

На правах рукописи



Бородина Александра Валентиновна

**РЕГЕНЕРАТИВНАЯ МОДИФИКАЦИЯ
МЕТОДА РАСЩЕПЛЕНИЯ
ДЛЯ ОЦЕНИВАНИЯ ВЕРОЯТНОСТИ ПЕРЕГРУЗКИ
В СИСТЕМАХ ОБСЛУЖИВАНИЯ**

Специальность 05.13.18 — математическое моделирование,
численные методы и комплексы программ

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Петрозаводск – 2008 г.

Работа выполнена в Институте прикладных математических исследований Карельского научного центра Российской академии наук.

Научный руководитель: доктор физико-математических наук,
профессор Евсей Викторович Морозов

Официальные оппоненты: доктор физико-математических наук,
профессор Владимир Васильевич Рыков
кандидат физико-математических наук,
доцент Ирина Валерьевна Пешкова

Ведущая организация: Московский государственный университет

Защита состоится "....."..... 200... г. в час. на заседании диссертационного совета Д 212.190.03 в Петрозаводском государственном университете по адресу: 185910, г. Петрозаводск, пр. Ленина, 33.

С диссертацией можно ознакомиться в библиотеке Петрозаводского государственного университета.

Автореферат разослан "....."..... 200... г.

Учёный секретарь
диссертационного совета
Д 212.190.03



В. В. Поляков

Общая характеристика работы

Актуальность темы. В настоящее время в связи с развитием процессов информатизации, информационно-вычислительных систем, систем связи и передачи данных актуальными становятся сетевые модели, так называемые сети массового обслуживания, которые являются обобщением систем массового обслуживания.

В качестве показателей эффективности функционирования системы обслуживания, т. е. степени приспособленности к выполнению задач, для которых она создана, чаще всего используются математические ожидания стационарного суммарного числа требований в системе, стационарного времени пребывания требования, стационарного времени ожидания, периода занятости и свободного периода.

Требования к повышению качества обслуживания (QoS) диктуются потребностью в надежных системах обслуживания с высокой устойчивостью к сбоям. Известным примером является сбой в сети AT&T в 1990 году, когда на протяжении 9 часов сеть была недоступна большей части зданий Нью-Йорка, что привело к большим финансовым потерям.

Существует много методов исследования характеристик современных коммуникационных систем и сетей. Одни включают аналитические и численные методы и применяются в основном для простых систем с независимыми данными. Другой подход - это методы имитационного моделирования, самым известным среди которых является метод Монте-Карло.

Целью имитационного моделирования систем (сетей) обслуживания обычно является построение оценки некоторых стационарных или переходных характеристик, от которых зависит качество системы (сети) с точки зрения пользователя. Примерами таких характеристик являются стационарное время ожидания заявок в очереди, вероятность потери заявок при перегрузке буфера требований, вероятность превышения интервала ожидания заявки в очереди, вероятность переполнения буфера до того, как система станет пустой и т.д. При этом оцениваемые характеристики очень малы (порядка 10^{-9} и ниже), следовательно применение метода Монте-Карло требует слишком много времени для получения результата с высокой точностью.

Поэтому актуальной задачей является разработка методов ускоренного имитационного моделирования для оценивания вероятностей редких событий. Основная идея заключается в изменении статистического поведения исходного процесса так, чтобы редкое событие стало более вероятным. Существует два основных подхода:

1. *RESTART/метод расщепления:* вводится последовательность вложенных подмножеств множества состояний процесса и уровни, с которых достижение редкого множества становится более вероятным. При достижении каждого уровня происходит запуск сразу нескольких копий процесса.

2. *Метод существенной выборки*: за счет изменения вероятностной меры редкое событие становится более вероятным.

Следует отметить, что несмотря на большой интерес к методам, основанным на ветвлении процесса (RESTART, метод расщепления), в настоящее время вопрос о статистических свойствах оценок остается в значительной степени открытым. Известно, что оценка вероятности переполнения очереди при использовании метода расщепления для марковских систем является несмещенной. Однако, только ограниченный класс систем вида $M/M/1$ может быть описан марковским процессом. Более того, состоятельность оценки и доверительное оценивание на основе метода расщепления в общем случае тоже является открытой проблемой. Исследование состоятельности и асимптотической нормальности оценки для адаптивного многоуровневого метода расщепления было лишь недавно (2005 г.) проведено для некоторого узкого класса марковских процессов.

В связи с этим, актуальной задачей является исследование стат. свойств оценки в методе расщепления. А также, исследование применимости метода для немарковских процессов обслуживания.

В данной работе основное внимание уделяется системам обслуживания, которые характеризуются пуассоновским входным потоком и произвольно распределенной длительностью обслуживания при одном обслуживающем канале, а также системам общего вида (т. е. системам $M/G/1$, $GI/G/1$ в символике Кендалла), где процесс очереди не является марковским.

Основными методами в диссертации являются метод регенеративного моделирования, который применим также в случае зависимых циклов регенерации, и метод расщепления.

Цель диссертационной работы заключается в разработке метода ускоренного регенеративного имитационного моделирования (УРИМ) и построении на его основе состоятельных и асимптотически нормальных оценок вероятности перегрузки в одноканальных системах обслуживания. В работе решаются следующие основные задачи.

1. Расширить область применения метода расщепления, модифицировав исходный алгоритм для моделирования процесса нагрузки.
2. Построить состоятельную оценку стационарной вероятности перегрузки в одноканальной системе общего вида, методом УРИМ для процессов очереди и нагрузки.
3. Обосновать асимптотическую нормальность оценки стационарной вероятности перегрузки, а также оценки вероятности превышения заданного уровня на цикле регенерации, построенных методом УРИМ, для процессов очереди и нагрузки.
4. Исследовать дисперсию оценки при замене немарковского процесса очереди в системе $M/G/1$ на вложенную цепь Маркова.

