМО с экспоненциальными распределениями.

Апробация результатов исследования

Основные положения и результаты диссертации докладывались и обсуждались на следующих научных конференциях: XV Международная конференция им. А.Ф. Терпугова «Информационные технологии и математическое моделирование», Катунь, 2016; VI Международная научно-техническая конференция «Динамика систем, механизмов и машин», Омск, 2017; VI Международная конференция «Математика, ее приложения и математическое образование», Улан-Удэ, 2017; XVI Международная конференция им. А.Ф. Терпугова «Информационные технологии и математическое моделирование», Казань, 2017; XVII Международная конференция

им. А.Ф. Терпугова «Информационные технологии и математическое моделирование», Томск, 2018; VII Международная научно-техническая конференция «Динамика систем, механизмов и машин», Омск, 2018; І Всероссийская конференция «Системы управления, информационные технологии и математическое моделирование», Омск, 2019; XVIII Международная конференция им. А.Ф. Терпугова «Информационные технологии и математическое моделирование», Саратов, 2019.

Публикации по теме исследования

Основные результаты диссертационного исследования изложены в 20 работах, среди которых 4 статьи в журналах, включенных в «Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук», 5 статей, проиндексированных международной системой цитирования Scopus, 2 свидетельства о регистрации электронных ресурсов, 9 работ, опубликованных в сборниках материалов конференций.

Связь работы с научными проектами

В основу диссертационной работы положены результаты научных исследований, выполненных в Омском государственном техническом университете по теме «Разработка методов, алгоритмов и программных средств для снижения вероятности потерь в сетях передачи данных с фрактальным трафиком», включенной в НИР ОмГТУ № 17307В (приказ №896/1 от 01.11.2017г.). В том числе работа была выполнена в рамках научного проекта на тему «Разработка и исследование ускоренных аналитико-имитационных методов расчета и оптимизации очередей в телекоммуникационных сетях с фрактальным трафиком» с регистрационным номером АААА-А16-116081910053-3.

Результаты работы используются в учебном процессе на факультете информационных технологий и компьютерных систем (ФИТиКС) Омского государственного технического университета при проведении практических и лабораторных работ для студентов, обучающихся по направлению 09.03.03 «Прикладная информатика» (см. Приложение Г).

Структура и объем диссертации. Диссертация состоит из введения, пяти глав, заключения, списка литературы и трех приложений. Объем диссертации с приложениями составляет 158 страниц, без приложений – 147 страниц. Диссертация содержит 8 таблиц и 43 рисунка. Список литературы включает 119 наименований.

В первой главе приводятся основные причины широкого распространения систем и сетей массового обслуживания со степенными хвостами распределений как моделей современных сетей передачи данных и сетевых узлов. Представлены основные теоретические сведения по теме диссертационного исследования, в том числе существующие методы исследования СМО и СеМО со степенными хвостами распределений. Приводится обзор программных средств имитационного моделирования.

Во второй главе рассматривается проблема корректной реализации распределений с тяжелыми хвостами (в разделе 2.1). Исследуются проблемы организации и особенности планирования имитационных экспериментов с рассматриваемыми СМО и СеМО (разделы 2.2 и 2.3).

В третьей главе разрабатывается ускоренный метод расчета вероятности потерь в системах с очередями, который используется для анализа систем со степенными хвостами распределений, позволяющий определять виды зависимостей вероятности потерь от размера буфера (раздел 3.1). Разрабатывается эффективный метод обеспечения низкой вероятности потерь сообщений в СеМО со степенными хвостами распределений за счет оптимального распределения каналов по узлам сетей (раздел 3.2), основанный на аппроксимациях вероятностей потерь конечнолинейных СМО стационарными вероятностями состояний и хвостами распределений вероятностей состояний соответствующих бесконечнолинейных СМО (раздел 3.3).

В четвертой главе разрабатывается метод обеспечения низкой вероятности потерь заявок в СМО с бесконечной дисперсией времени обслуживания, основанный на введении дисциплины абсолютных приоритетов с дообслуживанием и разделении входящих заявок на бесконечное число приоритетных классов.

В пятой главе приводится описание комплекса программ, предназначенного для исследования рассматриваемых СМО и СеМО, реализующего разработанные в диссертационной работе методы расчета и оптимизации систем и сетей с очередями со степенными распределениями.

Глава 1 Системы массового обслуживания со степенными распределениями и современные сети передачи данных

Теория массового обслуживания (ТМО) зародилась ещё в первой половине XX века и предназначалась для описания телефонных систем и сетей, а её методы применялись при решении задач в области телефонии. В настоящее время методы ТМО применяются для исследования различного рода реальных систем, в первую очередь производственных, дорожных и компьютерных. Стохастический характер процесса передачи данных в сетях передачи данных (СПД) и наличие буферов для хранения блоков данных сделал ТМО удобным инструментом исследования СПД. Методы ТМО занимают одно из центральных мест в исследованиях СПД на системном уровне [1, 5, 18, 30, 43, 84, 119], так как практически каждый узел сети может быть представлен соответствующей СМО.

В результате выявления фрактальной природы сетевого трафика, характеризуемой долговременными зависимостями и РТХ [100], стало понятно, что традиционные модели и методы ТМО стали непригодны для исследования и проектирования сетей с фрактальным трафиком. Это привело к интенсивной разработке новых математических моделей и методов [119], которые позволяли бы адекватно учитывать особые свойства фрактального трафика.

1.1 Современный сетевой трафик

Первые компьютерные сети появились еще в начале 60-х годов прошлого века. Технология коммутации пакетов, разработанная Л. Клейнроком, П. Бэраном и Д. Дэвисом, послужила основой для создания первой компьютерной сети с коммутацией пакетов – ARPAnet, которая стала прообразом Интернета. В 70-х годах начался бурный рост числа локальных сетей, таких как ALOHANet, Cyclades и Telnet. Рост сетей сопровождался появлением протоколов TCP, IP, UDP, Ethernet и др., обеспечивающих взаимодействие отдельных сетей, а также разработкой новых стандартов и рекомендаций для СПД [37, 43, 55]. Возникшая впоследствии глобальная компьютерная сеть, объединяющая локальные сети превратилась в сеть

сетей — Интернет [37]. Методы, используемые для анализа трафика и проектирования компьютерных сетей, опирались на известные в то время положения ТМО, установленные в работах А. К. Эрланга, Т. Энгсета, А. Я. Хинчина, К. Пальма, К. Якобеуса.

В начале 90-х годов произошло событие, которое оказало значительное влияние на рост и развитие Интернета — появление Всемирной паутины (WWW — World Wide Web). Стремительное разрастание WWW сопровождалось экспоненциальным ростом числа веб-серверов, хостов (конечных узлов) [55] и объема трафика [43]. В это же время группа ученых — У. Лелланд, М. Такк, У. Уиллингер и Д. Уилсон по результатам исследования трафика LAN сетей в корпорации Bellcore Center, построенных по стандарту Ethernet, в докладе [100] заключили, что поведение агрегированных данных сетевого трафика обладает статистическим самоподобием. Именно после публикации данного доклада в теории сетей появился новый термин — «фрактальный трафик» [45]. Статистические характеристики фрактального трафика являются неизменными при варьировании масштаба времени, и для него характерны пульсации на больших масштабах времени (см. Рисунок 1.1, заимствованный из [18]).

Ученые пришли к выводу, что при исследовании фрактального трафика потоки данных сети нельзя аппроксимировать пуассоновскими потоками [100, 103] и, следовательно, модели фрактального трафика должны существенно отличаться от классических моделей. Тем самым были установлены причины ряда инженерных неудач, случившихся при проектировании СПД, когда расчеты объемов буферов в узлах, значений вероятностей потерь пакетов и времени их задержек всё ещё осуществлялись с помощью классических методов ТМО [45], а сетевой трафик приобрёл фрактальные свойства. Перечисленные события привели к разработке новых математических методов для описания и анализа трафика СПД [100] и появлению большого количества соответствующих работ, посвященных изучению и анализу фрактального трафика [1, 4, 11, 18, 28, 36, 38, 39, 44, 45, 56, 57, 59, 61, 62, 64, 65, 72, 74-77, 79, 80-83, 91, 93, 99-101, 103 и др.].

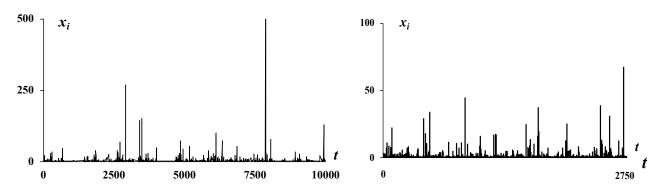


Рисунок 1.1 — Пульсирующая структура трафика. Время x_i (вертикальные линии) обслуживания поступающих заявок распределено по закону Парето с параметрами $\alpha = 1,5, K = 1$. Расстояние между вертикальными линиями вдоль оси t соответствует времени между приходами заявок, распределенному по закону Парето с параметрами $\alpha = 1,2, K = 1$. Справа на рисунке представлена в укрупненном масштабе начальная часть процесса, изображенного слева

Следует заметить, что основополагающая работа «On Self-Similar Nature of Ethernet Traffic» [100] в 1994 году получила награду У. Беннетта в области сетей связи. Среди западных ученных, занимающихся данной тематикой помимо выше-упомянутых, следует выделить М. Кровеллу [75, 77], К. Парка [101], В. Паксона [103], С. Флойд и др., а среди отечественных ученых – О. И. Шелухина [44, 61, 62], В. И. Неймана, Б. С. Цыбакова [59, 93] и др.

Для описания сетевого трафика в настоящее время используются разнообразные модели, включая следующие:

- потоки заявок, время обслуживания которых распределено по закону Парето с бесконечной дисперсией [93, 97, 103];
- потоки заявок, образуемые суммированием большого числа потоков, генерируемых ON/OFF источниками, ON периоды которых имеют PTX с бесконечной дисперсией [109];
- временные ряды (например, авторегрессионные модели ARIMA и FARIMA) [100];
 - хаотические отображения [81];
 - фрактальное дробное броуновское движение (ФБД) [99, 100];
 - мультифракталы [44] и др.

Основные проблемы в СПД, обусловленные самоподобной структурой трафика, концентрируются вокруг очередей. Это делает СеМО и СМО адекватными моделями сетей и сетевых устройств, а ТМО – эффективным инструментом исследования сетей с фрактальным трафиком. Прежде чем переходить к исследованию сетей с фрактальным трафиком в терминах ТМО, рассмотрим понятия и определения, необходимые для корректного обсуждения фрактальной природы сетевого трафика.

1.2 Самоподобие, долговременная зависимость и распределения с тяжелыми хвостами

1.2.1 Понятия фрактала и фрактальной случайной величины

Родоначальником теории фракталов считается франко-американский математик Бенуа Мандельброт, который в своем математическом эссе [40] ввел термин «фрактал». На интуитивном уровне под фракталами понимаются такие объекты, которые в результате их масштабных преобразований не изменяются («масштабно-инвариантные объекты», объекты, у которых «части подобны целому», «самоподобные объекты»). Широко известными примерами таких объектов являются снежинка Коха, салфетка Серпинского (см. Рисунок 1.2) [40]. Термин «фрактал» образован от латинского слова *fractus* — ломанный. Б. Мандельброт предлагает несколько неэквивалентных определений фрактала. В одном из них фрактал (множество точек евклидова пространства) определяется как «... множество, размерность Хаусдорфа-Безиковича для которого строго больше его топологической размерности». Размерность *D* Хаусдорфа-Безиковича еще называют *дробной* или *фрактальной* размерностью, она играет ведущую роль в теории фракталов. Точные определения топологической и фрактальной размерностей даны в [40]. Фрактальная размерность геометрического фрактала определяется соотношением (1.1):

$$N = r^D, (1.1)$$

где N — количество уменьшенных в r раз копий, необходимых для заполнения исходного объекта. Например, для салфетки Серпинского, которая получается последовательным вырезанием равносторонних треугольников, фрактальная размерность равна $D = \ln(N)/\ln(r) = \ln(3)/\ln(2) = 1,58496$.

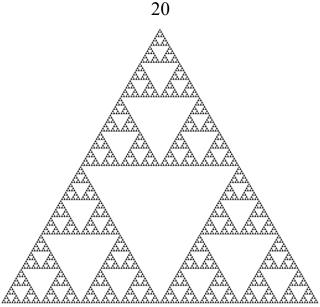


Рисунок 1.2 – Салфетка Серпинского

Мандельброт Б. разделяет фракталы на неслучайные, которые являются «детерминированными конструкциями», и случайные. Основным инструментом исследования случайных фракталов является теория вероятностей. Отличительной особенностью случайных фракталов является неизменность их вероятностных характеристик при различных масштабных преобразованиях. Современный сетевой трафик относится к случайным фракталам [44].

Особым видом случайных фракталов является масштабно-инвариантная случайная величина: «случайная величина X масштабно-инвариантна при преобразовании T(X), если распределения X и T(X) тождественны во всех отношениях, кроме масштаба. Термин "преобразование" понимается здесь в широком смысле, например, сумма двух независимых реализаций случайной величины X рассматривается как результат преобразования X. . . . при многих других преобразованиях, как выясняется, для инвариантности требуется асимптотически гиперболическое распределение» [40].

В [98] приводится доказательство того, что степенные (гиперболические) распределения являются единственным видом распределений, которые обладают свойством масштабной инвариантности. В [18] приводятся примеры, иллюстрирующие масштабную инвариантность случайной величины (с.в.), имеющей распределение Парето.

1.2.2 Распределение Парето как частный

случай распределений с тяжелыми хвостами

Распределение Парето – одно из самых распространенных степенных распределений, используемых в моделировании фрактального трафика. Оно определяется функцией распределения (ф.р.):

$$F(t) = 1 - \left(\frac{K}{t}\right)^{\alpha}, t \ge K, K > 0, \alpha > 0, \tag{1.2}$$

где K — параметр масштаба, α — параметр формы (и в то же время — фрактальная размерность; о показателе степени степенного распределения как о размерности см. [40]). Плотность вероятностей (п.в.) распределения Парето, согласно (1.2), имеет вид:

$$f(t) = F'(t) = \frac{\alpha K^{\alpha}}{t^{\alpha+1}}, \ t \ge K.$$
 (1.3)

С помощью (1.3) можно выразить первый и второй моменты с.в. X, распределенной по закону Парето $Pa(K, \alpha)$:

$$M(X) = \int_{t=K}^{\infty} t \cdot f(t) dt = \int_{t=K}^{\infty} t \cdot \frac{\alpha K^{\alpha}}{t^{\alpha+1}} dt = \begin{cases} \frac{\alpha K}{\alpha - 1}, \text{ при } \alpha > 1, \\ \infty, \text{ при } \alpha \leq 1, \end{cases}$$
 (1.4)

$$M(X) = \int_{t=K}^{\infty} t \cdot f(t) dt = \int_{t=K}^{\infty} t \cdot \frac{\alpha K^{\alpha}}{t^{\alpha+1}} dt = \begin{cases} \frac{\alpha K}{\alpha - 1}, & \text{при } \alpha > 1, \\ \infty, & \text{при } \alpha \leq 1, \end{cases}$$

$$D(X) = M(X^{2}) - M^{2}(X) = \begin{cases} \frac{\alpha K^{2}}{(\alpha - 1)^{2}(\alpha - 2)}, & \text{при } \alpha > 2, \\ \infty, & \text{при } \alpha \leq 2. \end{cases}$$

$$(1.4)$$

На практике интересен случай, когда математическое ожидание (м.о.) является конечным, а дисперсия – бесконечной, т.е. когда $1 < \alpha \le 2$.

Распределение Парето является представителем распределений с тяжелыми хвостами, которые также используются при моделировании фрактального трафика (например, распределение Вейбулла [110] с параметром формы $0 < \beta < 1$). Хвостом ф.р. F(t) называют ее дополнение $P(X > t) = \overline{F}(t) = 1 - F(t)$. Наличие общих свойств у степенных распределений и других РТХ позволяет при исследовании очередей со степенными распределениями приходить к результатам, справедливым для СМО с PTX [18].

Распределение F(t) с.в. ξ имеет тяжелый хвост, если для любой константы $\lambda > 0$ выполняется условие [18, 74, 119, и др.]

$$M[e^{\lambda\xi}] = \infty,$$

или, что эквивалентно:

$$\lim_{t \to \infty} \frac{P(\xi > t)}{e^{-\lambda t}} = \frac{\overline{F}(t)}{e^{-\lambda t}} = \infty.$$
 (1.6)

Как видно из (1.6), хвост РТХ характеризуется более медленным убыванием, чем хвост экспоненциального распределения. Тем не менее, в большинстве работ, посвященных исследованию фрактального трафика [1, 44, 45, 62, 75, 76, 100, 101], РТХ определяются как распределения, хвост которых подчиняется степенному закону следующего вида:

$$P(\xi > t) = \overline{F}(t) \sim ct^{-\alpha}, \tag{1.7}$$

где c > 0 — некоторая константа, α — параметр формы. Определение (1.7) сужает понятие РТХ, так как включает не все подклассы РТХ. Подробнее с подклассами РТХ можно ознакомиться в работах [18, 95, 119]. Вообще говоря, (1.7) является частным случаем правильно меняющихся на бесконечности функций, которые определяются как:

$$P(\xi > t) = F(t) \sim L(t) \cdot t^{-\alpha}, \qquad (1.8)$$

при $t \to \infty$, где $\alpha > 0$, а L(t) — медленно меняющаяся на бесконечности функция, определяемая свойством $\lim_{t\to\infty} \frac{L(\lambda t)}{L(t)} = 1$ для любого $\lambda > 0$. В некоторых работах формулу (1.8) также можно встретить в качестве определения РТХ (напр., см. [83]). Видно, что, согласно (1.6), определение (1.8) тоже неточное. Правильно меняющиеся функции являются подклассом РТХ. Имеется и такая точка зрения, что единого подхода к определению РТХ вообще не существует, о чем говорит автор работы [95]. В данной диссертационной работе под РТХ будем понимать распределения, обладающие свойством (1.6).

1.2.3 Самоподобие и долговременные зависимости

Известно [62, 76], что распределения с тяжелыми хвостами могут являться причиной такого свойства фрактального трафика, как долговременные зависимости (ДВЗ). Именно ДВЗ существенно влияет на процесс образования очередей в буферах сетевых устройств [100, 101]. В работе [79] отмечается, что игнорирование

наличия ДВЗ в сетях в первую очередь приводит к недооценке требуемой производительности сети и неэкономичному распределению сетевых ресурсов.

Рассмотрим стационарный случайный процесс с дискретным временем X_t , $t \in N$, математическим ожиданием $M(X) < \infty$ и дисперсией $\sigma^2 < \infty$, имеющий автокорреляционную функцию вида:

$$r(k) = M[(X_{t+k} - M(X))(X_t - M(X))]/\sigma^2 = R(k)/\sigma^2$$
,

где k — длительность интервала времени между наблюдаемыми значениями X_t и X_{t+k} . Обозначим через $X_k^{(m)}$ агрегированный, т.е. усредненный по блокам (интервалам) длины m, процесс:

$$X_k^{(m)} = \frac{1}{m} \sum_{i=(k-1)m+1}^{km} X_i, \quad k, m \in \mathbb{N}.$$
 (1.9)

Стационарный в широком смысле процесс X_t , $t \in N$ с дискретным временем является *строго самоподобным в широком смысле* с параметром Херста H тогда и только тогда, когда автокорреляционная функция данного процесса задается следующим образом:

$$r(k) = \frac{1}{2} \left[|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H} \right] \forall k \in \mathbb{N}.$$
 (1.10)

Стационарный в широком смысле процесс X_t , $t \in N$ с дискретным временем является асимптотически самоподобным в широком смысле с параметром Херста H, если

$$\lim_{m \to \infty} r^{(m)}(k) = \frac{1}{2} \left[(k+1)^{2H} - 2(k)^{2H} + (k-1)^{2H} \right],\tag{1.11}$$

где $r^{(m)}(k)$ автокорреляционная функция агрегированного процесса $X_k^{(m)}$.

Для таких процессов характерна убывающая по степенному закону автокорреляционная функция, т.е. говорят, что процесс обладает *долговременной зависимостью*. Для строго самоподобных в широком смысле процессов автокорреляционная функция при $k \to \infty$ имеет вид $r(k) \sim H(2H-1)k^{-\beta}$, для асимптотически самоподобных в широком смысле $-r(k) \sim ck^{-\beta}$, где c > 0 некоторая константа, $0 < \beta < 1$. Более подробное описание самоподобных процессов можно найти в [44, 59, 61, 62, 119]. Для выявления на практике самоподобия в сетевом трафике, как правило, достаточно знать вид корреляционной функции на больших масштабах времени [57].

Параметр Херста H (0,5 < H < 1) является мерой статистического самоподобия фрактального трафика [55], и он связан с параметром распределения Парето α следующим соотношением: α = 3 – 2H .

Следует заметить, что в литературных источниках можно встретить различные определения фрактального трафика. В некоторых работах, например, [44, 56, 101] фрактальным трафиком называется тот трафик, который обладает самоподобием и ДВЗ. Данного понимания фрактального трафика будем придерживаться в настоящей работе. В другой части работ, например [59], фрактальным именуется только тот трафик, который описывается моделью фрактального дробного броуновского движения (ФБД). Данный факт объясняется тем, что, когда впервые И. Норросом была предложена модель ФБД [99] для описания самоподобного трафика, такой трафик был назван автором фрактальным броуновским трафиком (fractional Brownian trafic).

1.3 Системы массового обслуживания со степенными хвостами распределений и существующие методы исследования

Системы и сети массового обслуживания представляют собой модели, описывающие реальные системы с очередями. Системы массового обслуживания задаются параметрами, включающими [30]:

- входящий поток заявок, поступающих в систему, который, как правило, задается распределением вероятностей для времени между приходами заявок или распределением числа заявок, приходящих за определенный промежуток времени;
 - распределение времени обслуживания заявок;
- структурную организацию системы обслуживания: количество обслуживающих приборов (каналов) и объем очереди (буфера) для хранения поступающих заявок;
- дисциплину обслуживания, т.е. порядок выбора заявок из очереди для дальнейшего их обслуживания.

Сети массового обслуживания являются более сложными моделями и представляют собой объединение нескольких СМО [30, 70]. Различают два типа СеМО

– открытые (или разомкнутые) и замкнутые. В первом случае обслуживаемые заявки приходят на вход в СеМО, а после обслуживания покидают сеть. В замкнутых СеМО число заявок постоянно, и они не покидают ее после обслуживания. В данной диссертационной работе рассматриваются СеМО только открытого типа. В обоих случаях к описанию структуры сети обычно добавляется маршрутная матрица, содержащая переходные вероятности.

Существует общепринятое сокращенное обозначение для СМО, введенное Дж. Кендаллом, которое состоит из четырех составляющих: A/B/n/m. Символы A и B обозначают входящий поток требований и функцию распределения вероятностей интервалов времени обслуживания заявок, символ n ($n \ge 1$) обозначает число каналов в СМО, а m ($m \ge 0$) — число мест в очереди. Если $m = \infty$, то тогда его указание часто опускается, и запись в нотации Кендалла выглядит как A/B/n. Введем следующие общепринятые обозначения для A и B:

- М экспоненциальное распределение;
- Ра распределение Парето;
- G, GI распределение произвольного типа (символ I говорит о том, что интервалы поступления заявок или, соответственно, длительности их обслуживания, независимы). Также в работе будут использоваться обозначение Γ для гамма-распределения, Logn для логнормального распределения, W для распределения Вейбулла.

В замкнутых СМО и СеМО входящий поток не задается.

В данной работе СМО и СеМО, описание которых включает РТХ или, в частности, степенные распределения, будут также именоваться *фрактальными* системами.

Основными методами исследования СМО и СеМО являются аналитическое, численное и имитационное моделирование.

В большом числе современных работ, посвященных исследованию и описанию компьютерных сетей в терминах ТМО, входящие потоки в СМО и СеМО и время обслуживания рассматриваются как экспоненциальные или как общего вида, но не как имеющие, например, бесконечные дисперсии, и используются методы классической, в частности, элементарной [30] ТМО [43, 84]. Особенностью СМО с РТХ при моделировании систем с фрактальным трафиком является то, что соответствующие распределения часто имеют бесконечную дисперсию. В таких случаях аналитические и/или численные методы ТМО удается применить крайне редко. Например, мы можем аналитические результаты ТМО использовать для изучения СМО типа М/GI/1/∞

(формулы Полячека-Хинчина), M/GI/∞ (формула для числа занятых приборов в системе), M/GI/n/0 (формула потерь Эрланга, обобщенная на случай систем данного вида Б. А. Севастьяновым) или $GI/M/n/\infty$ [30]. Для изучения вероятностей потерь в СМО со степенными хвостами распределений с бесконечной дисперсией, мы можем использовать результаты для систем M/GI/n/0. Но в основной нашей задаче буфер ограничен и его размер не равен нулю. Одним из самых распространенных приемов для вычисления вероятностей потерь в системах с конечным буфером является их аппроксимация вероятностями превышения некоторого значения k числа заявок в соответствующих системах с бесконечным числом обслуживающих приборов и/или бесконечным буфером [41, 80, 92, 94, 106]. Однако соответствующие аппроксимации для вероятностей потерь должны быть исследованы дополнительно для случаев с бесконечной дисперсией времени поступления и/или обслуживания заявок. Соответствующие точные аналитические результаты имеются в ТМО только для систем M/M/n/m, а они уже не являются системами с PTX, адекватно описывающими реальные системы, функционирующие в условиях фрактального трафика. Поэтому основным методом исследования рассматриваемых СМО и СеМО является имитационное моделирование или аналитико-имитационное моделирование (АИМ).

Тем не менее, на практике существует ряд трудностей [18, 49, 50, 67, 68, 76, 105], с которыми неизбежно приходится сталкиваться и при ИМ систем с РТХ:

- низкая скорость сходимости оценок к искомым значениям вследствие наличия долговременных зависимостей;
- смещение моментов у реализуемых в ИМ РТХ и, в частности, у степенных распределений;
 - длительные переходные процессы;
- высокая сложность решения оптимизационных задач, обусловленная проблематичностью вычисления градиентов в ИМ [48].

Эти трудности приводят к использованию АИМ, комбинирующего имитационные и аналитические подходы [17, 18].

Иногда СМО с РТХ, у которых дисперсия бесконечна, аппроксимируются аналогичными системами, но с РТХ или РЛХ, имеющими конечные первые два момента [64, 82]. Например, в [82] СМО с РТХ аппроксимируются соответствующими СМО с РЛХ.

Необходимость разработки более сложных и адекватных моделей, отличных от «классических» СМО и СеМО, отмечается также в работе [104].

1.4 Обзор программных средств имитационного моделирования

В практике ИМ СМО и СеМО изучаются в рамках парадигмы дискретно-событийного моделирования. Эта парадигма предложена Д. Гордоном еще в 60-х годах XX века. Он разработал одну из наиболее удачных сред ИМ – GPSS – которая до сих пор пользуется большой популярностью в исследованиях СМО и СеМО.

В целом в настоящее время распространено большое количество программных сред ИМ [51]. Так, обширный список наиболее популярных средств дискретнособытийного ИМ на 2016 год приводится в статье [78].

Часть программных средств ИМ предназначена для исследования моделей конкретной предметной области [49]. Например, для моделирования сетей передачи данных используются такие программные системы, как OMNeT++, OPNET, Comnet [87]. Эти системы имеют собственные базы данных о существующих аппаратных средствах сетевых устройств, что делает модели максимально приближенными к реальным системам. Другая часть программных средств ИМ содержит более универсальные системы и предназначена для ИМ систем общего вида, без привязки к конкретной области.

В настоящее время наиболее популярными такими средами являются AnyLogic [53], Arena, GPSS World [3, 86]. Последняя является некоммерческим продуктом. С каждым годом свою популярность увеличивает среда ИМ AnyLogic. Ее отличает универсальное сочетание всех четырех парадигм ИМ (моделирование динамических систем, системная динамика, дискретно-событийное ИМ и агентное ИМ).

Среда GPSS World является самой распространенной и общедоступной. Она имеет два встроенных языка — GPSS (высокоуровневый язык ИМ) и PLUS (низко-уровневый язык ИМ). При более детальном изучении GPSS World был выявлен ряд недостатков, подробно описанных в разделе 2.1 и ограничивающих использование GPSS World при ИМ СМО и CeMO с PTX.

Инструмент ИМ AnyLogic свободен от недостатков, обнаруженных в GPSS World. Среда AnyLogic содержит большое количество библиотек, встроенных инструментов моделирования и анализа СМО и СеМО. Функционал AnyLogic крайне

широк. Но AnyLogic – коммерческий программный продукт, поэтому его функционал ограничен приобретенной версией AnyLogic. И основным недостатком AnyLogic является именно его относительно высокая стоимость, обусловливающая приобретение версий с ограниченным функционалом.

Совместное использование GPSS и AnyLogic позволяет разрабатывать такие модели СМО и СеМО с РТХ, которые, с одой стороны, являются в совокупности экономичными, с другой – позволяют эффективно и точно решать задачи исследования СМО и СеМО. Поэтому в настоящем диссертационном исследовании используются как AnyLogic, так и GPSS.

Выводы

Наличие фрактальных свойств у трафика современных СПД является неоспоримым фактом, на что указывает большое количество работ, например, [75, 77, 79, 100, и др.], содержащих статистический анализ измеренных трафиковых трасс. Характерные для фрактального трафика ДВЗ приводят на практике к ряду инженерных проблем [80], связанных с определением размеров буферов и организацией буферизации данных, управлением соединениями, управлением трафиком сетей передачи данных и др. Одной из первых инженерных проблем, обративших внимание специалистов на важность вопросов буферизации данных, стала проблема больших потерь ячеек в АТМ сети, где потери ячеек были значительно выше, чем это ожидалось [80].

Вероятность потерь пакетов является одной из важнейших характеристик, влияющей на качество обслуживания (QoS) современных СПД [89]. Поэтому обеспечение требуемой достаточно низкой вероятности потерь или даже сведение её к нулю является одной из наиболее актуальных задач проектировании сетей.

Таким образом, практическая значимость и недостаточная теоретическая разработанность проблем обеспечения низкой вероятности потерь обусловливают актуальность комплексного исследования, направленного на разработку эффективных математических и численных методов снижения вероятности потерь в СМО и СеМО со степенными хвостами распределений и бесконечными дисперсиями времени обслуживания.

Глава 2 Разработка методов планирования и организации имитационных экспериментов при моделировании систем с РТХ

Известно, что при математическом моделировании основными видами погрешностей являются погрешности математической модели, погрешности исходных данных, погрешности используемых методов и округления при вычислениях [2]. При ИМ систем с РТХ особое внимание следует уделять погрешностям реализации с.в. с заданными распределениями, относящимся к погрешностям используемых методов.

В настоящей главе в разделе 2.1 рассматривается проблема корректной реализации РТХ при ИМ СМО и СеМО с РТХ. В разделах 2.2–2.3 решаются проблемы организации имитационных экспериментов при исследовании систем с асимптотически степенными хвостами распределений. Материал данной главы основан на опубликованных автором статьях [13, 23, 114].

2.1 Особенности реализации РТХ в имитационных экспериментах

Как показывает практика, уже на этапе генерации случайных чисел необходимо следить за точностью реализации РТХ, числовые характеристики которых могут сильно смещаться и тем самым приводить к ложным результатам моделирования исследуемых СМО и СеМО.

Так, например, в работе [105] авторы пришли к выводу, что возможная причина недооценки среднего значения распределения Парето кроется в недостаточном количестве значащих цифр в используемом генераторе стандартных псевдослучайных чисел. В работах [16, 18] подробно исследуются и выявляются причины смещения моментов РТХ, реализуемых в ИМ. Основной причиной является то, что обычный шаг дискретизации у используемых ГСПЧ, лежащий в пределах от $\varepsilon = 10^{-15}$ до $\varepsilon = 10^{-6}$, достаточно малый для реализации распределений с легкими хвостами (РЛХ), при реализации РТХ часто оказывается неприемлемо большим. Приводятся также результаты многочисленных экспериментов, демонстрирующих влияние некорректной реализации РТХ на вычисляемые в ИМ оценки различных

показателей СМО с РТХ, в частности, в СМО вида M/Pa/n, M/Pa/n/0, Pa/M/n. Даются рекомендации по устранению искажений РТХ, сводящиеся, в основном, к рекомендации уменьшать шаг дискретизации є используемых стандартных псевдослучайных чисел (СПЧ).

2.1.1 Генераторы случайных чисел

Библиотеки современных компиляторов и систем ИМ, как правило, содержат свои ГСПЧ. При этом, поскольку требования к качеству генерируемых псевдослучайных последовательностей определяются особенностями решаемых задач, исследователям приходится самостоятельно анализировать пригодность предоставляемых ГСПЧ для решения этих задач. Существуют основные критерии, которым должен удовлетворять всякий ГСПЧ: равномерность, случайность, отсутствие автокорреляции [60].

Рассмотрим ГСПЧ, встроенные в GPSS World и в AnyLogic, с помощью которых выполнена часть экспериментов, представленных в диссертации.

Согласно справочному руководству GPSS World [86], используемый в GPSS мультипликативный конгруэнтный алгоритм Д. Лемера генерирует последовательность целых псевдослучайных чисел в открытом интервале от 0 до 2 147 483 647 с максимальным периодом $T_{\rm max} = 2^{31} - 1$. При вычислении случайных функций используются СПЧ, покрывающие диапазон от 0 до 0,999999 с шагом дискретизации $\varepsilon = 10^{-6}$.

Для реализации равномерно распределенной случайной величины в GPSS World можно использовать произвольное число ГСПЧ Uniform (i, 0, 1), где i — номер генератора. В [21] экспериментальным путем показано, что генераторы Uniform (i, 0, 1), где $i = \overline{1, 4}$, имеют период меньший, чем максимальный, а именно T = 2 047 999 999. При аналогичном исследовании автором диссертации первых семи ГСПЧ Uniform (i, 0, 1), полученные в [21] результаты подтвердились, и выяснилось, что все семь проверенных генераторов Uniform (i, 0, 1) имеют длину периода меньшую, чем $T_{\text{max}} = 2^{31} - 1$.

В AnyLogic различные распределения вероятностей реализуются на основе ГСПЧ класса *java.util.Random*. Данный ГСПЧ основан на линейном конгруэнтном

алгоритме с максимальным периодом $T_{\text{max}} = 2^{48}$, который выдает равномерно распределенные СПЧ в промежутке [0,1) и имеет шаг дискретизации $\varepsilon = 2^{-48} \approx 10^{-15}$ (меньший, чем в GPSS World). Так как в схеме $X_{n+1} = (aX_n + c) \mod m$ линейного конгруэнтного алгоритма параметры a = 2521490397, c = 11 и $m = 2^{48}$ удовлетворяют следующим условиям:

- 1) c и m являются взаимно простыми;
- 2) a-1 кратно p для каждого простого p, являющегося делителем m;
- 3) a-1 кратно 4, если m кратно 4, то период генерации является максимальным, т.е. $T=2^{48}$ [32].

Генераторы СПЧ рассматриваемых средах ИМ обладают необходимыми требованиями для проведения имитационных экспериментов. Однако, как будет показано в разделе 2.1.3, при моделировании СМО и СеМО с РТХ эти необходимые требования не являются достаточными [19, 20, 118]. К этим требованиям необходимо добавить требование малого шага дискретизации СПЧ.

2.1.2 Реализация случайных величин с РТХ

Значение с.в. ξ , распределенной в соответствии с некоторой ф.р. F(t), можно реализовать по следующей формуле:

$$\xi = F^{-1}(1-r), \tag{2.1}$$

где F^{-1} — функция, обратная функции F, r — базовая случайная величина (БСВ), равномерно распределённая на интервале (0, 1). Такой метод получения значений с.в. называется *методом обращения*. Конкретные выражения в формуле (2.1) можно получить за два простых шага:

- 1) записываем формальное уравнение $F(\xi) = r$;
- 2) решаем его относительно ξ;

Например, при выводе формулы для реализации экспоненциально распределенной с.в. ξ , ф.р. которой имеет вид $F(t) = 1 - e^{-\lambda t}$, эти два шага выполняются следующим образом:

1) записываем уравнение $1 - e^{-\lambda \xi} = r$;

2) решая его относительно ξ , получаем формулу для реализации экспоненциальной с.в. ξ путем преобразования БСВ r:

$$\xi = -\frac{1}{\lambda} \ln(1-r)$$
. (2.2)

По формуле (2.1), как и по формуле (2.2) при наименьшем значении БСВ r, которое равно нулю, реализуется и наименьшее значение с.в. ξ (в случае формулы (2.2) экспоненциальной с.в. это наименьшее значение ξ равно нулю), а при наибольшем значении r, равном единице, реализуется и наибольшее значение ξ (для экспоненциальной с.в. равное бесконечности). При r, близких к единице, реализуются большие значения ξ , описываемые асимптотикой хвоста $\overline{F}(t)$ распределения.

Заметим, что формулы (2.1), (2.2) часто переписываются в более простом виде:

$$\xi = F^{-1}(r),$$
 (2.3)

$$\xi = -\frac{1}{\lambda} \ln(r), \qquad (2.4)$$

на том простом основании, что распределения с.в. r и (1-r) совпадают. При использовании формул (2.3) и (2.4) большие значения ξ реализуются при выпадении r, близких к нулю.

Вместе с тем формулы (2.3), (2.4) могут быть получены таким же методом обращения, каким получены формулы (2.1), (2.2), с той лишь разницей, что обращается в этом случае не ф.р. F(t), а ее хвост $\overline{F}(t) = 1 - F(t)$. Поэтому формулы (2.3), (2.4) можно назвать формулами, полученными методом обращения хвоста распределения. В результате такого обращения сразу получается выражение с аргументом r:

$$\xi = \overline{F}^{-1}(r) \tag{2.5}$$

последующей замены одного аргумента другим не требуется).

Например, используя (2.5) для экспоненциальной с.в., сразу получаем формулу:

$$\xi = -\frac{1}{\lambda} \ln(r) \,. \tag{2.6}$$

При реализации любых с.в. преобразованием СПЧ r, реализуемых любым из распространенных стандартных ГСПЧ, не имеет особого значения, какой из методов обращения использовать — метод обращения ф.р. или метод обращения хвоста ф.р. Но при реализации с.в. с РТХ на основе метода ARAND, как будет показано в разделе 2.1.5, между этими двумя подходами возникают существенные различия, обусловленные особенностями метода ARAND и особенностями машинной арифметики. Эти особенности делают при реализации РТХ на основе метода ARAND предпочтительным метод обращения хвоста.

Выведем формулы для реализации с.в. со следующими часто используемыми РТХ – распределением Парето, распределением Бурра, распределением Вейбулла и логнормальным распределением [95].

1. Распределение Парето $Pa(K, \alpha)$

Для с.в. $\xi \in \text{Pa}(K, \alpha)$ (2.6) запишем исходное формальное уравнение в виде:

$$\overline{F}(\xi) = \left(\frac{K}{\xi}\right)^{\alpha} = r.$$

Выразив ξ через r, получаем следующую формулу для реализации с.в. $\xi \in \operatorname{Pa}(K, \alpha)$

$$\xi = \frac{K}{\alpha \sqrt{r}} = Kr^{-1/\alpha}.$$
 (2.7)

2. Распределение Бурра $Burr(\alpha, K, \tau)$

Хвост ф.р. Бурра имеет вид $\overline{F}(t) = \left(\frac{K}{K+t^{\tau}}\right)^{\alpha}$, где все три параметра (α,K,τ)

принимают только положительные значения. При $\tau=1$ распределение Бурра принимает одну из форм распределений Парето. При $\alpha \tau \downarrow 1$ хвост распределения Бурра становится особенно тяжелым.

