МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории функций и стохастического анализа

МОДЕЛИРОВАНИЕ И АНАЛИЗ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С ПОТЕРЯМИ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 412 группы
направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета
Криушиной Алины Юрьевны

Научный руководитель
старший преподаватель ______________ Н. В. Сергеева

Заведующий кафедрой
д. ф.-м. н. ______________ С. П. Сидоров

Саратов 2021

Актуальность темы. Актуальность темы, такой как моделирование и анализ систем массового обслуживания(СМО), обуславливается тем, что моделирование является прекрасным инструментом для рассматрения реальных СМО в информационно-вычислительных системах, с целью улучшения характеристик обслуживания или в силу удобства их применения.

Целью бакалаврской работы является моделирование и анализ СМО с потерями с пуассоновским входящим потоком заявок и экспоненциальным законом распределения для повышения их эффективности и качества обслуживания.

Объектом исследования являются СМО с потерями с пуассоновским входящим потоком заявок и экспоненциальным законом распределения длительности их обслуживания.

Предмет исследования – поступление заявок в систему массового обслуживания с потерями и обслуживание данных заявок.

В рамках указанной цели были поставлены следующие задачи:

1. Изучить основные понятия, связанные с СМО;
2. Выделить основные элементы СМО;
3. Привести классификации СМО;
4. Более детально изучить СМО с пуассоновским входящим потоком требований и эскпоненциальной длительностью их обслуживания с потерями;
5. Рассмотреть основные теоретические характеристики СМО с потерями;
6. Построить имитационную модель данной системы и получить характеристики в ходе моделирования;
7. На основе вычисленных характеристик сделать выводы о функционировании построенной СМО с потерями.

Практическая значимость проводимого численного эксперимента состоит в том, что построенная имитационная модель максимально приближает нас к оптимизированию и регулированию случайного потока поступления заявок и времени обслуживания на реально существующих обслуживающих системах.

Структура и содержание бакалаврской работы. В первом разделе рассмотрены основные элементы моделей массового обслуживания, а также приведена классификация систем массового обслуживания и входных потоков. Во втором разделен описаны основные распределения в системах массового обслуживания, такие как: эскпоненциальное и распределение Пуассона. В третьем разделе приведена необходимая для изучения СМО теоритическая информация о параметрах, характеристиках и структуре СМО. В четвертом разделе описаны общие принципы построения имитационной модели системы массового обслуживания и принципы ее функционирования. В шестом разделе описывается необходимая нам система обслуживания типа $(M/M/c) : (GD/N/\infty)$. В пятом, заключительном разделе построена имитационная модель СМО с потерями, на основании которой были рассчитаны основные характеристики данной системы. А также был проведен сравнительный анализ практических и теоретических характеристик. Список использованных источников содержит 20 наименований.

Основное содержание работы

Во введении обосновывается актуальность темы работы, формулируется цель работы и решаемые задачи, отмечается практическая значимость полученных результатов.

В первом разделе рассмотрены основные элементы теории массового обслуживания, такие как клиент (заявка или требование на обслуживание, либо просто "объект обслуживания") и сервис (обслуживающее устройство, средства обслуживания и т.п.).

Каждая СМО включает в свою структуру некоторое число обслуживающих устройств, называемых каналами обслуживания (к их числу можно отнести лиц, выполняющих те или иные операции, - кассиров, операторов, менеджеров и т.п.), обслуживающих некоторый поток заявок (требований), поступающих на ее вход в случайные моменты времени. Обслуживание заявок происходит за неизвестное, обычно случайное время и зависит от множества самых разнообразных факторов. После обслуживания заявки канал освобождается и готов к приему следующей заявки.

Случайный характер потока заявок и времени их обслуживания приводит к неравномерности загрузки СМО - перегрузке с образованием очередей заявок или недогрузке - с простаиванием каналов.

Следует отметить, что время обслуживания заявки зависит от характера самой заявки или требований клиента и от состояния и возможностей обсуживающей системы. В ряде случаев приходится также учитывать вероятность выхода обслуживающего прибора по истечении некоторого ограниченного интервала времени.

Источник, генерирующий клиентов, подлежащих обслуживанию, может иметь конечную или бесконечную мощность.

При изучении систем массового обслуживания выделяются наиболее важные критерии функционирования обслуживающей системы: среднее время ожидания в очереди и средняя длина очереди. Эта информация используется затем для выбора надлежащего уровня обслуживания.

Результаты исследования системы обслуживания также можно использовать для оптимизации модели со стоимостными характеристиками, в которой минимизируется сумма затрат, связанных с предоставлением услуг, и потерь, обусловленных задержками в их предоставлении.

Также в данном разделе приведена общая классификация СМО. Системы массового обслуживания разделяются на три подсистемы.