5. Исследовать влияние зависимости циклов регенерации в методе УРИМ на точность доверительного оценивания.

Научная новизна работы заключается в том, что для исследования свойств оценки вероятности перегрузки в методе расщепления впервые применяется теория регенерации.

При построении оценок вероятности перегрузки в немарковской системе обслуживания $M/G/1$ методом расщепления применен метод вложенной цепи Маркова.

Обоснована состоятельность и асимптотическая нормальность оценки, полученной в традиционном методе расщепления, для регенерирующих процессов.

Расширена область применения метода расщепления на случай оценки стационарной вероятности перегрузки, а также для проведения моделирования процесса нагрузки. Разработано программное обеспечение для оценивания вероятности перегрузки для процессов очереди и нагрузки в системах $M/M/1$, $M/G/1$, $GI/G/1$, а так же интервального оценивания методом УРИМ.

Практическую ценность в работе представляет построенная регенеративная модель траекторий процесса, полученных методом расщепления, а также метод построения состоятельных и асимптотически нормальных оценок вероятности достижения высокого уровня на цикле регенерации и стационарной вероятности перегрузки при моделировании процессов очереди и времени ожидания.

Положения, выносимые на защиту. На защиту выносятся следующие положения.

1. Разработан метод ускоренного регенеративного имитационного моделирования (УРИМ).
2. Построена состоятельная оценка стационарной вероятности перегрузки, получаемая методом УРИМ.
3. Доказана состоятельность оценок, построенных методом УРИМ.
4. Разработан алгоритм построения состоятельной оценки вероятности перегрузки и вероятности достижения заданного уровня на цикле регенерации процессом очереди и процессом нагрузки в системах $M/G/1$, $GI/G/1$. Разработан комплекс программ для построения оценок.
5. Предложен метод уменьшения дисперсии оценки для процесса очереди в $M/G/1$ на основе вложенной цепи Маркова.
6. Разработан алгоритм доверительного оценивания стационарной вероятности перегрузки и вероятности достижения заданного уровня на цикле с учетом зависимости циклов регенерации, полученных методом УРИМ. Показано, что зависимость циклов существенно влияет на точность доверительного оценивания.

Связь работы с научными программами, темами. Работа поддержана Российским Фондом Фундаментальных исследований, гранты 04-01-00671, 04-07-90115, 07-07-00088.

Апробация работы. Результаты работы докладывались и обсуждались на следующих конференциях:

1. 23 - 25 ноября 2005. Международная школа-конференция "Информационно-телекоммуникационные системы" МГИЭТ. Москва. Ускоренные методы моделирования в регенерирующих процессах обслуживания.
2. 23 - 25 августа 2006. Annual International Workshop. Advances in Methods of Information and Communication Technology (AMICT'06). ПетрГУ. Regenerative rare event estimation based on permutations.
3. 26 - 31 августа 2006. Russian-Scandinavian Symposium "Probability theory and applied probability"(PTAP'06). Using of regenerative sequences in Splitting method.
4. 21 - 22 августа 2007. Annual International Workshop. Advances in Methods of Information and Communication Technology (AMICT'07). ПетрГУ. Regeneration cycles dependence in Confidence Estimation by Splitting.
5. 10 - 12 сентября 2007. International workshop. Distributed Computer and Communication Networks: Theory and Application (DCCN'2007). Институт проблем передачи информации им А. А. Харкевича, Москва.
6. 22 - 26 октября 2007. XXVI International Seminar Stability Problems for Stochastic Models, Nahariya, Israel. Speed-up consistent estimation of a high workload probability in M/G/1 queue.
7. 20 - 22 мая 2008. Annual International Workshop. Advances in Methods of Information and Communication Technology (AMICT'07). ПетрГУ. On Modification of Regenerative Splitting for Embedded Markov Chain.
8. 1 - 6 июня 2008. VII Международная Петрозаводская конференция "Вероятностные методы в дискретной математике". Регенеративная модификация метода расщепления.
9. 24 - 26 сентября 2008. RESIM, International Workshops on Rare Event Simulation, Rennes, France. A regenerative modification of the multilevel splitting.

По материалам диссертации опубликовано 10 работ, из них 8 статей [1, 2, 3, 4, 5, 6, 7, 8] и тезисы 2 докладов [9, 10]

Структура и объем работы.

Диссертационная работа состоит из введения, четырех глав, заключения и списка литературы. Общий объем диссертации составляет 117 страниц. Список литературы включает 72 наименования.

Содержание работы

Во введении отражена актуальность работы, поставлена цель исследования, обоснована новизна работы, сформулированы положения, выносимые на защиту, показана практическая ценность полученных результатов.

В первой главе рассматривается проблема оценивания вероятности редких событий. Требуется построить оценку вероятности

$$\gamma := \int I(x \in A)f(x)dx = EI(X \in A) = P(X \in A),$$

где X - случайная величина с плотностью $f(x)$, принимающая значения из множества E , $A \subset E$, $I(X \in A)$ - индикатор. Если $\gamma \approx 0$, то событие $\{X \in A\}$ называют *редким*, а множество A *редким множеством*.

Дается обзор существующих методов имитационного моделирования, применяемых для оценивания вероятностей перегрузки очень малого порядка. К ним относятся метод классического имитационного моделирования Монте-Карло (МК), а также современные методы *ускоренного моделирования*: метод существенной выборки, RESTART, метод расщепления.