Записывая уравнение
$$\overline{F}(\xi) = \left(\frac{K}{K+\xi^{\tau}}\right)^{\alpha} = r$$
 , получаем расчетную формулу

$$\xi = \sqrt[\tau]{\frac{K}{\sqrt[\alpha]r} - K} \ . \tag{2.8}$$

3. Распределение Вейбулла W(\lambda,\beta)

Ф.р. Вейбулла имеет вид $F(x) = 1 - e^{-(t/\lambda)^{\beta}}$, где β – параметр формы, $\lambda > 0$ – параметр масштаба. При $\beta < 1$ распределение Вейбулла имеет тяжелый хвост, т.к. при любом $\lambda_1 > 0$ имеем $\lim_{t \to \infty} \frac{\overline{F}(t)}{e^{-\lambda_1 t}} = \lim_{t \to \infty} \frac{e^{-(t/\lambda)^{\beta}}}{e^{-\lambda_1 t}} = \lim_{t \to \infty} e^{\lambda_1 t - (t/\lambda)^{\beta}} = \infty$. При $\beta > 1$ данный предел будет равен нулю для любого β . При $\beta = 1$ распределение Вейбулла превращается в экспоненциальное распределение, т.е. имеет легкий хвост. Для реализации с.в. с распределением Вейбулла методом обращения хвоста получаем расчетную формулу $\xi = \lambda^{\beta} \sqrt{-\ln(r)}$.

4. Логнормальное распределение $Logn(\mu, \sigma)$

Заметим, что не любые с.в. удобно генерировать методом обращения. Нередко более подходят другие методы. Например, для реализации с.в. ξ с логнормальным распределением удобнее вначале разыграть нормальную с.в. x с требуемыми параметрами μ , σ и затем преобразовать ее в логнормальную с.в. ξ (которая будет иметь те же μ , σ) по формуле $\xi = e^x$.

Используя критерий (1.6), покажем, что логнормальное распределение имеет тяжелый хвост. Логнормальное распределение задается ф.р. $F_{Logn}(t) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[\frac{\ln(t) - \mu}{\sigma \sqrt{2}} \right], \text{ где erf(z)} - «функция ошибок»:}$

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_{0}^{z} e^{-t^{2}} dt$$
.

Нам следует найти предел отношения:

$$L = \lim_{t \to \infty} \frac{1 - F_{Logn}(t)}{e^{-\lambda t}}.$$
 (2.9)

Функция erf(z) не выражается в элементарных функциях в замкнутом виде. Поэтому применим к (2.9) правило Лопиталя — продифференцируем по t числитель и знаменатель находящегося под знаком предела выражения. При этом, дифференцируя числитель, на шаге вычисления производной от F_{Logn} мы можем использовать тот факт, что эта производная является известной функцией (это, очевидно, п.в. логнормального распределения):

$$L = \lim_{t \to \infty} \frac{1 - F_{Logn}(t)}{e^{-\lambda t}} = \lim_{t \to \infty} \frac{-f_{Logn}(t)}{-\lambda e^{-\lambda t}} = \lim_{t \to \infty} \frac{f_{Logn}(t)}{\lambda e^{-\lambda t}} = \lim_{t \to \infty} \frac{\frac{1}{t\sigma\sqrt{2\pi}}e^{\frac{-(\ln t - \mu)^2}{2\sigma^2}}}{\lambda e^{-\lambda t}} = \frac{1}{\lambda\sigma\sqrt{2\pi}}\lim_{t \to \infty} \frac{e^{\frac{-(\ln t - \mu)^2}{2\sigma^2}}}{te^{-\lambda t}} = \frac{1}{\lambda\sigma\sqrt{2\pi}}\lim_{t \to \infty} \frac{e^{\frac{-(\ln t - \mu)^2}{2\sigma^2}}}{t} = \infty.$$

Итак, применяя критерий (1.6), устанавливаем, что логнормальное распределение при любых значениях его параметров μ , σ является РТХ.

2.1.3 Смещение моментов при реализации РТХ в имитационном моделировании

Непрерывная БСВ, реализуемая распространенными ГСПЧ, в силу ограниченности разрядной сетки вычислительных машин реализуется как n-разрядное дискретное СПЧ r, множество возможных значений которого представляет собой на числовой оси набор равноотстоящих точек, расположенных в промежутке между нулем и единицей. В силу этого реализуемые путем преобразования БСВ непрерывные с.в. ξ также оказываются дискретными — см. формулы (2.6)—(2.8). Обозначим эти дискретные реализации непрерывных с.в. ξ через ξ^* . В связи с дискретностью СПЧ r равномерный закон распределения вероятностей СПЧ r называют иногда r равномерным законом распределения [5]. У такого закона распределения все возможные значения r следуют друг за другом с определенным шагом r следуют друг за другом с определенным r следуют r следуют друг за другом с определенным r следуют r следу

Одной из основных проблем при моделировании СМО и СеМО с РТХ является смещение моментов РТХ из-за шага дискретизации $\varepsilon > 0$ у СПЧ r. При реализации РТХ «обычный» шаг дискретизации оказывается иногда слишком большим. Вследствие смещения моментов РТХ при ИМ СМО и СеМО с РТХ получаются смещенными и значения рассчитываемых показателей.

Смещение моментов РТХ, обусловленное дискретностью СПЧ, можно точно выразить через шаг ε их дискретизации. Псевдослучайные числа r принимают равновероятные числовые значения $\{r_1, r_2, r_3, ..., r_N\} = \{\varepsilon, 2\varepsilon, 3\varepsilon, ..., N\varepsilon\}$, где N — число возможных значений r (N не превосходит длину T периода ГСПЧ).

Рассмотрим множество значений с.в. ξ^* , генерируемой по формуле (2.7) реализации распределения Парето. Эта с.в. имеет множество равновероятных значений $\xi_1, \xi_2, ..., \xi_N$, получаемых подстановкой в формулу (2.7) значений $\{r_1, r_2, r_3, ..., r_N\} = \{\varepsilon, 2\varepsilon, 3\varepsilon, ..., N\varepsilon\}$ ПСЧ r. Исходя из равновероятности значений $\xi_1, \xi_2, ..., \xi_N$ мы можем найти начальные моменты с.в. ξ^* следующим образом:

$$M[(\xi^*)^k] = \frac{1}{N} \sum_{i=1}^N \left(\frac{K}{\sqrt[\alpha]{r_i}}\right)^k = \frac{1}{N} \sum_{i=1}^N \left(\frac{K}{\sqrt[\alpha]{i\varepsilon}}\right)^k = \frac{1}{N} \sum_{i=1}^N K^k \left(\frac{i}{N}\right)^{-k/\alpha}.$$
 (2.10)

Сумма в этой формуле содержит большое число N слагаемых — например, миллион (как в GPSS), или 10^{15} (как в AnyLogic). Такие суммы хорошо рассчитываются в современных математических пакетах, в том числе в WolframAlpha.com, а правильность расчетов легко контролируется — например, путем приближенной замены сумм или их частей соответствующими интегралами.

Рассчитаем, например, м.о. рассматриваемой с.в. ξ^* , реализующей распределение Парето при K=1, $\alpha=1,01$ (это случай высокой тяжести хвоста распределения Парето, поскольку α близко к единице). Для ГСПЧ GPSS имеем $N=10^6$ и, согласно формуле (2.10),

$$M(\xi^*) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{1,01 \sqrt{\frac{i}{N}}} = \frac{1}{10000000} \sum_{i=1}^{10000000} \left(\frac{i}{10000000}\right)^{-1/1,01} = 13,4151.$$
 (2.11)

В действительности м.о. у распределения Парето при K=1, $\alpha=1,01$ составляет величину $K\alpha/(\alpha-1)=101$. Сравнивая с этим точным м.о. результат (2.11), видим, что в GPSS произойдет значительное смещение м.о. с.в. ξ^* в сторону уменьшения.

Аналогичный расчет для ГСПЧ AnyLogic, где $N = 10^{15}$, дает

$$M(\xi^*) = \frac{1}{10^{15}} \sum_{i=1}^{10^{15}} \left(\frac{i}{10^{14}}\right)^{-1/1,01} = 29,6621.$$
 (2.12)

Таким образом, в AnyLogic м.о. реализованной с.в. ξ^* также сильно смещается, хотя и немного меньше, чем в GPSS.

Подробное исследование смещения моментов распределения Парето в зависимости от шага ε при различной степени тяжести хвоста приведено в [18]. Проведем аналогичное исследование числовых характеристик распределений Вейбулла и Бурра. Для них зависимость реализуемых первых двух моментов (м.о. и дисперсии) от шага дискретизации ГСПЧ рассчитана в Таблицах 2.1, 2.2. В Таблице 2.1 истинные значения этих моментов обозначены как $M(\xi)$, $D(\xi)$, а фактически реализуемые моменты, рассчитанные с учетом шага дискретизации ε , – как $M(\xi^*)$ и $D(\xi^*)$.

Таблица 2.1 – Числовые характеристики с.в. $\xi \in W(\lambda = 1, \beta)$ и ее реализации ξ^*

β		$M(\xi^*)$		$M(\xi)$ $D(\xi^*)$		D(ξ)		
	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$	3)	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$	3)
0,1	3 086 800	3 625 840	3 628 800	3 628 800	1,43316·10 ¹⁷	2,2698·10 ¹⁸	2,26988·10 ¹⁸	2,4328·10 ¹⁸
0,2	119,5150	120,0000	120,0000	120,0000	3 072 550	3 614 140	3 614 400	3 614 000
0,3	9,2557	9,2605	9,2605	9,2605	2458,64	2507,81	2507,81	2507,81
0,4	3,3229	3,3234	3,3234	3,3234	108,473	108,955	108,955	108,955
0,5	1,9999	2,0000	2,0000	2,0000	19,9690	20,0000	20,0000	20,0000
0,6	1,5045	1,5046	1,5046	1,5046	6,9921	6,9968	6,9968	6,9969
0,7	1,2658	1,2658	1,2658	1,2658	3,4256	3,4268	3,4268	3,4268
0,8	1,1330	1,1330	1,1330	1,1330	2,0392	2,0397	2,0397	2,0397
0,9	1,0522	1,0522	1,0522	1,0522	1,3713	1,3715	1,3715	1,3715

Для распределения Вейбулла значения $M(\xi)$, $D(\xi)$, $M(\xi^*)$ и $D(\xi^*)$ рассчитаны с помощью следующих формул:

$$M(\xi) = \lambda \Gamma (1 + 1/\beta), \quad D(\xi) = \lambda^{2} (\Gamma (1 + 2/\beta) - \Gamma^{2} (1 + 1/\beta)),$$

$$M(\xi^{*}) = \frac{1}{N} \sum_{i=1}^{N} \lambda (-\ln(i \cdot \varepsilon))^{1/\beta},$$

$$D(\xi^{*}) = \frac{1}{N} \sum_{i=1}^{N} (\lambda (-\ln(r))^{1/\beta})^{2} - \left(\frac{1}{N} \sum_{i=1}^{N} \lambda (-\ln(r))^{1/\beta}\right)^{2}.$$
(2.13)

Таблица 2.2 — Числовые характеристики с.в. $\xi \in Burr(\alpha,K,\tau)$ при $K=1,\ \tau=5$ и ее реализации ξ^*

α	α $M(\xi^*)$		Μ(ξ)	α	$D(\xi^*)$		D(ξ)		
	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$. (3)		$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-12}$	$\varepsilon = 10^{-15}$	(3)
0,21	10,3602	15,4602	16,9960	20,9399	0,41	8,5085	16,7602	19,9550	37,3646
0,23	6,4308	7,4087	7,5235	7,6019	0,43	5,5733	9,0019	9,0019	11,0037
0,25	4,6507	4,9130	4,9263	4,9307	0,45	4,1021	5,5252	5,5252	5,9165
0,27	3,6916	3,7809	3,7830	3,7834	0,47	3,0486	3,7324	3,7324	3,8327
0,29	3,1076	3,1437	3,1441	3,1442	0,49	2,3463	2,7007	2,7007	2,7312
0,31	2,7191	2,7359	2,7359	2,7359	0,51	1,8591	2,0549	2,0549	2,0654
0,33	2,4436	2,4521	2,4522	2,4522	0,53	1,5094	1,6236	1,6269	1,6276

Числовые характеристики с.в $\xi \in Burr(\alpha, K, \tau)$ рассчитаны с помощью следующих формул:

$$M(\xi) = \frac{K^{1/\tau} \Gamma(1+1/\tau) \Gamma(\alpha-1/\tau)}{\Gamma(\alpha)},$$

$$D(\xi) = \left(\Gamma(1+2/\tau) \Gamma(\alpha-2/\tau) - \frac{\left(\Gamma(1+1/\tau) \Gamma(\alpha-1/\tau)\right)^{2}}{\Gamma(\alpha)}\right) \frac{K^{2/\tau}}{\Gamma(\alpha)},$$

$$M(\xi^{*}) = \frac{1}{N} \sum_{i=1}^{N} \sqrt[\tau]{\frac{K}{\sqrt[\alpha]{r}} - K} = \frac{1}{N} \sum_{i=1}^{N} \sqrt[\tau]{\frac{K}{\sqrt[\alpha]{i \cdot \varepsilon}} - K},$$

$$D(\xi^{*}) = \frac{1}{N} \sum_{i=1}^{N} \left(\sqrt[\tau]{\frac{K}{\sqrt[\alpha]{i \cdot \varepsilon}} - K}\right)^{2} - \left(\frac{1}{N} \sum_{i=1}^{N} \sqrt[\tau]{\frac{K}{\sqrt[\alpha]{i \cdot \varepsilon}} - K}\right)^{2}.$$
(2.14)

Следует отметить, что для распределения Бурра k-й момент существует только когда $k < \alpha \tau$.

Из Таблиц 2.1 и 2.2 видно, что шаг $\varepsilon > 0$ дискретизации ГСПЧ напрямую влияет на точность получаемых при реализации моментов с.в. с РТХ. Самые большие смещения моментов распределений возникают в критических зонах, в которых хвост особенно тяжел. При этом истинные значения моментов не обязательно принимают очень большие значения (см. Таблицу 2.2).

Если при моделировании с.в. с распределением Вейбулла еще можно использовать шаг дискретизации СПЧ $\epsilon = 10^{-15}$, то для распределения Бурра такой шаг

дискретизации оказывается неприемлемо большим (как и для распределения Парето, см. [18]). При реализации на основе стандартных ГСПЧ любого РТХ следует учитывать особенности его поведения и выявлять «критические зоны» значений его параметров, где искажения РТХ становятся наиболее существенными.

При реализации любого распределения F(t) с.в. ξ на основе стандартных ГСПЧ извлекаются выборки из множества N равновероятных значений $\{\xi_1, \xi_2, ..., \xi_N\}$. Нужная форма распределения F(t) достигается за счет подходящей неравномерной расстановки этих равновероятных значений вдоль числовой оси t.

2.1.4 Реализация РТХ в имитационном моделировании

Рассмотрим реализацию с.в. $\xi \in Pa(1, 1, 1)$ в GPSS World и рассчитаем путем ИМ смещение первого момента $M(\xi)$. При K=1, $\alpha=1,1$ м.о. $M(\xi)=11$, в то время как для фактически реализованной с.в. ξ^* м.о. $M(\xi^*)$ получается существенно меньшим.

Для расчета $M(\xi^*)$ используем ИМ. Чтобы повысить точность результатов ИМ воспользуемся методом стратификации [29]. Для этого разобьем множество Ω значений СПЧ r на страты Ω_i :

$$\Omega_i = \{r : 10^{-i} < r \le 10^{-(i-1)}\}, \quad i = 1, 2, \dots$$
(2.15)

(при этом, как и в GPSS, будем использовать СПЧ r, содержащие 6 десятичных разрядов после запятой, т.е. шаг ε дискретизации СПЧ r будет равен 10^{-6}).

По формуле (2.7) определим границы соответствующих страт Θ_i для с.в. ξ^* . Страты для r и для ξ^* приведены в Таблице 2.3.

Страты Ω_i для СПЧ r Страты Θ_i для $c.в. \xi^* \in Pa(1, 1, 1)$ $0,1 < r \le 1,$ $1^{-1/\alpha} \le \xi^* < 0,1^{-1/\alpha},$ $0,01 < r \le 0,1,$ $0,001 < r \le 0,01,$ $0,0001 < r \le 0,001,$ $0,0001 < r \le 0,0001,$ $0,00001 < r \le 0,0001,$ $0,0001^{-1/\alpha} \le \xi^* < 0,0001^{-1/\alpha},$ $0,0001^{-1/\alpha} \le \xi^* < 0,00001^{-1/\alpha},$ $0,0001^{-1/\alpha} \le \xi^* < 0,00001^{-1/\alpha},$ $0,0001^{-1/\alpha} \le \xi^* < 0,00001^{-1/\alpha},$ $10^{-6} < r \le 10^{-5},$ $10^{5/\alpha} \le \xi^* < 10^{6/\alpha},$ $10^{6/\alpha} < \xi^* < 10^{7/\alpha}$		
$\begin{array}{lll} 0,1 < r \leq 1, & 1^{-1/\alpha} \leq \xi^* < 0,1^{-1/\alpha}, \\ 0,01 < r \leq 0,1, & 0,1^{-1/\alpha} \leq \xi^* < 0,01^{-1/\alpha}, \\ 0,001 < r \leq 0,01, & 0,01^{-1/\alpha} \leq \xi^* < 0,001^{-1/\alpha}, \\ 0,0001 < r \leq 0,001, & 0,001^{-1/\alpha} \leq \xi^* < 0,0001^{-1/\alpha}, \\ 0,0001^{-1/\alpha} \leq \xi^* < 0,0001^{-1/\alpha}, \\ 10^{-6} < r \leq 10^{-5}, & 10^{5/\alpha} \leq \xi^* < 10^{6/\alpha}, \end{array}$	Страты Ω_i для СПЧ r	Страты Θ_i для
$\begin{array}{lll} 0,01 < r \leq 0,1, & 0,1^{-1/\alpha} \leq \xi^* < 0,01^{-1/\alpha}, \\ 0,001 < r \leq 0,001, & 0,0001 < r \leq 0,0001, \\ 0,00001 < r \leq 0,0001, & 0,0001^{-1/\alpha} \leq \xi^* < 0,0001^{-1/\alpha}, \\ 10^{-6} < r \leq 10^{-5}, & 10^{5/\alpha} \leq \xi^* < 10^{6/\alpha}, \end{array}$		c.b. $\xi^* \in Pa(1, 1, 1)$
$\begin{array}{lll} 0,001 < r \leq 0,01, \\ 0,0001 < r \leq 0,001, \\ 0,00001 < r \leq 0,0001, \\ 10^{-6} < r \leq 10^{-5}, \end{array} \qquad \begin{array}{lll} 0,01^{-1/\alpha} \leq \xi^* < 0,001^{-1/\alpha}, \\ 0,001^{-1/\alpha} \leq \xi^* < 0,0001^{-1/\alpha}, \\ 0,0001^{-1/\alpha} \leq \xi^* < 0,00001^{-1/\alpha}, \\ 0,0001^{-1/\alpha} \leq \xi^* < 0,00001^{-1/\alpha}, \\ 10^{5/\alpha} \leq \xi^* < 10^{6/\alpha}, \end{array}$	$0,1 < r \le 1,$	$1^{-1/\alpha} \le \xi^* < 0, 1^{-1/\alpha},$
$\begin{array}{c} 0,0001 < r \le 0,001, \\ 0,00001 < r \le 0,0001, \\ 10^{-6} < r \le 10^{-5}, \\ \end{array} \qquad \begin{array}{c} 0,01 \le \zeta < 0,001, \\ 0,001^{-1/\alpha} \le \xi^* < 0,0001^{-1/\alpha}, \\ 0,0001^{-1/\alpha} \le \xi^* < 0,00001^{-1/\alpha}, \\ 10^{5/\alpha} \le \xi^* < 10^{6/\alpha}, \end{array}$	$0.01 < r \le 0.1$,	$0.1^{-1/\alpha} \le \xi^* < 0.01^{-1/\alpha}$
$\begin{array}{c c} 0,00001 < r \le 0,0001, \\ 10^{-6} < r \le 10^{-5}, \\ 0,0001 \le \zeta < 0,0001, \\ 0,0001^{-1/\alpha} \le \xi^* < 0,00001^{-1/\alpha}, \\ 10^{5/\alpha} \le \xi^* < 10^{6/\alpha}, \end{array}$	$0.001 < r \le 0.01$,	$0.01^{-1/\alpha} \le \xi^* < 0.001^{-1/\alpha}$
$10^{-6} < r \le 10^{-5}, \qquad 10^{5/\alpha} \le \xi^* < 10^{6/\alpha},$	$0,0001 < r \le 0,001,$	$0.001^{-1/\alpha} \le \xi^* < 0.0001^{-1/\alpha}$
	$0.00001 < r \le 0.0001,$	$0.0001^{-1/\alpha} \le \xi^* < 0.00001^{-1/\alpha}$
$10^{-7} < r \le 10^{-6}$. $10^{6/\alpha} < \xi^* < 10^{7/\alpha}$	$10^{-6} < r \le 10^{-5},$	$10^{5/\alpha} \le \xi^* < 10^{6/\alpha}$
	$10^{-7} < r \le 10^{-6},$	$10^{6/\alpha} \le \xi^* < 10^{7/\alpha}$,

Таблица 2.3 – Страты для СПЧ r и с.в. ξ^*

Благодаря такой стратификации можно не только обеспечить высокую точность имитационной оценки м.о. $M(\xi^*)$ (для определения его смещения), но и проследить за изменением смещений условных м.о. при удалении страт в направлении больших значений с.в. ξ^* .

Безусловное м.о. $M(\xi^*)$ выражается через условные м.о. следующим образом:

$$M(\xi^*) = \sum_{i=1}^{\infty} w_i M_i$$
, (2.16)

где M_i^* — условное м.о. с.в. ξ^* в страте Θ_i , т.е. $M_i^* = M(\xi^* \mid \xi^* \in \Theta_i)$, а w_i^* — вероятность попадания с.в. ξ^* в страту Θ_i : $w_i^* = P(\xi^* \in \Theta_i)$. В качестве примера рассчитаем вероятность w_1^* попадания с.в. ξ^* в страту Θ_1 :

$$w_1^* = P(1^{-1/\alpha} \le \xi^* < 0, 1^{-1/\alpha}) = P(0, 1 < r \le 1) = (1000000 - 100000) \cdot \varepsilon =$$

$$= (1000000 - 100000) \cdot 10^{-6} = 0, 9.$$
(2.17)

Аналогично рассчитываются вероятности попадания с.в. ξ^* в другие страты, поэтому для них получаем

$$w_2^* = 9 \cdot 10^{-2}, \ w_3^* = 9 \cdot 10^{-3}, \ w_4^* = 9 \cdot 10^{-4}, \ w_5^* = 9 \cdot 10^{-5}, \ w_6^* = 9 \cdot 10^{-6},$$
 $w_7^* = 10^{-6}, \ w_i^* = 0 \text{ при } i > 7.$ (2.18)

Вероятность w_7^* по виду отличается от предыдущих вероятностей из-за того, что в страту Θ_7 попадает единственное из миллиона значений с.в. ξ^* , и расчетная формула для вероятности этой страты принимает вид $w_7^* = ... = (1-0) \cdot 10^{-6} = 10^{-6}$.

В системе ИМ GPSS проведено h=100 млн. опытов с оптимальным распределением их по стратам, т.е. число h_i опытов со значениями $r\in\Omega_i$ определялось по формуле [33]

$$h_i = \frac{w_i^* \sigma_i}{\sum_{j:w_{i>0}^*} w_j^* \sigma_i} h, \qquad \forall i: w_i^* > 0,$$

где $\sigma_i = \sqrt{D_i}$, $D_i = D(\xi^* \mid \xi^* \in \Theta_i)$ — условная дисперсия с.в. ξ^* , определенная при условии, что $\xi^* \in \Theta_i$, или, что то же самое, при условии, что $r \in \Omega_i$.

В результате такого стратифицированного ИМ получены оценки $\hat{\mathbf{M}}_i$ условных м.о. \mathbf{M}_i^* для всех страт (Таблица 2.4). Для получения каждой оценки $\hat{\mathbf{M}}_i$ разыгрывались только те значения r, которые относятся к страте Ω_i . Такая реализация значений r в стратах осуществлялась с помощью встроенного в GPSS генератора Duniform(1, c, d), где 1 — номер генератора, c и d — наименьшее и наибольшее значения целочисленной с.в. Например, значения с.в. ξ^* , относящиеся к первой страте (см. Таблицу 2.3), реализовывались следующим выражением:

где ALPHA – параметр α ф.р. Парето.

Точные значения условных м.о. M_i^* , учитывающие шаг дискретизации $\varepsilon = 10^{-6}$ встроенного ГСПЧ, рассчитываются с учетом следующего соображения. Поскольку все значения с.в. ξ^* равновероятны, то они остаются равновероятными и при условии их принадлежности той или иной страте. Потому для точного расчета условного м.о. M_i^* достаточно найти среднее арифметическое тех значений с.в. ξ^* , которые принадлежат страте Θ_i . Таким образом, для 1-й страты получаем расчетную формулу:

$$\mathbf{M}_{1}^{*} = \frac{1}{N_{1}} \sum_{r \in \Omega_{1}} r^{-1/\alpha} = \frac{1}{N_{1}} \sum_{r \in \Omega_{1}} r^{-1/1,1} = \frac{1}{900000} \sum_{j=100001}^{1000000} (j\varepsilon)^{-1/1,1}, \qquad (2.19)$$

где $N_1=900000$ — число значений СПЧ r в страте Ω_1 . В следующих стратах $N_2=100000-10000=90000$, ..., $N_6=10-1=9$, $N_7=1-0=1$. Границы суммирования в страте Ω_2 — от 10001 до 100000, в страте Ω_3 — от 1001 до 10000, ..., в страте Ω_6 — от 2 до 10, в страте Ω_7 — от 1 до 1.

Результаты соответствующих вычислений приведены в столбце M_i^* .

Заметим, что условное м.о. квадрата с.в. ξ^* для страты Θ_1 по формуле (2.10) можно рассчитать следующим образом:

и далее, вычитая из него квадрат м.о. (2.19), можно найти условную дисперсию с.в. ξ^* для этой страты. Этим способом и рассчитывались условные дисперсии для приведенной выше формулы оптимального распределения h опытов по стратам.

Таблица 2.4 – Условные и безус	ловные м.о. с.в. ٤ [*] и «	$\xi \in \text{Pa}(1,$	1.1) в стратах Θ_i
1 000111140 = 1 1 000 11 000 1	1021210 11100 0120 5 11	5 (-)	1,1) 2 01pului 01

i	$\hat{ ext{M}}_i$	M_i^*	$M_i(\xi)$
1	2,3083	2,3084	2,3084
2	18,7220	18,7238	18,7242
3	151,8362	151,8514	151,8774
4	1229,6927	1229,8194	1231,9242
5	9824,1359	9824,4155	9992,5167
6	69394,2800	69383,8378	81052,3836
7	284803,5868	284804,5868	657440,8724
8	0	0	5332705,6105
Итог	8,0292	8,0297	11

Для с.в. ξ (теоретической) вероятности w_i страт Θ_i рассчитываются следующим образом:

$$w_1 = P(1^{-1/\alpha} \le \xi < 0, 1^{-1/\alpha}) = F(0, 1^{-1/1, 1}) - F(1^{-1/1, 1}) = \overline{F}(1^{-1/1, 1}) - \overline{F}(0, 1^{-1/1, 1}) = \overline{F}(1^{-1/1, 1}) = \overline{F}(1^{-1/1, 1}) - \overline{F}(0, 1^{-1/1, 1}) = \overline{F}(1^{-1/1, 1}) = \overline{F}(1^{-1/1, 1}) - \overline{F}(1^{-1/1, 1}) = \overline{F$$

$$= \left(\frac{1}{1^{-1/1,1}}\right)^{1,1} - \left(\frac{1}{0,1^{-1/1,1}}\right)^{1,1} = 0,9,$$
(2.20)

$$w_2 = 9 \cdot 10^{-2}, \dots, w_i = 9 \cdot 10^{-i}, \dots$$
 (2.21)

Начиная с вероятности w_7 последовательность (2.20), (2.21) отличается от последовательности (2.17), (2.18).

Чтобы рассчитать условные м.о. $M_i(\xi)$, выведем выражение для соответствующего условного хвоста ф.р. Для этого воспользуемся известной формулой условной ф.р., определяемой при условии попадания с.в. в заданный промежуток (a, b):

$$F(t \mid a \le \xi < b) = \begin{cases} 0, & \text{при} \quad t < a, \\ \frac{F(t) - F(a)}{F(b) - F(a)}, & \text{при} \quad a \le t < b, \\ 1, & \text{при} \quad b \le t. \end{cases}$$

Отсюда находим условный хвост распределения:

$$\overline{F}(t \mid a \le \xi < b) = 1 - F(t \mid a \le \xi < b) = \begin{cases} 1, & \text{при} \quad t < a, \\ \frac{F(b) - F(t)}{F(b) - F(a)}, & \text{при} \quad a \le t < b, \\ 0, & \text{при} \quad b \le t. \end{cases}$$

Принимая во внимание то, что левее a хвост равен единице, правее b – нулю, находим условное м.о. $M(\xi | a \le \xi \le b)$ следующим образом:

$$M(\xi \mid a \le \xi < b) = \int_{0}^{\infty} \overline{F}(\xi \mid a \le \xi < b) dt = \int_{0}^{a} dt + \int_{a}^{b} \frac{F(b) - F(t)}{F(b) - F(a)} dt =$$

$$= a + \int_{a}^{b} \frac{F(b) - F(t)}{F(b) - F(a)} dt.$$
(2.22)

Для ф.р. $Pa(1, \alpha)$ имеем $F(b) = 1 - b^{-\alpha}$, $F(t) = 1 - t^{-\alpha}$, $F(a) = 1 - a^{-\alpha}$, поэтому для данной ф.р. расчетная формула (2.22) принимает вид:

$$M(\xi \mid a \le \xi < b) = a + \int_{a}^{b} \frac{1 - b^{-\alpha} - (1 - t^{-\alpha})}{1 - b^{-\alpha} - (1 - a^{-\alpha})} dt =$$

$$= a + \frac{b^{1 - \alpha} - a^{1 - \alpha}}{a^{-\alpha} - b^{-\alpha}} \cdot \frac{1}{1 - \alpha} - \frac{(b - a)b^{-\alpha}}{a^{-\alpha} - b^{-\alpha}}.$$

Поскольку в i-ой страте $a = 10^{i-1/\alpha}$, $b = 10^{i/\alpha}$, то:

$$M_{i}(\xi) = 10^{\frac{i-1}{\alpha}} + \frac{10^{i \cdot \frac{1-\alpha}{\alpha}} - 10^{(i-1) \cdot \frac{1-\alpha}{\alpha}}}{10^{1-i} - 10^{-i}} \cdot \frac{1}{1-\alpha} + 10^{-i} \cdot \frac{10^{\frac{i}{\alpha}} - 10^{\frac{i-1}{\alpha}}}{10^{1-i} - 10^{-i}}.$$
 (2.23)

Рассчитывая значения $M_i(\xi)$ при $\alpha=1,1$ и далее, с учетом известных вероятностей слоев (2.21) вычисляя по формуле $M(\xi)=\sum_{i=1}^{\infty}w_iM_i(\xi)$ сумму первых 100 сла-

Точные теоретические значения условных м.о. $M_i(\xi)$ непрерывной с.в. ξ приведены в последнем столбце Таблицы 2.4.

В последней строке таблицы (строка «Итог») приведены следующие значения безусловных м.о.: (1) оценка $\hat{\mathbf{M}}(\xi^*) = \sum_i w_i^* \hat{\mathbf{M}}_i = 8,0292$, полученная с помощью ИМ на GPSS, (2) точно рассчитанное значение $\mathbf{M}(\xi^*) = \sum_i w_i^* \mathbf{M}_i^* = 8,0297$ и (3) истинное значение безусловного м.о. $\mathbf{M}(\xi) = 11$ теоретической с.в. ξ , которая должна была бы быть реализована без искажений. Совпадение значений точно рассчитанного м.о. $\mathbf{M}(\xi^*)$ и оценки $\hat{\mathbf{M}}(\xi^*)$ с точностью до 4-х значащих цифр свидетельствует о высокой точности моделирования, достигнутой благодаря применению метода стратификации. И одновременно повышает доверие к результатам точного расчета $\mathbf{M}(\xi^*)$.

Из Таблицы 2.4 видно, что с удалением страт в область больших значений с.в. ξ^* ее условные м.о. все больше смещаются относительно истинных условных м.о. с.в. ξ . Значения фактически реализованной с.в. ξ^* не превышают 284803,6 изза шага дискретизации $\varepsilon = 10^{-6}$, который, таким образом, в данном случае – случае РТХ Парето – оказывается слишком большим. Этот вывод справедлив и для других РТХ. В случае РЛХ с.в. ξ ее условные м.о. убывают столь быстро, что при шаге дискретизации $\varepsilon = 10^{-6}$ заметной разницы между ее условными м.о. и условными м.о. с.в. ξ^* не возникает, и м.о. $M(\xi^*)$, как известно, практически не смещаются.

Смещение моментов РТХ может приводить к существенным ошибкам при ИМ СМО и СеМО. Рассмотрим следующий пример. Пусть в СМО М/Ра/1 время обслуживания ξ принадлежит распределению Ра(K, α). Тогда, если $\alpha \in (1, 2]$, то стационарная средняя длина L очереди заявок будет бесконечной при любом коэффициенте загрузки $\rho \in (0, 1)$. Действительно, согласно формуле Полячека-Хинчина, здесь при любом $\rho \in (0, 1)$ и при любой интенсивности входящего потока $\lambda > 0$ получаем:

$$L = \frac{\lambda \xi^{(2)}}{2(1-\rho)} = \infty, \tag{2.24}$$

так как второй момент $\xi^{(2)}$ времени обслуживания ξ при $\alpha \in (1, 2]$ бесконечен. Но если эту же СМО исследовать методом ИМ, то вместо времени обслуживания $\xi \in \text{Pa}(K, \alpha)$ реализуется время ξ^* со смещенными моментами. Например, при K=1, $\alpha=2$ (самый легкий в диапазоне $1<\alpha\leq 2$ хвост) и $\epsilon=10^{-15}$ (как в системе AnyLogic), согласно формуле (2.10), реализуется второй момент

$$\xi^{*(2)} = M[(\xi^*)^2] = \frac{1}{10^{15}} \sum_{i=1}^{10^{15}} \left[\left(10^{-15} i \right)^{-1/2} \right]^2 = 35,116,$$

вследствие чего, согласно формуле Полячека-Хинчина, при достаточно большой длине прогона модели получим:

$$L \approx \frac{35,116 \cdot \lambda}{2(1-\rho)} \approx \frac{17,5 \cdot \lambda}{1-\rho}.$$
 (2.25)

Этот вывод подтверждается и результатами имитационных экспериментов. Особенностью рассмотренного случая является принципиальное отличие имитационного решения (2.25) от правильного решения (2.24).

Распространенные значения шага ε дискретизации СПЧ, лежащие в пределах от $\varepsilon = 10^{-15}$ до $\varepsilon = 10^{-6}$, приводят к значительным смещениям моментов РТХ, реализуемых в ИМ. Значительные смещения моментов РТХ приводят, в свою очередь, к существенным, иногда принципиальным погрешностям результатов моделирования СМО и СеМО. Этим определяется актуальность разработки методов для кардинального уменьшения шага ε дискретизации БСВ r. При этом желательно, чтобы разработанный метод уменьшения искажений РТХ основывался на существующих ГСПЧ, используемых в языках и системах ИМ, и не требовал увеличения длины n существующей разрядной сетки ЭВМ.

2.1.5 Meтод ARAND

В качестве методов, позволяющих уменьшить є и тем самым уменьшить искажения реализуемых РТХ, можно использовать методы «длинной арифметики» и «длинных взвешенных сумм» [12]. Эти методы основаны на программной эмуляции увеличения длины разрядной сетки, используемой стандартным ГСПЧ. При этом

для реализации одного значения r с уменьшенным шагом дискретизации требуется несколько раз обратиться к стандартному ГСПЧ, из-за чего скорость генерации псевдослучайных чисел в несколько раз уменьшается и, соответственно, в несколько раз сокращается длина периода генерируемой последовательности. К тому же, как видно из расчетов (2.11), (2.12) и из Таблицы 2.1, увеличение длины разрядной сетки в 2–3 раза, равносильное уменьшению ε приблизительно на 10–15 порядков, может не приводить к достаточному уменьшению смещения моментов у реализуемых РТХ.

Наиболее универсальным и эффективным методом генерации СПЧ, устраняющим искажения РТХ, является метод ARAND, предложенный в [10, 116].

Метод ARAND [116] преобразует реализуемые существующим стандартным ГСПЧ n-разрядные равномерно распределенные на [0, 1] числа r с постоянным шагом дискретизации ε в равномерно распределенные на [0, 1] числа r с переменным шагом дискретизации $\varepsilon(r)$, таким, что шаг ε всегда остается примерно пропорциональным значениям r. Благодаря этому, метод ARAND позволяет получать достаточно большие значения реализуемых с.в. ξ^* , необходимые для несмещенной реализации моментов PTX.

Метод ARAND является предельным обобщением предложенного в [111] каскадного метода, устраняющего смещения моментов степенных распределений. Алгоритм реализующей метод ARAND процедуры, написанной автором диссертации для AnyLogic на языке Java, представлен на Рисунке 2.1 (листинг кода программы представлен в Приложении A).

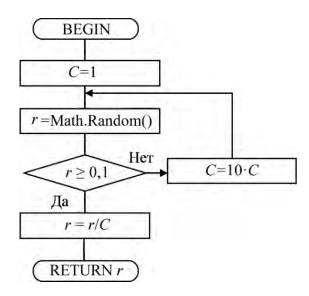


Рисунок 2.1 – Схема алгоритма процедуры ARAND на Java

В базовой версии алгоритма ARAND первое обращение к стандартному ГСПЧ (см. Рисунок 2.1) производится при масштабе C=1 и, если сгенерированное значение r попало в промежуток $0,1 \le r < 1$, то оно содержит n значащих десятичных цифр и возвращается вызывающей программе в качестве реализованного процедурой значения СПЧ. В этом случае шаг дискретизации $\varepsilon = 1/N$: на Java $N \approx 10^{15}$, в GPSS $N=10^6$. Число возможных значений r в этом промежутке составляет 0,9N, т.е. около $9\cdot10^{14}$ для Java и $9\cdot10^5$ для GPSS.

Если при первом обращении к ГСПЧ значение r попадает в диапазон $0 \le r < 0,1$, т.е. содержит менее чем n значащих десятичных цифр, то оно отбрасывается, масштаб C увеличивается на порядок и значение r разыгрывается снова путем второго обращения к ГСПЧ. Новое значение $r \in [0,1]$, но с учетом масштаба C интерпретируется как принадлежащее промежутку $0 \le r < 0,1$. В этом промежутке новое значение r распределено также равномерно, как и отброшенное предыдущее значение. Если это новое значение попадает в промежуток $0,1 \le r < 1$ (с учетом масштаба C интерпретируемый как промежуток $0,01 \le r < 0,1$), то оно содержит n значащих десятичных цифр, делится на C и возвращается вызывающей программе. Деление на C сохраняет n значащих десятичных цифр у числа r, поскольку деление и его результат реализуются в машинной арифметике с плавающей запятой. В этом случае шаг дискретизации $\varepsilon = 1/(10N)$ на порядок уменьшается. Число возможных значений r в промежутке $0,01 \le r < 0,1$ такое же, как и в предыдущем промежутке.

Аналогично происходят переходы r в последующие промежутки, укорачивающиеся на порядок до тех пор, пока соответствующие значения r могут быть представлены в форме с плавающей запятой. При двойной точности вычислений последний вычисляемый промежуток лежит левее промежутка $10^{-301} \le r < 10^{-300}$. Но при равномерном распределении r вероятность дойти до промежутка $10^{-301} \le r < 10^{-300}$, приблизительно равная 10^{-300} , мала настолько, что это событие, происходящее в среднем лишь один раз за 10^{300} обращений к ГСПЧ, согласно пределу Бреммермана-Эшби при любом достижимом быстродействии компьютера не случится и за миллиарды лет. Заметим, что в промежутке $10^{-301} \le r < 10^{-300}$ число возможных значений r составляет по-прежнему 0,9N, как и в промежутке $0,1 \le r < 1$.