Первая подсистема - это система массового обслуживания без потерь. Под термином система без потерь (с полным ожиданием) понимают систему, в которой, если все приборы заняты, требование становится в очередь и не покидает ее до тех пор, пока не будет обслужено.

Вторая подсистема - это система с частичными потерями. Подобная подсистема характеризуется тем, что требование либо не становится в очередь, если эта очередь превышает по длине некоторую величину (система с ограниченной длиной очереди), либо становится в очередь, но покидает ее, если время пребывания в ней превышает определенную величину (система с ограниченным временем пребывания), или, если время ожидания в очереди начала обслуживания превышает определенную величину (система с ограниченным временем ожидания начала обслуживания).

Третья подсистема - это система без очередей. Под этим термином понимают систему, в которой требование покидает систему, если все обслуживающие устройства (приборы) заняты. В такой системе, очевидно, очереди не может быть. Системы, имеющие очередь, подразделяются на системы с одной очередью и системы с несколькими очередями.

Все системы массового обслуживания делятся на системы с одним каналом и системы с конечным числом каналов обслуживания. В тех случаях, когда приборов много удобно (математически более просто) считать, что их бесконечное число.

Входные потоки по характеру требований разделяется на детерминированный поток требований и стохастический.

Детерминированный входной поток может быть двух видов. В первом случае требования поступают через равные промежутки времени. Другим видом детерминированного потока является поток, в котором требования поступают по известной программе - расписанию, когда моменты поступления новых требований известны заранее.

Если промежутки времени между поступлениями требований случайны, то это будет стохастический процесс. Стохастический поток требований подразделяется на три вида: поток с произвольными стохастическими свойствами, рекуррентный поток и совершенно случайный или пуассоновский поток требований.

Во втором разделе приведена необходимая для изучения СМО теоритическая информация про экспоненциальное распределение и распределение Пуассона.

Экспоненциальное распределение в системах массового обслуживания.

В теории массового обслуживания большое внимание уделено задачам, в которых заявки поступают в систему согласно экспоненциальному (показательному) закону с интенсивностью входного потока заявок λ, не зависящей от времени t. Промежуток между поступлениями заявок ξ есть непрерывная случайная велечина, имеющая показательное распределение с параметром $\lambda > 0$.

Случайная величина ξ принимает только неотрицательные значения, а её функция распределения $F_\xi(x)$ имеет вид:

$$F_\xi(x) = \begin{cases} 0, x \leqslant 0, \\ 1 - e^{-\lambda x}, x > 0. \end{cases} \tag{1}$$

Соответствующая функция плотности распределения $f_\xi(x)$ имеет вид:

$$f_\xi(x) = \begin{cases} \lambda e^{\lambda x}, x \geqslant 0, \\ 0.x < 0. \end{cases} \tag{2}$$

Математическое ожидание случайной величины ξ, подчиненной показательному распределению:

$$M_\xi = \frac{1}{\lambda}, \tag{3}$$

где λ - интенсивность (количество событий в единицу времени) с которой появляются события.

Дисперсия случайной величины ξ:

$$D_\xi = \frac{1}{\lambda^2}. \tag{4}$$

Распределение Пуассона.

Вероятность ненаступления заявки в течение периода $t(n = 0)$:

$$p_0(t) = e^{-\lambda t}.$$

Вероятность наступления хотя бы одной заявки в течение периода t:

$$p_{n>0}(t) = 1 - e^{-\lambda t}.$$

Дифференциальное уравнение вида:

$$f(x) = y^{'} + a(x)y,$$

$$\mu(x) = e^{\int a(x)dx},$$

имеет общее решение вида:

$$y = \frac{1}{\mu(x)}[\int f(x)\mu(x)dx + C].$$

Формула для расчета вероятности наступления числа событий $n(n = 0, 1, 2 \ldots)$ для заданного периода времени t:

$$p_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!},\ n = 1, 2, 3 \ldots . \tag{5}$$

В данном случае мы получили дискретную плотность вероятности распределения Пуассона с математическим ожиданием:

$$M\{n|t\} = \lambda t \tag{6}$$

поступлений за время t.

Дисперсия распределения Пуассона также равна λt:

$$D\{n|t\} = \lambda t. \tag{7}$$

В третьем разделе введены основные значения и понятия общей системы массового обслуживания:

— n - число клиентов в системе обслуживания (в очереди и на обслуживании);
— λ_n - интенсивность поступления в систему клиентов при условии, что в системе уже находится n клиентов;
— μ_n - интенсивность выходного потока обслуженных клиентов при условии, что в системе находится n клиентов;
— p_n - вероятность того, что в системе находится n клиентов.

А также описаны следующие обозначения, для краткой характеристики СМО в теории массового обслуживания:$(a/b/c) : (d/e/f)$.