При оценивании вероятностей очень малого порядка (например, 10^{-20} и меньше), классический метод МК становится неэффективен. Оценка $\hat{\gamma}$ строится в форме выборочного среднего

$$\hat{\gamma} = \frac{1}{N} \sum_{n=1}^N I(X_n \in A).$$

Поэтому, при стремлении $\gamma \rightarrow 0$ относительная ошибка оценивания

$$\Lambda(\hat{\gamma}) = \frac{\sqrt{Var(\hat{\gamma})}}{E\hat{\gamma}} = \frac{\sqrt{\gamma - \gamma^2}}{\gamma\sqrt{N}} \sim \frac{1}{\sqrt{\gamma N}} \rightarrow \infty.$$

Таким образом, чтобы достигнуть заданной точности при малых значениях вероятности γ требуется огромный размер выборки N , и значительно возрастает время моделирования.

Метод существенной выборки основан на идее сделать редкое событие более вероятным, т. е. *ускорить* моделирование за счет изменения меры (выбора новой плотности $\tilde{f}(x)$). Новая плотность $\tilde{f}(x)$ выбирается с целью минимизировать дисперсию оценки.

Эффективность метода напрямую зависит от удачного выбора $\tilde{f}(x)$. Оптимальный выбор $\tilde{f}(x)$ возможен при известном γ (тогда $Var(\hat{\gamma}) = 0$), т. е. когда решение можно получить аналитически. Для сложных систем, где аналитическое решение получить невозможно, использование метода не всегда приводит к уменьшению дисперсии. Для некоторых $\tilde{f}(x)$, $Var(\hat{\gamma}_N(\tilde{f})) \rightarrow \infty$. В этой связи, иногда выгоднее воспользоваться стандартным методом моделирования МК, чем найти "хорошую" плотность $\tilde{f}(x)$.

Метод расщепления и метод RESTART являются методами ускоренного имитационного моделирования. Суть обоих методов состоит в разделении пространства состояний процесса на M вложенных подмножеств, границы которых называют *порогами (уровнями)* L_i , $i \in [0, M]$. При достижении некоторого уровня, траектория процесса расщепляется на несколько траекторий-потомков. Те из них, которые достигнут следующего уровня, тоже будут расщеплены. Между порогами траектория процесса моделируется стандартным методом МК. Таким образом, при увеличении количества траекторий, увеличивается шанс достижения редкого множества. В формулу оценок входит множитель, компенсирующий искусственное размножение траекторий.

В методе RESTART все траектории, кроме одной, обрываются при пересечении (сверху-вниз) уровня, с которого они стартовали. Предполагается, что все траектории пронумерованы. Далее моделируется только одна (как правило последняя) траектория, которая будет расщепляться при пересечении любого уровня.

В методе расщепления траектории не обрываются, однако расщепление возможно только при пересечении более высоких (по сравнению с уровнем старта) уровней.

Метод расщепления применяется для оценивания вероятности достижения процессом высокого уровня, до возвращения в нуль. Метод RESTART предназначен для оценивания стационарной вероятности превышения процессом заданного уровня.

Преимуществом метода RESTART и метода расщепления является существенное сокращение времени получения оценки по сравнению с методом МК. Однако, статистические свойства оценок (в первую очередь, состоятельность и асимптотическая нормальность) недостаточно исследованы.

Для марковского процесса доказана несмещенность оценки. В общем случае задача обоснования свойств оценок, как и задача оптимального выбора порогов и количества расщеплений на каждом пороге, остается открытой.

При моделировании немарковского процесса, траектории после расщепления не являются независимыми. Это может быть источником неустойчивости оценки. Также при доверительном оценивании необходимо учесть зависимость между траекториями, возникающую вследствие расщепления.

Во **второй главе** даны необходимые определения, свойства и теоремы теории регенерирующих процессов. Описан метод регенеративного моделирования для построения состоятельных и асимптотически нормальных оценок стационарных характеристик регенерирующего процесса на основе регенеративной центральной предельной теоремы (РЦПТ).

Описаны достаточные условия применимости регенеративного метода и доверительное регенеративное оценивание в неклассическом случае k -зависимых циклов регенерации.

В **третьей главе** описан метод ускоренного регенеративного имитационного моделирования (УРИМ) малых вероятностей. Рассматриваются два сле-

дующих случайных процесса, непрерывных справа.

- Процесс очереди $\{\nu(t), t \geq 0\}$, где $\nu(t)$ - количество заявок в системе в момент времени t , включая заявку, которая обслуживается сервером. Для стационарной очереди обозначим слабый предел $\nu(t) \Rightarrow \nu$ (что означает слабую сходимоссть $P(\nu(t) \in \cdot) \Rightarrow P(\nu \in \cdot)$ при $t \rightarrow \infty$).
- Процесс нагрузки (времени ожидания) $\{W_n, n \geq 1\}$, где W_n - время ожидания в очереди заявки с номером n . Обозначим W_∞ стационарное время ожидания, когда существует слабый предел $W_n \Rightarrow W_\infty$.

Сначала рассмотрим процесс очереди, а затем, модифицировав алгоритм моделирования, исследуем процесс нагрузки. Рассмотрим систему обслуживания $GI/G/1$ с бесконечным буфером для ожидающих требований, в предположении, что коэффициент загрузки $\rho < 1$.