Например, в GPSS первые 10 чисел на выходе стандартного ГСПЧ Uniform(1,0,1), равные 0,842366, 0,777717, 0,880991, 0,260493, 0,463553, **0,083898**, **0,022383**, **0,948592**, 0,344245, 0,929906 преобразуются методом ARAND (базовая версия) в числа 0,842366, 0,777717, 0,880991, 0,260493, 0,463553, **0,00948592**, 0,344245, 0,929906.

Из-за потери части чисел исходной последовательности длина периода у ГСПЧ ARAND в среднем в 1/0,9 раз меньше, чем у исходного ГСПЧ Uniform(j,0,1). Период ГСПЧ ARAND, преобразующего выходную последовательность ГСПЧ Uniform(1,0,1), содержит 1 843 200 000 чисел, наименьшее из них равно $6,54026\cdot10^{-11}$ [18]. Генерация СПЧ замедляется методом ARAND в среднем также в 1/0,9 раз.

При выходе из цикла после i-го его прохождения (см. Рисунок 2.1) и деления r на масштаб C СПЧ r равномерно распределено на отрезке $\Omega_i = [10^{-i}, 10^{-(i-1)}], i = 1, 2, \dots$

Вероятность того, что выход из цикла произойдет после i-го прохождения тела цикла, равна $0,9\cdot 0,1^{-(i-1)}$ (см. Рисунок 2.1), т.е. равна длине отрезка Ω_i . Следовательно, вероятность попадания r в любой из отрезков $\{\Omega_i\}$ равна длине соответствующего отрезка. Вероятность попадания в любую часть отрезка Ω_i равна длине этой части. Следовательно, вероятность попадания в любую часть [a,b] отрезка [0,1] равна длине (b-a) этой части. И, следовательно, число r на выходе процедуры ARAND равномерно распределено на отрезке [0,1].

Заметим, что выбор конкретных точек (конкретных значений r) на отрезке [0,1] происходит с вероятностями, зависящими от положения этих точек. Любая из 0.9N точек на отрезке Ω_i выбирается с вероятностью $0.9\cdot0.1^{-(i-1)}/0.9N = 0.1^{-(i-1)}/N$, где типичные значения N в разных программных продуктах лежат в промежутке от 10^6 до 10^{15} . На каждом следующем из промежутков Ω_i плотность выбираемых точек возрастает на порядок.

Выше было показано, что при использовании традиционных ГСПЧ нужная форма *неравномерных* распределений F(t) с.в. ξ достигается за счет подходящей неравномерной расстановки *равновероятных* значений ξ . Теперь можно видеть,

что метод ARAND реализует *равномерное* на отрезке [0, 1] распределение с.в. r за счет подходящего неравномерного распределения *не равновероятных* значений r.

Рассмотренный метод реализации БСВ r с переменным шагом дискретизации $\varepsilon(r)$ приводит к корректной программной реализации РТХ, поскольку, в отличие от традиционных ГСПЧ не отсекает хвосты распределений в области больших значений с.в. И, следовательно, обеспечивает правильную реализацию характеристик СМО, зависящих от моментов распределений [116].

По аналогии с базовой версией метода ARAND может быть построена любая другая его версия, отличающаяся от базовой тем, что отрезок [0, 1] делится в ней на части $[0, \theta)$ и $[\theta, 1]$ при некотором произвольно выбранном $\theta \in (0, 1)$. И далее каждый новый отрезок $[0, \theta)$, в который может попасть очередное значение $r \in [0, 1]$ при i-м прохождении цикла, интерпретируется как увеличенный в $1/\theta^i$ раз отрезок $[0, \theta^i)$. При выходе из цикла (см. Рисунок 2.1) значение r масштабируется делением на $1/\theta^i$, сохраняющем число значащих точных цифр результата за счет его представления в форме числа с плавающей запятой. Наиболее естественно использовать значения $\theta \in \{1/2, 1/3, ..., 1/10\}$, при которых количество сохраняемых точных значащих цифр СПЧ понимается как количество цифр в двоичной, троичной, ..., десятичной системах счисления соответственно.

2.2 Организация последовательных и параллельных прогонов при моделировании СМО с РТХ

При имитационном моделировании систем с РТХ следует учитывать длительные переходные процессы, обусловленные присущими этим системам ДВЗ. Когда переходные процессы практически заканчиваются, т.е. влияние переходных процессов становится практически несущественным, выборочные оценки характеристик СМО с РТХ сходятся к точным значениям стационарных характеристик чрезвычайно медленно [76]. Причем чем ближе показатель фрактальности трафика с к критическому нижнему пределу, равному единице, тем длительнее переходные процессы и тем медленнее в стационарном режиме оценки стационарных характеристик сходятся к точным значениям этих характеристик.

Метод одного длинного прогона при моделировании фрактальных очередей, как правило, оказывается неэкономичным из-за низкой скорости сходимости оценок. Поэтому он может приводить и к принципиально ошибочным заключениям, даже при весьма значительной длине прогона. Так, например, на Рисунке 2.2 представлен хвост эмпирического распределения времени ожидания заявок в системе Ра/Ра/1 с параметрами Парето-распределения интервалов поступления заявок $\alpha_1 = 1, 1, K_1 = 1$ и параметрами Парето-распределения времени обслуживания $\alpha_2 = 1, 1, K_2 = 0, \bar{\tau} = 11$. В этой системе число каналов n = 1, среднее время между приходами заявок $\bar{\tau} = 11$, среднее время обслуживания заявки $\bar{x} = 1,1$. Коэффициент загрузки $\rho = \overline{x} / (n\overline{\tau}) = 0.1$ не превосходит единицы, т.е. система находится в стационарном режиме функционирования. Эмпирическое распределение времени ожидания, хвост которого представлен на рисунке, получено в последовательном прогоне модели по значениям времени ожидания 100 млн. заявок. Спрямление графика на Рисунке 2.2 слева при логарифмическом масштабе по обеим осям координат, высокий коэффициент достоверности R^2 и значительная длина выборки приводят к заключению, что хвост является асимптотически степенным и что параметр $\alpha = 1,188$ хвоста аппроксимирующего уравнения на Рисунке 2.2 оценен с достаточно высокой точностью. И поскольку $\alpha > 1$, то распределение имеет конечное м.о., т.е. среднее время ожидания в моделируемой системе конечно. Но это – неверное заключение.

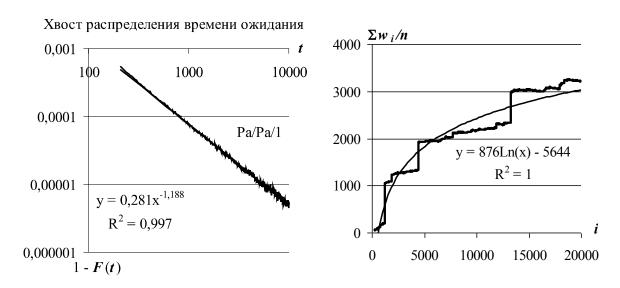


Рисунок 2.2 — Результаты обработки данных последовательного прогона (слева) и параллельных прогонов модели Ра/Ра/1 при $\alpha_1 = \alpha_2 = 1, 1, K_1 = 1, K_2 = 0, 1$

В действительности это не так. Выполнив и обработав несколько довольно коротких прогонов, можно получить наглядную информацию (Рисунок 2.2, справа) о том, что среднее время обслуживания заявок сходится к бесконечности. На этом рисунке показана зависимость от номера i заявки времени ее обслуживания, усредненного по n=10 тыс. независимых реализаций процесса (параллельным прогонам). Стационарный режим в системе существует ($\rho < 1$), но за конечное время не достигается. График роста среднего времени обслуживания имеет характерные скачки, обусловленные тяжелыми хвостами распределений Парето, задающих систему. При этом рост средней величины элементов w_i выборки с ростом их номера i имеет характерную логарифмическую скорость (логарифмическая линия тренда – гладкая кривая на правом рисунке).

Аналогичные ИЭ в AnyLogic также показали, что среднее время ожидания в рассматриваемой системе Pa/Pa/1 является бесконечным. Однако в AnyLogic к такому выводу можно прийти, обработав результаты и одного длинного прогона, что связано с более коротким, чем в GPSS шагом дискретизации $\varepsilon = 10^{-14}$ встроенного ГСПЧ (см. Рисунок А.1). Независимые параллельные прогоны позволяют сэкономить время моделирования. Так, для данной системы на один длинный прогон в 100 млн. заявок затрачено 2321,9 с. реального времени, на n = 10 тыс. параллельных прогонов по 20 тыс. заявок — 469,9 с. Описание имитационной модели в AnyLogic представлено в разделе 5.2.

Вместе с тем использование параллельных прогонов не обеспечивает автоматического получения адекватных результатов моделирования фрактальных очередей, если не решается проблема корректной реализации РТХ. При этом ошибочные выводы могут быть получены не только при $\alpha \downarrow 1$, когда хвосты РТХ «особенно тяжелы», но и при $\alpha \uparrow 2$, когда эти хвосты на интересующем нас интервале $1 < \alpha \le 2$ «наиболее легки». Как показано в [113], моделирование очереди М/Ра/1 c применением обычных стандартных ГСПЧ должно при $\alpha \uparrow 2$ приводить к принципиально ошибочным результатам расчета среднего времени ожидания. На Рисунке 2.3 слева приведен пример результатов такого моделирования системы М/Ра/1 с применением большого числа параллельных прогонов.

Каждая оценка \hat{w}_i на Рисунке 2.3 слева получена усреднением значений случайной величины w_i , полученных при данном конкретном i в большом числе независимых прогонов модели. Без применения метода ARAND моменты реализуемого здесь РТХ смещаются [113], и оценка \hat{w}_i с ростом i действительно сходится к 3,6.

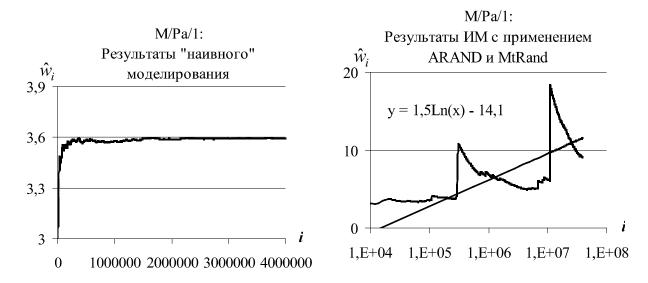


Рисунок 2.3 — Результаты ИМ системы M/Pa/1 при $\bar{\tau} = 4$, K = 1, $\alpha = 2$ в GPSS с использованием параллельных прогонов и встроенным ГПСЧ (слева); с использованием параллельных прогонов и применением ARAND (справа)

В таком же эксперименте получены и оценки, представленные кривой на графике справа, с той лишь разницей, что использовался не обычный стандартный ГСПЧ, а ARAND. Можно увидеть, что теперь зависимость \hat{w}_i от i не сходится ни к какому конечному значению. Кривая на правом рисунке имеет характерные для не сходящейся оценки скачки, которые приводят к росту оценки приблизительно с логарифмической скоростью (масштаб по оси абсцисс на этом рисунке логарифмический). Применение ARAND позволяет увидеть действительное положение дел, состоящее в том, что стационарное среднее время ожидания существует (поскольку коэффициент загрузки системы $\rho = 0,5 < 1$) и что оно равно бесконечности (поскольку оценка \hat{w}_i растет не пропорционально i, а как Ln(i)). Этот пример показывает, что метод параллельных прогонов моделей и метод ARAND комплементарны, дополняют друг друга, и только при совместном использовании обеспечивают реальную возможность расчета фрактальных очередей в системах со степенными хвостами распределений с контролем погрешностей и достижением заданной точности расчетов.

2.3 Доверительные интервалы при моделировании СМО со степенными хвостами распределений

Для получения оценки исследуемого параметра с приемлемой для практики точностью, необходимо тщательно спланировать эксперимент, в ходе которого должна быть получена выборка значений измеряемого показателя, имеющая объем, достаточный для оценки математического ожидания этого показателя с требуемой надежности.

Организация имитационных экспериментов при моделировании классических СМО и СМО со степенными хвостами распределений существенно различается. Такое различие обусловлено видом зависимости, возникающей между случайными величинами на выходе имитационной модели. Длительные переходные процессы, присутствующие в СМО со степенными распределениями с бесконечной дисперсией, приводят к резкому увеличению затрат компьютерного времени на достижение стационарного режима. Например, длительность переходного процесса у оценки вероятности потерь в системе Pa/Pa/1/100 при $\alpha_1 = \alpha_2 = 1, 1, K_1 = 1, K_2 = 0, 5,$ график которого приведен в Приложение А на Рисунке А.3, в десятки тысяч раз превосходит длительность переходного процесса в системе M/M/1/100 с такими же коэффициентом загрузки и размером буфера. Такие длительные переходные процессы обусловлены ДВЗ, возникающими в системах со степенными распределениями. При обработке стационарных выборок, получаемых «отбрасыванием» переходных режимов, ДВЗ приводят к определенным трудностям расчетов доверительных интервалов для оцениваемых по этим выборкам показателям. Например, при использовании одного из наиболее простых методов – метода групповых средних [29, 88, 102] – становится невозможным или весьма затруднительным формирование слабосвязанных групп элементов выборки. В работе [71] также отмечается актуальность разработки новых статистических методов для выборок с ДВЗ. Поэтому разработка новых формул, позволяющих строить доверительные интервалы по выборкам с ДВЗ, является весьма актуальной задачей. Эта задача решается в данном разделе диссертации.

2.3.1 Построение доверительных интервалов в зависимых испытаниях

На практике при анализе данных часто возникает задача статистического оценивания рассматриваемой случайной величины ξ , когда необходимо построить её точечную оценку $\hat{\xi}$ и доверительный интервал, в который она попадает с определенной вероятностью. Для этого строится выборочная ф.р., которая, как правило, является одной из табулированных [35]: t-распределение, F-, нормальное и др. Нормальное распределение с параметрами μ и $\sigma > 0$, является одним из наиболее рас-

пространенных, п.в. которого имеет вид $f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$. Доверительный интервал для нормально распределённой с.в. строится по правилу «трех сигм», которое говорит о том, что с вероятностью p=0,997 с.в. не будет отклоняться по абсолютной величине от своего м.о. больше чем на 3σ [63]. Применение предельной теоремы для сумм зависимых случайных величин возможно при условии слабой зависимости, в качестве которого могут выступать условия сильного перемешивания, равномерно сильного перемешивания, полной регулярности [8]. Наиболее слабым [58] условием сильного перемешивания, представляющим собой «длинную» память, является гиперболическая ограниченность размера $\lambda > 0$, когда значения ξ_N , ξ_{N-k} при $k \to \infty$ становятся независимыми очень медленно. Опишем общий принцип построения доверительных интервалов в экспериментах с зависимыми исходами.

В общем случае при моделировании очередей для расчета оценки $\hat{\xi}$ какоголибо интересующего нас м.о. $\overline{\xi} = M(\xi)$ («показателя») используется выборка ξ_1, \ldots, ξ_N зависимых реализаций случайной величины (с.в.) ξ на выходе имитационной модели. Оценка $\hat{\xi}$ рассчитывается как выборочное среднее:

$$\hat{\xi} = \frac{1}{N} \sum_{i=1}^{N} \xi_i. \tag{2.26}$$

Дисперсию $D(\hat{\xi})$ оценки $\hat{\xi}$ можно выразить через дисперсию $\sigma^2 = D(\xi)$ и коэффициенты корреляции между реализациями $\xi_1, ..., \xi_N$ следующим образом:

$$\begin{split} &\mathbf{D}(\hat{\boldsymbol{\xi}}) = \mathbf{D}\left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\xi}_{i}\right) = \frac{1}{N^{2}}\mathbf{D}\left(\sum_{i=1}^{N}\boldsymbol{\xi}_{i}\right) = \frac{1}{N^{2}}\left(\mathbf{D}\left(\sum_{i=1}^{N}\boldsymbol{\xi}_{i}\right) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}K_{ij}\right) = \\ &= \frac{1}{N^{2}}\left(\sum_{i=1}^{N}\mathbf{D}(\boldsymbol{\xi}_{i}) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}r_{ij}\boldsymbol{\sigma}_{i}\boldsymbol{\sigma}_{j}\right) = \frac{1}{N^{2}}\left(N\boldsymbol{\sigma}^{2} + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}r_{ij}\boldsymbol{\sigma}^{2}\right) = \end{split}$$

$$= \frac{\sigma^2}{N} \left(1 + \frac{2}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij} \right) = \frac{\sigma^2}{N} (1 + 2R_N), \tag{2.27}$$

где $\sigma^2 = D(\xi_i) = D(\xi)$ — дисперсия выборочного значения ξ_i (не зависит от i , так как рассматривается стационарная последовательность реализаций $\xi_1, ..., \xi_N$);

$$r_{ij} = \operatorname{corr}(\xi_i, \xi_j) = \frac{\operatorname{M}(\xi_i \xi_j) - \operatorname{M}(\xi_i) \operatorname{M}(\xi_j)}{\sqrt{\operatorname{D}(\xi_i) \operatorname{D}(\xi_j)}} \quad = \quad \frac{\operatorname{M}(\xi_i \xi_j) - \operatorname{M}^2(\xi)}{\sigma^2} \quad - \quad \text{коэффициент}$$

корреляции элементов ξ_i , ξ_j (зависит только от расстояния s=|j-i| между элементами ξ_i , ξ_j); $R_N=\frac{1}{N}\sum_{j=1}^{N-1}\sum_{j=i+1}^N r_{ij}$ — «коэффициент последействия» (выражает степень

«прямого» влияния реализованных элементов выборки $\xi_1, ..., \xi_N$ на элементы ее продолжения).

Обозначая r_{ij} через r(s), где s = |j-i|, выражение коэффициента R_N можно записать в виде:

$$R_N = \frac{1}{N} \sum_{j=1}^{N-1} \sum_{j=i+1}^{N} r_{ij} = \frac{1}{N} \sum_{s=1}^{N-1} (N-s)r(s), \qquad (2.28)$$

где r(s) — коэффициент корреляции пары элементов выборки, отстоящих друг от друга на s шагов; (N-s) — число таких пар в выборке.

Учитывая, что $|r(s)| \le 1$ и $D(\hat{\xi}) \ge 0$, из (2.27) нетрудно вывести, что коэффициент $(1+2R_N)$ повышения дисперсии лежит в пределах от 0 до N. Как правило, показатель $\bar{\xi}$ при моделировании очередей имеет такой конкретный смысл (как, например, среднее время \bar{w} ожидания заявки), что все коэффициенты корреляции r(s) оказываются положительными, и, соответственно, коэффициент повышения дисперсии $(1+2R_N)$ принимает значение, большее единицы. Поэтому, согласно (2.27), корреляции элементов выборки приводят к увеличению дисперсии

 $D(\hat{\xi}) = \sigma_{\hat{\xi}}^2$ и расширению доверительного интервала (по сравнению с независимыми ξ_i), построенного по правилу трех сигм:

$$\xi = \left(\hat{\xi} \pm 3\sigma_{\hat{\xi}}\right),\tag{2.29}$$

где $\sigma_{\hat{\xi}} = \sqrt{D(\hat{\xi})} = \sigma \sqrt{\left(1 + 2R_N\right)} / \sqrt{N}$, σ – среднеквадратичное отклонение (с.к.о.) элементов ξ_i выборки, $\sqrt{\left(1 + 2R_N\right)}$ – коэффициент расширения доверительного интервала.

Расчет доверительных интервалов (2.29) при больших объемах N выборок значительно упрощается благодаря тому, что асимптотика коэффициентов последействия R_N при расчете классических и фрактальных очередей не отличается разнообразием: для классических очередей она определяется асимптотически экспоненциальным убыванием коэффициентов корреляции r(s) с ростом s, для фрактальных очередей — асимптотически степенным убыванием r(s). Это позволяет по выборкам умеренного объема находить приближенные аналитические выражения, пригодные для построения доверительных интервалов как функций от N и, соответственно, планировать при последовательном прогоне [29] модели такую длину выборки, которая обеспечивает необходимую точность результатов.

2.3.2 Доверительные интервалы в расчетах классических очередей

В качестве примера построения доверительных интервалов при моделировании классических очередей рассмотрим расчет стационарного среднего времени ожидания \overline{w} заявок в системе M/M/1 с интенсивностью входящего потока $\lambda=1$ при коэффициентах загрузки $\rho=0.75$ и $\rho=0.99$.

Статистическую оценку $\hat{r}(s)$ коэффициента корреляции r(s) между временем w_i ожидания i-й заявки и временем $w_i + s$ ожидания (i + s)-й заявки можно рассчитать по стационарной выборке w_1, \ldots, w_N достаточно большого объема N, применяя формулу

$$\hat{r}(s) = \frac{\hat{M}(w_i w_{i+s}) - \hat{M}^2(w)}{\hat{D}(w)},$$
(2.30)

где
$$\hat{\mathbf{M}}(w_i w_{i+s}) = \frac{1}{N-s} \sum_{i=1}^{N-s} w_i w_{i+s}$$
 , $\hat{\mathbf{M}}(w) = \frac{1}{N} \sum_{i=1}^N w_i$, $\hat{\mathbf{D}}(w) = \hat{\mathbf{M}}(w^2) - \hat{\mathbf{M}}^2(w)$, $\hat{\mathbf{M}}(w^2) = \frac{1}{N} \sum_{i=1}^N w_i^2$.

Оценки $\hat{r}(s)$ коэффициентов корреляции $r(s) = r(w_i, w_{i+s})$, рассчитанные по формуле (2.30) при длине прогона модели N=10 млн. заявок, показаны в виде графиков на построенном средствами Excel Рисунке 2.4. Слева показана зависимость r(s) при коэффициенте загрузки $\rho=0.75$, справа – при $\rho=0.99$.

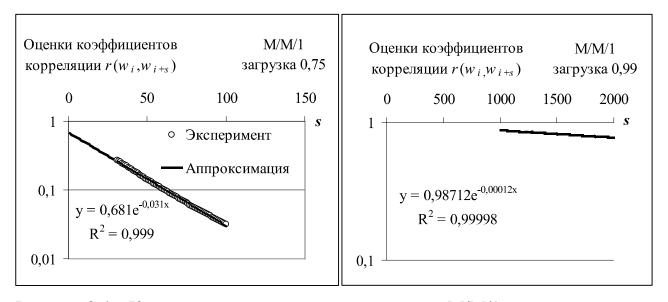


Рисунок 2.4 — Корреляция величин w_i и w_{i+s} в системе M/M/1 в зависимости от s

В общем случае в классических системах коэффициенты корреляции r(s) между двумя сдвинутыми на s шагов элементами выборки $w_1, ..., w_N$ описываются выражением

$$r(s) \sim ae^{-bs}. (2.31)$$

Отклонения функции r(s) от экспоненты (2.31) быстро уменьшаются с ростом s и на практике ими можно пренебречь. Учитывая (2.28) и (2.31), коэффициент последействия R_N можно при больших N выразить следующим образом:

$$R_{N} = \frac{1}{N} \sum_{s=1}^{N-1} (N-s)r(s) \sim \frac{1}{N} \sum_{s=1}^{N-1} (N-s)ae^{-bs} \sim \frac{a}{N} \int_{1}^{N-1} (N-s)e^{-bs} = \dots$$

$$= \frac{a}{Nb} \left\{ e^{-b(N-1)} \left(\frac{1}{b} - 1 \right) + e^{-b} \left(N - 1 - \frac{1}{b} \right) \right\} \rightarrow \frac{a}{b} e^{-b}. \tag{2.32}$$

Для прогонов длины N = 100 тыс. заявок и более при тех a, b, что представлены на Рисунке 2.4, можно использовать последнее (предельное) выражение из (2.32). Соответственно, при $\rho = 0,75$ получаем $R_N \approx 22$ и доверительный интервал (имеющий вероятность 0,997):

$$\overline{w} = \hat{w} \pm 3\sigma \sqrt{1 + 2R_N} / \sqrt{N} = \hat{w} \pm \frac{3\hat{\sigma}}{\sqrt{N}} \cdot 6,7, \qquad (2.33)$$

где \hat{w} и $\hat{\sigma}$ — выборочные оценки м.о. и с.к.о. времени ожидания заявки в стационарном режиме. Аналогично для $\rho = 0.99$ по формуле (2.32) находим, что при больших N коэффициент $R_N \approx 8000$ и доверительный интервал имеет вид:

$$\overline{w} = \hat{w} \pm \frac{3\hat{\sigma}}{\sqrt{N}} \cdot 130. \tag{2.34}$$

Выражения (2.33) и (2.34) доверительных интервалов подтверждаются дополнительными проверками.

В частности, при моделировании СМО М/М/1 с коэффициентом загрузки $\rho=0.99$ и $\lambda=1$ в результате 10 прогонов в АпуLogic длиной N=100 млн. заявок получены следующие оценки среднего времени ожидания: $\hat{w}_1=98.10$, $\hat{w}_2=97.29$, $\hat{w}_3=99.74$, $\hat{w}_4=97.47$, $\hat{w}_5=96.81$, $\hat{w}_6=99.41$, $\hat{w}_7=100.84$, $\hat{w}_8=96.55$, $\hat{w}_9=96.17$, $\hat{w}_{10}=100.71$; и соответствующие оценки с.к.о: $\hat{\sigma}_1=98.86$, $\hat{\sigma}_2=96.96$, $\hat{\sigma}_3=101.95$, $\hat{\sigma}_4=98.19$, $\hat{\sigma}_5=95.99$, $\hat{\sigma}_6=100.38$, $\hat{\sigma}_7=102.13$, $\hat{\sigma}_8=95.53$, $\hat{\sigma}_9=98.71$, $\hat{\sigma}_{10}=101.67$. Доверительные интервалы, построенные с помощью данных оценок, накрывают точное значение среднего времени ожидания, которое известно для системы $M/M/1-w=\rho^2/(1-\rho)/\lambda=98.01$ (см. Рисунок 2.5). Аналогичную проверку практическим моделированием прошли доверительные интервалы, построенные для случаев $\rho=0.90$ и $\rho=0.75$. Листинг кода и результаты ИМ системы M/M/1 при $\rho=0.75$ в среде АпуLogic приведены в Приложении А.

Принципиально важным является вывод, что при расчете очередей с РЛХ корреляция элементов выборки на выходе модели не влияет на скорость сходимости оценок к искомым точным значениям показателей. Погрешность оценки (половина длины доверительного интервала) уменьшается пропорционально $N^{0,5}$, и для

того, чтобы погрешность уменьшить в k раз, длину N прогона модели нужно увеличивать в k^2 раз.

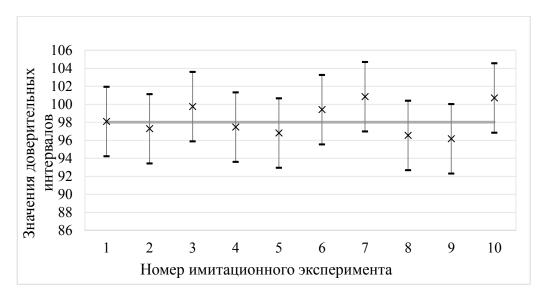


Рисунок 2.5 – Доверительные интервалы для M/M/1 при $\rho = 0.90$ и $\lambda = 1$

В то же время высокая корреляция элементов выборки может при одних и тех же требованиях к точности приводить к необходимости многократного увеличения длины выборки по сравнению со случаем независимых элементов.

2.3.3 Доверительные интервалы в расчете фрактальных очередей

При расчете СМО со степенными распределениями коэффициенты r(s) корреляции элементов ξ_i обрабатываемых выборок являются асимптотически степенными функциями от s. Это приводит к существенному отличию методов планирования экспериментов с фрактальными очередями по сравнению со случаем классических систем, где соответствующая асимптотика экспоненциальная. На Рисунке 2.6 показан график функции r(s), рассчитанной путем ИМ системы Pa/Pa/1 в стационарном режиме ее функционирования, наступающем после прохождения через систему 40 млн. заявок. Выборочными элементами здесь являются реализации $\gamma_i \in \{0,1\}$ индикаторов отказа заявок с номерами $i=1,2,\ldots$ Среднее значение $\bar{\gamma}$ индикаторов в стационарном режиме равно искомой вероятности P потери заявки. Длительность переходного процесса определялась посредством достаточно большого числа независимых прогонов модели.

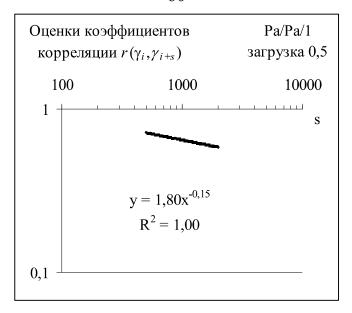


Рисунок 2.6 — Асимптотика коэффициентов корреляции r(s) между индикаторами отказа i-й и (i+s)-й заявок в системе Pa/Pa/1/100 при $\alpha_1=\alpha_2=1,1,\,K_1=1,\,K_2=0,5$ и размере буфера m=100

В общем случае при расчете любых показателей типа вероятности отказа, среднего времени ожидания и т.д. функция r(s) имеет при расчете фрактальных очередей следующий вид:

$$r(s) = as^{-b}$$
, (2.35)

где a и b — некоторые константы, определяемые с помощью пробных прогонов модели (a > 0 и 0 < b < 1).

Используя точную формулу (2.28) и асимптотическое представление (2.35), для доверительных интервалов находим:

$$R_{N} = \frac{1}{N} \sum_{s=1}^{N-1} (N-s)r(s) = \frac{1}{N} \sum_{s=1}^{N-1} (N-s)as^{-b} \sim \frac{1}{N} \int_{1}^{N-1} (N-s)as^{-b}ds =$$

$$= \frac{a}{1-b} (N-1)^{1-b} - \frac{a}{1-b} - \frac{a}{N} \frac{1}{2-b} (N-1)^{2-b} + \frac{a}{N} \frac{1}{2-b} \sim$$

$$\sim \left(\frac{a}{1-b} - \frac{a}{2-b}\right) N^{1-b}. \tag{2.36}$$

Поэтому, в соответствии с (2.29), доверительный интервал запишется в виде

$$\overline{\xi} = \hat{\xi} \pm 3\sigma \left(\frac{1 + 2R_N}{N}\right)^{1/2} \sim \hat{\xi} \pm 3\sigma \cdot CN^{-b/2}, \qquad (2.37)$$

где константа
$$C = \sqrt{2a\left(\frac{1}{1-b} - \frac{1}{2-b}\right)}$$
.

Как видим, величина полуинтервала убывает здесь медленнее, чем $N^{-1/2}$, (поскольку b < 1). Например, при расчете вероятности P потери заявки в системе, характеризуемой Рисунком 2.6, параметр b в (2.37) равен 0,15 (см. уравнение линии тренда на Рисунке 2.6), и, следовательно, величина полуинтервала убывает пропорционально $N^{-0,075}$. Иными словами, чтобы уменьшить погрешность (полуинтервал) в 10 раз, длину прогона N нужно увеличивать в $10^{1/0,075} \approx 10^{13}$ раз. Чаще всего такая цена повышения точности бывает практически неприемлемой, и этим обусловлена высокая актуальность разработки ускоренных методов расчета оценок вероятностей потерь (см. главу 3) при моделировании систем со степенными распределениями.

Принципиальной при любом b здесь является следующая особенность погрешностей оценок: скорость уменьшения погрешностей с ростом N всегда меньше, чем $cN^{-1/2}$. В случае же классических очередей корреляция элементов выборок приводит лишь к масштабному увеличению полуинтервала, но скорость его уменьшения с ростом N всегда остается равной $cN^{-1/2}$, как в случае независимых параллельных прогонов.

На практике в зависимости от требуемой точности необходимые формулы (2.32) и (2.36) могут быть скорректированы в зависимости от N.

Асимптотика коэффициентов корреляции r(s) между индикаторами отказов заявок для некоторых других фрактальных СМО вида M/Pa/1/100 приведена в Таблице 2.5. Следует отметить, что аналогичная асимптотика индикаторов отказа для систем вида Pa/M/1/m — экспоненциальная для любого размера буфера m, что согласуется с результатами, полученными в разделе 3.1.

Таблица 2.5 – Асимптотика коэффициентов корреляции r(s) между индикаторами отказа γ_i и γ_{i+s} для фрактальных СМО при ρ =0,5

СМО	Параметры	r(s)	Коэффициент детерминации R^2
M/Pa/1/100	$\lambda = 0.04545, K = 1, \alpha = 1.1$	$r(s) \sim 3.375x^{-0.257}$	$R^2 = 0.997$
	$\lambda = 0.16667, K = 1, \alpha = 1.5$	$r(s) \sim 3,245x^{-0,316}$	$R^2 = 0.999$

Выводы

1. Выполнено исследование реализуемых в GPSS World PTX на примере реализаций типичных представителей данного класса, таких, как распределение Парето, Вейбулла и Бурра. Показано, что смещение моментов PTX обусловлено исчезновением в генерируемой выборке тех маловероятных значений с.в., которые слишком велики для того, чтобы ими можно было пренебречь. Их исчезновение вызвано недостаточным шагом дискретизации ГСПЧ, встроенных в GPSS. Подтверждающий расчеты имитационный эксперимент проводился с помощью метода стратификации, который значительно ускорил расчет условных математических ожиданий внутри каждой страты и соответствующего безусловного математического ожидания [23].

Выявленные при моделировании РТХ проблемы можно устранить с помощью методов корректной реализации РТХ, рассмотренных в главе. Их более подробное описание можно найти в [10, 12, 18, 111, 114, 116].

- 2. Обоснована необходимость систематического использования многократных параллельных независимых прогонов моделей систем рассматриваемого класса. При этом необходимо использовать и методы корректной реализации РТХ, наиболее универсальным и эффективным из которых является метод ARAND. Кроме того, параллельные прогоны позволяют решать такие актуальные для моделирования фрактальных очередей задачи, как определение длительности переходного процесса и распознавание отсутствия стационарного режима функционирования у моделируемой системы, что является актуальным при моделировании СМО и СеМО со степенными хвостами распределений.
- 3. Предложена методика расчета асимптотических доверительных интервалов, которая позволяет планировать ИЭ с заданной точностью расчетов. Показано, что принципиальное различие в моделировании классических и фрактальных очередей заключается в том, что скорость сходимости оценок СМО со степенными хвостами является медленной из-за степенной асимптотики коэффициентов корреляции r(s) между элементами выборок. В то время как при моделировании классических СМО и СеМО доверительные интервалы убывают пропорционально квадратному корню из объема N накопленной выборки (как $cN^{-1/2}$), типичная для моделирования фрактальных очередей скорость убывания может выражаться, например, функцией $c'N^{-1/13}$.

Глава 3 Методы расчета и уменьшения вероятности потерь заявок в системах со степенными хвостами

В данной главе решается проблема повышения качества информационного обслуживания в сетях передачи данных с фрактальным трафиком за счет разработки метода, обеспечивающего низкую вероятность потерь сообщений. В терминах теории массового обслуживания формулируется и решается задача структурной оптимизации сетей с фрактальным трафиком (раздел 3.2). Разрабатывается ускоренный метод расчета вероятности потерь в системах с очередями (раздел 3.1), который используется для анализа фрактальных систем. Материал данной главы основан на опубликованных автором статьях [15, 24, 25, 115, 117].

Потери пакетов в сетях передачи данных

Процесс потери пакетов в сетях передачи данных обусловлен внутренним устройством пакетных коммутаторов, входящих в состав сетей передачи данных. Для простоты, процесс потери пакетов в сетях будет рассматриваться на примере маршрутизаторов, которые также являются пакетными коммутаторами. Архитектура маршрутизатора условно состоит из входных и выходных интерфейсов, совмещенных в случае двунаправленной связи (когда данные могут передаваться в обоих направлениях), коммутационной фабрики и процессора маршрутизации. Входные и выходные интерфейсы состоят из портов, в которых содержатся буферы для хранения пакетов, ожидающих либо перенаправления из входного в выходной порт, либо их передачи из выходного порта [37, с. 356].

В настоящее время наблюдается процесс конвергенции компьютерных и телекоммуникационных сетей, что приводит к созданию единой мультисервисной сети [36, 43, 54], которая предоставляет возможность передачи разнородного трафика. Современные мультисервисные сети строятся на основе принципа коммутации пакетов (когда необходимые для передачи сообщения разбиваются на так называемые «пакеты», которые далее передаются по сети). Отличительной особенностью коммутаторов в сетях с коммутацией пакетов является наличие в них буферов для хранения пакетов [43]. Использование буферов позволяет регулировать скорость продвижения данных по линиям связи, а также предотвратить потери пакетов. Однако существуют

ситуации, когда при высокой загрузке буфер переполняется и происходят потери пакетов.

На сетевом уровне [37] данные, передающиеся по сети, представляют собой пакеты (дейтаграммы). Вероятность потерь пакетов является одной из важнейших характеристик, влияющей на качество обслуживания (QoS) современных мультисервисных сетей [9, 89], в которых наблюдается тенденция увеличения объема трафика [73]. Поэтому обеспечение требуемой низкой вероятности потерь или даже сведение её к нулю является важной задачей при проектировании сетей.

3.1 Расчет вероятностей потерь во фрактальных системах ускоренным методом

В данном разделе разрабатывается ускоренный метод расчета вероятностей потерь P во фрактальных СМО вида GI/GI/n/m при $n,m < \infty$, основанный на регенеративном моделировании СМО. Ускоренный метод позволяет получить оценки вероятностей потерь с высокой точностью для сразу нескольких значений m СМО с конечным буфером (КБ) за один прогон модели соответствующей системы с бесконечным буфером (ББ). Данный метод применим при маленькой вероятности потерь (порядка 10^{-6} и ниже).

Регенеративный метод моделирования подразумевает наличие временных точек – *точек регенерации* β_i , в которых вновь начинается рассматриваемый случайный процесс. Между точками регенерации случайные процессы являются идентичными в вероятностном смысле, т.е. на каждом периоде регенерации (ПР). Каждый ПР начинается с периода непрерывной занятости, за которым следует период простоя. При использовании регенеративного моделирования можно получить строго состоятельную оценку и асимптотически точный доверительный интервал (при конечной дисперсии) для среднего значения исследуемой величины f(X) от регенерирующего процесса X, рассматривая одну реализацию модели [27]. Такой способ расчета оценки исследуемых параметров быстрее классического метода, при котором необходимо принимать во внимание переходный процесс и учитывать наблюдения, полученные только в стационарном режиме. Поэтому при моделировании фрактальных СМО, где остро стоит проблема длительных переходных процессов [76], регенеративный метод

[27, 42, 46, 69, 85, 96] позволяет избежать данную проблему. Случайные процессы, протекающие в произвольном ПР, не зависят от процессов, которые происходят в пределах другого ПР, поэтому при расчете вероятности потерь P, сумма числа потерянных заявок C_i , соответствующих i-му ПР, является суммой независимых одинаково распределенных случайных величин, что исключает ДВЗ, а сам процесс изначально является стационарным, т.е. отсутствует переходный процесс.

При классическом регенеративном моделировании оценка \hat{P} вероятности потерь P рассчитывается как отношение числа потерь к общему числу пришедших заявок:

$$\hat{P} = \sum_{i=1}^{n} C_i / \sum_{i=1}^{n} N_i , \qquad (3.1)$$

где C_i — число потерь, а N_i число пришедших заявок на i-ом периоде регенерации. В разрабатываемом ускоренном методе предлагается использовать не сами значения C_i , а их оценки \hat{C}_i , которые позволяют получить за один прогон имитационной модели значения вероятности потерь для произвольного диапазона значений размера буфера m. Такой подход позволяет быстро определять необходимый размер буфера m для обеспечения заданной вероятности потерь, а также исследовать зависимость P(m) для различных видов фрактальных СМО.

3.1.1 Ускоренный метод расчета вероятности потерь

В ускоренном методе для расчета числа потерянных заявок используется следующее приближенное соотношение [11]:

$$\hat{C}_i \sim \max\{Q_{\max i} - m, 0\},\tag{3.2}$$

 $Q_{\max i}$ — максимальная длина очереди на i-ом ПР в СМО с ББ. Таким образом, данная формула позволяет найти \hat{C}_i для малых значений P для систем с КБ объема m, используя значения оценок числа потерь \hat{C}_i , полученных при моделировании систем с ББ. Формула (3.2) является приближенной, но при малых значениях вероятности P порядка 10^{-2} её относительная погрешность составляет 1...3 % [11].

Ускорение данного метода заключается в том, что можно за один прогон модели получить число отказов сразу для нескольких значений m, используя полученную при моделировании для системы с ББ последовательность $Q_{\max i}$. Для этого

лишь необходимо поставить соответствующие значения m в (3.2). Такой подход сокращает машинное время на получение зависимости P(m) на два-три порядка.

Основанный на (3.2) ускоренный метод был реализован в виде программы Metod_AIM в среде GPSS и AnyLogic (см. раздел 5.3).

3.1.2 Проверка точности ускоренного метода

Широко известной (традиционной) оценкой для P, получаемой путем моделирования системы с ББ, является вероятность q_m того, что длина очереди (в системе с ББ) превысит m, например [90]. При проверке точности разработанного ускоренного метода выполнялось также сравнение точности оценки (3.2) с точностью традиционной оценки.

С целью проверки точности целесообразно сравнить решение, получаемое с помощью программы Metod_AIM, с точно известным решением. Поэтому для проверки точности решалась задача расчета вероятности потерь в системе M/M/n/m единственной системе в классе GI/GI/n/m, для которой известно точное решение. На Рисунке 3.1 показаны результаты расчета зависимости P(m) в системе M/M/1/m с коэффициентом загрузки $\rho = 0.8$.

Точное решение определено расчетом P по известной для систем M/M/1/m формуле:

$$P = \rho^{m+1} \frac{1 - \rho}{1 - \rho^{m+2}}. (3.3)$$

Точное решение изображено на Рисунке 3.1 кривой 1. Кривая 2 – это решение, полученное программой Metod_AIM. Решение, полученное применением традиционной оценки (кривая 3), рассчитано без моделирования системы с ББ. Вероятность q_m того, что длина l очереди будет больше, чем m, найдена как сумма известных вероятностей состояний системы M/M/1/m:

$$q_m = p(l > m) = \sum_{k=m+1}^{\infty} p_k = \sum_{k=m+1}^{\infty} (1 - \rho) \cdot \rho^k = \rho^{m+1},$$
 (3.4)

где p_k – вероятность того, что в системе находится k заявок

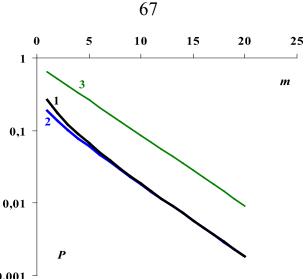


Рисунок 3.1 – Результаты проверки точности ускоренного метода: кривая 1 — точная зависимость P(m) в системе M/M/1/m с коэффициентом загрузки 0,8; 2 – зависимость, полученная выполнением программы Metod_AIM; 3 – зависимость, полученная применением традиционной оценки

Аналогично выглядят результаты расчета зависимости P(m) в системах M/M/2/m и M/M/3/m с коэффициентом загрузки 0,8. Результаты расчета системы M/M/3/m представлены на Рисунке 3.2.

Для расчета вероятности P в системах M/M/n/m при n > 1 использовалась известная формула:

$$P = \rho^{m+n} \frac{n^n}{n!} \pi_0,$$
пде
$$\pi_0 = \left[\sum_{j=0}^{n-1} \frac{(n\rho)^j}{j!} + \frac{(n\rho)^n}{n!} \cdot \frac{1-\rho^{m+1}}{1-\rho} \right]^{-1}.$$

Вероятности p_k состояний в соответствующих системах с ББ, в которых используется традиционная оценка (т.е. в системах $M/M/n/\infty$), определялись по формулам [30]:

$$p_k = \begin{cases} p_0 \frac{(n\rho)^k}{k!}, k \le n, \\ p_0 \frac{\rho^k n^n}{n!}, k \ge n, \end{cases}$$
 (3.6)
 где $p_0 = \left[\sum_{j=0}^{n-1} \frac{(n\rho)^j}{j!} + \frac{(n\rho)^n}{n!} \cdot \frac{1}{1-\rho} \right]^{-1}.$

Соответственно, для расчета традиционной оценки q_m при $m \ge n$ получаем формулу:

$$q_m = p(l > m) = \sum_{k=m+1}^{\infty} p_k = \sum_{k=m+1}^{\infty} p_0 \frac{n^n}{n!} \rho^k = p_0 \frac{n^n}{n!} \frac{\rho^{m+1}}{1 - \rho}.$$
 (3.7)

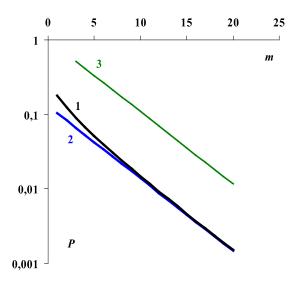


Рисунок 3.2 — Результаты расчета разными методами системы M/M/3/m с $\rho = 0.8$ (обозначения см. на Рисунке 3.1)

Выполненная проверка позволяет сделать вывод, что, во-первых, точность разработанного ускоренного метода является достаточной, и, во-вторых, применение оценки (3.2) вместо традиционной оценки является обоснованным. Эта проверка позволяет также составить представление о том, как быстро с убыванием P относительная погрешность расчета P этим методом сходится к нулю.

Далее естественно воспользоваться разработанным методом для исследования фрактальных систем, для расчета P в которых аналитические выражения отсутствуют, а расчет зависимостей P(m) непосредственным моделированием слишком трудоемок.

3.1.3 Расчет фрактальных систем М/Ра/n/m

Рассмотрим результаты расчета одноканальной системы M/Pa/1/m, параметры которой определены следующим образом: $\overline{\tau} = 10$, K = 1, $\alpha = 1,25$. В этой системе среднее значение распределенного по закону Парето времени обслуживания ко-

нечно: $\overline{x} = K\alpha / (\alpha - 1) = 5$, дисперсия времени обслуживания бесконечна и коэффициент загрузки системы $\rho = \overline{x}/\overline{\tau} = 0,5$. Для размера m буфера выбран диапазон от 1 до 100. Пятикратный прогон программы Metod_AIM по 500 000 ПНЗ дает 5 траекторий — оценок зависимости P(m) (см. Рисунок 3.3), м.о. которых является интересующей нас точной зависимостью. При этом, говоря «точная зависимость», учитываем, что ее точность ограничена той небольшой погрешностью оценки (3.2), которой при малых P можно пренебречь.

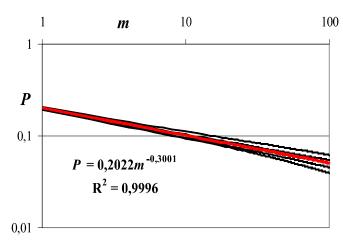


Рисунок 3.3 – Пять траекторий, полученных при пятикратном прогоне программы Metod_AIM, усредненная оценка для P(m) (по центру) и уравнение регрессии с коэффициентом его достоверности \mathbb{R}^2

Показанное на Рисунке 3.3 уравнение регрессии можно использовать для решения задачи определения m, обеспечивающего требуемую вероятность потерь P. Имея уравнение $P = 0.2022 m^{-0.3001}$, ы получаем, решая его относительно m, выражение:

$$m = \left(\frac{P}{0,2002}\right)^{-\frac{1}{0,3001}} = 4,995P^{-3,332},\tag{3.8}$$

которое можно использовать для оценки размера m буфера, обеспечивающего заданную вероятность потерь P, не прибегая к моделированию. Например, задавая $P = 10^{-6}$, из (3.8) получаем, что для обеспечения такой P требуется буфер размера $m \approx 5 \cdot 10^{17}$. Разумеется, буфер подобного размера технически нереализуем, а если бы и был реализуем, то задержки заявок в таком буфере были бы безумно велики. Заметим, что при пуассоновском входящем потоке с той же интенсивностью, как

следует из (3.3), достаточен буфер размера m = 18.

Этот пример показывает, что в общем случае при фрактальном трафике стратегия обеспечения малой вероятности потерь путем увеличения размера буфера может быть бесперспективной. Фрактальную систему, в которой обратная зависимость m(P), подобно зависимости (3.3), делает стратегию увеличения буфера технически бесперспективной, далее будем называть жесткой системой. Чем медленнее убывает вероятность P с ростом m, тем система жестче. Высокая жесткость рассмотренной системы M/Pa/1/m обусловлена близостью ее параметра α к нижней границе диапазона $1 < \alpha \le 2$. Параметр α определяется трафиком, поэтому его называют также параметром трафика.

В отношении предлагаемого ускоренного метода заметим, что возможность построения надежных уравнений регрессии, основанных на большом числе точек траекторий, и их использования для оценки зависимостей P(m) при малых порядка $P = 10^{-6}$ вероятностях решает в рамках рассматриваемой задачи классическую для моделирования проблему оценки малых вероятностей.

Перейдем теперь к системам M/Pa/n/m при n > 1.

На Рисунке 3.4 показаны шесть траекторий P(m), полученные для системы M/Pa/3/m, и уравнение регрессии. По-прежнему хорошей аппроксимацией зависимости P(m) является степенное уравнение. Увеличение числа каналов в системе привело к увеличению абсолютной величины показателя степени, уравнение регрессии приняло вид $P=0.0309m^{-0.5651}$, и решением задачи о буфере стал его размер $m(P)=m(10^{-6})\approx 8.8\cdot 10^7$. Размер буфера уменьшился на десять порядков.

Рассчитанная на Рисунке 3.4 система M/Pa/3/m сопоставима с исследованной выше одноканальной системой M/Pa/1/m. Здесь $\overline{\tau}=10$, K=3, $\alpha=1,25$, т.е. параметр формы распределения Парето не изменился. Произошло лишь масштабное преобразование времени обслуживания, и среднее значение этого времени стало равным $\overline{x}=K\alpha/(\alpha-1)=15$. Благодаря этому коэффициент загрузки $\rho=(\overline{x}/\overline{\tau})/n$ остался равным 0,5.

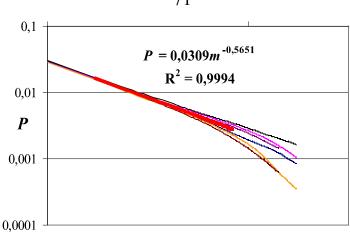


Рисунок 3.4 — Траектории P(m) и уравнение регрессии для системы M/Pa/3/m

Таким образом, выполненный с помощью разработанного ускоренного метода приближенный анализ моделей M/Pa/n/m приводит к следующему заключению. В условиях жесткого фрактального трафика, когда его параметр α близок к единице (на практике он ограничивается снизу значением $\alpha = 1,2$), стратегия обеспечения малой вероятности потерь за счет увеличения только объема буфера одноканальной системы становится бесперспективной. Более реалистичной она становится в сочетании со стратегией увеличения числа каналов в системе.

3.1.4 Влияние параметра формы а

Для оценки влияния параметра α на ресурсы, необходимые для обеспечения низкой вероятности потерь, его значение 1,25 можно заменить менее жестким значением 1,75. При этом для сохранения коэффициента загрузки $\rho = 0,5$ в системе M/Pa/1/m значение масштабного коэффициента K = 1 следует заменить значением K = 15/7, и, соответственно, в системе M/Pa/3/m значение K = 3 — значением K = 45/7. В результате прогона программы Metod_AIM для этих систем получаются следующие результаты, существенно отличающиеся от предыдущих.

Теперь для того, чтобы обеспечить в системе M/Pa/1/m вероятность потерь, равную 10^{-6} , требуется буфер размера $m \approx 700~000$. Это в определенных условиях может быть вполне приемлемым для сетевых устройств решением.

В системе M/Pa/3/m для решения той же задачи теперь требуется буфер размера $m \approx 400$.

Сравнивая полученные решения с предыдущими, нетрудно заметить, что у многоканальной системы характеристика m(P) значительно менее чувствительна к изменениям параметра α , чем у одноканальной системы.

3.1.5 Другие фрактальные системы

Для того чтобы лучше понимать степень общности тех или иных результатов, имеет смысл проверять их на различных фрактальных системах, которые могут рассматриваться как различные модели проектируемых сетевых устройств.

С помощью разработанного ускоренного метода можно увидеть, например, что в системах Pa/Pa/n/m с ростом n подходящая степенная аппроксимация зависимости P(m) быстро переходит в более подходящую экспоненциальную (в отличие от систем M/Pa/n/m). В качестве иллюстрации на Рисунке 3.5 показан усредненный результат пятикратного прогона программы $Metod_AIM$ для системы Pa/Pa/5/m, заданной следующими параметрами: $K_1 = 10/3$, $\alpha_1 = 1,5$ (параметры с.в. τ), $K_2 = 75/7$, $\alpha_2 = 1,75$ (параметры с.в. x). В этой системе $\overline{\tau} = 10$, $\overline{x} = 25$, $\rho = 0,5$.

На Рисунке 3.5 логарифмический масштаб используется только по оси ординат. Прямолинейный вид графика (асимптотически прямолинейный) в такой системе координат свидетельствует об экспоненциальном (асимптотически экспоненциальном) виде изображаемой графиком зависимости.

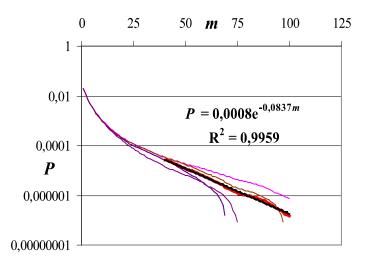


Рисунок 3.5 — Асимптотически-экспоненциальная зависимость P(m) в системе Pa/Pa/5/m

А в системах Pa/M/n/m характеристика P(m) похожа на экспоненциальную функцию при любом $n \ge 1$. На Рисунке 3.6 представлены результаты пяти прогонов программы $Metod_AIM$ для системы Pa/M/1/m, сопоставимой с предыдущими системами по загрузке и параметру формы.

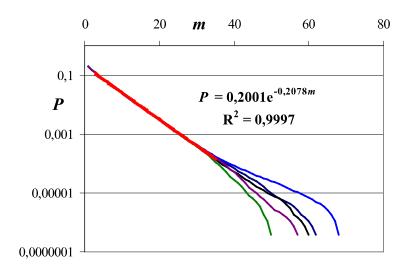


Рисунок 3.6 - Экспоненциальная зависимость <math>P(m) в системе Pa/M/1/m

Параметры системы имеют следующие значения: $K=2,~\alpha=1,25,~\overline{x}=5$. В этой системе $\overline{\tau}=10$, $\rho=0,5$.

3.1.6 Предварительная оценка точности и коэффициента ускорения метода при моделировании фрактальных систем

Оценку точности разработанного метода, выполненную в разделе 3.1.2 на примере экспоненциальных систем, нельзя механически распространять на фрактальные системы, в формирование характеристик которых вносят вклад последовательности случайных величин с бесконечными дисперсиями. Для фрактальных систем необходимо выполнить дополнительную проверку точности ускоренного метода. Используем для такой проверки систему M/Pa/1/m, рассмотренную в разделе 3.1.3 наиболее жесткую из рассчитанных систем. На Рисунке 3.7 красная прямая линия (расположенная между двумя другими сплошными линиями) является графиком усредненной траектории P(m) для указанной системы. Этот график визуально совпадает с линией тренда, уравнение которой также представлено на рисунке (эти дан-

ные имеются и на Рисунке 3.3). На получение сотни точек этой траектории ускоренным методом потребовалось 20 минут времени персонального компьютера. Усреднялись 5 траекторий, полученных по 500 000 ПНЗ. Для предварительной оценки точности и коэффициента ускорения разработанного метода выполнено непосредственное моделирование системы при $m \in \{1, 25, 50, 75, 100\}$. Для каждого из этих m выполнено по пять прогонов длиной в 500 000 ПНЗ. Полученные при каждом из перечисленных m значения p_l отмечены на Рисунке 3.7 пятью маркерами разной формы.

Расположение маркеров на рисунке подтверждает правильность траектории, найденной ускоренным методом. Синяя «провисающая вниз» линия на графике проведена через усредненные для каждого m результаты непосредственного ИМ.

На каждый результат-маркер затрачено 6 секунд компьютерного времени, на одну точку усредненной траектории — 30 секунд. На 100 точек усредненной траектории потребовалось бы 3000 секунд, т.е. 50 минут. Это незначительно — в два с половиной раза — превышает время, потраченное ускоренным методом. Однако, как видно из Рисунка 3.7, при непосредственном ИМ системы будет при этом получена траектория с большими стохастическими возмущениями. При сопоставимых затратах времени ускоренный метод демонстрирует значительно более высокую точность.

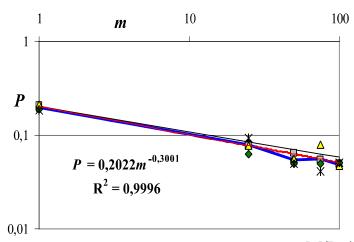


Рисунок 3.7 – Сравнение результатов расчета системы M/Pa/1/*m*, полученных ускоренным методом и непосредственным ИМ

Верхняя сплошная тонкая линия на Рисунке 3.7 проведена через три точки P(m), рассчитанные при m=1, m=75 и m=100 путем непосредственного ИМ с дли-

ной прогона 50 млн. ПНЗ. В каждой точке усреднялись результаты пяти таких длинных прогонов. Эти усредненные результаты уже могут конкурировать по точности с результатами ускоренного метода. Время компьютера, затраченное на одну усредненную точку, составляет 50 минут. На 100 точек требуется 5000 минут, т.е. в 250 раз больше, чем при использовании ускоренного метода. Таким образом, выполненная предварительная оценка разработанного метода позволяет рассчитывать, что при моделировании фрактальных систем метод позволяет достичь заданной точности результатов на два-три порядка быстрее, чем непосредственное ИМ.

Таким образом, выполненное исследование зависимости P(m) для фрактальных систем различного вида показало, что вероятность потерь можно эффективно снижать за счет увеличения числа каналов, особенно в системах вида M/Pa/1/m и Pa/Pa/1/m, где зависимость P(m) близка к степенной.

3.2 Метод уменьшения вероятностей потерь пакетов в сетях с фрактальным трафиком

C учетом полученных в разделе 3.1 результатов исследования зависимостей вероятности потерь P от размера буфера m, в данном разделе решается проблема повышения качества информационного обслуживания в сетях передачи данных с фрактальным трафиком за счет разработки метода, обеспечивающего низкую вероятности потерь сообщений.

В [11] показано, что при фрактальном трафике стратегия борьбы с потерями заявок за счет увеличения буферов и/или ускорения каналов крайне неэффективна. В качестве примера рассмотрим фрактальную систему M/Pa/1. Средняя длина L очереди в ней, согласно формуле Полячека-Хинчина [30], здесь при любом $\rho > 0$ составляет

$$L = \frac{\lambda^2 b^{(2)}}{2(1-\rho)} = \infty,$$

поскольку второй момент $b^{(2)}$ времени обслуживания, распределенного здесь по закону Парето при $1 < \alpha \le 2$, бесконечен. В частности, в работе [93] было доказано, что в системе M/Pa/1 средняя длина очереди становится бесконечной, и что является прямым следствием бесконечной дисперсии распределения Pa. Это поясняет,

почему при конечном буфере (т.е. в системе M/Pa/1/m) снижение вероятности потери заявки за счет выбора достаточно большого размера m буфера и/или за счет повышения быстродействия канала оказывается малоэффективным. Отсюда вытекает необходимость поиска более эффективных путей борьбы с потерями заявок в фрактальных системах и сетях.

Теоретическая задача, решаемая в данном разделе, состоит в исследовании эффективности снижения вероятности потерь в фрактальных сетях путем наращивания числа каналов в узлах таких сетей. Прикладная задача состоит в разработке эффективного метода минимизации потерь за счет оптимального распределения каналов по узлам фрактальных сетей. Ключевым звеном разрабатываемого метода является построение и применение асимптотически точных выражений, описывающих зависимость вероятности потерь от числа каналов фрактальной системы.

Эффективность разработанного метода снижения вероятности потерь заключается в ускорении, достигаемым за счет аппроксимации вероятности потерь P конечнолинейных СМО вероятностями состояний p_k и хвостом $P(k \ge n)$ соответствующих бесконечнолинейных систем, что позволяет получить за один прогон имитационной модели значения вероятностей потерь для целого диапазона значений числа каналов n.

3.2.1 Классические бесконечнолинейные СМО

В классической системе с бесконечным числом каналов GI/GI/ ∞ оба распределения A(t), B(t) имеют конечную дисперсию. Недавно доказано [41], что в таких СМО с ростом нагрузки распределение p_k вероятностей числа k занятых каналов сходится к гауссову распределению $N(\overline{k}, \sigma_k)$, т.е. $p_k \to g_k$,

$$g_k = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{(k-\overline{k})^2}{2\sigma_k^2}\right],\tag{3.9}$$

где
$$\bar{k} = \lambda b$$
, $\sigma_k^2 = \lambda b + \kappa \beta$, $\kappa = \lambda^3 (\sigma^2 - a^2)$, $\beta = \int_0^\infty [1 - B(\tau)]^2 d\tau$.

На практике возникает задача выбора такого наименьшего числа n каналов, которое при отсутствии буфера для хранения заявок (т.е. при m=0) обеспечивало бы малую вероятность потери заявки, не превосходящую заданной величины Q.

Иначе говоря, нужно найти наименьшее n, при котором многолинейная система GI/GI/n/0 с теми же A(t) и B(t) (т.е. система GI/GI/n/0, соответствующая исходной бесконечнолинейной системе $GI/GI/\infty$) имеет вероятность потери, не превосходящую Q. Назовем эту задачу задачей нахождения n(Q), подразумевая, что Q – достаточно малая вероятность (10^{-5} и менее). При известном для исходной системы $GI/GI/\infty$ распределении p_k задачу нахождения n(Q) можно переформулировать и решать как задачу нахождения наименьшего n, удовлетворяющего условию

$$P(k \ge n) = \sum_{k=n}^{\infty} p_k \le Q$$
, т.е. условию $1 - P(k \le n) \le Q$. Рассматривая $k, n, P(k \le n)$ как

непрерывные величины, можно просто найти такое n, при котором 1 - P(k < n) = Q, т.е. решить при заданном малом Q уравнение

$$1 - F(n) = Q, (3.10)$$

где $F(n) - \phi.р.$ с.в. k.

При большой загрузке (т.е. при большом λb) и малой вероятности потерь задача (3.10) нахождения n(Q) может решаться за счет применения аппроксимации (3.9) распределения p_k [41].

Рассмотрим конкретный пример — систему $\Gamma_1/\Gamma_2/\infty$, в которой гамма-распределение [35] Γ_1 имеет параметры $\alpha_1=1/3$, $\beta_1=2/3$, и гамма-распределение Γ_2 — параметры $\alpha_2=1/3$, $\beta_2=1/30$. Следовательно, м.о. интервалов поступления заявок $\lambda^{-1}=a=\alpha_1/\beta_1=1/2$, и дисперсия этих интервалов $\sigma^2=\alpha_1/\beta_1^2=3/4$. Среднее время обслуживания в этой системе $b=\alpha_2/\beta_2=10$. Далее, используя формулы (3.9), последовательно найдем $\overline{k}=\lambda b=20$, $\kappa=\lambda^3\Big(\sigma^2-a^2\Big)=4$, параметр β рассчитывался в WolframAlpha с помощью встроенной гамма-функции распределения Q следующим образом: $\beta=\int\limits_0^\infty \left[1-B(\tau)\right]^2 d\tau=\int\limits_0^\infty \left[1-Q_\Gamma\left(\alpha_2,0,x\cdot\beta_2\right)\right]^2 d\tau\approx 2,86826...$, дисперсия распределения $\sigma_k^2\approx\lambda b+\kappa\beta=31,47304...$

Параметры \overline{k} , σ_k гауссовой аппроксимации $q_k = N(\overline{k}, \sigma_k) = N(20, 5, 61008)$ распределения p_k найдены, и можно решать задачу (3.10). Одновременно с гауссовой аппроксимацией g_k посредством высокоточного моделирования в AnyLogic,

после прогона 100 млн. заявок, найдено действительное распределение p_k . Оба распределения показаны на Рисунке 3.8 (слева). На Рисунке 3.8 справа показаны найденные по распределениям p_k и g_k зависимости величины 1 - F(n) от n^2 , вместе с соответствующими линиями и уравнениями тренда. Близкие к единице значения коэффициентов достоверности \mathbb{R}^2 свидетельствуют о достаточно высокой точности полученных уравнений.

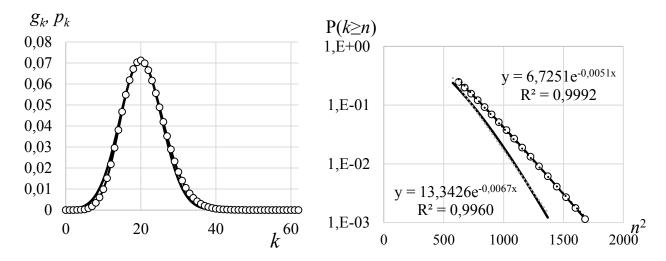


Рисунок 3.8 — Слева: распределение p_k (линия с маркерами) и его гауссова аппроксимация g_k ; справа — хвост $Q = 1 - F(n) = P(k \ge n)$ распределения F(n), рассчитанный на основе действительного распределения p_k (линия с маркерами) и на основе его гауссовой аппроксимации g_k

Из Рисунка 3.8 видно, что использование аппроксимации g_k приводит к занижению вероятности 1 - F(n) на несколько порядков.

Сравним теперь результаты решения задачи (3.10), получаемые на основе распределений p_k и g_k . Используя построенное на основе действительного распределения p_k уравнение тренда (см. Рисунок 3.8), запишем:

$$1 - F(n) = P(k \ge n) = 6,7251e^{-0,0051n^2}$$

С учетом этой формулы уравнение (3.10) принимает вид $6,7251e^{-0,0051n^2}=Q$. Решая его относительно n, получаем:

$$n = \sqrt{-200 \ln Q + 381,7} \ . \tag{3.11}$$

Аналогичное решение, основанное на гауссовой аппроксимации g_k (Рисунок 3.8), дает формулу:

$$n = \sqrt{-142,9 \ln Q + 370,1} \ . \tag{3.12}$$

Решения (3.11) и (3.12) сравниваются на Рисунке 3.9.

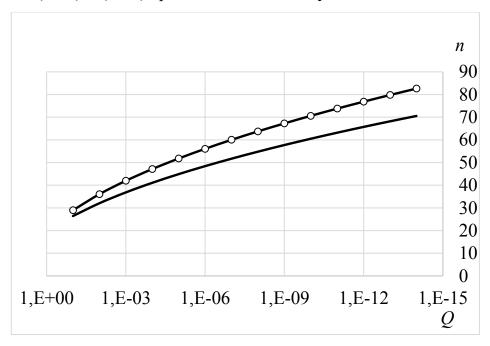


Рисунок 3.9 — Решение задачи (3.10), основанное на распределении p_k (линия с маркерами) и решение, основанное на гауссовой аппроксимация g_k

В целом, выполненное исследование конкретной классической многоканальной системы показывает, что, *во-первых*, можно эффективно обеспечить малую вероятность потерь за счет наращивания числа каналов, даже не используя буфер для хранения заявок; *во-вторых*, для решения задачи обеспечения малой вероятности потерь в общем случае следует предпочесть моделирование применению асимптотических аппроксимаций; *в-третьих*, уменьшение вероятности потерь на несколько порядков достигается за счет относительно небольшого повышения избыточности числа каналов (см. Рисунок 3.9).

Кроме того, проявившаяся в данном конкретном исследовании относительно высокая чувствительность решения задачи (3.10) к погрешностям используемых приближений приводит к выводу о необходимости анализа таких приближений, необходимый анализ приведен в разделе 3.3. В разрабатываемом методе такими приближениями являются аппроксимации вероятностей потерь P конечнолинейной

системы вероятностями состояний p_k и хвостом $P(k \ge n)$ соответствующей бесконечнолинейной системы.

Результаты выполненного в разделе 3.3 сравнения двух аппроксимаций учитываются в разрабатываемом далее методе оптимизации фрактальных сетей (см. раздел 3.2.4). Задача нахождения наименьшего числа n(Q) каналов, обеспечивающего для многолинейной системы вероятность потери заявки, не превосходящую Q, сформулированная выше в традиционной форме (3.10), формулируется теперь как задача решения уравнения:

$$p(n) = Q, \tag{3.13}$$

где вероятность $p_k = p(n)$ состояния n в соответствующей бесконечнолинейной системе, в связи с применяемыми аппроксимирующими выражениями рассматривается как непрерывная функция непрерывного n. Получаемое решение задачи (3.13) округляется вверх до ближайшего целого.

3.2.2 Фрактальные бесконечнолинейные СМО

Когда в системе $GI/GI/\infty$ фрактальной является лишь ф.р. B(t), мы также можем использовать гауссову аппроксимацию g_k вероятностей p_k состояний, поскольку параметр β в этом случае может быть конечным.

Пусть, например,
$$B(t) = 1 - \left(\frac{K}{t}\right)^{\alpha}$$
, $1 < \alpha \le 2$. Тогда, в соответствии с (3.9),

$$\beta = \int_{0}^{\infty} \left[1 - B(\tau) \right]^{2} d\tau = \int_{0}^{\infty} \left(\frac{K}{t} \right)^{2\alpha} d\tau = \frac{2\alpha K}{2\alpha - 1},$$

и конечно при любом $\alpha > 0,5$ (что выполняется автоматически, так как всегда $\alpha > 1$).

Далее для этой системы нетрудно определить и остальные параметры гауссовой аппроксимации (3.9):

$$b = \frac{K\alpha}{\alpha - 1}, \ \overline{k} = \lambda b, \ \kappa = \lambda^3 \left(\sigma^2 - a^2\right), \ \sigma_k^2 = \lambda b + \kappa \beta = \lambda b + \kappa \frac{2\alpha K}{2\alpha - 1}$$
 (3.14)

Таким образом, если A(t) имеет конечную дисперсию σ^2 , то в фрактальной системе GI/Pa/ ∞ распределение p_k числа k занятых каналов с ростом загрузки λb сходится к гауссову распределению $N(\bar{k}, \sigma_k)$ с параметрами (3.14).

В качестве примера рассмотрим систему Г/Ра/ ∞ с гамма-распределением с

параметрами $\alpha_1 = 2$, $\beta_1 = 2$ (его м.о. a = 1 и дисперсия $\sigma^2 = 0.5$) и с распределением Pa(2, 1.25) (его дисперсия бесконечна и м.о. b = 10). Здесь среднее число занятых каналов $\overline{k} = \lambda b = (1/a)b = 10$, и параметр σ_k^2 гауссовой аппроксимации, согласно (3.14), равен $25/3 \approx 8.3333$. На Рисунке 3.10 слева гауссова аппроксимация g_k (штриховая линия) сравнивается с действительным распределением p_k (сплошная линия), рассчитанным с помощью моделирования. При моделировании через систему прошло 100 млн. заявок. Как видим, уже при $\lambda b = 10$ гауссова аппроксимация g_k хорошо согласуется с действительным распределением с.в. k.

Задача (3.10) нахождения n(Q), если ее решать для этой системы $\Gamma/\text{Pa}/\infty$ на основе гауссовой аппроксимации распределения p_k , также приводит к неприемлемым погрешностям, как и в случае рассмотренной выше классической системы $\Gamma_1/\Gamma_2/\infty$. И так же, как в случае классической системы, действительное распределение p_n , рассчитываемое путем моделирования, позволяет решить задачу нахождения n(Q) для рассматриваемой фрактальной системы с достаточной точностью. На Рисунке 3.10 (справа) показаны соответствующие, полученные путем моделирования зависимости от n^2 вероятности p_n и хвоста 1 - F(n), характеризующие рассматриваемую систему $\Gamma/\text{Pa}/\infty$.

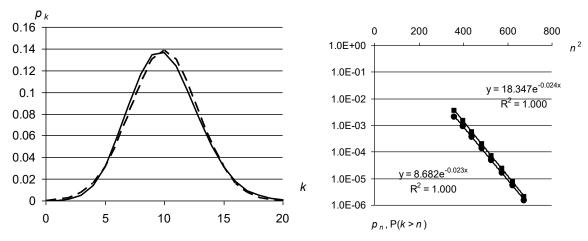


Рисунок 3.10 — Распределения p_k и g_k в системе GI/Pа/ ∞ при $\lambda b = 10$ (слева). Вероятности p_k (круглые маркеры) и их уравнение регрессии (внизу). Хвосты $P(k \ge n)$ (квадратные маркеры) и их уравнение регрессии (вверху). Буквой x в уравнениях обозначена величина n^2

Полученное на Рисунке 3.10 справа уравнение линии тренда для p_k

$$p_k = 8,682e^{-0.023n^2} (3.15)$$

дает для рассматриваемой системы Г/Ра/∞ следующее решение задачи (3.13):

$$n = \sqrt{-43,48 \ln p_k + 93,96} \ . \tag{3.16}$$

Эффективность стратегии наращивания числа каналов по сравнению со стратегиями наращивания длины буфера и/или производительности единственного канала можно проиллюстрировать на следующем примере. Пусть входящий поток заявок обслуживается одним каналом с производительностью, равной суммарной производительности 30 каналов (обеспечивающих, согласно (3.15), вероятность потерь $P \approx 1 \cdot 10^{-8}$). Пусть также эта одноканальная система имеет буфер, достаточный для хранения m = 10~000 заявок. Таким образом, речь идет о системе $\Gamma/\text{Pa}/\text{1/m}$ с тем же входящим потоком и со «сжатым» в 30 раз Парето-временем обслуживания Pa(2/30, 1.25). Моделирование этой системы показывает, что, несмотря на довольно большой объем m буфера и малый коэффициент загрузки ($\rho = 1/3$), вероятность потери заявки составляет $P \approx 0,007$. Таким образом, в общем случае борьба с потерями за счет повышения размера m буфера и/или производительности единственного канала оказывается в условиях фрактального трафика практически бесперспективной.

На Рисунке 3.11 (слева) приведена зависимость распределение p_k для системы $Pa/Pa/\infty$, в которой интервалы поступления заявок описываются законом Pa(1/5, 1,25), а время обслуживания — законом с менее тяжелым хвостом Pa(10/3, 1,5). В этой системе $\lambda = 1$, b = 10. При моделировании через нее прошло более 100 млн. заявок.

На Рисунке 3.11 справа для малых вероятностей p_k превышения заданного числа n каналов показана линия тренда, описываемая уравнением

$$p_k = 24,294e^{-0.0133n^2}$$
.

Используя его, находим решение задачи (3.13) в виде выражения: $n = \sqrt{-75,19 \ln Q + 239,87}$. В соответствии с этим для обеспечения, например, вероятности потерь $Q = 10^{-6}$ достаточно установить в системе 35,8, т.е. 36 каналов.

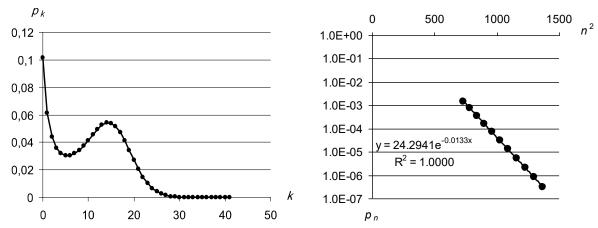


Рисунок 3.11 — Слева — распределение p_k в рассматриваемой системе $Pa/Pa/\infty$; справа — линия тренда зависимости p_k от n^2 в системе $Pa/Pa/\infty$

Многочисленные имитационные эксперименты повышенной точности с разнообразными фрактальными системами $GI/GI/\infty$ показывают, что зависимость p_k и, соответственно, P от n с ростом n в любой такой системе при достаточной (порядка $\lambda b = 10$) загрузке хорошо описывается формулой:

$$p_k \sim c_0 e^{-Cn^2}$$
, (3.17)

где c_0 , C – некоторые константы, свои для каждой конкретной системы. Закон (3.17) позволяет рекомендовать для борьбы с потерями заявок наращивание числа каналов как эффективную универсальную стратегию. При этом даже для весьма малых P избыточность числа каналов по сравнению со средним числом λb используемых каналов оказывается относительно небольшой.

Впрочем, в сетях с фрактальным трафиком можно сочетать стратегию наращивания числа каналов со стратегией буферизации. Выбрав число n каналов, обеспечивающее определенную достаточно малую вероятность P потерь, можно добавить затем и буфер для хранения заявок.

Результат (3.17) обладает признаками универсального закона, выполняющегося как для классических, так и для фрактальных систем. Этот закон хорошо согласуется и с теоретическим результатом (3.9), доказанным в [41] для классических систем. В связи с этим, учитывая взаимное подобие между зависимостью (3.17) и хвостовой частью гауссова распределения, можно переписать (3.17) в следующей теоретически более корректной форме:

$$P \sim p_k \sim c_0 e^{-C(n-\lambda b)^2}$$
 (3.18)

Следует предположить, что аппроксимации (3.17), (3.18) будут применимы и для узлов сетей с очередями (как фрактальных, так и классических).

3.2.3 Фрактальные СеМО с многоканальными узлами

Имитационные эксперименты с различными СеМО, содержащими многоканальные узлы без буферов, показывают, что приближения (3.17), (3.18) с высокой точностью выполняются и для узлов таких сетей.

В качестве примера на Рисунке 3.12 приведены распределения вероятностей состояний четырех узлов сети, полученной модификацией соответствующей четырехузловой сети, описанной в [41]. При модификации сохранены маршрутная матрица **M** сети:

$$\mathbf{M} = \begin{pmatrix} 0.1 & 0.4 & 0.4 & 0 \\ 0 & 0.4 & 0.1 & 0.2 \\ 0.1 & 0 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.5 & 0 \end{pmatrix}$$

и бесконечное число каналов в каждом узле. Входящий поток и время обслуживания в каждом узле изменены и заданы следующим образом. Извне в сеть поступают четыре входящих потока. В первый узел извне поступает регулярный поток с интенсивностью 2. Во второй узел — пуассоновский поток с интенсивностью 2. В третий узел — поток с интервалами поступления, распределенными по закону Pa(0,2,1,25). И в четвертый узел — поток с интервалами поступления, распределенными по закону Pa(1/15;1,5).

Время обслуживания в первом узле детерминированное и равно 0,5. Во втором узле оно распределено экспоненциально со средним 1. В третьем узле — по закону Pa(1/6, 1,5). В четвертом узле — по закону Pa(0,4, 1,25).

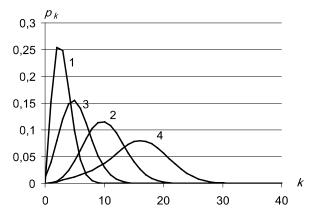


Рисунок 3.12 – Распределения вероятностей состояний узлов 1–4 сети

На Рисунке 3.13 показаны полученные с помощью моделирования зависимости вероятностей P от $x = (n - \lambda b)^2$, в форме соотношения (3.18), для всех четырех узлов в виде соответствующих линий тренда и их уравнений. Уравнения линий тренда в нижней части рисунка перечислены (в соответствии с изображениями линий) в порядке номеров 1, 3, 2, 4 соответствующих узлов сети. Исходные зависимости, полученные в имитационном эксперименте, изображены на Рисунке 3.13 линиями, практически совпавшими с линиями тренда.

Аппроксимации (3.17), (3.18) могут быть использованы в качестве основы для разработки различных методов структурной оптимизации фрактальных систем и сетей с целью обеспечения малой вероятности потерь. Весьма значимой особенностью законов (3.17), (3.18) является то, что наращивание числа каналов в узлах при относительно небольшой их избыточности приводит к резкому снижению вероятности потерь. В условиях фрактального трафика это кардинально отличает стратегию наращивания числа каналов от стратегии увеличения объема буферов и/или повышения производительности каналов.