Обозначим t_n и t_n^* - момент прихода и ухода заявки с номером n , соответственно. Определим для процесса очереди $X = \{\nu(t), t \geq 0\}$ последовательность моментов регенерации уровня 0 (моментов 0-регенерации) $\tau = (\tau_i, i \geq 1)$, где моменты τ_i определяются рекурсивно:

$$\tau_0 = 0, \tau_i = \min_{n \geq 1} (t_n > \tau_{i-1} : \nu(t_n - 0) = 0).$$

Обозначим $\alpha_i = \tau_i - \tau_{i-1}$, $i \geq 1$, длины циклов регенерации уровня 0.

Далее будем строить точечные оценки для следующих вероятностей:

1. вероятность того, что очередь достигнет (большого) значения L на цикле регенерации, т. е.

$$\gamma^{(1)} = P(\max_{0 \leq t < \alpha_1} \nu(t) \geq L),$$

2. вероятность того, что стационарная очередь превысит значение L , т. е.

$$\gamma^{(2)} = P(\nu \geq L).$$

При больших значениях уровня L вероятности $\gamma^{(1)}$ и $\gamma^{(2)}$ малы и более того $\gamma^{(1)} \rightarrow 0$, $\gamma^{(2)} \rightarrow 0$ при $L \rightarrow \infty$.

Во втором параграфе дана регенеративная интерпретация метода расщепления. Задача построения оценки вероятности редкого события может быть сведена к задаче построения (выделения) циклов регенерации с использованием техники ускоренного моделирования.

Метод построения циклов регенерации на основе расщепления приводится для двухуровневой модели (с единственным и множественными расщеплениями). Затем метод обобщается для общего случая многоуровневой модели расщепления с несколькими расщеплениями на каждом уровне. Построенная математическая модель циклов регенерации является основой метода УРИМ.

В методе УРИМ моделирование происходит по циклам регенерации, которые сгруппированы в пучки. Циклы регенерации принадлежат одному пучку,

если они имеют общую точку расщепления на первом уровне (L_1). Моделирование пучка начинается при старте одной траектории с уровня $L_0 = 0$. Расщепляя траекторию на заданных уровнях L_i , мы увеличим вероятность достижения уровня L в $D := R_1 R_2 \cdots R_M$ раз, где R_i - число расщеплений на уровне L_i .

С точки зрения теории регенерации это означает, что вместо одного цикла регенерации (как в классическом методе МК) одним запуском мы генерируем D (зависимых) циклов регенерации (пучок).

Если траектория, расщепившаяся на уровне L_i , не достигает уровня L_{i+1} , то она, очевидно, не может расщепиться и на последующих уровнях L_{i+2} , $L_{i+3} \dots$. В этом случае фактически мы теряем $D_i := R_{i+1} R_{i+2} \cdots R_M$ циклов регенерации. Чтобы число циклов регенерации совпадало с компенсирующим множителем (иначе оценка не будет компенсироваться полностью), мы считаем, что на самом деле имеется D_i идентичных циклов регенерации, совпадающих с данной траекторией.

Используя эту идею, для каждой траектории, стартующей с уровня L_0 , мы получим пучок циклов, состоящий из $D := R_1 \cdot R_2 \cdots R_M$ зависимых циклов регенерации.

Важно, что фактическое построение циклов регенерации при моделировании не требуется. Это процедура необходима лишь для формального описания регенеративной структуры моделируемого процесса с целью обоснования статистических свойств получаемых оценок.

Последовательность циклов регенерации, полученная методом УРИМ, представлена на рис. 1. Все циклы можно разбить на R_0 групп зависимых цик-

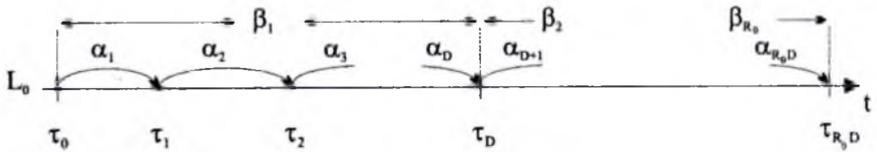


Рис. 1: Регенеративная последовательность уровня L_0

лов, каждая группа содержит ровно $D < \infty$ циклов. Для любого $j \in \{(i - 1)D + 1, iD\}$ длины циклов α_j являются одинаково распределенными, но зависимыми случайными величинами, $i \in [1, R_0]$. Всего генерируется $n := R_0 \cdot R_1 \cdots R_M = R_0 \cdot D$ циклов.

На рис. 1 видно, что последовательность зависимых циклов регенерации сама регенерирует в моменты $\tau_{i \cdot D}$, $i \in [0, R_0]$, когда стартует новый пучок траекторий с уровня L_0 . Поэтому последовательность зависимых циклов является регенерирующей. Это можно назвать вложенной регенерацией. Циклы

вложенной (по моментам старта пучков) регенерации являются независимыми. Длины таких циклов - независимые одинаково распределенные случайные величины (н. о. р. с. в.), и равны сумме из D длин исходных (зависимых) циклов, составляющих данный пучок:

$$\beta_i = \sum_{j=(i-1) \cdot D+1}^{i \cdot D} \alpha_j, \quad i = 1, \dots, R_0.$$

Вместо шкалы непрерывного времени, мы будем пользоваться дискретным временем, рассматривая траекторию в шкале "событий". Событием будем называть каждое изменение величины очереди (приход/уход).

В системах обслуживания с Пуассоновским входным потоком заявок ($M/\cdot/\cdot$) выполняется специальное свойство, состоящее в том, что доля заявок, которые в момент прихода наблюдают систему в некотором состоянии A , совпадает в пределе с долей времени, когда система находится в состоянии A . Это свойство носит название PASTA (Poisson Arrivals See Time Averages).

В третьем параграфе приведено собственное доказательство свойства PASTA, ориентированное на случай регенерирующих процессов обслуживания, имеющих скачки вверх только в моменты t_n прихода заявок. Для таких процессов известное сложное доказательство можно существенно упростить. В работе свойство PASTA используется при доказательстве равенства регенеративной оценки и оценки, полученной методом УРИМ.

При моделировании методом УРИМ в работе также использовано субэкспоненциальное распределение Парето в качестве распределения времени обслуживания в системе $M/G/1$ (а также распределения интервалов между приходами) требований в $GI/G/1$). Выбор данного распределения обусловлен результатами наблюдения за реальным трафиком в информационных сетях. Необходимые определения приводятся в четвертом параграфе.

В пятом параграфе описано построение состоятельных оценок вероятностей $\gamma^{(1)}$, $\gamma^{(2)}$ для процесса очереди в одноканальных системах обслуживания.

Рассмотрим построение оценки вероятности $\gamma^{(1)}$. Задача оценивания такой вероятности является традиционной для метода расщепления. Как правило, исследования свойств оценки ограничены обоснованием несмещенности в случае марковского процесса.

Для каждого пучка введем случайные величины

$$Y_i := \sum_{j=(i-1) \cdot D+1}^{i \cdot D} I^{(j)}, \quad i = 1, \dots, R_0,$$

где $I^{(j)} = 1$, если очередь достигает уровня L на j -м цикле регенерации. Последовательность $\{I^{(j)}, j \geq 1\}$ является регенеративной с постоянной длиной циклов $\beta_i = D$, $i \in [1, R_0]$ и моментами регенерации τ_{iD} , $i \in [0, R_0]$, $n = R_0 \cdot D$.

Регенеративный метод дает следующее выражение для $\gamma^{(1)}$, если $E\beta_1 < \infty$:

$$\gamma^{(1)} = \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N I^{(i)}}{N} = EY_1/E\beta_1 = P\left(\max_{0 \leq t < \alpha_1} \nu(t) \geq L\right).$$

Регенеративная оценка $\gamma_n^{(1)}$ вероятности $\gamma^{(1)}$ имеет стандартную форму выборочного среднего

$$\gamma_n^{(1)} = \frac{\sum_{j=1}^{R_0} Y_j}{R_0 D} = \frac{\sum_{j=1}^n I^{(j)}}{n}, \quad n = R_0 D.$$

Стандартная оценка $\hat{\gamma}^{(1)}(R_0)$ вероятности $\gamma^{(1)}$ методом расщепления равна

$$\hat{\gamma}^{(1)}(R_0) = \frac{N_{M+1}}{R_0 R_1 \cdots R_M}.$$

Теорема 1 Оценка $\hat{\gamma}^{(1)}(R_0)$ вероятности $\gamma^{(1)} = P\left(\max_{0 \leq t < \alpha_1} \nu(t) \geq L\right)$, полученная методом УРИМ, совпадает с регенеративной оценкой $\gamma_n^{(1)}$, построенной на $n = R_0 D$ циклах регенерации, т. е.

$$\hat{\gamma}^{(1)}(R_0) = \gamma_n^{(1)}.$$

Следствие 1 Оценка $\hat{\gamma}^{(1)}(R_0)$, построенная методом УРИМ, является сильно состоятельной оценкой вероятности $\gamma^{(1)}$.

Данный результат сильно упрощает процедуру оценивания. В частности, можно обойтись без фактического построения циклов регенерации. Использование регенеративной структуры впервые обосновывает состоятельность оценки $\hat{\gamma}^{(1)}$.

На основе методов теории случайных блужданий доказана

Теорема 2 Пусть в системе $M/M/1$ выполняется условие стационарности $\rho = \lambda/\mu < 1$. Тогда аналитическое выражение для $\gamma^{(1)}$ имеет вид:

$$\gamma^{(1)} = P\left(\max_{0 \leq t < \alpha_1} \nu(t) \geq L\right) = \frac{\rho^L - \rho^{L-1}}{\rho^L - 1}.$$

Моделирование системы $M/M/1$ показывает хорошее согласие метода УРИМ с аналитическим результатом. В области малых значений порога $L < 20$ численные значения также соответствуют формуле из теоремы 2.

Построим состоятельную оценку вероятности $\gamma^{(2)} = P(\nu \geq L)$ для непрерывного и дискретного случая (т. е. соответственно, оценку доли времени, когда очередь пребывает выше заданного уровня L , и оценку доли заявок, которые приходят в систему при величине очереди, большей L).

Теорема 3 Оценка $\gamma_n^{(2)}$ вероятности $\gamma^{(2)}$

$$\gamma_n^{(2)} = \frac{\sum_{i=1}^n \int_{\tau_{i-1}}^{\tau_i} I(\nu(t) > L) dt}{\sum_{i=1}^n \alpha_i},$$

построенная по $n = R_0 D$ циклам регенерации, полученным методом УРИМ, является строго состоятельной.

При переходе от непрерывного случая к дискретному, структура регенеративной последовательности не меняется. Мы делаем переход от шкалы непрерывного времени к шкале событий (*приходы/уходы*).

Пусть Z_j - число событий (приходов/уходов) когда очередь превышала порог L на j -м цикле регенерации, и α_j - длина j -го цикла, равная числу событий на j -м цикле, $j = 1, \dots, D \cdot R_0$. Последовательность $\{Z_j\}_{j=1}^{D \cdot R_0}$ (а также $\{\alpha_j\}_{j=1}^{D \cdot R_0}$) является регенеративной и регенерирует в моменты τ_{iD} , $i \in [0, R_0]$ (см. стр. ??). Число циклов регенерации равно числу пучков R_0 . Определим величины

$$Y_i := \sum_{j=(i-1) \cdot D + 1}^{i \cdot D} Z_j; \quad \beta_i := \sum_{j=(i-1) \cdot D + 1}^{i \cdot D} \alpha_j, \quad i = 1, \dots, R_0.$$

По построению, случайные величины $\{Y_i\}$ (а также $\{\beta_i\}$) являются н. о. р. Заметим, что в терминах событий величина Y_i равна числу событий $\{\nu(t) > L\}$ на пучке траекторий и $\sum_{j=1}^{R_0} Y_i = N_{M+1}$. Длина цикла β_i равна количеству всех событий на траекториях пучка с номером i .