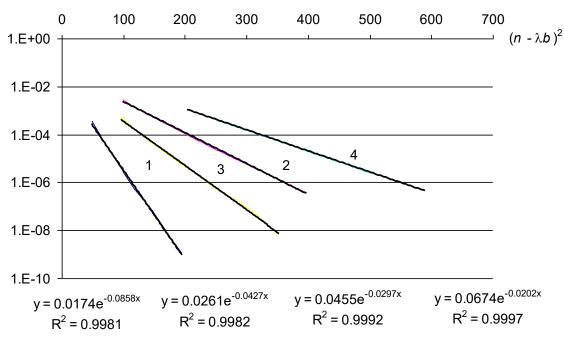


Рисунок 3.13 – Линии тренда зависимостей p_k от n^2 для узлов сети

Результаты эксперимента подтверждают гипотезу о хорошей точности приближения (3.17) не только в изолированных системах, но и в узлах сетей.

В качестве одного из практических применений результатов (3.17), (3.18) предложим относительно простой и быстрый метод оптимального распределения каналов по узлам фрактальной сети.

3.2.4 Задача и метод оптимального распределения каналов

Пусть заданы маршрутная матрица сети, распределения вероятностей $B_i(t)$ времени обслуживания в узлах i (i=1,...,M), входящий поток (входящие потоки) заявок. Буферы для хранения заявок в узлах сети отсутствуют. Требуется распределить N каналов (N >> M) по узлам сети таким образом, чтобы минимизировать сумму вероятностей потерь в узлах.

Задача оптимального распределения каналов возникает при построении (развитии) телекоммуникационной сети с фрактальным трафиком, если используется стратегия наращивания числа каналов в узлах. Какое бы число каналов ни имелось в нашем распоряжении, на практике оно всегда конечно, и от того, как мы распределим эти каналы по узлам, зависит эффективность их использования.

С учетом аппроксимаций (3.17), (3.18) формально задачу оптимального распределения каналов можно переписать следующим образом:

$$\sum_{i=1}^{M} c_{0i} e^{-C_i n_i^2} \to \min, \qquad (3.19)$$

или в виде

$$\sum_{i=1}^{M} c_{0i} e^{-C_i (n_i - \lambda_i b_i)^2} \to \min,$$
 (3.20)

используя в обоих случаях ограничения

$$\sum_{i=1}^{M} n_i = N,$$

$$n_i > 0, \qquad i = 1, \dots, M.$$

Все коэффициенты c_{0i} и C_i здесь известны. Действительно, на практике задача оптимального распределения каналов может решаться только с использованием моделирования, а как с его помощью определяются эти коэффициенты, было продемонстрировано выше.

В форме (3.19) или (3.20) решение задачи оптимального распределения каналов может быть легко получено любыми известными численными методами. Можно, например, рассматривать варьируемые переменные n_i как непрерывные величины, отыскивать их оптимальные значения любым известным градиентным методом и затем корректно эти значения округлять.

Решим, например, задачу оптимального распределения 100 каналов по узлам сети, описанной в разделе 3.2.3.

Используя данные ее моделирования, представленные на Рисунке 3.13, запишем эту задачу оптимизации в форме (3.20):

$$0.0174e^{-0.0858(n_1-3.0390)^2} + 0.0455e^{-0.0297(n_2-10.0800)^2} + 0.0261e^{-0.0427(n_3-5.2250)^2} + 0.0674e^{-0.0202(n_4-15.7150)^2} \rightarrow \min,$$
(3.21)

$$\sum_{i=1}^{4} n_i = 100,$$

$$n_i > 0, \qquad i = 1, ..., 4.$$
(3.22)

Решая задачу (3.21), (3.22) с помощью сервиса «Поиск решения» Excel, получаем: $n_1 = 13,5145, n_2 = 28,3098, n_3 = 20,1214, n_4 = 38,0543$ или, корректно округляя:

$$n_1 = 13, n_2 = 29, n_3 = 20, n_4 = 38.$$
 (3.23)

Здесь возможны шесть способов округления, сохраняющих равенство (3.22). Корректным считается то округление, которое при подстановке в целевую функцию (3.21) дает наименьшее ее значение.

Распределение каналов (3.23) проверено с помощью моделирования. При этом распределении суммарная вероятность отказа в узлах составляет **1,34·10**⁻⁶. Все «соседние» распределения (в которых одна из четырех координат решения (3.23) на единицу уменьшена, другая на единицу увеличена) характеризуются худшей суммарной вероятностью потери в узлах, чем распределение (3.23).

Заметим, что значения целевой функции (3.21), использующей аппроксимации, значительно отличаются от соответствующих величин, получаемых при моделировании. Тем не менее, результаты решения задачи (3.21), (3.22)получаются достаточно точными.

Задача, поставленная в форме (3.19) дает это же решение (3.23).

Если используются аппроксимации вероятностей потерь в виде хвостов $P(p_k \ge p_n)$, то получается менее точное решение $n_1 = 13$, $n_2 = 28$, $n_3 = 20$, $n_4 = 39$, при котором моделирование дает суммарную вероятность потери **1,85·10**⁻⁶.

Сравним оптимальное распределение (3.23) с равномерным распределением каналов по узлам и с распределением каналов, обеспечивающим одинаковые коэф-

фициенты ρ_i загрузки узлов. С помощью моделирования, задав равномерное распределение каналов, по 25 каналов на каждый узел, получаем суммарную вероятность потерь в узлах **2,4·10**⁻³. Если же распределить каналы так, чтобы коэффициенты загрузки узлов были одинаковы (тогда $n_1 = 9$, $n_2 = 30$, $n_3 = 16$, $n_4 = 45$ и все $\rho_i = 0.33$), то суммарная вероятность потерь составит **2,6·10**⁻⁴. В обоих случаях результат на несколько порядков хуже оптимального, что свидетельствует о практической значимости предложенного метода оптимизации.

При проектировании телекоммуникационной сети после оптимального распределения каналов по ее узлам можно в каждый узел добавить буфер для хранения заявок (сообщений). Таким способом можно свести вероятность потерь практически к нулю.

Например, при добавлении к узлам оптимизированной только что сети буферы размером $m_i = 100$, при многократном прохождении 10 млн. заявок, очереди длиной более 7 не наблюдались. Очевидно, потеря заявок в полученной сети практически исключена.

3.3 Аппроксимация вероятности потерь P в системах GI/GI/n/0 вероятностями состояний p_k и хвостом $P(k \ge n)$ систем $GI/GI/\infty$

Сама идея определения вероятностей потерь в многолинейных СМО с помощью бесконечнолинейных систем не является оригинальной [41, 92, 94, 106], однако рассматриваемые в литературе задачи аппроксимации вероятностей потерь предполагают наличие конечной дисперсии распределений интервалов времени поступления и обслуживания заявок, что не соответствует рассматриваемым случаям.

Далее исследуется метод определения вероятности потерь заявок P в многолинейных фрактальных системах массового обслуживания вида GI/GI/n/0, т.е. в системах с числом каналов n и без буфера для хранения заявок (размер буфера m=0), который позволяет получать значения вероятностей потерь P в многолинейной фрактальной СМО для любых n путем аппроксимации этих вероятностей соответствующими вероятностями состояний p_k или хвостом $P(k \ge n)$ бесконечнолинейной СМО GI/GI/ ∞ с аналогичными характеристиками, когда вероятности потерь малы (10^{-5} и менее).

Стационарная вероятность потерь P в будет определяться как предел отношения числа потерянных заявок C_i к общему числу приходов N.

Постановка задачи

Рассмотрим фрактальную многолинейную систему A(t)/B(t)/n/0, с ф.р A(t). времени поступления заявок с м.о. a; ф.р. B(t) времени обслуживания заявок с м.о. b. Определим интенсивность входящего потока заявок $\lambda = 1/\alpha$, интенсивность времени обслуживания заявок $\mu = 1/b$ и коэффициент загрузки $\rho = \lambda b$. Предполагается, что рассматриваемые СМО находятся стационарном режиме функционирования, когда вероятности их состояний $p_k(t) = p_k$ со временем не меняются.

Требуется показать, что вероятности потерь заявок в фрактальной многолинейной системе вида GI/GI/n/0 можно заменять стационарными вероятностями состояний p_k или хвостом $P(k \ge n)$ соответствующей бесконечнолинейной фрактальной СМО $GI/GI/\infty$, когда вероятности потерь достаточно малы.

Необходимость такой замены возникает при решении задач оптимизации числа каналов в сетях массового обслуживания. Одна из подобных задач рассмотрена в [15]. Основной недостаток такой оптимизации заключается в том, что для каждого отдельного узла рассматриваемой СеМО, вероятность потерь P приходится моделировать по отдельности для любого фиксированного n. Уже при небольшом количестве узлов СеМО, проектировщику сетей приходится тратить большое количество времени на имитационное моделирование. Для ускорения процесса поиска оптимального числа каналов в каждом узле СеМО конечнолинейную фрактальную СМО можно заменять на аналогичную бесконечнолинейную, стационарные вероятности состояний p_k и хвост $P(k \ge n)$ которой можно получить для всех n за один прогон имитационной модели.

3.3.1 Аппроксимация вероятности потерь P вероятностями состояний p_k

Аналитическое решение можно получить для фрактальных систем вида M/GI/n/0 и $M/GI/\infty$, для которых существуют формулы расчета стационарных вероятностей состояний. В многолинейной СМО M/GI/n/0 (2.20), которая запишется как

$$P = \frac{(\lambda/\mu)^n}{n!} \cdot \left(\sum_{i=0}^n \frac{(\lambda/\mu)^i}{i!}\right)^{-1} = \frac{\rho^n}{n!} \cdot \left(\sum_{i=0}^n \frac{\rho^i}{i!}\right)^{-1}, \text{ в бесконечнолинейной системе M/GI/∞}$$

$$p_k = e^{-\lambda/\mu} \frac{\left(\lambda/\mu\right)^n}{n!} = e^{-\rho} \frac{\rho^n}{n!}.$$
 (3.24)

Очевидно, что для систем с экспоненциальным распределением времени между приходами заявок при $n \to \infty$ предел отношения вероятности потери заявки P в многолинейной СМО к стационарной вероятности состояния p_k бесконечнолинейной СМО будет равен единице:

$$\lim_{n \to \infty} \frac{P}{p_k} = \lim_{n \to \infty} \frac{\frac{\rho^n}{n!} \cdot \left(\sum_{i=0}^n \frac{\rho^i}{i!}\right)^{-1}}{e^{-\rho} \frac{\rho^n}{n!}} = 1,$$
(3.25)

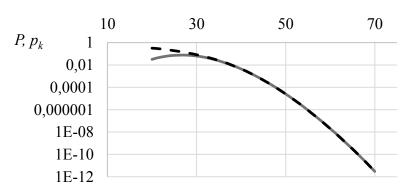
так как при увеличении числа каналов $n \to \infty$ сумма в числителе является разложением экспоненты в ряд Тейлора.

Фрактальные СМО с входящим пуассоновским потоком

Рассмотрим СМО М/Ра/n/0 и М/Ра/ ∞ с интенсивностью входящего потока $\lambda = 8,333$, средним временем обслуживания заявок b = 3,5. На Рисунке 3.14 показаны вероятности потерь заявок P и стационарные вероятности состояний p_k , рассчитанные по формулам (2.20) и (3.24). Видно, что при увеличении числа каналов указанные выше вероятности уменьшаются по значению и приближаются друг к другу. Значения вероятностей P и p_k для некоторых n приведены в Таблице 3.1.

Таблица 3.1 - 3начения P и p_k

n	47	48	49	50	51	52	53
P	0,000151	8,50964E-05	4,71165E-05	2,55664E-05	1,36011E-05	7,0965E-06	3,63283E-06
p_k	0,000151	8,50879E-05	4,7114E-05	2,55657E-05	1,36009E-05	7,09645E-06	3,63282E-06



n, число каналов

Рисунок 3.14 — Вероятности потерь P (пунктирная) и стационарные вероятности состояний p_k (сплошная) для систем M/Pa/n/0 и $M/Pa/\infty$

По данным Таблицы 3.1 можно видеть, что уже при n = 53 значения вероятностей потерь и стационарных вероятностей состояний совпадают с точностью до 4 значащих цифр.

Данный результат также подтверждается и имитационным моделированием. На Рисунке 3.15 для вероятностей P и p_k построены линии тренда, которые обозначены для P и p_k , как y и y^* соответственно.

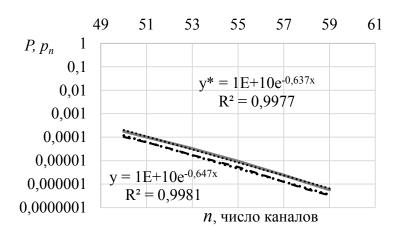


Рисунок 3.15 — Полученные путем ИМ вероятности потерь P (пунктирная) и стационарные вероятности состояний p_k (сплошная) для систем M/Pa/n/0 и $M/Pa/\infty$

Высокие значения коэффициента детерминации R^2 говорят о достаточно хорошем приближении значений линии тренда к исходным данным.

Фрактальные СМО общего вида

Имитационные эксперименты для различных фрактальных СМО вида GI/GI/n/0, у которых входящий поток не является пуассоновским, показали, что вероятности потерь P заявок и стационарные вероятности состояний p_k также сходятся при увеличении n, когда вероятности потерь малы.

На Рисунке 3.16 изображены промоделированные системы $Pa_1/Pa_2/n/0$ и $Pa_1/Pa_2/\infty$, у которых значения параметров масштаба и формы равны $K_1=0,043$, $\alpha_1=1,5$ для Pa_1 и $K_2=1$, $\alpha_2=1,4$ для Pa_2 , интенсивность входящего потока $\lambda=7,75$, а среднее время обслуживания заявок b=3,5.



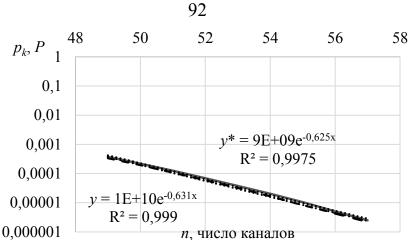


Рисунок 3.16 — Вероятности потерь P (пунктирная) и стационарные вероятности состояний p_k (сплошная) для систем $Pa_1/Pa_2/n/0$ и $Pa_1/Pa_2/\infty$

Стационарные вероятности состояний для фрактальных СМО определялись как отношение времени пребывания t_n системы в состоянии n, к общему времени моделирования T, что можно записать следующим образом:

$$p_n = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} I(k=n),$$
 (3.26)

где $I(\cdot)$ – ступенчатая функция Хэвисайда, которая принимает значение 1, если условие (\cdot) выполняется, и значение 0 – в противном случае; k – это случайная величина, выражающая число занятых каналов.

3.3.2 Аппроксимация вероятности потерь P хвостом $P(k \ge n)$

В литературе встречаются случаи [41, 94], когда в качестве аппроксимации вероятности потерь заявок P в многолинейных CMO используют хвост $P(k \ge n)$, который выражает вероятность того, что в системе будет занято не меньше чем n каналов. При пуассоновском входящем потоке такая аппроксимация дает хорошую сходимость хвоста к вероятности потерь при их малых значениях. На Рисунке 3.17 для рассмотренных выше систем M/Pa/n/0 и $M/Pa/\infty$, изображены вероятность потерь заявок P, рассчитанная по (3.24), и хвост $P(k \ge n)$ стационарных вероятностей состояний p_k , полученных по (2.20):

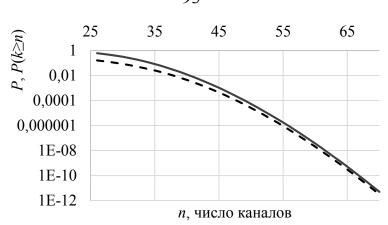


Рисунок 3.17 — Вероятности потерь P (пунктирная) и хвост $P(k \ge n)$ (сплошная) для систем M/Pa/n/0 и $M/Pa/\infty$

Данные выводы можно подтвердить аналитически. Рассмотрим сумму $\rho^n/n! + \rho^{n+1}/(n+1)! + \ldots = \sum_{k=n}^\infty \rho^k/k!$. При всяком фиксированном $\rho > 0$ и достаточно большом n, слагаемые данной суммы убывают столь быстро, что она может быть представлена в виде $\sum_{k=n}^\infty \rho^k/k! = (\rho^n/n!) \cdot (1+\epsilon)$, где ϵ очень маленькая величина. При $n \to \infty$ значение $\epsilon \to 0$ и тогда можно сказать, что вся сумма равна своему первому члену:

$$\lim_{n \to \infty} \frac{\left(\frac{\rho^n}{n!}\right)}{\sum_{k=n}^{\infty} \frac{\rho^k}{k!}} = 1. \tag{3.27}$$

Теперь, на основании вышеизложенного, вычислим предел отношения вероятности $P(k \ge n)$ к вероятности потерь P:

$$\lim_{n \to \infty} \frac{P(k \ge n)}{P(k = n)} = \lim_{n \to \infty} \frac{\left(\sum_{i=n}^{\infty} \frac{\rho^{i}}{i!}\right) \cdot e^{-\rho}}{\frac{\rho^{n}}{n!} \left[\sum_{i=0}^{n} \frac{\rho^{i}}{i!}\right]^{-1}} = \lim_{n \to \infty} \frac{\frac{\rho^{n}}{n!} (1 + \varepsilon) \cdot e^{-\rho}}{\frac{\rho^{n}}{n!}} \cdot \sum_{i=0}^{n} \frac{\rho^{i}}{i!} =$$

$$= \lim_{n \to \infty} (1 + 0) \cdot e^{-\rho} \cdot e^{\rho} = 1.$$
(3.28)

Таким образом, при больших значениях n хвост $P(k \ge n)$ сходится к вероятности P = P(k=n) с относительной погрешностью ноль.

На Рисунке 3.18 представлены результаты ИМ для M/Pa/n/0 и $M/Pa/\infty$, подтверждающие данные выводы.



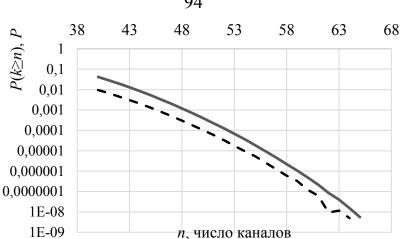


Рисунок 3.18 — Вероятности потерь P (пунктирная) и хвост $P(k \ge n)$ (сплошная) для систем M/Pa/n/0 и $M/Pa/\infty$, полученных путем ИМ

Как показали имитационные эксперименты, для фрактальных систем общего вида, когда распределение интервалов времени между поступлениями заявок не экспоненциальное, при увеличении n хвост $P(k \ge n)$ также сходится к вероятности P. На Рисунке 3.19 представлены результаты ИМ для рассмотренных выше систем $Pa_1/Pa_2/n/0$ и $Pa_1/Pa_2/\infty$, а на Рисунке 3.20- для систем $Pa/\Gamma/n/0$ и $Pa/\Gamma/\infty$ с параметрами K=0.04, $\alpha=1.5$ для Парето распределения, $\beta=1$ и $\alpha_G=3.6$ для гамма-распределения, и хвост $P(k \ge n)$ с интенсивностью входящего потока $\lambda=7.75$ и средним временем обслуживания b=3.6. Рассматриваемое гамма-распределение имеет следующий вид

плотности вероятностей $f(t) = \frac{t^{\alpha_G}}{\beta^{\alpha_G} \Gamma(\alpha_G)} \cdot e^{-\frac{t}{\beta}}$, где $\Gamma(\alpha_G)$ – гамма-функция.

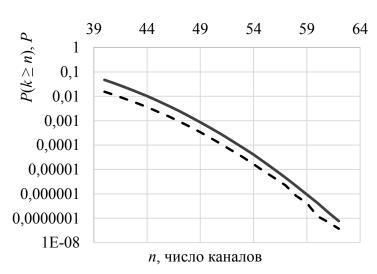


Рисунок 3.19 — Вероятности потерь P (пунктирная линия) и хвост $P(k \ge n)$ (сплошная) для систем $Pa_1/Pa_2/n/0$ и $Pa_1/Pa_2/\infty$.

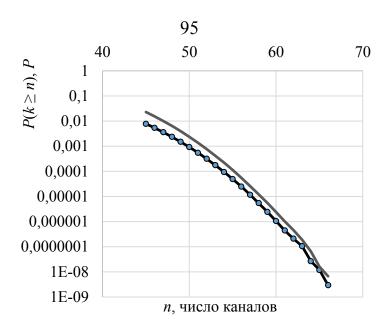


Рисунок 3.20 — Вероятности потерь P (маркированная сплошная) и хвост $P(k \ge n)$ (сплошная) для систем $Pa/\Gamma/n/0$ и $Pa/\Gamma/\infty$

3.3.3 Погрешность аппроксимации потерь P вероятностями состояний p_k и хвостом $P(k \ge n)$

При приближенной замене вероятностей потерь P вероятностями состояний p_k или хвостом $P(k \ge n)$ появляется погрешность, о которой можно судить по отношению p_k/P и $P(k \ge n)/P$. Обозначим отношение p_k/P как $\delta_1(n)$ и $P(k \ge n)/P$ как $\delta_2(n)$. Изучение поведения отношений $\delta_1(n)$, $\delta_2(n)$ с ростом числа каналов n позволяет выработать определенные рекомендации по выбору аппроксимирующего выражения для вероятности потерь P. На Рисунке 3.21 представлены графики $\delta_1(n)$, $\delta_2(n)$ для СМО вида M/GI/n/0 с коэффициентом загрузки $\rho = 5$, полученные аналитически с помощью (2.20) и (3.24).

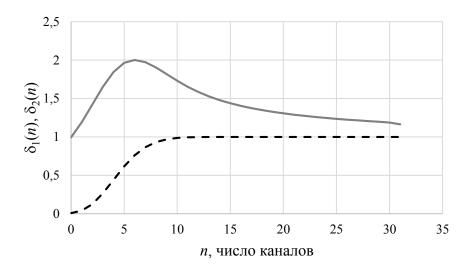


Рисунок 3.21 — Отношения $\delta_1(n)$ (пунктирная линия) и $\delta_2(n)$ (сплошная) для СМО вида M/GI/n/0 с ρ = 5

Поведение отношений $\delta_1(n)$, $\delta_2(n)$ для других фрактальных СМО, как показывают имитационные эксперименты, имеют схожий характер. В [24] в ряде имитационных экспериментов показано, что для фрактальных систем общего вида также можно использовать аппроксимации p_k и $P(k \ge n)$, и что аппроксимации p_k получаются более точными. На Рисунке 3.22 показаны результаты одного из таких экспериментов.

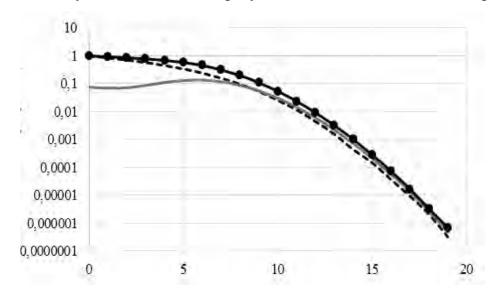


Рисунок 3.22 — Вероятности потерь P (пунктирная линия), состояний p_k (сплошная без маркеров) и хвосты $P(k \ge n)$ для систем $Pa_1/Pa_2/n/0$, $Pa_1/Pa_2/\infty$ с параметрами $\alpha_1 = 1.6$, $K_1 = 0.1$ (для Pa_1) и $\alpha_2 = 1.5$, $K_2 = 0.445$ (для Pa_2)

Необходимо заметить, что аппроксимации p_k и более просты. Расчет системы $GI/GI/\infty$ путем ее моделирования дает непосредственные оценки всех стационарных p_k (как отношения соответствующих накопленных времен пребывания в состояниях n ко всему времени моделирования стационарного процесса), а для расчета хвостов необходимо еще рассчитать через эти p_k соответствующие суммы.

Выводы

1. С помощью разработанного ускоренного метода расчета вероятности P потери заявок в системах с очередями установлено, что в системах M/Pa/1/m и Pa/Pa/1/m зависимость P(m) близка к степенной, с увеличением числа n каналов в системах Pa/Pa/n/m степенная зависимость P(m) переходит в близкую к экспоненциальной. Это позволяет за счет увеличения числа каналов в системах Pa/Pa/n/m

кардинально снижать вероятность потерь и ее зависимость от такого нестабильного фактора фрактального трафика, как параметр формы α.

Предварительная оценка точности и коэффициента ускорения, достигаемых разработанным методом при расчете фрактальных систем, показала, что он позволяет достичь заданной точности результатов на два-три порядка быстрее, чем непосредственное имитационное моделирование.

2. Показано, что при борьбе с потерями сообщений в сетях с фрактальным трафиком стратегия повышения числа каналов обладает кардинальными преимуществами перед стратегиями увеличения буферов и производительности каналов.

Основным результатом исследования являются соотношения (3.17), (3.18), которые дают ключи к решению проблемы потерь сообщений в сетях с фрактальным трафиком и позволяют с хорошей точностью решать задачи анализа и оптимизации таких сетей. Как правило, (3.17), (3.18) можно использовать при загрузке многолинейных узлов $\lambda b = 10$ и более.

3. Показано, что для определения вероятностей потерь в многолинейных фрактальных СМО, когда эти вероятности малы (от 10^{-5} и менее), можно использовать стационарные вероятности состояний p_k или хвосты $P(k \ge n)$ аналогичных бесконечнолинейных СМО. Аппроксимация вероятностями состояний и хвостом $P(k \ge n)$ позволяет избегать многократных прогонов имитационных моделей фрактальных СМО для отдельных n, и получать необходимые значения вероятностей потерь за один прогон имитационной модели, что существенно сокращает время моделирования СеМО, особенно когда число узлов в сети является большим.

Глава 4 Использование абсолютных приоритетов с дообслуживанием

В предыдущей главе рассматривался эффективный метод снижения вероятностей потерь, основанный на увеличении числа каналов. Этот путь связан с оправданными затратами на дополнительное оборудование. Настоящая глава посвящена разработке другого эффективного метода борьбы с потерями, но уже не ориентированного на дополнительные аппаратные затраты. Разрабатываемый метод состоит во введении приоритетного обслуживания. Приоритетные СМО можно разделить на два основных класса в зависимости от дисциплины обслуживания: системы с абсолютными и с относительными приоритетами [31]. При использовании относительных приоритетов заявка с наибольшим приоритетом, дожидается обслуживания заявки с низшим приоритетом, но занимает наилучшее положение в очереди. При использовании абсолютных приоритетов заявка с низшим приоритетом удаляется из обслуживающего прибора при поступлении заявки с более высоким приоритетом. Далее заявки, имеющие одинаковый класс, обслуживаются в порядке поступления. В работе [112] с помощью разработанной метамодели было показано, что для многопотоковых СМО при определенных условиях, эффективность введения абсолютных приоритетов растет с увеличением числа потоков.

Особенностью разрабатываемого метода является использование абсолютных приоритетов с дообслуживанием для входящих заявок в зависимости от их трудоёмкости (времени обслуживания) для бесконечного числа приоритетных классов, кардинально меняющего свойства систем рассматриваемого вида. Такой способ решает проблему «блокировки» единственного обслуживающего канала заявкой с высокой трудоемкостью, являющейся следствием присутствия РТХ в современном сетевом трафике. Введение абсолютных приоритетов для систем вида M/Pa/1/m и Pa/Pa/1m с бесконечным вторым моментом времени обслуживания позволило уменьшить вероятность потерь. Следует отметить, что для расчета вероятностей потерь P в таких системах нет точного аналитического результата, существует аналитическое решение только для среднего времени ожидания W для одноканальных СМО вида M/G/1 [31] с конечным вторым моментом распределения P0. В работе [52] отмечается, что полноценное исследование СМО с приоритетным обслуживанием, требующих варьирование законов распределения времен поступления и обработки заявок, возможно только при помощи имитационного моделирования.

Материал данной главы основан на опубликованных автором статьях [14, 22, 112].

4.1 Формирование приоритетных классов

Разобьём диапазон $K \le t < \infty$ возможных значений трудоёмкости x, используя последовательность точек (разметку):

$$\{t_k\} = t_0, t_1, ..., t_k, ...$$
 (4.1)

на промежутки $[t_0,t_1), \quad [t_1,t_2), \quad ..., \quad [t_{k-1},t_k), \quad ...,$ где $t_0=K$. И, если входящая в систему заявка имеет трудоёмкость, принадлежащую k-му промежутку, т.е. если $x\in [t_{k-1},t_k)$, то отнесём эту заявку к k-му приоритетному классу. При этом будем считать приоритет заявки тем меньшим, чем выше номер k её приоритетного класса. Назовём такое назначение абсолютных приоритетов назначением, обусловленным разметкой оси трудоёмкостей. Если число промежутков $[t_{k-1},t_k)$ конечно, то в последнем из них $[t_{N-1},t_N)$ имеем $t_N=\infty$.

Определим среднее время W ожидания в системе M/Pa/1 при таком разделении входящего потока на приоритетные составляющие. Согласно [31], среднее время U_k пребывания в этой системе заявки k-го приоритетного класса можно выразить следующим образом:

$$U_k = \frac{b_k}{1 - \sigma_{k-1}} + \frac{\sum_{i=1}^k \lambda_i b_i^{(2)}}{2(1 - \sigma_k)(1 - \sigma_{k-1})}, \quad k = 1, 2, ...,$$
(4.2)

где b_k — среднее время обслуживания заявок k-го приоритетного класса, λ_i — интенсивность поступления заявок i-го приоритетного класса, $b_i^{(2)}$ — второй момент времени обслуживания заявок i-го приоритетного класса, $\sigma_k = \sum_{i=1}^k \rho_i$ — сумма коэффициентов загрузки системы заявками приоритетных классов с 1-го по k-й, $\rho_i = \lambda_i b_i$, $\sigma_{k-1} = \sigma_k - \rho_k$.

Тогда среднее время U пребывания заявок в системе можно определить как сумму

$$U = \sum_{k} p_k U_k \tag{4.3}$$

(в которой $p_k = P(t_{k-1} \le x < t_k)$ — вероятность того, что входящая заявка отнесена к k-му приоритетному классу), а среднее время ожидания — как разность

$$W = U - b, \tag{4.4}$$

где b = M(x).

Показатели (4.3) и (4.4) рассматриваемой системы зависят от разметки $\{t_k\}=t_0,t_1,...,t_k,...$ оси трудоёмкостей на промежутки, которыми определяются приоритетные классы заявок. Найдём для заданной разметки $\{t_k\}$ значения параметров системы M/Pa/1, входящих в правую часть формулы (4.2).

Поскольку вероятность $p_k = P(t_{k-1} \le x < t_k)$ промежутка $[t_{k-1}, t_k)$ при распределении Парето (1.2) трудоёмкости x составляет

$$p_k = \left(\frac{K}{t_{k-1}}\right)^{\alpha} - \left(\frac{K}{t_k}\right)^{\alpha} = K^{\alpha} \left(t_{k-1}^{-\alpha} - t_k^{-\alpha}\right),\tag{4.5}$$

условная функция распределения трудоёмкости x, принадлежащей k-му промежутку, имеет вид:

$$F_k(t) = F(t \mid t_{k-1} \le x < t_k) = \frac{F(t) - F(t_{k-1})}{p_k} = \frac{t_{k-1}^{-\alpha} - t^{-\alpha}}{t_{k-1}^{-\alpha} - t_k^{-\alpha}}, \quad t_{k-1} \le x < t_k.$$
 (4.6)

Поэтому для формулы (4.2) получаем:

$$b_k = \int_{t_{k-1}}^{t_k} t dF_k(t) = \int_{t_{k-1}}^{t_k} t \frac{\alpha t^{-\alpha - 1}}{t_{k-1}^{-\alpha} - t_k^{-\alpha}} dt = \frac{\alpha}{\alpha - 1} \cdot \frac{t_{k-1}^{1 - \alpha} - t_k^{1 - \alpha}}{t_{k-1}^{-\alpha} - t_k^{-\alpha}},$$
(4.7)

$$\lambda_k = p_k \lambda \,, \tag{4.8}$$

(все приоритетные составляющие входящего потока пуассоновские),

$$b_{k}^{(2)} = \int_{t_{k-1}}^{t_{k}} t^{2} dF_{k}(t) = \begin{cases} \frac{\alpha}{\alpha - 2} \cdot \frac{t_{k-1}^{2-\alpha} - t_{k}^{2-\alpha}}{t_{k-1}^{2-\alpha} - t_{k}^{2-\alpha}}, & \alpha \neq 2, \\ \frac{2}{t_{k-1}^{2-\alpha} - t_{k}^{2-\alpha}} \ln\left(\frac{t_{k}}{t_{k-1}}\right), & \alpha = 2, \end{cases}$$
(4.9)

(здесь неопределённость типа 0/0 при $\alpha = 2$ раскрыта по правилу Лопиталя),

$$\rho_k = \lambda_k b_k,$$

$$\sigma_{k} = \sum_{i=1}^{k} \rho_{i} = \sum_{i=1}^{k} \lambda_{i} b_{i} = \lambda \sum_{i=1}^{k} p_{i} b_{i} = \lambda \sum_{i=1}^{k} K^{\alpha} (t_{i-1}^{-\alpha} - t_{i}^{-\alpha}) \cdot \left[\frac{\alpha}{\alpha - 1} \cdot \frac{t_{i-1}^{1-\alpha} - t_{i}^{1-\alpha}}{\alpha - 1 \cdot t_{i-1}^{-\alpha} - t_{i}^{-\alpha}} \right],$$

$$\frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \sum_{i=1}^{k} (t_{i-1}^{1-\alpha} - t_{i}^{1-\alpha}) = \lambda b K^{\alpha - 1} (t_{0}^{1-\alpha} - t_{1}^{1-\alpha} + t_{1}^{1-\alpha} - t_{2}^{1-\alpha} + \dots + t_{k-1}^{1-\alpha} - t_{k}^{1-\alpha}) =$$

$$= \rho K^{\alpha - 1} (K^{1-\alpha} - t_{k}^{1-\alpha}) = \rho (1 - K^{\alpha - 1} t_{k}^{1-\alpha}),$$

$$\sigma_{k-1} = \rho (1 - K^{\alpha - 1} t_{k-1}^{1-\alpha}).$$
(4.10)

Нетрудно убедиться, что полученные параметры приоритетных составляющих, на которые разделен входящий поток, отвечают проверочным условиям

$$\sum_{k} p_{k} = 1, \sum_{k} p_{k} F_{k}(t) = F(t), \sum_{k} p_{k} b_{k} = b,$$

$$\sum_{k} \lambda_{k} = \lambda, \sum_{k} p_{k} b_{k}^{(2)} = b^{(2)}, \sum_{k} \rho_{k} = \rho.$$
(4.11)

Исследуем теперь среднее время W ожидания (4.4) при различных разметках $\{t_k\}$. Сразу заметим, что при $1 < \alpha \le 2$ любые конечные разметки $\{t_k\} = \{t_k\}_1^N$ оставляют среднее время ожидания бесконечным из-за того, что в последнем, N-м полубесконечном промежутке $[t_{N-1},t_N)=[t_{N-1},\infty)$ второй момент $b_N^{(2)}$ трудоёмкости, представленный равенством (4.9), бесконечен. Бесконечный момент $b_N^{(2)}$ приводит к бесконечному среднему времени пребывания U_N – см. (4.2) – заявок младшего, N-го приоритетного класса, к бесконечному безусловному среднему времени пребывания U – см. (4.3), и бесконечному среднему времени ожидания W – см. (4.4). Исследуем возможность получения за счет абсолютных приоритетов при $1 < \alpha \le 2$ конечного среднего времени W. Далее будут рассматриваться только бесконечные разметки $\{t_k\} = \{t_k\}_1^\infty$, не оставляющие на оси трудоёмкостей, распределённых по закону Парето, правого полубесконечного промежутка.

4.2 Регулярная разметка оси трудоемкостей

Покажем, что при регулярной разметке (РР), в которой

$$t_k - t_{k-1} = \Delta = \text{const}, \qquad (4.12)$$

т.е. в которой длины всех промежутков одинаковы, введение дисциплины абсолютных приоритетов с дообслуживанием, описанное в предыдущем разделе, делает среднее время ожидания конечным.

Теорема. Введение в систему M/Pa/1 при ρ < 1, 1 < α ≤ 2 дисциплины абсолютных с дообслуживанием приоритетов, определяемых бесконечной регулярной разметкой (4.12) с положительным шагом Δ < ∞ , делает среднее время ожидания конечным.

Доказательство. Среднее время пребывания в системе, определяемое выражениями (4.2) и (4.3), определяется следующим образом:

$$U = \sum_{k} p_{k} U_{k} = \sum_{k=1}^{\infty} \frac{p_{k} b_{k}}{1 - \sigma_{k-1}} + \sum_{k=1}^{\infty} \frac{p_{k} \sum_{i=1}^{k} \lambda_{i} b_{i}^{(2)}}{2(1 - \sigma_{k})(1 - \sigma_{k-1})}.$$
(4.13)

Здесь первая сумма в правой части конечна в силу следующих соотношений:

$$\sum_{k=1}^{\infty} \frac{p_k b_k}{1 - \sigma_{k-1}} \le \sum_{k=1}^{\infty} \frac{p_k b_k}{1 - \rho} = \frac{1}{1 - \rho} \sum_{k=1}^{\infty} p_k b_k = \frac{b}{1 - \rho}$$

(см.(4.11)), ρ < 1 – коэффициент загрузки системы.

Вторую сумму ограничим сверху выражением:

$$\sum_{k=1}^{\infty} \frac{p_k \sum_{i=1}^k \lambda_i b_i^{(2)}}{2(1-\sigma_k)(1-\sigma_{k-1})} \le \sum_{k=1}^{\infty} \frac{p_k \sum_{i=1}^k \lambda_i b_i^{(2)}}{2(1-\rho)(1-\rho)} = \frac{1}{2(1-\rho)^2} \sum_{k=1}^{\infty} p_k \sum_{i=1}^k p_i \lambda b_i^{(2)} = \frac{\lambda}{2(1-\rho)^2} \sum_{k=1}^{\infty} p_k \sum_{i=1}^k p_i b_i^{(2)},$$

и, таким образом, остаётся доказать конечность суммы $S = \sum_{k=1}^{\infty} p_k \sum_{i=1}^{k} p_i b_i^{(2)}$.

Подставляя в неё соответствующие выражения (4.5) и (4.9), получаем:

$$S = \sum_{k=1}^{\infty} \left[\left(\frac{K}{t_{k-1}} \right)^{\alpha} - \left(\frac{K}{t_{k}} \right)^{\alpha} \right] \sum_{i=1}^{k} \left[\left(\frac{K}{t_{i-1}} \right)^{\alpha} - \left(\frac{K}{t_{i}} \right)^{\alpha} \right] \left[\frac{\alpha}{\alpha - 2} \cdot \frac{t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}}{t_{i-1}^{2-\alpha} - t_{i}^{-\alpha}} \right] =$$

$$= \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} \left[\left(\frac{1}{t_{k-1}} \right)^{\alpha} - \left(\frac{1}{t_{k}} \right)^{\alpha} \right] \sum_{i=1}^{k} \left[\left(\frac{1}{t_{i-1}} \right)^{\alpha} - \left(\frac{1}{t_{i}} \right)^{\alpha} \right] \left[\cdot \frac{t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}}{t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}} \right] =$$

$$= \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) \sum_{i=1}^{k} (t_{i-1}^{-\alpha} - t_{i}^{-\alpha}) \left(\frac{t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}}{t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}} \right) =$$

$$= \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) \sum_{i=1}^{k} (t_{i-1}^{2-\alpha} - t_{i}^{2-\alpha}) < \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) \sum_{i=1}^{k} (t_{i-1} - t_{i}) =$$

$$= \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) \sum_{i=1}^{k} \Delta = \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) k \Delta =$$

$$= \frac{\alpha K^{2\alpha}}{\alpha - 2} \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha}) k. \qquad (4.14)$$

Вычислим последнюю из полученных сумм, учитывая, что координаты РР $(t_0,t_1,...,t_k,...)$ имеют вид $(K,K+\Delta,K+2\Delta,...,K+k\Delta,...)$:

$$S_{1} = \sum_{k=1}^{\infty} (t_{k-1}^{-\alpha} - t_{k}^{-\alpha})k = \sum_{k=1}^{\infty} (kt_{k-1}^{-\alpha} - kt_{k}^{-\alpha}) =$$

$$= (1t_{0}^{-\alpha} - 1t_{1}^{-\alpha}) + (2t_{1}^{-\alpha} - 2t_{2}^{-\alpha}) + (3t_{2}^{-\alpha} - 3t_{3}^{-\alpha}) + \dots + (kt_{k-1}^{-\alpha} - kt_{k}^{-\alpha}) + \dots = \sum_{k=0}^{\infty} t_{k}^{-\alpha} =$$

$$= \sum_{k=0}^{\infty} (K + k\Delta)^{-\alpha} = \Delta^{-\alpha} \sum_{k=0}^{\infty} \left(k + \frac{K}{\Delta}\right)^{-\alpha} = \Delta^{-\alpha} \zeta(\alpha, K/\Delta), (4.15)$$

где $\zeta(s,q)$ – дзета-функция Гурвица [66]:

$$\zeta(s,q) = \sum_{k=0}^{\infty} (k+q)^{-s}$$
 (4.16)

Применяя к ряду (4.16) интегральный критерий сходимости Коши убеждаемся, что при вещественных s > 1 и q > 0 этот ряд сходится (при комплексных s, q дзета-функция определена для Re(s) > 1, Re(q) > 0). Следовательно, сумма S_1 (4.15) при $\alpha > 1$ конечна и, соответственно, конечны сумма S (4.14), среднее время пребывания (4.13) и среднее время W ожидания (4.4).