Общее число событий на всех циклах регенерации равно сумме событий, произошедших на каждом цикле, т. е.

$$\sum_{j=1}^{R_0} \beta_j = N_{g_0}(R_1 \cdots R_{M+1}) + N_{g_1}(R_2 \cdots R_{M+1}) + \cdots + N_{g_M}, \quad R_{M+1} = 1.$$

Оценка вероятности $\gamma^{(2)} = P(\nu > L)$, построенная методом УРИМ, в дискретном случае имеет вид:

$$\hat{\gamma}^{(2)}(N_{M+1}) = \frac{N_{M+1}}{\sum_{i=0}^M [N_{g_i} \prod_{k=i+1}^{M+1} R_k]}, \quad R_{M+1} = 1.$$

Теорема 4 В системе $M/G/1$ выполняется равенство

$$\gamma_n^{(2)} = \hat{\gamma}^{(2)}(N_{M+1}),$$

где $n = R_0 D$.

Следствие 2 Оценка $\hat{\gamma}^{(2)}(N_{M+1})$ вероятности $\gamma^{(2)}$, построенная методом УРИМ, является строго состоятельной.

В системе $M/G/1$ оценку $\hat{\gamma}^{(2)}(N_{M+1})$ можно использовать как для непрерывного, так и для дискретного времени. В случае системы более общего вида $GI/G/1$ оценку $\hat{\gamma}^{(2)}(N_{M+1})$, построенную методом УРИМ, можно использовать только для дискретного времени.

Результаты моделирования системы $M/M/1$ показали, что метод УРИМ дает оценку, которая очень хорошо согласуется с аналитической формулой $\gamma^{(2)} = \rho^{L+1}$, в том числе в области малых значений L .

Если процесс не является марковским, то траектории после расщепления являются зависимыми. В системе $M/G/1$ процесс очереди не является марковским. В момент прихода заявки остается незавершенное время обслуживания с неизвестным распределением и, поэтому, после расщепления траектории процесса будут зависимыми. Эта зависимость может значительно повлиять на дисперсию оценки.

Для устранения данной проблемы при моделировании системы $M/G/1$, можно применить метод вложенной цепи Маркова. Вместо немарковского процесса $\nu(t)$, $t \geq 0$, моделируется цепь Маркова $\{\nu_n^*\}$, $\nu_n^* = \nu(t_n^*)$, где t_n^* - моменты ухода. Для сохранения вида оценок $\hat{\gamma}^{(1)}(R_0)$, $\hat{\gamma}^{(2)}(N_{M+1})$, требуется внести изменения в алгоритм метода УРИМ.

Условие расщепления: Если в момент ухода заявки траектория, стартовая в области $G_i = [L_i, L_{i+1})$, пересекла некоторый уровень L_{i+k} , $i+k \leq M+1$, то она расщепляется на $\prod_{j=1}^k R_{i+j}$ траекторий ($R_{M+1} = 1$).

Изменение условий расщепления позволяет легко подсчитать компенсирующий множитель в формулах для оценок $\hat{\gamma}^{(1)}(R_0)$ и $\hat{\gamma}^{(2)}(N_{M+1})$, независимо от расщеплений на том или ином уровне, а общее число циклов по-прежнему равно $n = R_0 D$.

Для сравнения результатов эксперимента в системе $M/G/1$ с субэкспоненциальным распределением интервалов обслуживания S используется асимптотическая формула

$$P(\nu > k) \sim \frac{\rho}{ES(1-\rho)} \int_0^k P(S > y) dy, \quad k \rightarrow \infty,$$

где λ - интенсивность приходов, интенсивность загрузки $\rho < 1$. Результаты имитационных экспериментов показывают хорошую близость оценок, полученных методом УРИМ с использованием вложенной цепи Маркова к результатам, полученным по формуле, а также более стабильное поведение оценки по сравнению с обычным методом расщепления. При этом выигрыш в дисперсии оценки составляет 2-3 порядка.

Метод расщепления применяется преимущественно для целочисленного процесса, где скачки имеют значения ± 1 . В таком случае расщепления можно проводить точно на каждом заданном уровне L_i .

Алгоритм метода можно модифицировать для оценки вероятностей, аналогичных $\gamma^{(1)}, \gamma^{(2)}$, но определенных для процесса времени ожидания (нагрузки).

Такая задача рассмотрена в шестом параграфе. В терминах процесса нагрузки $\gamma^{(1)}$ - вероятность того, что нагрузка превысит фиксированный уровень L в течение цикла регенерации, т. е. $\gamma^{(1)} = P(\max_{0 \leq n < \alpha_1} W_n > L)$. Вероятность $\gamma^{(2)} = P(W_\infty > L)$ - вероятность превышения стационарной нагрузкой значения L , где $W_n \Rightarrow W_\infty$.