Теорема доказана.

Дзета-функция Гурвица хорошо рассчитывается существующими математическими программами. Например, вычисляя сумму S_1 (4.15) при $\alpha = 8/7$, K = 0,0625, $\Delta = 0,2$ средствами сайта wolframalpha.com [108], получаем $S_1 = 68,767911...$, при этом число значащих цифр результата легко увеличивается до нескольких сотен.

4.3 Оптимизация шага регулярной разметки

Когда мы вводим в систему M/Pa/1/m абсолютные приоритеты, обусловленные бесконечной PP, то при любом Δ , например, при $\Delta = M(x)$, достигаем принципиального сокращения вероятности P потерь, и, кроме того, как показывают имитационные эксперименты, с ростом размера m буфера P(m) убывает не медленнее, чем с экспоненциальной скоростью. Эти качественные перемены могут объясняться тем, что в СМО с бесконечным буфером введение обусловленных бесконечной разметкой абсолютных приоритетов приводит к уменьшению среднего времени ожидания W, а, следовательно, и средней длины очереди L, с бесконечной величины до конечной.

В то же время конкретное достигаемое значение W зависит от параметра Δ разметки, поэтому естественно выбирать не любое, а оптимальное Δ . Значение Δ будем называть оптимальным, если оно минимизирует величину W. Т.е для определения оптимального Δ необходимо решить задачу

$$W(\Delta) = U - b = \sum_{k=1}^{\infty} p_k U_k - b = \sum_{k=1}^{\infty} \frac{p_k b_k}{1 - \sigma_{k-1}} + \sum_{k=1}^{\infty} \frac{p_k \sum_{i=1}^{k} \lambda_i b_i^{(2)}}{2(1 - \sigma_k)(1 - \sigma_{k-1})} - b \to \min_{\Delta}, \quad (4.17)$$

где входящие в минимизируемое выражение показатели зависят от используемой PP $(t_0, t_1, ..., t_k, ...) = (K, K + \Delta, K + 2\Delta, ..., K + k\Delta, ...)$, определяемой параметром Δ . Параметры λ, K, α системы M/Pa/1/ ∞ , для которой решается задача (4.17), считаются заданными. Показатели, входящие в (4.17), определяются по разметке $(t_0, t_1, ..., t_k, ...)$ формулами (4.5)–(4.10).

Решить задачу (4.17) стандартными методами математического анализа, т.е. решением уравнения $\partial W/\partial \Delta = 0$, не удаётся, а при использовании любых численных методов оптимизации основной трудностью становится вычисление значений $W(\Delta)$

при заданном шаге Δ PP. И поскольку разрабатываемый метод оптимизации систем с бесконечной дисперсией времени обслуживания ориентирован на применение инженерами — проектировщиками сетей передачи данных, потребовалась разработка вычислительного метода расчёта $W(\Delta)$, доступного инженерам.

4.3.1 Вывод расчетной формулы

После постановки выражений (4.5)–(4.10) в (4.17) первая в (4.17) сумма с учетом формулы регулярной разметки $t_k = K + k\Delta$ принимает вид:

$$S_{1} = \sum_{k=1}^{\infty} \frac{p_{k}b_{k}}{(1-\sigma_{k-1})} = \sum_{k=1}^{\infty} \frac{\alpha K^{\alpha}}{\alpha - 1} \cdot \frac{t_{k-1}^{1-\alpha} - t_{k}^{1-\alpha}}{1-\rho(1-K^{\alpha-1}t_{k-1}^{1-\alpha})} =$$

$$= \sum_{k=1}^{\infty} bK^{\alpha-1} \frac{t_{k-1}^{1-\alpha} - t_{k}^{1-\alpha}}{1-\rho(1-K^{\alpha-1}t_{k-1}^{1-\alpha})} = b\sum_{k=1}^{\infty} \frac{[K+\Delta(k-1)]^{1-\alpha} - [K+k\Delta]^{1-\alpha}}{(1-\rho)K^{1-\alpha} + \rho[K+\Delta(k-1)]^{1-\alpha}} =$$

$$= b\sum_{k=1}^{\infty} \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho(k+\gamma-1)^{1-\alpha}},$$

$$(4.18)$$

где $\gamma = \frac{K}{\Lambda}$ — нормированная обращённая форма шага Δ PP.

Вторая сумма в (4.17) с учётом (4.5)-(4.10) преобразуется следующим образом:

$$\begin{split} S_2 &= \sum_{k=1}^{\infty} \frac{p_k \sum_{i=1}^k \lambda_i b_i^{(2)}}{2(1-\sigma_k)(1-\sigma_{k-1})} = \sum_{k=1}^{\infty} \frac{\lambda \alpha K^{\alpha}}{\frac{\alpha-2}{2}} p_k (K^{2-\alpha} - t_k^{2-\alpha})}{2(1-\sigma_k)(1-\sigma_{k-1})} = \\ &= \sum_{k=1}^{\infty} \frac{\lambda \alpha K^{\alpha}}{2(\alpha-2)} \frac{K^{\alpha} (t_{k-1}^{-\alpha} - t_k^{-\alpha})(K^{2-\alpha} - t_k^{2-\alpha})}{(1-\sigma_k)(1-\sigma_{k-1})} = \\ &= \frac{\lambda \alpha K^{2\alpha}}{2(\alpha-2)} \sum_{k=1}^{\infty} \frac{(t_{k-1}^{-\alpha} - t_k^{-\alpha})(K^{2-\alpha} - t_k^{2-\alpha})}{[1-\rho(1-K^{\alpha-1}t_k^{1-\alpha})] \cdot [1-\rho(1-K^{\alpha-1}t_{k-1}^{1-\alpha})]} = \\ &= \frac{\lambda \alpha K^{2\alpha}}{2(\alpha-2)} \sum_{k=1}^{\infty} \frac{[(K+\Delta(k-1))^{-\alpha} - (K+k\Delta)^{-\alpha}] \cdot [K^{2-\alpha} - (K+k\Delta)^{2-\alpha}]}{[(1-\rho) + \rho K^{\alpha-1}(K+k\Delta)^{1-\alpha})] \cdot [(1-\rho) + \rho K^{\alpha-1}(K+\Delta(k-1))^{1-\alpha})]} = \\ &= \frac{\lambda \alpha K^2}{2(\alpha-2)} \sum_{k=1}^{\infty} \frac{[(K+\Delta(k-1))^{-\alpha} - (K+k\Delta)^{-\alpha}] \cdot [K^{2-\alpha} - (K+k\Delta)^{2-\alpha}]}{[(1-\rho)K^{1-\alpha} + \rho(K+k\Delta)^{1-\alpha})] \cdot [(1-\rho)K^{1-\alpha} + \rho(K+\Delta(k-1))^{1-\alpha})]}. \end{split}$$

Деля числитель и знаменатель дроби под знаком суммы на $\Delta^{2-2\alpha}$ и получаем:

$$S_{2} = \frac{\lambda \alpha K^{2}}{2(\alpha - 2)} \sum_{k=1}^{\infty} \frac{\left[(\gamma + (k-1))^{-\alpha} - (\gamma + k)^{-\alpha} \right] \cdot \left[\gamma^{2-\alpha} - (\gamma + k)^{2-\alpha} \right]}{\left[(1 - \rho)\gamma^{1-\alpha} + \rho(\gamma + k)^{1-\alpha} \right] \cdot \left[(1 - \rho)\gamma^{1-\alpha} + \rho(\gamma + (k-1))^{1-\alpha} \right]} = \frac{\lambda \alpha K^{2}}{2(\alpha - 2)} \sum_{k=1}^{\infty} \frac{\left[(k + \gamma - 1)^{-\alpha} - (k + \gamma)^{-\alpha} \right] \cdot \left[\gamma^{2-\alpha} - (k + \gamma)^{2-\alpha} \right]}{\left[(1 - \rho)\gamma^{1-\alpha} + \rho(k + \gamma)^{1-\alpha} \right] \cdot \left[(1 - \rho)\gamma^{1-\alpha} + \rho(k + \gamma - 1)^{1-\alpha} \right]}.$$
 (4.19)

Неопределённость, возникающая здесь при $\alpha = 2$, легко раскрывается по правилу Лопиталя:

$$S_2\big|_{\alpha=2} = \lim_{\alpha \to 2} S_2 = \lambda K^2 \sum_{k=1}^{\infty} \frac{[(k+\gamma-1)^{-2} - (k+\gamma)^{-2}] \cdot [\ln(k+\gamma) - \ln(\gamma)]}{[(1-\rho)\gamma^{-1} + \rho(k+\gamma)^{-1})] \cdot [(1-\rho)\gamma^{-1} + \rho(k+\gamma-1)^{-1}]}.$$

С учётом (4.18) и (4.19) зависимость $W(\Delta)$, определяемая соотношением (4.17), может быть представлена следующей явной расчётной формулой:

$$W(\gamma) = S_1 + S_2 - b = b \sum_{k=1}^{\infty} \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho(k+\gamma-1)^{1-\alpha}} + \frac{\lambda \alpha K^2}{2(\alpha-2)} \sum_{k=1}^{\infty} \frac{[(k+\gamma-1)^{-\alpha} - (k+\gamma)^{-\alpha}] \cdot [\gamma^{2-\alpha} - (k+\gamma)^{2-\alpha}]}{[(1-\rho)\gamma^{1-\alpha} + \rho(k+\gamma)^{1-\alpha})] \cdot [(1-\rho)\gamma^{1-\alpha} + \rho(k+\gamma-1)^{1-\alpha}]} - b. \quad (4.20)$$

4.3.2 Упрощение расчетной формулы с контролем погрешностей

При расчете $W(\gamma)$ по формуле (4.20) приходится суммировать достаточно большое число начальных членов рядов, поскольку эти ряды сходятся очень медленно. Например, при значениях α , достаточно далеких от единицы (например, при $\alpha \ge 1,8$), еще можно рассчитать $W(\gamma)$ с точностью до нескольких значащих цифр, суммируя несколько миллионов первых слагаемых соответствующих рядов, входящих в (4.20). Но при близких к единице значениях α (которые при моделировании реального трафика иногда приближаются к $\alpha = 1,2$ и даже к $\alpha = 1,1$) эти ряды сходятся настолько медленно, что их расчет через частичные суммы становится неприемлемо затратным.

Пример 4.1. Вычисляя часть суммы S_1 (4.18) при параметрах $\lambda = 1$, $\alpha = 8/7$, K = 0.0625, $b = \rho = 0.5$, $\Delta = 0.2$, $\gamma = K/\Delta = 0.3125$, суммируя в рядах большое число N первых слагаемых, получаем следующие частичные суммы $S_1(N)$ (округлённые до 6 значащих цифр): $S_1(10^6) = 0.576005$, $S_1(10^7) = 0.605963$, $S_1(10^8) = 0.628093$, $S_1(10^9) = 0.644327$, $S_1(10^{10}) = 0.656177$, $S_1(10^{11}) = 0.664792$,

 $S_1(10^{12}) = 0,671039$. С задачей расчета суммы $S_1 = S_1(\infty)$ такие мощные математические сервисы, как сайт wolframalpha.com, не справляются.

Поэтому для приближенного расчёта W предлагается к частичным суммам $S_1(N)$, $S_2(N)$ (взятым при некотором достаточно большом N) добавлять интегральные оценки $I_1(N+1)$, $I_2(N+1)$ остаточных членов рядов:

$$W(\gamma) \approx S_1(N) + I_1(N+1) + S_2(N) + I_2(N+1) - b$$
,

где:

$$\begin{split} S_1(N) = b \sum_{k=1}^N \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho \, (k+\gamma-1)^{1-\alpha}}, \\ I_1(N+1) = b \int\limits_{N+1}^\infty \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho \, (k+\gamma-1)^{1-\alpha}} dk \,, \\ S_2(N) = \frac{\lambda \alpha K^2}{2(2-\alpha)} \sum_{k=1}^N \frac{[(k+\gamma-1)^{-\alpha} - (k+\gamma)^{-\alpha}] \cdot [(k+\gamma)^{2-\alpha} - \gamma^{2-\alpha}]}{[(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma)^{1-\alpha})] \cdot [(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}]}, \\ I_2(N+1) = \frac{\lambda \alpha K^2}{2(2-\alpha)} \int\limits_{N+1}^\infty \frac{[(k+\gamma-1)^{-\alpha} - (k+\gamma)^{-\alpha}] \cdot [(k+\gamma)^{2-\alpha} - \gamma^{2-\alpha}]}{[(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma)^{1-\alpha})] \cdot [(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}]} dk \,. \end{split}$$

Чтобы несобственный интеграл I_1 мог быть вычислен средствами математических пакетов, аппроксимируем числитель подынтегрального выражения путем разложения его в ряд Тейлора по степеням k:

$$(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha} = \left[k^{1-\alpha} - (\gamma-1)(\alpha-1)k^{-\alpha} + \frac{(\gamma-1)^2}{2}\alpha(\alpha-1)k^{-\alpha-1} - \dots \right] - \left[k^{1-\alpha} - \gamma(\alpha-1)k^{-\alpha} + \frac{\gamma^2}{2}\alpha(\alpha-1)k^{-\alpha-1} - \dots \right] \approx (\alpha-1)k^{-\alpha}.$$
(4.21)

Пример 4.2. При параметрах, перечисленных в примере 4.1, сумма $S_1(10^8) + I_1(10^8 + 1)$ совпадает с суммой S_1 с точностью как минимум до 4-х значащих цифр. Вычисляя ее с помощью сайта wolframalpha.com, находим, что:

$$\begin{split} S_1 \approx S_1(10^8) + I_1(10^8+1) &= b \sum_{k=1}^{1000000} \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}} + \\ &\quad + b \int_{1000001}^{\infty} \frac{(\alpha-1)k^{-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}} dk \approx 0,687258 \,. \end{split}$$

Из этого примера видно, что частичная сумма $S_1(10^{12}) = 0,671039$, вычисленная в Примере 4.1, является плохим приближением суммы S_1 .

Аналогично во втором интеграле заменяем в числителе разность $[(k+\gamma-1)^{-\alpha}-(k+\gamma)^{-\alpha}] \ \text{величиной} \ \alpha k^{-\alpha-1}.$

Кроме того, в знаменателе во втором интеграле второй сомножитель заменяем первым, поскольку их разность пренебрежимо мала: она представляет собой разность двух степенных функций, легко оцениваемую при помощи их разложения в ряд Тейлора.

Таким образом, среднее время ожидания можно с контролируемыми погрешностями рассчитывать по формуле:

$$W(\gamma) \approx b \sum_{k=1}^{N} \frac{(k+\gamma-1)^{1-\alpha} - (k+\gamma)^{1-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}} + b \int_{N+1}^{\infty} \frac{(\alpha-1)k^{-\alpha}}{(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}} dk + \frac{\lambda \alpha K^{2}}{2(2-\alpha)} \sum_{k=1}^{N} \frac{[(k+\gamma-1)^{-\alpha} - (k+\gamma)^{-\alpha}] \cdot [(k+\gamma)^{2-\alpha} - \gamma^{2-\alpha}]}{[(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma)^{1-\alpha})] \cdot [(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma-1)^{1-\alpha}]} + \frac{\lambda \alpha^{2} K^{2}}{2(2-\alpha)} \int_{N+1}^{\infty} \frac{(k+\gamma)^{2-\alpha} k^{-1-\alpha} - \gamma^{2-\alpha} k^{-1-\alpha}}{[(1-\rho)\gamma^{1-\alpha} + \rho (k+\gamma)^{1-\alpha})]^{2}} dk - b.$$

$$(4.22)$$

Погрешности, получаемые при использовании формулы (4.22), проще всего контролировать путем пробных расчетов с постепенным увеличением числа N слагаемых в частичных суммах S_1 и S_2 , поскольку с ростом N эти погрешности уменьшаются достаточно быстро.

4.3.3 Пример решения задачи оптимизации РР

В Таблице 4.2 представлен расчет зависимости $W(\Delta)$ по формуле (4.22) при параметрах исходной системы M/Pa/1, перечисленных в Примере 4.1.

Часть графика этой зависимости, содержащая малую окрестность искомого оптимального шага Δ_{opt} , показана на Рисунке 4.1.

Таблица 4.2 – Расчет	$W(\Delta)$ π	о частям суммы	(26).
----------------------	-------------------	----------------	-------

N	Δ	γ	$S_1(N)$	$I_1(N+1)$	$S_2(N)$	$I_2(N+1)$	W
_	0	-	_	_	_	_	0,206297
10 10	0,01	6,25	0,645499	0,047295	0,013239	2,29E-05	0,206056
10°	0,02	3,125	0,633278	0,059165	0,013379	3,58E-05	0,205858
	0,05	1,25	0,619756	0,071669	0,013938	5,22E-05	0,205415
	0,1	0,625	0,624742	0,065127	0,015142	4,32E-05	0,205055
10 8	0,15	0,416667	0,626916	0,061573	0,016432	3,87E-05	0,204960
	0,2	0,3125	0,628093	0,059165	0,017778	3,58E-05	0,205072
	0,3	0,208333	0,629215	0,055920	0,020566	3,20E-05	0,205733
	0,4	0,15625	0,629632	0,053734	0,023422	2,95E-05	0,206817
	0,5	0,125	0,583584	0,098250	0,026030	9,73E-05	0,207961
10 ⁶	1	0,0625	0,586893	0,089389	0,040552	8,08E-05	0,216915
	1,5	0,041667	0,587976	0,084565	0,054862	7,24E-05	0,227475
	10	0,00625	0,586728	0,065127	26,41610	4,32E-05	26,56800
	100	0,000625	0,576914	0,047295	206,8180	2,29E-05	206,9422
	1000	6,25E-05	0,563815	0,034262	1604,210	1,21E-05	1604,308

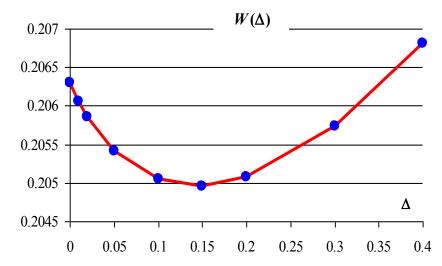


Рисунок 4.1 – График рассчитанной зависимости $W(\Delta)$

Как видно из Рисунка 4.1, $\Delta_{\rm opt}\approx 0,15$. В окрестности оптимума график $W(\Delta)$ достаточно пологий, и при завышении или занижении величины $\Delta_{\rm opt}$ даже вдвое время W существенно не изменится.

Таким образом, в рассмотренном примере введение в систему M/Pa/1 абсолютных приоритетов с оптимальной бесконечной PP уменьшает среднее время ожидания и среднюю длину очереди с бесконечности до приблизительно 0,2.

4.4 Расчет $W(\Delta)$ при $\Delta \rightarrow 0$

При $\Delta \to 0$ происходит переход от дискретной бесконечной разметки к непрерывной, поэтому нумерация промежутков становится неподходящим средством. Передём к представлению промежутков в виде [t, t+dt). Нужные для вычисления W по формуле (4.17) показатели превращаются в показатели соответствующих бесконечно малых промежутков, а суммы, входящие в (4.17), превращаются, соответственно, в интегралы. Найдём вначале необходимые показатели бесконечно малых промежутков.

Вероятность (4.5) конечного промежутка превращается в вероятность

$$p(t,t+dt) = f(t)dt = \frac{\alpha K^{\alpha}}{t^{\alpha+1}}dt, \qquad (4.23)$$

где f(t) – плотность вероятностей распределения Парето.

Среднее значение трудоёмкости х на бесконечно малом промежутке составляет

$$b(t, t+dt) = t. (4.24)$$

Аналогично определяется среднее значения квадрата трудоёмкости на промежутке:

$$b^{(2)}(t,t+dt) = t^2. (4.25)$$

Интенсивность входящего подпотока с трудоёмкостью заявок, лежащей в бесконечно малом промежутке:

$$\lambda(t, t + dt) = \lambda p(t, t + dt) = \frac{\lambda \alpha K^{\alpha}}{t^{\alpha + 1}} dt.$$
 (4.26)

Аналогично можно определить соответствующие коэффициенты загрузки и их частичные суммы:

$$\rho(t,t+dt) = \lambda(t,t+dt)b(t,t+dt) = \frac{\lambda \alpha K^{\alpha}}{t^{\alpha+1}}tdt = \frac{\lambda \alpha K^{\alpha}}{t^{\alpha}}dt, \qquad (4.27)$$

$$\sigma(K,t) = \int_{K}^{t} \rho(s,s+ds) = \int_{K}^{t} \frac{\lambda \alpha K^{\alpha}}{s^{\alpha}} ds = \frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \left(K^{1-\alpha} - t^{1-\alpha} \right). \tag{4.28}$$

Частичные суммы, взятые до последнего и до предпоследнего промежутка, совпадают.

Входящему в (4.17) выражению $p_k \sum_{i=1}^k \lambda_i b_i^{(2)}$ теперь будет соответствовать вы-

ражение
$$p(t,t+dt) \int_{K}^{t} \lambda(s,s+ds)b^{(2)}(s,s+ds) = \frac{\alpha K^{\alpha}}{t^{\alpha+1}} dt \int_{K}^{t} \frac{\lambda \alpha K^{\alpha}}{s^{\alpha+1}} s^{2} ds =$$

$$= \frac{\lambda \alpha^{2} K^{2\alpha}}{t^{\alpha+1}} \frac{1}{2-\alpha} \left(t^{2-\alpha} - K^{2-\alpha}\right) dt. \tag{4.29}$$

Выражению $p_k b_k$ соответствует $\frac{\alpha K^{\alpha}}{t^{\alpha+1}} t dt = \frac{\alpha K^{\alpha}}{t^{\alpha}} dt$.

Подставляя найденные выражения и показатели в (4.17), окончательно получаем:

$$W(0) = \int_{K}^{\infty} \frac{\frac{\alpha K^{\alpha}}{t^{\alpha}}}{\left[1 - \frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \left(K^{1 - \alpha} - t^{1 - \alpha}\right)\right]} dt + \int_{K}^{\infty} \frac{\frac{\lambda \alpha^{2} K^{2\alpha}}{t^{\alpha + 1}} \frac{1}{2 - \alpha} \left(t^{2 - \alpha} - K^{2 - \alpha}\right)}{2\left[1 - \frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \left(K^{1 - \alpha} - t^{1 - \alpha}\right)\right]^{2}} dt - b = \alpha K^{\alpha} \int_{K}^{\infty} \frac{t^{-\alpha}}{\left[1 - \frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \left(K^{1 - \alpha} - t^{1 - \alpha}\right)\right]} dt + \frac{\lambda \alpha^{2} K^{2\alpha}}{2(2 - \alpha)} \int_{K}^{\infty} \frac{t^{-1 - \alpha} \left(t^{2 - \alpha} - K^{2 - \alpha}\right)}{\left[1 - \frac{\lambda \alpha K^{\alpha}}{\alpha - 1} \left(K^{1 - \alpha} - t^{1 - \alpha}\right)\right]^{2}} dt - b \cdot (4.30)$$

Путём вычисления (4.30) и были рассчитаны значения W(0) для Рисунка 4.1. Интегралы в (4.30) хорошо вычисляются математическими пакетами (например, средствами WolframAlpha [108]) при заданных значениях числовых коэффициентов. Вычисление W(0) позволяет осуществлять дополнительный контроль точности расчётов $W(\Delta)$ в области малых Δ , где необходимое количество вычисляемых слагаемых суммы (4.17) становится особенно велико.

4.5 Введение приоритетов при $\alpha > 2$

Регулярная разметка может использоваться и для реализации абсолютных приоритетов в таких системах $M/Pa/1/\infty$, в которых распределение Парето имеет параметр $\alpha > 2$. При этом оптимизация шага PP также может выполняться численным методом. Например, на Рисунке 4.2 показан расчёт зависимости $W(\Delta)$ при $\lambda = 1$, K = 0.48, $\alpha = 2.5$ (в этой системе b = 0.8, $\rho = 0.8$, $b^{(2)} = 1.152$).

Рассчитанная по формуле (4.17) зависимость показана красной кривой, круглые маркеры на ней показывают соответствующие значения, полученные имитационным

моделированием. Верхняя горизонтальная зелёная линия соответствует значению W_0 , которое характеризует изначальную систему (систему без приоритетов). Нижняя горизонтальная линия соответствует значению W(0), которое система имеет при введении абсолютных приоритетов, обусловленных PP с бесконечно малыми промежутками, определяющими приоритетные классы. Тот факт, что кривая начинается практически от верхней горизонтали, а заканчивается практически нижней горизонталью, свидетельствует о том, что рассмотрен достаточно широкий диапазон значений Δ . Можно заметить ещё, что при малых Δ рассчитанная кривая несколько превышает точное предельное значение W(0), к которому она, судя по всему, должна асимптотически приближаться снизу. Это свидетельствует о том, что 60-ти тысяч слагаемых, при которых рассчитывались по формуле (4.17) значения $W(\Delta)$, оказалось при малых Δ не вполне достаточно. Относительно небольшое число слагаемых в данном случае было выбрано потому, что при $\alpha > 2$ суммы в (4.17) сходятся относительно быстро.

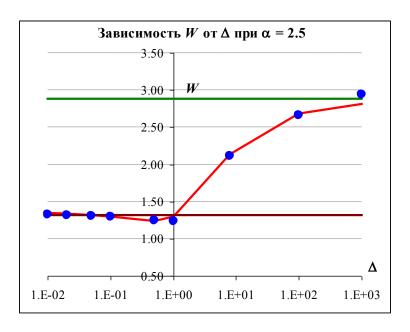


Рисунок 4.2 — Расчёт зависимости $W(\Delta)$ в системе M/Pa/1/ ∞ с параметром $\alpha=2,5$

Оптимальное значение Δ составило здесь величину $\Delta_{\rm opt} \approx 0,4653$, при таком Δ достигается значение $W(\Delta_{\rm opt}) \approx 1,236$. Оценивая эффективность введения абсолютных приоритетов при $\alpha > 2$ коэффициентом сокращения $\xi = W_0 / W(\Delta_{\rm opt})$, получаем $\xi = 2,88 / 1,236 \approx 2,33$.

На Рисунке 4.3 показан расчёт зависимости $W(\Delta)$ в системе с параметрами $\lambda=1,~K=0,48,~\alpha=3,5$ (в этой системе $b=\rho=0,672,~b^{(2)}=0,5376$). Здесь $\Delta_{\rm opt}\approx 0,4046,~W(\Delta_{\rm opt})\approx 0,6541,$ показатель эффективности $\xi=0,8195$ / $1,236\approx 1,25$.

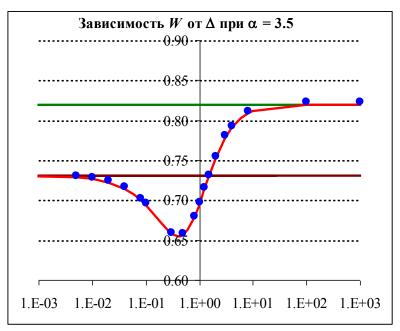


Рисунок 4.3 — Расчёт зависимости $W(\Delta)$ в системе M/Pa/1/ ∞ с параметром $\alpha=3.5$

На Рисунке 4.3 можно видеть, что при малых Δ рассчитанные по конечным суммам значения W практически совпадают с точно рассчитанным по формуле (4.17) значением W(0), хотя суммы (4.17) при расчете усекались до тех же 60 000 слагаемых, что и в расчёте, показанном на Рисунке 4.2. Это объясняется тем, что с ростом α скорость сходимости сумм (4.17) возрастает.

Использование 60 000 слагаемых в суммах (4.17) позволило реализовать их расчёт в Excel, и, соответственно, для определения $\Delta_{\rm opt}$ и $W(\Delta_{\rm opt})$ использовать надстройку «Поиск решения». Поиск решения осуществляется при этом в течение пары десятков секунд. Поэтому расчёт зависимостей на Рисунках 4.2 и 4.3 выполнялся лишь с целью убедиться в единственности искомого локального минимума.

В численных и имитационных экспериментах установлено, что показатель эффективности ξ быстро растёт по мере приближения α сверху к значению $\alpha = 2$, достигая при $\alpha = 2$ бесконечности (и, разумеется, оставаясь таким на всём диапазоне $1 < \alpha \le 2$).

4.6 Экспоненциальная разметка

Численные эксперименты с большим числом других разметок, отличных от PP, показали, что наиболее эффективной из всех рассмотренных оказалась бесконечная экспоненциальная разметка $(t_0, t_1, ..., t_k, ...)$, в которой точки t_k определяются следующим образом:

$$t_0 = K$$
, $t_k = K + ce^{ak}$, $(k = 1, 2, ...)$, (4.31)

где c, a — коэффициенты, которые можно оптимизировать с целью минимизации среднего времени ожидания W при заданных λ , K, α .

Кроме того, что экспоненциальная разметка (4.32) позволяет снижать время W до конечных величин, она обладает тем преимуществом, что, вследствие быстрого роста длин интервалов между точками t_k , частичные суммы в (4.17), вычисленные при нескольких десятках начальных слагаемых, совпадают с соответствующими суммами бесконечных рядов с точностью до нескольких значащих цифр. Благодаря этому расчёт W по формуле (4.17) можно выполнять в небольшом числе строк таблицы Excel, и при этом осуществлять высокоточную оптимизацию параметров c, a разметки градиентным методом, встроенным в надстройку «Поиск решения». На такую оптимизацию тратится несколько секунд компьютерного времени. Опыты с оптимальными экспоненциальными разметками показывают, что они не уступают по эффективности снижения W оптимальным PP.

Небольшое число практически реализуемых уровней бесконечной экспоненциальной разметки (4.32) выгодно отличают ее от РР при практической реализации метода бесконечных разметок.

4.7 Эксперименты с вероятностями потерь при конечных размерах буферов

4.7.1 Основная задача, решаемая введением абсолютных приоритетов

Конечной целью выполняемых в данной главе исследований является радикальное снижение вероятностей потерь в СМО с тяжёлыми хвостами распределений. Разрабатываемый метод состоит во введении в системах с конечным буфером абсолютных приоритетов, обусловленных бесконечными разметками. Метод, разрабатываемый в данной работе, учитывает результаты известных эмпирических исследований систем, функционирующих в условиях фрактального трафика [79] и достижения классической теории оптимизации приоритетных дисциплин [6, 112]. Вместе с тем предлагаемый метод существенно отличается от других известных методов применением бесконечных разметок. Оптимизировать параметры используемых разметок предлагается численными методами, основанными на точных формулах теории массового обслуживания. Поскольку точных формул для расчёта рассматриваемых систем с конечным буфером не существует, в работе используется идея оптимизировать параметры разметки на системах с бесконечным буфером, т.е. выбирать такие параметры разметок, которые максимально сокращают среднее время обслуживания (среднюю длину очереди) в системе с бесконечным буфером. Тогда последующее ограничение буфера происходит в условиях, когда очередь становится в среднем максимально короткой, и, следовательно, вероятность потерь должна снизиться максимально. Это эвристическое обоснование разрабатываемого метода объясняет также независимость подбираемой разметки от длины буфера и, следовательно, косвенным образом обосновывает расчёт на быстрое снижение вероятности потерь с ростом размера буфера.

В данном разделе описываются выполненные с различными системами с конечным буфером имитационные эксперименты, которые самым убедительным образом подтверждают эффективность разработанного метода снижения вероятности потерь.

4.7.2 Эксперименты с системами M/Pa/1/m при 1 < $\alpha \le 2$

На Рисунке 4.4 показаны результаты шести имитационных экспериментов, в которых определены зависимости вероятности потерь P от размера буфера m в различных системах M/Pa/1/m при $1 < \alpha \le 2$. Горизонтальная координатная ось соответствует размеру m буфера, вертикальная — вероятности P потерь. Во всех экспериментах через систему пропускалось 10 млн. заявок.

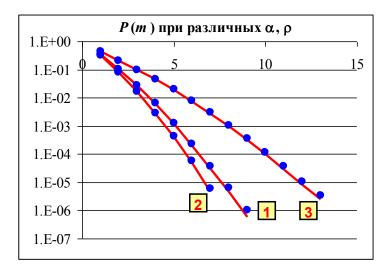


Рисунок 4.4 — Зависимости P(m) в различных системах M/Pa/1/m с конечным буфером

Непрерывными красными линиями изображены графики зависимостей P(m), полученных в условиях применения оптимальных экспоненциальных разметок. Зависимости P(m), изображенные круглыми синими маркерами, получены в условиях применения оптимальных регулярных разметок.

Рисунок 4.4 иллюстрирует, что оптимальные экспоненциальные разметки не уступают оптимальным регулярным разметкам. Но экспоненциальные разметки значительно экономичнее, так как требуют применения лишь нескольких десятков приоритетных классов. На Рисунке 4.4 показаны результаты экспериментов, в которых использовалось 68 приоритетных классов на экспоненциальных разметках, в то время как оптимальные регулярные разметки при $1 < \alpha \le 2$ приводят к использованию десятков миллионов приоритетных классов. Несколько лучшие результаты минимизации P у экспоненциальных разметок на Рисунке 4.4 можно объяснить тем, что регулярные разметки оптимизированы менее точно.

Траектория 1 на Рисунке 4.4 соответствует экспериментам с системой M/Pa/1/m, в которой $\alpha=1,5$, $\rho=0,5$. Оптимальное значение Δ соответствующей регулярной разметки, определенное с точностью до двух десятичных цифр, составляет 0,30. Оптимальные значения параметров экспоненциальной разметки таковы: a=1,055, c=0,08424.

Траектория 2 получена для системы с тем же $\rho = 0.5$, но с гораздо более тяжелым хвостом распределения Парето: здесь $\alpha = 8 \ / \ 7 \approx 1,1428$. Из сравнения траекторий 1 и 2 можно сделать вывод, что утяжеление хвоста распределения только повышает эффективность предлагаемого метода, в то время как в обычных подходах (простое увеличение объема буфера или повышение производительности канала) утяжеление хвоста приводит к практически непреодолимым проблемам. В экспериментах, результаты которых показаны на траектории 2, $\Delta_{\rm opt} = 0.15$, оптимальные значения параметров a и c экспоненциальной разметки составили, соответственно, 1,01007 и 0,03896.

Траектория 3 соответствует системе, в которой $\alpha = 1,5$, но повышена нагрузка на систему, здесь $\rho = 0,8$. Параметры используемых оптимальных разметок здесь следующие: $\Delta_{\rm opt} = 0,45$, a = 0,8980, c = 0,1424.

Ещё одна пара экспериментов проведена для сравнения эффективности предлагаемого метода и обычно рекомендуемого метода снижения вероятности потерь,

состоящего в простом увеличении размера буфера без введения абсолютных приоритетов. Имитационные эксперименты проведены с той же системой, которая рассматривалась первой в предыдущей серии экспериментов: это система M/Pa/1/m, у которой $\alpha = 1,5$, $\rho = 0,5$. Результаты показаны на Рисунке 4.5 (оси координат здесь такие же, как на Рисунке 4.4).

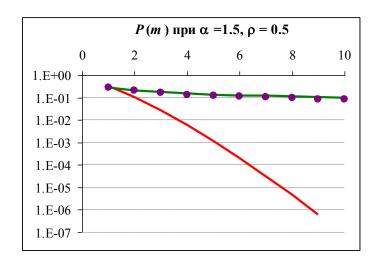


Рисунок 4.5 — Сравнение эффективности наращивания буфера в обычном и приоритетных режимах. Горизонтальная координатная ось соответствует размеру m буфера, вертикальная — вероятности P потерь

Непрерывная красная, резко уходящая вниз линия — это результат наращивания буфера при использовании абсолютных приоритетов, обусловленных оптимальной экспоненциальной разметкой (эта линия взята из траектории 1 на Рисунке 4.4). Линия сходится к прямой, что при логарифмическом масштабе оси ординат свидетельствует о том, что представляемая этой линией зависимость P(m) асимптотически экспоненциальная.

Зелёная сплошная линия в верхней части диаграммы рассчитана при имитационном моделировании исходной системы, не использующей приоритетов. Поскольку соответствующая зависимость степенная, здесь линия уходит вниз с замедлением. При размере буфера m=100 вероятность P снижается здесь лишь до 0,029, при m=1000 имеем P=0,00764, при m=1000 получаем P=0,00125. Как можно видеть из Рисунка 4.5, при использовании абсолютных приоритетов такая вероятность потерь достигается уже при размере буфера m=5. Кроме того, разумеется,

использовать буферы, рассчитанные на хранение 10 000 пакетов, в реальных сетевых устройствах не имеет смысла из-за больших расходов на соответствующее оборудование и больших задержек, возникающих в соответствующих очередях.

Маркерами на верхней линии обозначены результаты, полученные при использовании относительных приоритетов (с тем же составом приоритетных классов, что и при абсолютных приоритетах). Поскольку применение относительных приоритетов уменьшает вероятность, достигаемую в бесприоритетном режиме, лишь на 5%—10%, визуально маркеры совпали со сплошной зелёной линией.

Следует наблюдаемым при моделировании эффектам дать какое-то объяснение на уровне физического смысла. Представляется, что необычайная эффективность абсолютных приоритетов как в плане снижения вероятностей потерь, так и в плане трансформации в среднем бесконечного времени ожидания в конечное объясняется тем, что в рассматриваемых системах время обслуживания х имеет бесконечную дисперсию. Абсолютные приоритеты «перепланируют» порядок обслуживания поступающих сообщений, а менять порядок обслуживания имеет особенный смысл в тех случаях, когда трудоёмкости поступающих заявок представляют собой значительное разнообразие. Однако, сравнивая высокую эффективность абсолютных приоритетов с низкой эффективностью относительных, необходимо конкретизировать объяснение следующим замечанием. При бесконечной дисперсии трудоёмкости х в систему изредка поступают заявки с очень высокой, «катастрофической» трудоёмкостью. Если такая заявка поступает в систему и занимает канал в бесприоритетном режиме или режиме относительных приоритетов, то за время её обслуживания на входе накапливается длинная очередь, что приводит к переполнению даже очень большого буфера. В режиме абсолютных приоритетов поступающие «некатастрофические» заявки, имеющие более высокий приоритет, «не замечают» катастрофическую заявку, просто вытесняют её в очередь и обслуживаются так, как если бы её не было. В результате, накопления длинных очередей не происходит.

Эти рассуждения приводят к предположению, что предлагаемый метод будет столь же эффективен в любых системах GI/GI/1/m, в которых распределение времени обслуживания x является PTX, имеющим бесконечную (или просто очень большую) дисперсию.

4.7.3 Другие системы с РТХ

На рисунке 4.6 показаны результаты моделирования двух систем Pa/Pa/1/m. Нижняя линия — это график зависимости P(m) в системе со следующими параметрами. Оба распределения Парето имеют параметр формы $\alpha = 1,5$. Параметр K у первого распределения Парето равен 4 / 15, у второго — вдвое меньше. Поэтому коэффициент загрузки у системы $\rho = 0,5$. В системе используются абсолютные приоритеты с приоритетными классами, определяемыми экспоненциальной разметкой. Разметка не оптимизирована — просто взята первая экспоненциальная разметка из предыдущих шести экспериментов (разметка с параметрами a = 1,055, c = 0,8424). Верхняя линия — это результаты моделирования второй системы Pa/Pa/1/m, полученной из первой системы заменой обоих параметров формы $\alpha = 1,5$ на $\alpha = 1,25$. Здесь использовалась та же экспоненциальная разметка. Рисунок демонстрирует тот факт, что, несмотря на использование неоптимизированных разметок, предлагаемый метод обеспечивает низкие вероятности потерь при малых m. С ростом m вероятности потерь убывают по асимптотически экспоненциальному закону.

При необходимости можно оптимизировать разметки в таких системах с помощью имитационного моделирования.

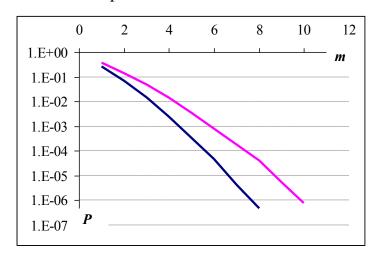


Рисунок 4.6 – Результаты моделирования двух систем Ра/Ра/1/м

На Рисунке 4.7 показаны результаты моделирования системы, в которой входящий поток задан гамма-распределением интервалов, а время обслуживания x имеет логнормальное распределение.

Гамма-распределение имеет математическое ожидание 1 и относительно высокую дисперсию, равную 16 (параметры β и α распределения выбраны равными 16 и 1/16 соответственно). Логнормальное распределение имеет параметры μ = -10, σ = 4.3. При таких параметрах логнормального распределения у времени x обслуживания имеем M(x) = b = 0.47, $D(x) = 5.938 \cdot 10^{42}$

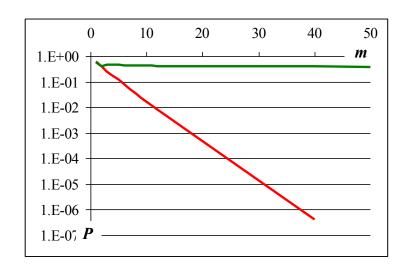


Рисунок 4.7 – Результаты моделирования системы $\Gamma/Logn/1/m$

Верхняя практически горизонтальная линия на Рисунке 4.7 показывает зависимость P(m) в бесприоритетном режиме: P с ростом m убывает очень медленно. При $m=10\ 000$ здесь получена вероятность P=0,25436, при $m=100\ 000$ — вероятность P=0,234502. На Рисунке 4.7 эта последовательность вероятностей начинается значениями порядка 0,5...0,4.

Нижняя линия на диаграмме соответствует полученной в эксперименте зависимости P(m) при использовании абсолютных приоритетов с бесконечной разметкой трудоёмкостей.

Выводы

Для решения проблем, к которым на практике приводит бесконечная дисперсия D(x) времени x обслуживания, в текущей главе разрабатывается и исследуется метод, основный на введении абсолютных приоритетов заявок с бесконечным числом приоритетных классов. Выполненные исследования приводят к следующим

основным выводам.

- 1) Для регулярной разметки полубесконечного диапазона значений времени обслуживания, в которой все промежутки $[t_{k-1},t_k)$ имеют одинаковую конечную длину Δ , справедлива теорема, что если в систему $M/Pa/1/\infty$ с бесконечным средним временем ожидания (т.е. при $1 < \alpha \le 2$) ввести абсолютные приоритеты, определяемые такой регулярной разметкой, то среднее время ожидания W становится конечным.
- 2) Задача оптимизации регулярной разметки по критерию $W \to \min$ эффективно решается разработанным численным методом, использующим расчетную формулу, полученную для среднего времени ожидания W при $\Delta \to 0$.
- 3) Среди большого числа исследованных нерегулярных разметок наиболее эффективна экспоненциальная разметка, в которой длина следующих друг за другом интервалов растёт как экспонента с двумя постоянными коэффициентами. Численные эксперименты показали, что по степени сокращения среднего времени ожидания она не уступает регулярной разметке. В то же время экспоненциальная разметка значительно экономичнее: при $1 < \alpha \le 2$, когда по оптимальной регулярной разметке на практике реализуются десятки миллионов приоритетных классов, при оптимальной экспоненциальной разметке их число не превышает 50.
- 4) Введение абсолютных приоритетов, определяемых рассмотренными бесконечными разметками, кардинально снижает вероятность потерь в системах M/Pa/1/m с конечным буфером и $D(x) = \infty$. При этом медленное со степенной скоростью убывание P(m) превращается в убывание с экспоненциальной скоростью.
- 5) Имитационные эксперименты с разнообразными системами GI/GI/1/m с бесконечной или очень большой дисперсией D(x) показали, что метод введения абсолютных приоритетов с бесконечной регулярной или экспоненциальной разметкой в таких системах столь же эффективен, как и в системах M/Pa/1/m.

Глава 5 Комплекс программ для исследования СМО и СеМО с РТХ

В настоящей главе представлено описание моделирующих программ для исследования СМО и СеМО со степенными хвостами распределений в средах ИМ GPSS World и AnyLogic. Правильность работы моделирующих программ проверена с помощью известных аналитических решений. Материал данной главы основан на опубликованных автором статьях [7, 26, 34].

5. 1 Реализация метода ARAND

В среде ИМ AnyLogic можно создавать свой собственный ГПСЧ, который должен быть подклассом Java класса *Random*, поэтому метод ARAND, который предназначен для высокоточной качественной генерации СПЧ с РТХ, в частности со степенными хвостами распределений, реализован в виде отдельного класса (см. панель проекта на Рисунке 5.1). Реализация ГПСЧ в виде отдельного Java класса позволяет подключаться в любой из сценариев экспериментов в AnyLogic. При работе с ГСПЧ ARAND для каждой модели необходимо создавать и инициализировать новый экземпляр класса ARAND.

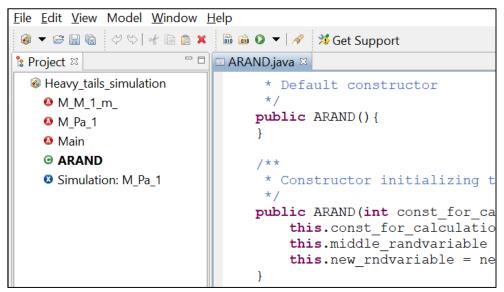


Рисунок 5.1 – Класс ARAND в AnyLogic

В среде ИМ GPSS World метод ARAND реализован в виде самостоятельной процедуры на языке PLUS, принимающей в качестве аргумента номер генератора GPSS World. Листинги кода ARAND на языках программирования Java и PLUS приведены в Приложении A.

При моделировании СМО и СеМО с РТХ в GPSS Word максимальная длина периода генерации СПЧ, равная $T_{\rm max} = 2^{31} - 1$, является недостаточной (см. раздел 2.1.1). Совместно с методом ARAND в GPSS Word использовалась библиотека mtrand.dll, предназначенная для увеличения максимального периода генерации СПЧ. Библиотека mtrand.dll генерирует целые случайные числа на основе алгоритма «Вихрь Мерсенна», реализующего длину периода генерации, равную $2^{19937} - 1$, что позволяет решить проблему недостаточной длины периода генерации СПЧ.

Метод ARAND зарегистрирован в Реестре программ для ЭВМ [7]; расширяемая библиотека GPSS процедур для генерации случайных величин на основе алгоритма «Вихрь Мерсенна» зарегистрирована в объединенном фонде электронных ресурсов «Наука и образование» [34]. Свидетельства о регистрации электронных ресурсов приведены в Приложении В.

5.2 Организация последовательных и параллельных прогонов

Для исследования среднего времени ожидания в СМО Ра/Ра/1, рассмотренной в разделе 2.2, в среде *AnyLogic* создана имитационная модель, реализующая два режима моделирования: стандартное моделирование, необходимое для проведения одного длинного прогона, и моделирование с варьированием переменных — для независимых параллельных прогонов. При проведении последовательных прогонов использовался режим моделирования *Simulation* в рабочей области *Experiment*, а при проведения параллельных прогонов — *Parameters Variation*. Количество параллельных экспериментов (replications per iteration) задавалось в поле *Replications* сценария *Parameters Variation*.

Среднее время ожидание \overline{w}_i рассчитывалось с помощью дополнительно созданной переменной startWaiting подкласса заявок Packets, унаследованного от базового класса заявок Entity.

Вывод результатов осуществлялся в отдельный *.txt файл, который создавался методом *createNewFile*() класса *File*. Для работы с файлами в AnyLogic в свойствах модели во вкладке *import* необходимо подключить java.io*.

В результате одного длинного прогона имитационной модели (см. Рисунок 5.2) сохраняются файлы со значениями среднего времени ожидания \overline{w}_i , а в окне

эксперимента отображаются значения переменных на момент завершения имитационного прогона: выборочные средние $\bar{\tau}$, \bar{x} ф.р. времени поступления и времени обслуживания заявок (см. значения «Mean» в полях middle_Pa_1_statistics и middle_Pa_2 statistics), выборочное среднее (поле Mean) и с.к.о (deviation) в окне «data».

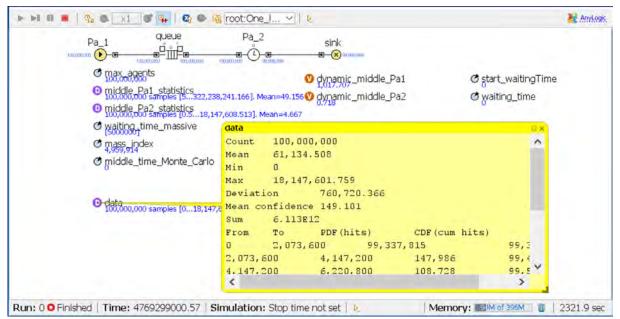


Рисунок 5.2 – Результаты выполнения одного длинного прогона модели Ра/Ра/1 с параметрами $\alpha_1 = \alpha_2 = 1,1$ и $K_1 = 1, K_2 = 0,1$

Обработка результатов ИЭ в Excel для определения момента выхода оценки в установившийся режим с помощью метода групповых средних требовала большие затраты времени на однотипные операции. Для их автоматизации и снижения затрачиваемого времени на обработку данных созданы макросы на языке Visual Basic for Applications. Листинг кода одного из них представлен на Рисунке 5.3, где представлен расчет дисперсии оценок вероятностей потерь для каждой группы.

Рисунок 5.3 – Листинг кода макроса для вычисления дисперсии оценки в Excel

В GPSS организация параллельных прогонов производилась с помощью команды START со значением операнда [В] равным «NP», позволяющим не формировать отчет после каждого прогона, и с помощью команды CLEAR с операндом [А] равным значению «OFF», сохраняющим содержимое ячеек, логических ключей и матриц. Для сохранения данных, полученных на каждом прогоне, создавались объекты MATRIX, изменения в которых осуществлялись с помощью блока MSAVEVALUE. Вывод значений производился в отдельный текстовый файл с помощью блоков OPEN, WRITE, CLOSE.

5.3 Программная реализация имитационных экспериментов для исследования вероятностей потерь

В данном разделе приводится описание программной реализации методов, разработанных в главах 3—4, а именно: ускоренного метода расчета вероятностей потерь (см. раздел 3.1), метода оптимального распределения каналов по узлам сети (раздел 3.2), метода уменьшения вероятности потерь с использованием абсолютных приоритетов (глава 4).

Ускоренный метод расчета вероятностей потерь. Программная реализация ускоренного метода расчета вероятностей потерь P состоит из двух частей, реализованных в AnyLogic на Java. Первая часть содержит программный код для расчета оценки числа потерянных заявок C_i по периодам непрерывной занятости с использованием аппроксимации (3.2), что реализовано перед входом заявок в очередь. Контроль за потерями заявок из-за переполнения очереди осуществляется с помощью элемента *Select Output*. Условием дальнейшего прохождения заявок в СМО является наличие мест в очереди, которое проверяется с помощью встроенной функции *queue.canEnter*(). Если очередь в системе полностью заполнена, то заявки переходят по ветке *false*, после чего они «теряются», то есть удаляются из системы.

Вторая часть программного кода расположена в области *Experiment*: *Parameters Variation*, где реализована организация независимых параллельных прогонов. Основные вероятностные распределения были реализованы с помощью функций, основанных на объектах класса ARAND.

В Приложении Б приведен листинг части кода, реализующий ускоренный метод расчета вероятности потерь для СМО вида М/Ра/1/*m*. Реализацию ускоренного

метода расчета.

Реализация метода оптимального распределения каналов. Алгоритм поиска оптимального распределения числа каналов по узлам состоит из двух этапов.

На *первом* этапе отыскивается приближенно-оптимальное распределение каналов, полученное в ходе ИМ аппроксимирующей СеМО с бесконечнолинейными СМО в её узлах. Процесс нахождения приближенно-оптимального распределения каналов состоит из следующих шагов.

- *Шаг* 1. Задаются начальные данные: функции распределения вероятностей $A_i(t)$, $B_i(t)$, i = 1, ..., M; M число узлов СеМО; суммарное число каналов N; маршрутная матрица \mathbf{M} ; интенсивности входящих в СеМО потоков Λ_i ; вычисление интенсивностей потоков λ_i , входящих в узлы.
- *Шаг* 2. Моделируется аппроксимирующая СеМО, узлы которой представляют собой бесконечнолинейные СМО. В результате ИМ находятся распределения вероятностей состояний p_{ki} для каждого узла.
- *Шаг* 3. Восстанавливается зависимость p_{ki} от n, на основе которой определяются коэффициенты C_i , c_{0i} .
- *Шаг* 4. Полученные на шаге 3 коэффициенты подставляются в (3.19) или (3.20), после чего находится распределение числа каналов n_i .
- *Шаг* 5. Значения n_i округляются до целых чисел таким образом, чтобы целевая функция (3.19) или (3.20) была минимальной.

На *втором* этапе производится ИМ СеМО с конечнолинейными узлами для всех распределений каналов, отличающихся от найденного на первом этапе на единицу в двух позициях, при этом общее число каналов N остается неизменным. Точным оптимальным решением будет то, которое даст наименьшее значение суммарной вероятности потерь P. Точное сравнение различных вариантов распределений n_i . обеспечивается методом общих случайных чисел.

Рассмотрим пример поиска оптимального распределения каналов n_i в СеМО с шестью узлами M = 6, приведенную на Рисунке 5.4.

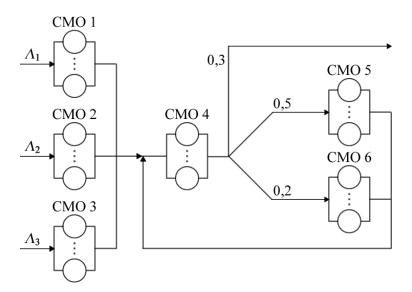


Рисунок 5.4 – Шестиузловая СеМО

Пусть из внешней среды в сеть поступает три потока заявок: в первый узел поступает пуассоновский поток с интенсивностью $\Lambda_1=0,01$, во второй узел Парето-поток с интенсивностью $\Lambda_2=0,03$, в третий узел поступает поток, время между заявками которого распределено равномерно с интенсивностью $\Lambda_3=0,2$. Время обслуживания заявок в первом узле имеет распределение Парето Pa (1,5, 1000), во втором — экспоненциальное с интенсивностью $\mu_2=0,001875$, в третьем — нормальное с $b_3=100$ и $\sigma_3^2=1$, в четвертом — Парето Pa (1,8, 150/18), в пятом — гамма распределение Γ (α , β) = Γ (60, 2), в шестом — Парето Pa (1,5, 33). Вероятности перехода представлены на Рисунке 5.4 и они равны $p_{40}=0,3$, $p_{45}=0,5$, $p_{46}=0,2$. Оставшиеся интенсивности поступления потоков в узлы СеМО найдены путем решения системы уравнений баланса потоков: $\lambda_4=0,8$, $\lambda_5=0,4$, $\lambda_6=0,16$. Общее число каналов в СеМО N=210, буферы для хранения заявок в узлах отсутствуют. Необходимо распределить число каналов по узлам СеМО таким образом, чтобы суммарная вероятность потерь P была минимальной.

Имитируя СеМО с бесконечнолинейными узлами получаем распределения p_{ki} (см. Рисунок Б.1) для каждого узла СеМО, после чего находим зависимость p_k от n. На Рисунке 5.5 представлены значения зависимостей p_{ki} от $(n-\lambda b)^2$ для рассматриваемой сети.

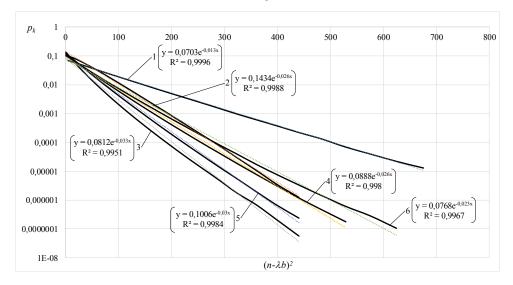


Рисунок 5.5 — Линии тренда зависимостей распределений вероятностей состояний p_k от $(n-\lambda b)^2$

Подставляя коэффициенты c_{0i} и C_i из уравнений линий тренда (Рисунок 5.5) в (3.20), получаем: $n_1 = 51,43$, $n_2 = 33,38$, $n_3 = 34,20$, $n_4 = 31,06$, $n_5 = 27,31$, $n_6 = 32,56$. Округление полученных значений до целых чисел привело к следующему приближенно-оптимальному распределению числа каналов:

$$n_1 = 52, n_2 = 33, n_3 = 34, n_4 = 31, n_5 = 27, n_6 = 33.$$
 (5.1)

Имитационное моделирование показало, что для распределения (5.1) суммарная вероятность потерь P составляет $2,478\cdot10^{-4}$ (см. первую строчку Таблицы Б.1). Исследование точности найденного решения проводилось на втором этапе с помощью ИМ СеМО, в узлах 1–6 которой были заданы значения n_i , отличающиеся от приближенно-оптимального (5.1) на плюс/минус единицу в двух позициях. После проведения $M\cdot(M-1)=6\cdot 5=30$ дополнительных имитационных экспериментов (см. Таблицу Б.1), приближенное решение было уточнено и получилось следующим: $n_1=51$, $n_2=33$, $n_3=34$, $n_4=32$, $n_5=27$, $n_6=33$.

При распределении 5.1 вероятность потерь P составляет **1,586·10**-4. Для дальнейшего уточнения результатов необходимо провести еще $M \cdot (M-1) = 6 \cdot 5 = 30$ экспериментов при значениях n_i , отличающихся от приближенно-оптимального (5.1) на плюс/минус двойку в двух позициях.

Уточнения распределения каналов необходимо продолжать до тех пор, пока новое распределение каналов не достигнет улучшенного значения суммарной вероятности отказа \boldsymbol{P} .

Точное оптимальное решение для рассматриваемого примера было достигнуто за три дополнительных ступени уточнения (см. Таблицу 5.1). Из Таблицы 5.1 видно, что уже на четвертой ступени самое эффективное распределение каналов дает значение вероятности **P** хуже, чем на предыдущей. В первой колонке Таблицы 5.1 римскими цифрами обозначены два основных этапа оптимизации, а арабскими обозначены ступени уточнения полученного на первом этапе приближеннооптимального решения.

Таблица 5.1 – Этапы поиска точного оптимального распределения каналов

Этап	Оптимальное число каналов				лов		P	
Jian	n_1	n_2	<i>n</i> ₃	<i>n</i> ₄	n_5	n_6		
I	52	33	34	31	27	33	Аппроксимация: 6,848·10 ⁻⁴	
	32	33	J -	31 21	<i>J</i> 1	21 33	21 33	Имитационное моделирование: 2,478·10-4
II.1	51	33	34	32	27	33	1,586·10 ⁻⁴	
II.2	50	33	34	33	27	33	1,375·10 ⁻⁴	
II.3	49	33	34	34	27	33	1,262·10-4	
II.4	48	33	34	35	27	33	1,427·10 ⁻⁴	

Таким образом, точным оптимальным распределением каналов будет распределение:

$$n_1 = 49, n_2 = 33, n_3 = 34, n_4 = 34, n_5 = 27, n_6 = 33,$$
 (5.2) при котором $\mathbf{P} = \mathbf{1,262 \cdot 10^{-4}}$.

Другое распределение каналов, а именно равномерное $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 35$ дает вероятность потерь, равную **0,002=2·10**⁻², которая хуже суммарной вероятности (5.2) на два порядка.

Алгоритм поиска оптимального распределения каналов по критерию минимумы вероятности потерь P реализован в среде ИМ AnyLogic. Для каждой отдельной сети необходимо создавать новую модель на панели проекта. При составлении сети (см. Рисунок 5.4) использовались следующие основные элементы палитры: Source, Delay, Sink, Select Output.

Для расчета вероятностей состояний p_{ki} в каждом узле CeMO сбор статистики осуществлялся с помощью элементов палитры *Analysis*: *Statistics* и *Histogram Data*. При входе заявки в узел сети текущее состояние узла — числа занятых каналов сохранялось в созданный заранее параметр, после чего полученное значение переда-

валось объекту $Histogram\ Data$. При прохождении необходимого числа заявок через сеть значения плотностей вероятностей p_{ki} , вычисленные с помощью созданных объектов $Histogram\ Data$, сохранялись в отдельные текстовые файлы.

Метод общих случайных чисел, позволяющий корректно сравнивать значения для различных вариантов распределений n_i , обеспечивался с помощью встроенной опции *Fixed seed* в области *Simulation*.

Реализация метода уменьшения вероятностей потерь с помощью введения дисциплины абсолютных приоритетов с дообслуживанием. В GPSS каждой заявке в имитационной модели можно задать свой собственный приоритет с помощью блока PRIORITY. При моделировании систем с абсолютными приоритетами необходимо использовать специальный блок PREEMPT. При использовании стандартных средств GPSS дисциплина абсолютных приоритетов с дообслуживанием реализуется неправильно и решение данной проблемы предложено автором работы [52]. Рассмотренный в [52] подход правильной реализации абсолютных приоритетов с дообслуживанием, заключающийся в назначении «четных» и «нечетных» приоритетов, применялся при разработке имитационных моделей, используемых в диссертационной работе.

На примере исследования эффективности введения абсолютных приоритетов для уменьшения вероятности потерь в системе с конечным буфером M/Pa/1/m рассмотрим функционирование одной из разработанных программ в GPSS. Рассчитаем с помощью программы, листинг кода которой приведен в Приложении Б, значения вероятности потерь в бесприоритетном режиме, режиме относительных и абсолютных приоритетов. Для системы M/Pa/1/10 с параметрами $1/\lambda = 10$, K = 8/3, $\alpha = 1,5$ получаем следующие значения вероятности потерь:

- 1) P = 0,106000, в бесприоритетном режиме;
- 2) P = 0.078000, с использованием относительных приоритетов;
- 3) P = 0.000169, с использованием абсолютных приоритетов.

Число приоритетных классов составляет N = 10000; моделировалось 10^6 заявок.

С помощью разработанной программы можно увидеть, что зависимость P(m) в рассматриваемой системе при введении абсолютных приоритетов имеет экспоненциальную асимптотику (см. Рисунок 5.6), о чем свидетельствует спрямление

графика значений вероятности потерь в полулогарифмических шкалах. Ось ординат соответствует значениям вероятности потерь, абсцисс — размеру буфера m.

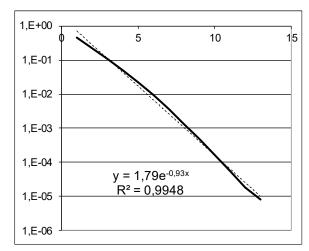


Рисунок 5.6 — Зависимость P(m) в системе M/Pa/1/m с абсолютными приоритетами с дообслуживанием

При программной реализации относительных приоритетов достаточно задать приоритеты заявок в блоке PRIORITY.

Выводы

Приведены описание разработанного программного комплекса и примеры проведения экспериментов в средах имитационного моделирования GPSS и AnyLogic. Программный комплекс реализует:

- разработанные методы организации и планирования имитационных экспериментов, моделирующих СМО и СеМО с РТХ;
 - метод корректной реализации PTX ARAND;
- метод расчета вероятностей потерь, основанный на аппроксимации конечнолинейных систем со степенными хвостами распределений соответствующими бесконечнолинейными;
- метод уменьшения вероятностей потерь в CeMO со степенными хвостами распределений с бесконечной дисперсией, основанный на оптимальном распределении каналов по узлам сети, и метод уменьшения вероятностей потерь, основанный на введении абсолютных приоритетов с бесконечным числом приоритетных классов, определяемых специальными разметками.

Использование GPSS в сочетании с AnyLogic позволило эффективно исследовать проблему правильной организации ИЭ СМО с РТХ и разработать аналитико-имитационные методы, существенно снижающие вероятности потерь заявок.

Заключение

В работе представлено комплексное исследование СМО и СеМО со степенными распределениями времени x обслуживания заявок при $D(x) = \infty$, проводимое с целью решения проблемы больших вероятностей потерь в современных сетях передачи данных с фрактальным трафиком. В ходе исследования получены следующие основные результаты.

- 1. Разработаны математические методы и программные средства для высокоточного имитационного моделирования СМО и СеМО со степенными распределениями и бесконечной дисперсией времени обслуживания, включающие:
- методы корректной (без смещения моментов) реализации в имитационном моделировании распределений с тяжелыми хвостами, частным случаем которых являются степенные распределения;
- специальные генераторы случайных чисел, отвечающие повышенным к ним требованиям со стороны корректности реализации РТХ и моделирования СМО и СеМО рассматриваемых классов, которые при имитационном моделировании характеризуются долговременными зависимостями;
- выражения для построения доверительных интервалов для показателей, оцениваемых по выборкам с долговременными зависимостями, отличающиеся от известных формул отсутствием необходимости предварительного разбиения выборки на большие слабо зависимые группы элементов;
- установленный общий характер сокращения доверительных интервалов при увеличении объема выборки с долговременными зависимостями, происходящего с асимптотически степенной скоростью с показателем степени (-b/2), где b < 1/2 и может принимать значения, сколь угодно близкие к нулю.
- 2. Разработан ускоренный метод регенеративного моделирования СМО рассматриваемых классов, отличающийся возможностью за один прогон модели рассчитывать зависимость P(m) вероятности P потерь заявок от размера m буфера на больших отрезках возможных значений m при малых дополнительных затратах компьютерного времени. С помощью этого метода найдены виды зависимостей P(m) для различных классов СМО.

- 3. Разработан оригинальный ускоренный аналитико-имитационный метод кардинального сокращения вероятностей потерь в СМО и СеМО рассматриваемых классов за счет оптимального (по критерию минимума вероятности потерь) распределения заданного числа каналов по узлам СеМО без буферов. В методе используется приближенное аналитическое выражение вероятности потерь в узле через число каналов в нем, которое строится по результатам пробных прогонов имитационной модели на основе известных асимптотических соотношений, исследованных и обощенных в диссертации. Рассмотрены простые модификации метода, позволяющие получать точные решения оптимизационной задачи и/или дополнительно на порядки снижать вероятность потерь за счет введения в узлы небольших буферов.
- 4. Предложена модель сетевого устройства в виде СМО с бесконечной дисперсией времени обслуживания и дисциплиной абсолютных приоритетов с дообслуживанием, в которой все входящие заявки разбиваются на бесконечное число приоритетных классов, определяемых специальными разметками диапазона возможных значений времени обслуживания.
- 5. Для системы M/Pa/1/∞ с бесконечной дисперсией времени обслуживания сформулирована и доказана теорема о том, что введение абсолютных приоритетов с приоритетными классами, определяемыми бесконечной регулярной разметкой диапазона значений времени обслуживания, снижает бесконечное стационарное среднее время ожидания в этой системе до конечной величины.
- 6. Разработан численный метод расчета зависимости времени ожидания W от шага Δ регулярной разметки, позволяющий оптимизировать ее. Показано, что при $\Delta \to 0$ разметка становится не оптимальной. Рассчитанные для ряда конкретных систем $M/Pa/1/\infty$ оптимальные регулярные разметки проверены и подтверждены высокоточным имитационным моделированием.
- 7. Предложен и исследован метод кардинального снижения вероятностей потерь в СМО M/Pa/1/m с бесконечной дисперсией времени обслуживания. Метод заключается в предварительной оптимизации применяемой разметки при $m = \infty$ и применении полученной оптимальной разметки при конечных объемах m буфера.

Показано, что применение предложенного метода кардинально снижает вероятность P потерь, а степенная скорость убывания с ростом m вероятности P(m) становится после введения приоритетов, определяемых бесконечной оптимальной разметкой, экспоненциальной скоростью убывания.

- 8. Предложена и исследована экспоненциальная разметка заявок на приоритетные классы, определяемая двумя параметрами и легко оптимизируемая градиентными численными методами. Экспериментально установлено, что оптимальные экспоненциальные разметки не уступают по эффективности снижения среднего времени ожидания и вероятности потерь оптимальным регулярным разметкам, но значительно превосходят их по простоте практической реализации.
- 9. В имитационных экспериментах проверена и подтверждена столь же высокая эффективность метода бесконечных разметок при его распространении на многие другие системы класса GI/GI/1/*m* с бесконечной или очень большой дисперсией времени обслуживания, что значительно расширяет область возможных применений предложенной модели и разработанных на ее основе методов.

Разработанные методы снижения вероятностей потерь заявок и их задержек в очередях являются достаточно простыми в практическом применении и могут быть непосредственно использованы проектировщиками сетей с фрактальным трафиком. Все методы реализованы в программном комплексе, разработанном для моделирования СМО и СеМО рассматриваемых классов.

Список сокращений и условных обозначений

АИМ – аналитико-имитационное моделирование

БСВ – базовая случайная величина

ГСПЧ – генератор стандартных псевдослучайных чисел

д.с.в. – дискретная случайная величина

ИМ – имитационное моделирование

ИЭ – имитационный эксперимент

КБ – конечный буфер

м.о. – математическое ожидание

н.с.в. – непрерывная случайная величина

п.в. – плотность распределения вероятностей

ПНЗ – период непрерывной занятости

ПР – период регенерации

РЛХ – распределение с легким хвостом

РР – регулярная разметка

РТХ – распределение с тяжелым хвостом

с.в. – случайная величина

СеМО – сеть массового обслуживания

с.к.о. – среднеквадратическое отклонение

СМО – система массового обслуживания

СПД – сеть передачи данных

СПЧ – стандартные псевдослучайные числа

ТМО – теория массового обслуживания

ф.р. – функция распределения

ФБД – фрактальное дробное броуновское движение

ЦПТ – центральная предельная теорема

QoS – качество обслуживания (Quality of Service)

Список использованных источников

- 1. Бахарева, Н. Ф. Аппроксимативные методы и модели массового обслуживания. Исследование компьютерных сетей : монография / Н. Ф. Бахарева, В. Н. Тарасов. Самара : СНЦ РАН, 2017. 328 с.
- 2. Бахвалов, Н. С. Численные методы / Н. С. Бахвалов, Н. П. Жидков, Г. М. Кобельков. 7-е изд. Москва : БИНОМ, 2012. 636 с.
- 3. Боев, В. Д. Моделирование систем. Инструментальные средства GPSS World / В.Д. Боев. Санкт-Петербург : БХВ-Петербург, 2004. 368 с.
- 4. Буранова, М. А. Исследование влияния статистических свойств мультимедийного IP-трафика на характеристики качества обслуживания : дис. ... канд. техн. наук : 05.12.13 / Буранова Марина Анатольевна. Самара, 2013. 137 с.
- 5. Бусленко, Н. П. Метод статистических испытаний (Монте-Карло) и его реализация на цифровых вычислительных машинах / Н. П. Бусленко, Ю. А. Шрейдер. Москва : Физматгиз, 1961. 226 с.
- 6. Вишневский, В. М. Теоретические основы проектирования компьютерных сетей: монография / В. М. Вишневский. Москва : Техносфера, 2003. 512 с.
- 7. Генератор случайных чисел для распределений с тяжелыми хвостами : свидетельство о государственной регистрации программы для ЭВМ. Российская Федерация / В. Н. Задорожный, Т. Р. Захаренкова ; заявитель и правообладатель Ом. гос. техн. ун-т. № 2018661141 ; заявл. 13.08.2018 ; опубл. 03.09.2018. Зарегистрировано в Реестре программ для ЭВМ.
- 8. Гринь, А. Г. О минимальном условии слабой зависимости в центральной предельной теореме для зависимых случайных величин / А. Г. Гринь // Математические структуры и моделирование. 2007. № 17. С. 13–18.
- 9. Еременко, В. Т. Методы управления потоками в сетях передачи данных на основе резервирования ресурсов / В. Т. Еременко, А. И. Офицеров // Методы и устройства передачи и обработки информации. 2009. № 11. С. 340–346.
- 10. Задорожный, В. Н. Метод ARAND / В. Н. Задорожный // Имитационное моделирование. Теория и практика: труды 7-й всерос. науч.-практ. конф. Москва, 2015. С. 239–244.

- 11. Задорожный, В. Н. Метод ускоренного расчета буферов для фрактальных систем с очередями / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. 2013. № 1 (117) С. 216–220.
- 12. Задорожный, В. Н. Методы имитационного моделирования фрактальных систем с очередями / В. Н. Задорожный, О. И. Кутузов // Информационные технологии и автоматизация управления: материалы VI всерос. науч.-практ. конф. Омск, 2015. С. 9–26.
- 13. Задорожный, В. Н. Методы планирования имитационных экспериментов при моделировании фрактальных очередей / В. Н. Задорожный, Т. Р. Захаренкова // Омский научный вестник. Омск, 2016. № 3 (147) С. 87–92.
- 14. Задорожный, В. Н. Методы снижения вероятности потерь в системах с бесконечной дисперсией времени обслуживания / В. Н. Задорожный, Т. Р. Захаренкова // Системы управления, информационные технологии и математическое моделирование: материалы I Всерос. научн.-практ. конф. с междунар. участием. Омск: ОмГТУ, 2019. С. 7–26.
- 15. Задорожный, В. Н. Минимизация риска потери сообщений в сетях с фрактальным трафиком / В. Н. Задорожный, Т. Р. Захаренкова // Омский научный вестник. Сер. Приборы, машины и технологии. 2016. № 5 (149). С. 125 130.
- 16. Задорожный, В. Н. О качестве программных генераторов случайных чисел / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. 2009. № 2 (80). С. 199–205.
- 17. Задорожный, В. Н. Оптимизация однородных немарковских сетей массового обслуживания / В. Н. Задорожный // Проблемы управления. -2009. -№ 6. С. 68–75.
- 18. Задорожный, В. Н. Очереди и степенные распределения : монография / В. Н. Задорожный. Омск : ОмГТУ, 2016. 162 с.
- 19. Задорожный, В. Н. Повышение точности GPSS-моделей путем применения генератора случайных чисел «Вихрь Мерсенна» / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. 2016. № 1 (145) С. 90—94.

- 20. Задорожный, В. Н. Проблемы генерации случайных величин с фрактальными распределениями / В. Н. Задорожный, О. И. Кутузов // Омский научный вестник. Сер. Приборы, машины и технологии. 2012. № 3 (113) С. 20–24.
- 21. Задорожный, В. Н. Реализация больших выборок при моделировании систем массового обслуживания на GPSS World / В. Н. Задорожный // Имитационное моделирование. Теория и практика (ИММОД-2015) : седьмая всерос. науч.-практ. конф., 21-23 окт. 2015 г. Москва, 2015. С. 225–230.
- 22. Задорожный, В. Н. Численный метод оптимизации регулярной разметки приоритетных классов, вводимых в систему M/Pa/1 с бесконечной дисперсией времени обслуживания / В. Н. Задорожный, М. Радапо, Т. Р. Захаренкова // Системы управления, информационные технологии и математическое моделирование: материалы I Всерос. научн.-практ. конф. с междунар. участием. Омск: ОмГТУ, 2019. С. 161–169.
- 23. Задорожный, В. Н. Экспериментальное исследование влияния дискретности случайных датчиков чисел на смещение моментов моделировании распределений с тяжелыми хвостами / В. Н. Задорожный, Т. Р. Захаренкова // Информационные технологии и автоматизация управления : материалы IX Всерос. науч.-практ. конф. студентов, аспирантов, работников образования и промышленности, 19 мая 2018 г. – Омск, 2018. – C. 50–58.
- 24. Захаренкова, Т. Р. О вероятности потерь в многолинейных фрактальных системах массового обслуживания / Т. Р. Захаренкова // Омский научный вестник. Сер. Приборы, машины и технологии. 2017. № 3 (153). С. 110-114.
- 25. Захаренкова, Т. Р. О методе расчета вероятности потерь в многолинейных фрактальных системах массового обслуживания / Т. Р. Захаренкова // Математика, ее приложения и математическое образование (МПМО17) : материалы VI Междунар. конф., 26 июня 1 июля 2017 г. Улан-Удэ, 2017. С. 190–194.
- 26. Захаренкова, Т. Р. Поиск оптимального распределения каналов в сетях массового обслуживания по критерию минимума вероятности потерь / Т. Р. Захаренкова // Системы управления, информационные технологии и математическое моделирование : материалы I Всерос. научн.-практ. конф. с междунар. Участием, 21–22 мая 2019 г. Омск, 2019. С. 156–160.

- 27. Иглхарт, Д. Л. Регенеративное моделирование сетей массового обслуживания: пер. с англ. / Д. Л. Иглхарт, Д. С. Шедлер. Москва : Радио и связь, 1984. 136 с.
- 28. Карташевский, И. В. Исследование и разработка методов анализа непуассоновских моделей трафика мультисервисных сетей: дис. ... канд. техн. наук: 05.12.13 / Карташевский Игорь Вячеславович.— Самара, 2010. 128 с.
- 29. Клейнен, Дж. Статистические методы в имитационном моделировании: пер. с англ. / Дж. Клейнен. Москва : Статистика, 1978. 556 с.
- 30. Клейнрок, Л. Теория массового обслуживания: пер. с англ. / Л. Клейнрок; под ред. В. И. Нейман. Москва: Машиностроение, 1979. 432 с.
- 31. Клейнрок, Л. Вычислительные системы с очередями: пер. с англ. / Л. Клейнрок; под ред. Б. С. Цыбаков. Москва: Мир, 1979. 595 с.
- 32. Кнут, Д. Э. Искусство программирования: в 4 т. / Д. Э. Кнут. 3-е изд. Москва : Вильямс, 2007. Т. 2 : Получисленные алгоритмы. 788 с.
- 33. Коваленко, И. Н. Методы расчета высоконадежных систем / И. Н. Коваленко, И. Ю. Кузнецов. Москва : Радио и связь, 1988. 176 с.
- 34. Комплекс GPSS-процедур для генерации случайных величин на основе генератора «Вихрь Мерсенна» : свидетельство о регистрации электронного ресурса №22124. / В. Н. Задорожный, С. В. Мясищев, Т. Р. Захаренкова ; организацияразработчик Ом. гос. техн. ун-т ; опубл. 01.09.2016.
- 35. Корн, Γ . Справочник по математике для научных работников и инженеров. Определения. Теоремы. Формулы / Γ . Корн, Γ . Корн. 6-е изд., стер. Санкт-Петербург : Лань, 2003. 831 с.
- 36. Кузьмин, В. В. Модели и процедуры управления трафиком в мультисервисной сети оператора связи: дис. ... канд. техн. наук: 05.13.01 / Кузьмин Васильевич. Нижний Новгород, 2015. 189 с.
- 37. Куроуз, Д. Компьютерные сети: Нисходящий подход / Д. Куроуз, К. Росс. 6-е изд. Москва: Эксмо, 2016. 912 с.
- 38. Кутузов, О. И. К оцениванию и сопоставлению классических и фрактальных систем массового обслуживания / О. И. Кутузов, Т. М. Татарникова // Информационно-управляющие системы. 2016. № 2(81). С. 48–55.