Процесс нагрузки даже в системе $GI/G/1$ является цепью Маркова и удовлетворяет рекуррентному соотношению Линдли

$$W_{n+1} = (W_n + X_n)^+, \quad n \geq 1, \quad W_1 = 0,$$

где $X_n = S_n - \tau_n$, $\tau_n = t_n - t_{n-1}$ - интервал между приходами n -й и $n - 1$ -й заявок, S_n - время обслуживания n -й заявки. За один скачок процесс может пересечь несколько порогов L_i между двумя последовательными расщеплениями.

Для построения оценок используется метод УРИМ с таким же условием расщепления, как и для случая вложенной цепи Маркова, только расщепление происходит в момент прихода.

В системе $M/M/1$ верны соотношения

$$\begin{aligned} \gamma^{(2)} &= P(W_\infty > L) = \rho e^{-(\lambda - \mu)L}, \\ \gamma^{(1)} &= P(\max_{0 \leq n < \alpha_1} W_n > L) = \frac{(1 - \rho)e^{-(\mu - \lambda)L}}{1 - \rho e^{-(\mu - \lambda)L}}. \end{aligned} \quad (1)$$

Сравнение оценок, полученных при моделировании процесса нагрузки методом УРИМ, и значений, вычисленных по аналитическим формулам (1) показало хорошую точность метода УРИМ.

Для более сложных систем $M/G/1$, $GI/G/1$ было проведено сравнение оценок, полученных методом УРИМ, с оценками, построенными классическим методом МК. Численные эксперименты подтверждают близость двух оценок, однако метод МК дает большую дисперсию. При этом время моделирования методом МК существенно (на несколько порядков) при достаточно большом L превышает время моделирования методом УРИМ.

В четвертой главе исследуется асимптотическая нормальность состоятельных оценок, полученных в главе 3. Приведены формулы для построения доверительного интервала вероятности достижения заданного уровня на цикле регенерации и стационарной вероятности переполнения для процессов очереди и нагрузки.

В первом параграфе строится доверительная оценка вероятностей на циклах регенерации, которые моделируются методом УРИМ. При построении оценки дисперсии необходимо учитывать зависимости, возникающие между

циклами в пучке. Число пучков равно R_0 , в каждом пучке $D = R_1 \cdots R_M$ циклов, общее количество циклов при моделировании равно $n = R_0 \cdot D$.

При доверительном оценивании вероятности

$$P(\max_{0 \leq t < \alpha_1} \nu(t) \geq L) = \gamma^{(1)} = EY_1/D,$$

(точечная состоятельная оценка $\hat{\gamma}^{(1)} = \hat{\gamma}^{(1)}(R_0)$ которой строится методом УРИМ при $n = R_0 \cdot D$) для каждого пучка вводится случайная величина

$$Y_i := \sum_{j=(i-1) \cdot D+1}^{i \cdot D} I^{(j)}, \quad i = 1, \dots, R_0.$$

Пусть $EY_1^2 < \infty$. Для R_0 циклов регенерации (при $\alpha_i = D$ и $\bar{\alpha}_i = D$) несмещенная оценка v_{R_0} дисперсии $Var(Y_1 - D \cdot \gamma^{(1)})$ равна:

$$v_{R_0} = 1/(R_0 - 1) \sum_{i=1}^{R_0} (Y_i - D \cdot \hat{\gamma}^{(1)})^2, \quad R_0 \geq 1.$$

Выражение для $100(1 - \delta)\%$ доверительного интервала имеет вид:

$$\left[\hat{\gamma}^{(1)} \pm \frac{z_\delta \sqrt{\sum_{i=1}^{R_0} (Y_i - D \cdot \hat{\gamma}^{(1)})^2}}{D \sqrt{R_0(R_0 - 1)}} \right],$$

где z_δ удовлетворяет условию $P(N(0, 1) \leq z_\delta) = 1 - \delta/2$, а точечная оценка $\hat{\gamma}^{(1)}$ строится методом УРИМ.

Для доверительного оценивания вероятности $\gamma^{(2)} = P(\nu > L)$ определяются величины

$$Y_i := \sum_{j=(i-1) \cdot D+1}^{i \cdot D} \xi_j; \quad \beta_i := \sum_{j=(i-1) \cdot D+1}^{i \cdot D} \alpha_j, \quad i = 1, \dots, R_0,$$

где ξ_j - число событий когда очередь превышала порог L на j -м цикле. Выражение для $100(\delta - 1)\%$ доверительного интервала имеет вид:

$$\left[\hat{\gamma}^{(2)} \pm \frac{z_\delta \sqrt{\sum_{i=1}^{R_0} (Y_i - \hat{\gamma}^{(2)} \beta_i)^2}}{\sqrt{R_0(R_0 - 1)} \cdot \bar{\beta}_{R_0}} \right],$$

а точечная оценка $\hat{\gamma}^{(2)} = \hat{\gamma}^{(2)}(N_{M+1})$ строится методом расщепления.

Для системы $M/G/1$ доверительный интервал можно записать в более удобной для вычислений форме:

$$\left[\hat{\gamma}^{(2)} \pm \frac{z_\delta \cdot \hat{\gamma}^{(2)} \sqrt{R_0 \cdot \sum_{i=1}^{R_0} (Y_i - \hat{\gamma}^{(2)} \beta_i)^2}}{N_{M+1} \sqrt{(R_0 - 1)}} \right].$$

Во втором параграфе исследована роль зависимости циклов регенерации при построении доверительных интервалов с использованием метода УРИМ для вероятности $\gamma^{(1)} = P(\max_{0 \leq t < \alpha_1} \nu(t) \geq L)$. Циклы в пучках зависимы и дисперсия $\sigma^2 = Var Z_1 = Var(Y_1 - D\gamma^{(1)})$ имеет вид

$$\sigma^2 = D \cdot Var(I^{(1)}) + 2 \sum_{j=2}^D \sum_{i=1}^{D-(j-1)} E(I^{(i)} I^{(i+j-1)}) - (\gamma^{(1)})^2 \cdot D(D-1).$$

Пусть a_k - число редких событий на i -м пучке.