- 39. Кутузов, О. И. Инфокоммуникационные сети. Моделирование и оценка вероятностно-временных характеристик : монография / О. И. Кутузов, Т. М. Татарникова. Санкт-Петербург : ГУАП, 2015. 382 с.
- 40. Мандельброт, Б. Фрактальная геометрия природы / Б. Мандельброт. Москва : Изд-во ИКИ, 2010. 656 с.
- 41. Моисеев, А. Н. Бесконечнолинейные системы и сети массового обслуживания / А. Н. Моисеев, А. А. Назаров. Томск : Изд-во НТЛ, 2015. 240 с.
- 42. Некрасова, Р. С. Регенеративное оценивание и его применение к системам с конечным буфером: дис. ... канд. физ.-мат. наук: 05.13.18 / Некрасова Руслана Сергеевна. Петрозаводск, 2012. 124 с.
- 43. Олифер, В. Компьютерные сети. Принципы, технологии, протоколы : учебник / В. Олифер, Н. Олифер. 5-е изд., Санкт-Петербург : Питер, 2016. 992 с.
- 44. Осин, А. В. Самоподобие и фракталы. Телекоммуникационные приложения / О. И. Шелухин, А. В. Осин, С. М. Смольский. Москва : ФИЗМАЛИТ, 2008. 368 с.
- 45. Петров, В. В. Структура телетрафика и алгоритм обеспечения качества обслуживания при влиянии эффекта самоподобия : дис. ... канд. техн. наук: 05.12.13 / Петров Виталий Валерьевич. Москва, 2004. 199 с.
- 46. Пешкова И. В. Методы регенеративного моделирования для анализа многосерверных систем обслуживания / И. В. Пешкова, А. С. Румянцев // Труды Карельского научного центра РАН. 2018. № 7. С. 68–82.
- 47. Полляк, Ю. Г. Оценка малых вероятностей при статистическом моделировании систем / Ю. Г. Полляк // Изв. АН СССР, Техн. кибернетика. -1973. -№ 2. C. 197–202.
- 48. Рыжиков, Ю. И. Имитационное моделирование. Теория и технологии / Ю. И. Рыжиков. Санкт-Петербург: КОРОНА, 2004. 384 с.
- 49. Рыжиков, Ю. И. Имитационное моделирование и метод квантилей / Ю. И. Рыжиков // Имитационное моделирование. Теория и практика : сборник докладов восьмой всероссийской научно-практической конференции (ИММОД-2017). Санкт-Петербург, 2017. С. 157—161.

- 50. Рыжиков, Ю. И. Теория очередей и распределение Парето / Ю. И. Рыжиков // Труды военно-космической академии им. А.Ф. Можайского. 2015. № 648. С. 28—43.
- 51. Сидоренко, В. Н. Имитационное моделирование в науке и бизнесе: подходы, инструменты, применение / В. Н. Сидоренко, А. В. Красносельский // Бизнес-информатика. 2009. №2. С. 52–57.
- 52. Соснин, В. В. Моделирование дисциплины обслуживания с абсолютными приоритетами в GPSS World / В. В. Соснин // Имитационное моделирование. Теория и практика : сборник докладов третьей всероссийской научно-практической конференции, ИММОД-2007. Санкт-Петербург, 2007. Том 1. С. 224—229.
- 53. Справка AnyLogic [Электронный ресурс] // Source: The AnyLogic Company. [S.I.], 2019. Режим доступа: https://help.anylogic.ru/index.jsp. (дата обращения 01.03.2019).
- 54. Степанов, С. Н. Основы телетрафика мультисервисных сетей / С. Н. Степанов. Москва : Эко-Трендз, 2010. 392 с.
- 55. Столлингс, В. Современные компьютерные сети / В. Столлингс. 2-е изд., Санкт-Петербург : Питер, 2003. 783 с.
- 56. Треногин, Н. Г. Фрактальные свойства сетевого трафика в клиентсерверной информационной системе / Н. Г. Треногин, Д. Е. Соколов // Вестник СИБГУТИ. -2017. -№ 4 (40). - C. 97-103.
- 57. Ушанев, К. В. Показатели своевременности обслуживания трафика массового обслуживания РА/М/1 на основе аппроксимации результатов имитационного моделирования / К. В. Ушанев, С. И. Макаренко // Системы управления, связи и безопасности. − 2016. − №1. − С. 42–65.
- 58. Хилл, Дж. Зависимость случайных процессов и теория стохастических пределов / Дж. Хилл // Квантиль. 2012. № 10. С.1–31.
- 59. Цыбаков, Б. С. Модель телетрафика на основе самоподобного случайного процесса / Б. С. Цыбаков // Радиотехника. 1999. № 5. С. 24—31.
- 60. Шевченко, Д. Н. Методика тестирования и использования генераторов псевдослучайных последовательностей / Д. Н. Шевченко, С. В. Кривенков. // Проблемы физики, математики и техники. 2014. Т. 19, №2. С. 89–95.

- 61. Шелухин, О. И. Моделирование информационных систем / О. И. Шелухин, А. М. Тенякшев, А. В. Осин. Москва : Сайнс-пресс, 2005. 327 с.
- 62. Шелухин, О. И. Причины самоподобия телетрафика и методы оценки показателя Херста / О. И. Шелухин // Электротехнические и информационные комплексы и системы. 2007. Т. 3, № 1. С. 5–14.
- 63. Ширяев, А. Н. Вероятность : в 2 кн. / А. Н. Ширяев. 4-е изд., перераб. и доп. Москва : 2007. Кн. 1 552 с. ; Кн. 2 416 с.
- 64. Яновский, Γ . Γ . Анализ характеристик сетей NGN с учетом свойств самоподобия трафика / Γ . Γ . Яновский, O. A. Симонина, A. M. Галкин // Электросвязь. 2007. № 12. C. 23–26.

- 65. A novel framework for G/M/1 queuing system based on scheduling-cum-polling mechanism to analyze multiple classes of self-similar and LRD traffic / M. Iftikhar [et al.] // Wireless Networks. -2016. V. 22, No 4. P. 1269-1284.
- 66. Adamchik, V. S. Some Series of the Zeta and Related Functions / V. S. Adamchik, H. M. Srivastava // Analysis. 1998. –V. 18. P. 131–144.
- 67. Asmussen, S. Rare events simulation for heavy-tailed distributions / S. Asmussen, K. Binswanger, B. Hojgaard // Bernoulli. 2000. V. 6. P. 303-322.
- 68. Boots, N. K. Simulating GI/GI/1 queues and insurance risk processes with subexponential distributions / N. K. Boots, P. Shahabuddin // ACM SIGMETRICS Performance Evaluation Review. 2001. V. 29. P. 38–39. DOI: 10.1145/507553. 507569.
- 69. Bratley, P. A Guide to Simulation / P. Bratley, B. L. Fox, L. E. Schrage. 2nd ed. New York : Springer, 1987. 397 p.
- 70. Breuer, L. An Introduction to Queueing Theory and Matrix-Analytic Methods / L. Breuer, D. Baum. Netherlands : Springer, 2005. 274 p.
- 71. Beran, J. Statistical Methods for Data with Long-Range Dependence / J. Beran // Statistical Science. 1992. V. 7, № 4. P. 404–416.
- 72. Cheng, Y. Calculation of Loss Probability in a Partitioned Buffer with Self-Similar Input Traffic / Y. Cheng, W. Zhuang // IEEE Global Telecommunications Conference. Dallas, 2004. P. 1453–1457.

- 73. Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper [Электронный ресурс] // Source: Cisco VNI. 2018. Режим доступа: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html. (дата обращения 26.03.2019).
- 74. Clegg, R. G. A critical look at power law modeling of the Internet / R. G. Clegg, C. Di Cairano-Gifedder, S. Zhou // Computer Communications. 2010. V. 33, № 3. P. 259–268.
- 75. Crovella, M. E. Tailed-Probability distributions in the World Wide Web / M. E. Crovella, M. Taqqu, A. Bestavros // Practical Guide to Heavy Tails. 1997. V. 5, №. 6. P. 835–846.
- 76. Crovella, M. Long-Lasting transient, conditions in simulation with heavy-tailed workloads / M. Crovella, L. Lipsky // Proc 1997 Winter Simulation Conference. Atlanta, 1997. P. 1005–1013.
- 77. Crovella, M. Self-similarity in world wide web traffic: evidence and possible causes / M. Crovella, A. Bestavros // Proceedings of the 1996 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems. Philadelphia, 1996. P. 160–169.
- 78. Discrete Simulation Software Ranking a Top list of the Worldwide most Popular and Used Tools / Luís M. S. Dias, António A. C. Vieira, Guilherme A. B. Pereira, José A. Oliveira // Proceedings of the 2016 Winter Simulation Conference. (WSC). Braga, 2016. P. 1060–1071. DOI: 10.1109/WSC.2016.7822165.
- 79. Erramilli, A. Experimental queueing analysis with long range dependent packet traffic / A. Erramilli, O. Narayan, W. Willinger // IEEE/ACM Transactions on Networking. 1996. –V. 4, №. 2. P. 209–223.
- 80. Erramilli, A. Modelling and management of self-similar traffic flows in high-speed net-works / A. Erramilli, W. Willinger, J. L. Wang // Network Systems Design. 1st. ed. Amsterdam : CRC Press, 1999. P. 69–97.
- 81. Erramilli, A. Modeling Packet Traffic with Chaotic Maps / A. Erramilli, R. P. Singh, P. Pruthi // Proceedings of ITC-14. Amsterdam, 1996. P. 329–338.

- 82. Feldmann, A. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models / A. Feldmann, W. Whitt // Performance evaluation. -1998. V. 31. P. 245-258.
- 83. Geong, H-D. Modeling of Self-Similar Teletraffic for Simulation [Электронный ресурс] // Source : UC Research Repository. Christchurch, 2002. Режим доступа: https://ir.canterbury.ac.nz/handle/10092/5454. (дата обращения 02.09.2019).
- 84. Giambene, G. Queuing Theory and Telecommunications: Networks and Applications / G. Giambene. 2nd ed. New York: Springer, 2014 516 p.
- 85. Glynn, P. W. Conditions for the applicability of the regenerative method / P. W. Glynn, D. L. Iglehart // Management Science. 1993. V. 39, №. 9. P. 1108–1111.
- 86. GPSS World reference manual [Электронный ресурс] // Source: Minuteman Software. 5th ed. [S.L.], 2009. Режим доступа: http://www.minutemansoftware.com/reference/rpreface.htm. (дата обращения 26.03.2019).
- 87. Gross, J. Modeling and Tools for Network Simulation / J. Gross, K. Wehrle, M. Gunes. Berlin: Springer, 2010. 256 p.
- 88. Heidelberger, P. Adaptive spectral methods for simulative output analysis / P. Heidelberger, P. D. Welch // IBM J. Research Develop. − 1981. − V. 25, №. 6. − P. 860–876.
- 89. ITU-T Recommendation Y.2011 Global information infrastructure, internet protocol aspects and next generation networks General principles and general reference model for Next Generation Networks [Электронный ресурс] // Source: ITU-T. [S.L.], 2004. Режим доступа: https://www.itu.int/rec/T-REC-Y.2011-200410-I/en. (дата обращения: 10.05.2017).
- 90. Kim, H. S. Loss probability calculations and asymptotic analysis for finite buffer multiplexers / H. S. Kim, N. Shroff // IEEE/ACM Transactions on Networking. -2001. -V. 9, N<math>0. 6. -P. 755–767.
- 91. Kulikovs, M. Remarks Regarding Queuing Model and Packet Loss Probability for the Traffic with Self-Similar Characteristics // M. Kulikovs, E. Petersons // World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering. − 2008. − V. 2, № 4. − P. 631–637.

- 92. Li, A. A. Approximate Blocking Probabilities in Loss Models with Independence and Distribution Assumptions Relaxed / A. A. Li, W. Whitt // Performance Evaluation. -2014.-V.~80.-P.~82-101.
- 93. Likhanov, N. Analysis of an ATM Buffer with Self-Similar («Fractal») Input Traffic / N. Likhanov, B. Tsybakov, N. Georganas // IEEE INFOCOM'95. 1995. V. 3. P. 985–992.
- 94. Massey, W. A. An analysis of the modified offered-load approximation for the nonstationary Erlang model / W. A. Massey, W. Whitt // The annals of applied probability. -1994. V. 4, N = 4. P. 1145-1160.
- 95. Mikosch, T. Regular Variation, Subexponentiality and Their Applications in Probability Theory / T. Mikosch. Groningen: Groningen University, 1999. 57 p.
- 96. Morozov E. Weak regeneration in modeling of queueing processes / E. Morozov // Queueing Systems. 2004. V. 46. P. 295–315.
- 97. Neame, T. A practical approach for multi-media traffic modeling / T. Neame, M. Zukerman, R. Addie // Broadband Communications: Convergence of Network Technologies. Hong Kong, 1999. P. 73–82.
- 98. Newman, M. E. J. Power laws, Pareto distributions and Zipf's law / M. E. Newman // Contemporary Physics. 2005. V. 46, № 5. P. 323–351.
- 99. Norros, I. On the Use of Fractional Brownian Motion in the Theory of Connectionless Networks / I. Norros // IEEE Journal on Selected Areas in Communications. 1995. V. 13, № 6. P. 953–962.
- 100. On the Self-Similar Nature of Ethernet Traffic / W.E. Leland [et al.] // ACM SIGCOMM'93. San Francisco, 1993. P. 183–193.
- 101. Park, K. Self-Similar Network Traffic and Performance Evaluation / K. Park, W. Willinger. New York: Wiley-Interscience, 2000 558 p.
- 102. Pawlikowski, K. Steady-State Simulation of Queueing Processes: a Survey of Problems and Solutions / K. Pawlikowski //ACM Computer Surveys. − 1990. − V. 22, №. 2. − P. 123–170.
- 103. Paxson, V. Wide-Area Traffic: The Failure of Poisson Modeling / V. Paxson, S. Floyd // IEEE/ACM Transactions on Networking. 1995. V. 3, № 3. P. 226–244.

- 104. Ramirez-Cobo, P. Bayesian analysis of a queueing system with a long-tailed arrival process / P. Ramirez-Cobo, R. E. Lillo, M. P. Wiper // Communication in Statistics. 2008.– V. 37. P. 697–712.
- 105. Simulation input analysis: difficulties in simulating queues with Pareto service / D. Gross [et al.] // Proceedings of the 34th Winter Simulation Conference: Exploring New Frontiers. San Diego, 2002. P. 407–415. DOI: 10.1145/1030453.1030510.
- 106. Srikant, R. Simulation run lengths to estimate blocking probabilities / R. Srikant, W. Whitt // ACM Transactions on modeling and computer simulation. 1996. V. 6, N 1. P. 7–52.
- 107. Systems with Multiple Servers under Heavy-tailed Workloads / K. Psounis [et. al.] // Performance Evaluation. 2005. V. 62, № (1–4). P. 456–474.
- 108. Wolfram Alpha [Электронный ресурс] // Source: Wolfram Alpha LLC. [S.I.], 2019. Режим доступа: https://www.wolframalpha.com. (дата обращения 29.04. 2019).
- 109. Taqqu, M. Proof of a Fundamental Results in Self-Similar Traffic Modeling / M. Taqqu, W. Willinger, R. Sherman // ACM SIGCOM Computer Communication Review. 1997. V. 27, № 2. P. 5–23.
- 110. The role of the Weibull Distribution in Internet traffic modeling / M. A. Arfeen, K. Pawlikowski, D. McNickle, A. Willig // 25th International Teletraffic Congress (ITC). Shanghai, 2013. P. 1–8. DOI: 10.1109/ITC.2013.6662948.
- 111. Zadorozhnyi, V. N. Cascade Method of Realization of Heavy-Tailed Distributions in Data Network Modelling / V. N. Zadorozhnyi // 2015 International Siberian conference on control and communications SIBCON, sec. Control of the Large Scale Systems. Omsk, 2015. P. 21–23.
- 112. Zadorozhnyi, V. N. Estimation of Prioritized Disciplines Efficiency Based on the Metamodel of Multi-flows Queueing Systems / V. N. Zadorozhnyi, T. R. Zakharenkova, D. A. Tulubaev // ITMM 2018. CCIS. 2018. V. 912. P. 290–304.
- 113. Zadorozhnyi, V. N. Fractal Queues Simulation Peculiarities / V. N. Zadorozhnyi // Communication in Computer and Information Science. 2015. V. 564. P. 415–432. DOI: 10.1007/978-3-319-25861-4 35.

- 114. Zadorozhnyi, V. N. Methods of Simulation Queueing Systems with Heavy Tails / V. N. Zadorozhnyi, T. R. Zakharenkova // Communications in Computer and Information Science. 2016. V. 638. P. 382–396.
- 115. Zadorozhnyi, V. N. Minimization of Packet Loss Probability in Network with Fractal Traffic / V. N. Zadorozhnyi, T. R. Zakharenkova // Communications in Computer and Information Science. 2017. V. 800. P. 168–183
- 116. Zadorozhnyi, V. N. Peculiarities and methods of fractal queues simulation / V. N. Zadorozhnyi // 2016 International Siberian Conference on Control and Communications (SIBCON). Moscow, 2016. P. 1–5. DOI: 10.1109/SIBCON.2016.7491713.
- 117. Zadorozhnyi V. N. Rapid Technique for the Calculation of Loss Probabilities in Queueing Systems / V. N. Zadorozhnyi, T. R. Zakharenkova, M. Pagano // Dynamics of Systems, Mechanisms and Machines (Dynamics). Omsk, 2018. P. 8601455-1–8601455-6. DOI: 10.1109/Dynamics.2018.8601455
- 118. Zadorozhnyi, V. N. Simulation modeling of fractal queues / V. N. Zadorozhnyi // Dynamics of Systems, Mechanisms and Machines (Dynamics). Omsk, 2014. P. 1–4. DOI: 10.1109/Dynamics. 2014.7005703.
- 119. Zwart, A. P. Queueing Systems with Heavy Tails / A. P. Zwart. Eindhoven: Eindhoven University of Technology, 2001. 227 p.

ПРИЛОЖЕНИЕ А

(обязательное)

ОРГАНИЗАЦИЯ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ ПРИ МОДЕЛИРОВАНИИ СМО СО СТЕПЕННЫМИ РАСПРЕДЕЛЕНИЯМИ

Листинг кода программы, рассчитывающей условное м.о. для первой страты в GPSS

```
****N=100 000 000
****страта 1
                    ((Duniform(1,100001,1000000)#10^-6)^(-1/ALPHA))1:90 000 000
                    ((Duniform(1,10001,100000)#10^-6)^(-1/ALPHA)) 2:90 000 00
****страта 2
                    ((Duniform(1,1001,10000)#10^-6)^(-1/ALPHA))
                                                                 3:90 000 0
****страта 3
****страта 4
                    ((Duniform(1,101,1000)#10^-6)^(-1/ALPHA))
                                                                  4:90 00 0
                    ((Duniform(1,11,100)#10^-6)^(-1/ALPHA))
****страта 5
                                                                  5:90 00
****страта 6
                    ((Duniform(1,2,10) #10^-6)^(-1/ALPHA))
                                                                  6:90 0
****страта 7
                    ((Duniform(1,1,1) #10^-6)^(-1/ALPHA))
                                                                  7:10
ALPHA
          EQU
                    1.1
                                            ;коэффициент альфа
DISTR
          TABLE
                    X1,1,10,100
                                            ;таблица для расчета
          GENERATE 1
          SAVEVALUE 1, ((Duniform(1,100001,1000000)#10^-6)^(-1/ALPHA))
          TABULATE DISTR
                                            ; составление эмп. распр. и вывод м.о.
          TERMINATE
          GENERATE 9000000
          TERMINATE
```

Последовательные и параллельные прогоны при моделировании систем со степенными распределениями

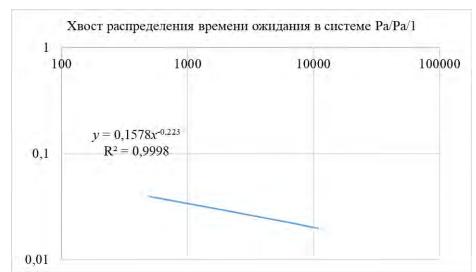


Рисунок А.1 — Хвост среднего времени ожидания системы Pa/Pa/1 с параметрами $\alpha_1 = \alpha_2 = 1,1$ и $K_1 = 1$, $K_2 = 0,1$ построенный по результатам моделирования в среде ИМ AnyLogic

Построение асимптотических доверительных интервалов

Ниже приводится часть кода на языке программирования Java, для вычисления корреляционной функции в системе M/M/1 с параметром λ = 1 и коэффициентом загрузки ρ = 0,75 .

```
transact counter++;//фиксируем число заявок
if(transact counter!=transact number+1)/до 10 000 000 заявок, сохраняем время ожи-
дания
{
      double current time=time()-entity.start wait time;
      waiting times array[0][transact counter-1]=transact counter;
      waiting times array[1][transact counter-1]=current time;
else //если у нас уже 10 000 000 заявок, считаем коррел. функцию, выводим её и
заканчиваем моделирование
      double mid mean=0;
      double mid sq mean=0;
      for(int ik=0;ik<transact counter-1;ik++)</pre>
            mid mean+=waiting times array[1][ik];
            mid sq mean+=waiting times array[1][ik]*waiting times array[1][ik];
      mean=mid mean/(double) transact number;
      mean squares=mid sq mean/(double) transact number;
      variance=(mean squares-(mean*mean));
      for (int i=0;i<lag;i++)//i+1-длина k
                  autocorr array[0][i]=i+1;
                  int cycle=transact number-i-1;
                  double cur summ=0;
                        for(int j=0;j<cycle;j++)</pre>
cur summ+=((waiting times array[1][j]-mean)*(waiting times array[1][j+i+1]-mean));
                  autocorr_array[1][i]=(cur summ/(variance*cycle));
            //вывод в файл
            try{
            File exp file=new File("correl function M M 1.txt");
            exp file.createNewFile();//создаем файл
            FileWriter writer ex=new FileWriter(exp file);//поток записи
                  for(int i=0;i<lag;i++)</pre>
      writer ex.append(autocorr array[0][i]+";"+autocorr array[1][i]+"\r\n");
      Autocorr function.add(i,autocorr array[1][i]);
            writer_ex.flush();//закрываем
            writer ex.close();
            finishSimulation();
            catch (IOException exeption ) { } }
```

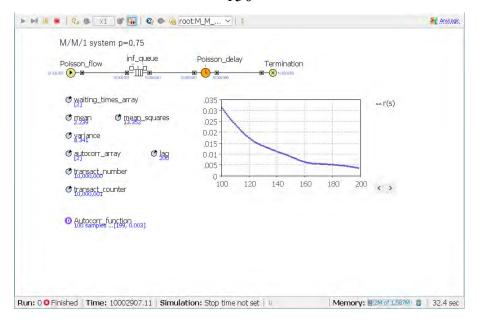


Рисунок А.2 — Результаты ИМ системы М/М/1 при λ = 1, ρ = 0,75 в AnyLogic; r(s) — корреляционная функция между значениями i-го и (i+s)

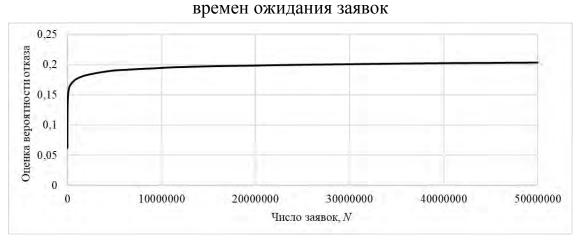


Рисунок А.3 — Определение переходного процесса для оценки вероятности отказа в Pa/Pa/1/100 системе с параметрами $\alpha_1=\alpha_2=1,1$ и $K_1=1,K_2=0,5$

Метод корректной реализации с.в. с РТХ в AnyLogic

приложение Б

(обязательное)

РАСЧЕТ И УМЕНЬШЕНИЕ ВЕРОЯТНОСТЕЙ ПОТЕРЬ В СМО И СЕМО СО СТЕПЕННЫМИ РАСПРЕДЕЛЕНИЙМИ

Листинг кода программы, рассчитывающей вероятности потерь в СМО Pa/Pa/54/0

```
run lght
          EQU
                    60000
                                     ; длина прогона модели
          run num
                    EQU
                                      10001
CAP1
          STORAGE
          TABLE
                    X1,1,1,3
Χ
                                     ;Х1 - частотное распределение
                    ,11000,<del>10001</del>
         MATRIX
                                    ;индикаторы отказа в каждом прогоне
TND
MVTransC MATRIX
                    ,9995,1
                                     ;матрица с вероятностями
MVTransCm MATRIX
                    ,9995,1
                                     ;усредненная верхняя по прогонам
          INITIAL X$N of runs,1
*----- Моделирование системы Pal/Pa2/n/0, рассчитывается Potk на дискретном t, в данном
блоке формируется матрица индикаторов
          GENERATE (MtArandPareto(0.043,1.5))
          TEST G
                    AC1,59000,KKK
          SAVEVALUE trunz num+,1
          TEST L S$CAP1,54,LOSS
          SAVEVALUE 1,0
          TABULATE X
          MSAVEVALUE IND, X$trunz num, X$N of runs, X1 ;значение индикатора отказа в
случае обслуживания
          ENTER
                   CAP1
          ADVANCE (MtArandPareto(1.1,1.4))
          LEAVE
                    CAP1
          TERMINATE
LOSS
          SAVEVALUE 1,1
          TABULATE
          MSAVEVALUE IND, X$trunz_num, X$N of runs, X1
                                                     ;значение индикатора отказа в
случае отказа = 1
         TERMINATE
         TEST G
                   AC1,50000,KKK2
         TEST L
                   S$CAP1,54,LOSS2
                                    ; функционирование при стационарном режиме
          ENTER
          ADVANCE (MtArandPareto(1.1,1.4))
          LEAVE
                    CAP1
          TERMINATE
LOSS2
         TERMINATE
         TEST L S$CAP1,54,END_
KKK2
                                     ; функционирование при стационарном режиме
          ENTER
                    CAP1
          ADVANCE
                    (MtArandPareto(1.1,1.4))
                   CAP1
         LEAVE
END
         TERMINATE
*Графическое отображение вероятности отказа во времени
          GENERATE ,,50005,1 ;через 50005 м.в. генерируется заявка, который проверяет
число потерь заявок
         ASSIGN
                    1.1
                    2,9995
          ASSIGN
                 3, (N$LOSS+N--
4, (N3+N$BLTEST)
CYCLE_
                    3, (N$LOSS+N$LOSS2)
          ASSIGN
          ASSIGN
          TEST NE P4,0,NULL
          MSAVEVALUE MVTransC+,P1,1,(P3/P4); записывается текущее значение вероятности отказа
          ADVANCE
```

```
1+,1
          ASSIGN
                                               ;переход к следующей строке матрицы
          LOOP
                     2,CYCLE
          TERMINATE
          MSAVEVALUE MVTransC+, P1, 1, 0
NULL
          ADVANCE 1
          ASSIGN
                    1+,1
                                               ;переход к следующей строке матрицы
          LOOP
                    2,CYCLE
          TERMINATE
*Таймер
          GENERATE run_lght,,,1
          ;Расчет числовых характеристик
          SAVEVALUE MO+, TB$X
          SAVEVALUE SKO+, TD$X
           *Если прогон последний:
          TEST E
                  X$N of runs, run num, MET2
          *1) усредняем значения p(t) для конечнолин СМО
                   3,1
          ASSIGN
                    4,9995
          ASSIGN
MET3
          MSAVEVALUE MVTransCm, P3, 1, (MX$MVTransC(P3, 1)/run num)
          ASSIGN 3+,1
          LOOP
                    4,MET3
          *2) запись в файл
          OPEN ("54.TXT"),1,Done; открытие файла и создание потока данных
file
          ASSIGN 1,1
ASSIGN 9,9995
                                       ; число членов корреляционной ф
                                       ; длина цикла ѕ
          WRITE MX$MVTransCm(P1,1); запись в открытый файл значение члена кор ф ASSIGN 1+,1 LOOP 9,DoWrite
DoWrite
                  Error_Parm,1 ; закрытие потока
          CLOSE
Done
          SAVEVALUE File_Error,P$Error_Parm
          *3) усреднее мо и ско
          SAVEVALUE MO, (X$MO/run num)
          SAVEVALUE SKO, (X$SKO/run num)
          *4) зануление
          SAVEVALUE N of runs+,1
MET2
          SAVEVALUE trunz_num,0
          SAVEVALUE 1,0
            TERMINATE
```

Листинг кода ускоренного метода расчета вероятности потерь в AnyLogic для CMO вида M/Pa/1/m

Расчет оценки числа потерянных заявок.

```
test=RNG object.ParetoDistributon(1,1.25);//для проверки среднего значения
total agent number=PoissonFlow.out.count();//число сгенерированных заявок
if(Queue m finite.size()==0 && ParetoDelays.size()==0) //если система пуста
            if (busy periods==0)
             \{//если ПНЗ первый, тогда
                   busy periods++;//увеличиваем текущее значение ПНЗ
                   fixed losses after PR=0;
                   fixed agents after PR=0;
            else//если ПНЗ не первый, тогда увеличиваем число ПР на 1
                   number losses in PR=Outputs.outF.count()-fixed losses af-
ter PR;//число потерь
      fixed losses after PR=Outputs.outF.count();//зафиксировали
      number agents in PR=Outputs.in.count()-fixed agents after PR;//число всех заявок
      fixed agents after PR=Outputs.in.count();//зафиксировали
      mass N[busy periods-1][0]=busy periods;
      mass N[busy periods-1][1]=number losses in PR;
      mass N[busy periods-1][2]=number agents in PR;
      busy periods++;//увеличиваем текущее значение ПНЗ
      if(N loss max<number losses in PR)</pre>
             {N loss max=number losses in PR;}
      if (N max<number agents in PR)</pre>
```

```
{N_max=number_agents_in_PR;}}
```

Реализация параллельных прогонов.

Метод оптимального распределения каналов по узлам СеМО с РТХ

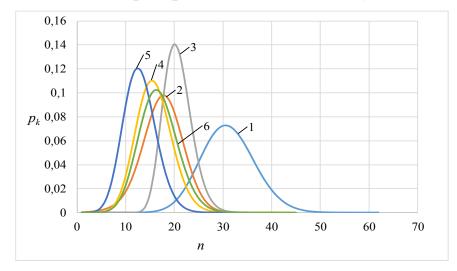


Рисунок Б.1 – Распределение вероятностей состояний в узлах СеМО

Таблица Б.1 – Результаты имитационного моделирования шестиузловой СеМО для различных распределений n_i

No॒	Узел 1	Узел 2	Узел 3	Узел 4	Узел 5	Узел 6	P	P
							(аппроксимация)	(ИМ)
-	52	33	34	31	27	33	6,848·10 ⁻⁴	2,478·10-4
1	51	33	34	31	27	34	-	2,194·10-4
2	53	33	34	31	27	32	-	2,792·10 ⁻⁴
3	51	33	34	31	28	33	-	2,254·10 ⁻⁴
4	53	33	34	31	26	33	-	3,184·10-4
5	51	33	34	32	27	33	-	1,586·10-4
6	53	33	34	30	27	33	-	4,136·10-4
7	51	33	35	31	27	33	-	2,404·10 ⁻⁴
8	53	33	33	31	27	33	-	2,812·10-4
9	51	34	34	31	27	33	-	2,428·10-4
10	53	32	34	31	27	33	-	2,839·10 ⁻⁴
11	52	32	34	31	27	34	-	2,371·10 ⁻⁴
12	52	34	34	31	27	32	-	2,689·10-4
13	52	32	34	31	28	33	-	2,289·10-4
14	52	34	34	31	26	33	-	3,234·10 ⁻⁴

Продолжение таблицы Б.1

15	52	32	34	32	27	33	-	1,649·10 ⁻⁴
16	52	34	34	30	27	33	-	4,157·10 ⁻⁴
17	52	32	35	31	27	33	-	2,461·10 ⁻⁴
18	52	34	33	31	27	33	-	2,814·10 ⁻⁴
19	52	33	33	31	27	34	-	2,584·10 ⁻⁴
20	52	33	35	31	27	32	-	2,752·10 ⁻⁴
21	52	33	33	31	28	33	-	2,413·10 ⁻⁴
22	52	33	35	31	26	33	-	3,103·10 ⁻⁴
23	52	33	33	32	27	33	-	1,959·10 ⁻⁴
24	52	33	35	30	27	33	-	4,201·10 ⁻⁴
25	52	33	34	30	27	34	-	3,859·10 ⁻⁴
26	52	33	34	32	27	32	-	2,008·10 ⁻⁴
27	52	33	34	30	28	33	-	3,817·10-4
28	52	33	34	32	26	33	-	2,412·10-4
29	52	33	34	31	26	34	-	3,231·10 ⁻⁴
30	52	33	34	31	28	32	-	2,409·10 ⁻⁴

Листинг кода программы, для расчета вероятности потерь в системе M/Pa/1/10 с абсолютными приоритетами

```
****** Модель CMO c КБ ******
N_
              10000
          EQU
                                     ;Число приоритетных классов
Κ_
          EQU
                    (8/3)
ALPHA
         EQU
                    1.5
* Формирование потока заявок
         GENERATE (MtExponential(0,10))
* Обслуживание заявки и корректировка содержимого ячеек QMAX, КРNZ
         ASSIGN time left, (MtArandPareto(K , ALPHA))
          TEST GE P$time_left,(K_+N_-1),PING
          ASSIGN Class, \overline{0}
          TRANSFER , PONG
PING
         ASSIGN Class, (N -Int(P$time left-K))
         PRIORITY (2#P$Class+1)
PONG
          SAVEVALUE CLL_+, (P$CLASS#(P$CLASS<0))
          SAVEVALUE CLB_+, (P$CLASS#(P$CLASS>10000))
          SAVEVALUE PRL_+, (PR#(PR<0))
          SAVEVALUE PRB +, (PR# (PR>20001))
          TEST L Q1,10, TERM
WX
          QUEUE
                    1
         PREEMPT 1, PR, inc_prior, time_left, RE
Kanal
          TEST L P$time_left,0,adv_
          SAVEVALUE Ost Time+, P$time left
         ASSIGN time left, 0
adv_
         ADVANCE P$time left
          RETURN 1
          DEPART
                    1
          TERMINATE
TERM
          TERMINATE
inc prior TEST E
                    (PR@2),1,Kanal
          PRIORITY
                    (PR+1)
                                             ;увеличиваем приоритет на 1
          TRANSFER , Kanal
                                             ;возвращаемся в очередь перед прибором
*** TIMER
          GENERATE 1000000
          SAVEVALUE P OTK, (N$TERM/N1)
          TERMINATE 1
```

приложение в

(справочное)

СВИДЕТЕЛЬСТВА О ГОСУДАРСТВЕННОЙ РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ

POCCHÜCKASI DELIEPAUUS





МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное бюджетное научное учреждение ИНСТИТУТ УПРАВЛЕНИЯ ОБРАЗОВАНИЕМ РОССИЙСКОЙ АКАДЕМИИ ОБРАЗОВАНИЯ

ОБЪЕДИНЕННЫЙ ФОНД ЭЛЕКТРОННЫХ РЕСУРСОВ "НАУКА И ОБРАЗОВАНИЕ" (основан в 1991 году)

СВИДЕТЕЛЬСТВО О РЕГИСТРАЦИИ ЭЛЕКТРОННОГО РЕСУРСА

Nº 22124

Настоящее свидетельство выдано на электронный ресурс, отвечающий требованиям новизны и приоритетности:

Комплекс GPSS-процедур для генерации случайных величин на основе генератора «Вихрь Мерсенна»

Дата регистрации: 01 сентября 2016 года

Авторы: Задорожный В.Н., Захаренкова Т.Р., Мясищев С.В.

Организация-разработчик: ФГБОУ ВО «Омский государственный технический университет»

Директор ФГБНУ ИУО РАО, доктор экономических наук

С.С. Неустроев

Руководитель ОФЭРНиО, почетный работник науки и техники России

осожина Галкина



ПРИЛОЖЕНИЕ Г

(справочное)

АКТЫ О ВНЕДРЕНИИ РЕЗУЛЬТАТОВ ДИССЕРТАЦИОННОЙ РАБОТЫ

Казенное учреждение города Омска

«УПРАВЛЕНИЕ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ ТЕХНОЛОГИЙ»

Юридический адрес: Гагарина ул., д. 34, Омск, 644099 Почтовый адрес: Чернышевского ул., д. 6, Омск, 644007 Тел. (3812) 78-79-00

Or 08.04 2019 No Clo-98151/0

УТВЕРЖДАЮ

Директор Казенного учреждения города Омска у правыше информационно-

омилинкационных технологий»

_И.Н. Катунин

прессея 2019 г

AKT

о внедрении результатов диссертационной работы Захаренковой Татьяны Романовны на тему «Аналитико-статистические методы расчёта и оптимизации систем и сетей массового обслуживания со степенными хвостами распределений», представленной на соискание учёной степени кандидата технических наук по специальности 05.13.18 — Математическое моделирование, численные методы и комплексы программ

Комиссия в составе:

председатель - А. А. Волуйко, начальник отдела телекоммуникаций,

члены комиссии - А. Л. Зайцева, ведущий инженер отдела телекоммуникаций,

И. В. Ерофеев, ведущий инженер отдела телекоммуникаций

составили настоящий акт о том, что результаты диссертационной работы «Аналитико-статистические методы расчёта и оптимизации систем и сетей массового обслуживания со степенными хвостами распределений» Захаренковой Т.Р. использованы в отделе телекоммуникаций Казенного учреждения города Омска «Управление информационно-коммуникационных технологий» в качестве практических рекомендаций по применению эффективных способов снижения потерь сообщений.

В целом разработанные в диссертации методы позволяют сократить временные и финансовые затраты при проектировании сетей передачи данных и сетевых узлов, а также резко снизить вероятности потерь пакетов в сетевых узлах. Ежегодный непрерывный рост объемов сетевого трафика обусловливает высокую актуальность полученных в диссертации результатов.

Начальник отдела телекоммуникаций

Ведущий инженер отдела телекоммуникаций

Ведущий инженер отдела телекоммуникаций

А. А. Волуйко

А. Л. Зайцева

И. В. Ерофеев

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «ОМСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

УТВЕРЖЛАЮ

Проректор по учебной работе Омского государственного технического университета, д.т.н.

О.В. Кропотин

2019 г.

AKT

о внедрении результатов диссертационной работы на тему «Аналитико-статистические методы расчёта и оптимизации систем и сетей массового обслуживания со степенными хвостами распределений» Захаренковой Татьяны Романовны в учебный процесс ОмГТУ

Комиссия в составе:

председатель комиссии – Макаров В.В., к.т.н., декан факультета информационных

технологий и компьютерных систем;

 Бояркин Г.Н., д.э.н., зав. кафедрой «Математические методы и члены комиссии

информационные технологии в экономике»;

- Потапов В.И., д.т.н., зав. кафедрой «Информатика и

вычислительная техника»

составили настоящий акт о следующем:

Результаты диссертационной работы Т.Р. Захаренковой используются при изучении дисциплины «Имитационное моделирование» студентами факультета информационных технологий и компьютерных систем направлений подготовки 09.03.03 «Прикладная информатика». Разработанные в диссертационной работе программные модули в среде имитационного моделирования GPSS World и AnyLogic применяются при проведении практических и лабораторных работ по указанной дисциплине для исследования систем и сетей массового обслуживания со степенными хвостами распределений с конечным первым и бесконечным вторым моментами.

Председатель комиссии к.т.н., декан факультета информационных технологий и компьютерных систем

В.В. Макаров

Члены комиссии:

д.э.н., зав. кафедрой «Математические методы и информационные технологии в экономике»

д.т.н., зав. кафедрой «Информатика и вычислительная техника»

Б. И. Потапов