Теорема 5 При $n = R_0 D \rightarrow \infty$ имеет место сходимость с в. 1

$$v_n \rightarrow \frac{\sigma^2}{D},$$

где

$$v_n = \hat{\gamma}^{(1)} - D \cdot (\hat{\gamma}^{(1)})^2 + \frac{2}{n} \sum_{k=1}^{R_0} \frac{a_k(a_k - 1)}{2},$$

и оценка $\hat{\gamma}^{(1)}$ вероятности $\gamma^{(1)}$ строится методом УРИМ.

Для исследования вклада зависимости циклов внутри пучка в длину доверительного интервала вводятся величины: Δ - половина длины доверительного интервала с учетом зависимости и Δ_1 - половина доверительного интервала без учета зависимости (т. е. если считать циклы в пучке независимыми).

Пусть $x = \Delta 10^2 / \gamma^{(1)}$ - точность вычисления доверительно интервала ($x\%$ от $\gamma^{(1)}$), и $y := |\Delta - \Delta_1| 10^2 / \gamma^{(1)}$ - погрешность оценивания (в $\%$ от $\gamma^{(1)}$) в случае, если пренебречь зависимостью циклов регенерации.

Теорема 6 Пусть z_δ удовлетворяет условию $P(N(0, 1) \leq z_\delta) = 1 - \delta/2$, и выполняется $n > 1/\hat{\gamma}^{(1)}$, тогда справедливо соотношение

$$\frac{\Delta_1}{\Delta} < \sqrt{\frac{10^4 z_\delta^2}{x^2 \hat{\gamma}^{(1)} n}}.$$

Следствие 3 При условии $\hat{\gamma}^{(1)} n > \frac{10^6 z_\delta^2}{x^2}$, где $x \in (0, 100]$ выполняется

$$\frac{y}{x} \in (0.9, 1].$$

Численные эксперименты показали, что зависимость существенно влияет на длину доверительного интервала. (точность вычислений x практически совпадает со значением y). При n , определяемом условием достижения x процентов для Δ от значения $\gamma^{(1)}$, значение Δ_1 близко к нулю (т. е. $\frac{\Delta_1}{\Delta} \approx 0$). В случае

пренебрежения зависимостью внутри пучка точность вычисления практически совпадает с ошибкой оценивания. Например, в проведенных экспериментах $\hat{\gamma}^{(1)} n \approx 10^5$. Если $x = 10$, $z_\delta = 1.96$, тогда $\Delta_1/\Delta < 0.061$ и $y/x \approx 0.938$.

В заключении представлен краткий перечень результатов, полученных в ходе исследований в рамках диссертационной работы.

Список опубликованных работ по теме диссертации

Статьи

- [1] Бородина А. В., Морозов Е. В. Доверительное оценивание вероятности переполнения буфера на основе ускоренного регенеративного моделирования системы М/М/1. Труды ИПМИ КарНЦ РАН, 2006, т. 7, с. 125-135, (вклад диссертанта 50%).
- [2] Borodina A., Morozov E. Simulation of rare events with speed-up techniques: Splitting and RESTART. Proceedings of Finnish Data Processing Week at the Petrozavodsk State University (FDPW'2005), 2006, vol. 7, pp. 152-173, (вклад диссертанта 50%).
- [3] A. Borodina. Rare Events Regenerativ Estimation of Queues Based on Splitting. //Proceedings of International workshop. Distributed Computer and Communication Networks: Theory and Application (DCCN'2007), V. 1. Moscow: ИИТ РАС, 2007, pp. 50-55.
- [4] Бородина А. В., Морозов Е. В. Ускоренное регенеративное моделирование вероятности перегрузки односерверной очереди. ОПиПМ, 2007, т. 14, в. 3, с. 385-397, (вклад диссертанта 50%).
- [5] А. В. Бородина. Влияние зависимости циклов, полученных методом расщепления, при доверительном оценивании вероятности перегрузки в системе М/Г/1. Труды ИПМИ КарНЦ РАН, 2007, т. 8, с. 76-83.
- [6] Alexandra V. Borodina, Evsey V. Morozov. Speed-up consistent estimation of a high workload probability in M/G/1 queue. // Transactions of XXVI International Seminar Stability Problems for Stochastic Models, Nahariya, Israel, 2007, V. I, pp. 36-42, (вклад диссертанта 50%).
- [7] Бородина А. В., Морозов Е. В. Ускоренное состоятельное оценивание вероятности большой загрузки в системах М/Г/1, GI/G/1. Сб. Статистические методы оценивания и проверки гипотез, Пермь, 2007, с. 124-140, (вклад диссертанта 50%).
- [8] A. Borodina, E. Morozov. A regenerative modification of the multilevel splitting. // Proceedings of 7th International Workshop on Rare Event Simulation (RESIM 2008), September 24-26, 2008, Renn, с. 276-282, (вклад диссертанта 50%).

Тезисы докладов

- [9] Alexandra V. Borodina. Using of regenerative sequences in Splitting method. Extended abstracts of Russian-Scandinavian symposium, Probability theory and applied probability (PTAP'06, August 26-31, 2006, Petrozavodsk), pp. 18-20.

- [10] А. В. Бородина. Ускоренные методы моделирования в регенерирующих процессах обслуживания. Сборник тезисов Международной школы-конференции "Информационно-телекоммуникационные системы МГИ-ЭТ, 2005, с. 45.