Там же приведен конкретный пример СМО, такой как: $(M/M/c) : (GD/N/\infty)$.

В соответствии с принятыми обозначениями здесь речь идет о системе (и, соответственно, модели) массового обслуживания с пуассоновским входным потоком (или экспоненциальным распределением интервалов времени между моментами последовательных поступлений клиентов) и временем обслуживания, и c параллельно функционирующими каналами. При этом дисциплина очереди не регламентирована, и максимальное количество допускаемых в систему клиентов равно N. Наконец, источник, порождающий клиентов, имеет неограниченную емкость.

В этой модели емкость системы ограничена сверху значением N, тогда максимальная длина очереди равна $N - c$. Интенсивности поступления и обслуживания клиентов равны λ и μ соответственно. Эффективная интенсивность поступления заявок в систему обслуживания $\lambda_{эфф}$, соответсвенно меньше λ в силу ограниченности емкости системы.

Параметры λ_n, μ_n общей модели обслуживающей системы в данной модели определяются следующим образом:

$$\lambda_n = \begin{cases} \lambda,\ 0 \le n \le N, \\ 0,\ n > N. \end{cases} \tag{8}$$

$$\mu_n = \begin{cases} n\mu,\ 0 \le n \le c, \\ c\mu,\ c < n \le N. \end{cases} \tag{9}$$

Вероятности p_n - вероятности, что в системе находятся n клиентов, можно найти, подставляя λ_n, μ_n в общее выражение для p_n(**??**), используя интенсивность потока поступающих в систему заявок:

$$p_n = \begin{cases} \frac{\rho^n}{n!} p_0,\ 0 \le n \le c, \\ \frac{\rho^n}{c! c^{n-c}} p_0,\ c < n \le N, \end{cases} \tag{10}$$

при

$$p_0 = \begin{cases} \left[\sum\limits_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c \left(1 - \left(\frac{\rho}{c}\right)^{N-c+1}\right]}{c! \left(1 - \frac{\rho}{c}\right)} \right]^{-1},\ \frac{\rho}{c} \ne 1, \\ \left[\sum\limits_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c}{c!} (N - c + 1) \right]^{-1},\ \frac{\rho}{c} = 1, \end{cases} \tag{11}$$

где

$$\rho = \frac{\lambda}{\mu}. \tag{12}$$

В четвертом разделе описываются общие принципы построения имитационной модели.

Моделированием в общем случае называется замещение одного объекта, называемого системой, другим объектом, называемым моделью, и проведение экспериментов с моделью (или на модели), исследование свойств модели, опираясь на результаты экспериментов с целью получения информации о системе.

При исследовании систем со стохастическим характером функционирования (СМО являются системами такого типа) результаты, полученные при единичном "прогоне" имитационной модели (при единичном воспроизведении на имитационной модели процесса функционирования системы), носят случайный характер.

Следовательно, для того, чтобы найти одну оценку (одно значение) какой-либо характеристики функционирования системы необходимо многократно "прогонять" имитационную модель (необходимо получить множество результатов) с последующей статистической обработкой полученных данных.

В данном разделе также приведена классификация имитационных моделей:

1. Непрерывная модель.

   Данная модель предполагает развитие имитационной ситуации в модели, как в непрерывной (аналоговой) среде, т.е. поведение системы изменяется непрерывно во времени. Типичным примером является - изучение динамики народонаселения мира;

2. Дискретная имитационная модель.

   Она представляет собой систему, в которой процессы моделируются как дискретные, т.е. проявляют свои функции и свойства в определенные моменты времени. Можно предполагать, что это происходит через равные интервалы времени, но в некоторых моделях могут существовать и асинхронные объекты, которые проявляют себя в случайные моменты времени.

Далее описывается, что представляет собой работа СМО – дискретно - непрерывный процесс. Работу всякой СМО можно представить как процесс ее перехода из одного состояния в другое, в общем случае в случайные моменты времени $t_0, t_1, \dots t_k$, когда в СМО происходят события, предусмотренные дисциплиной обслуживания. Каждое событие приводит к изменениям каких-либо переменных состояния системы.

Состояние случайного процесса скачком изменяется в некоторые моменты времени $t_0, t_1, \dots t_k$ наступления событий и остаются неизменными на интервалах времени между этими моментами.

Длительности интервалов между моментами очередных событий, связанных с прибытием заявки, определяются временем между поступлениями клиентов, а события, связанные с уходом - временем обслуживания.

Время наступления этих событий может быть либо детерменированным, либо случайным. В первом случае процедура определения времени наступления событий проста. В данной работе будет рассматриваться второй случай - когда будет использоваться специальная процедура для получения выборочных значений времени между событиями в системе, соответствующей заданному распределению.

В последнем **пятом разделе** строится имитационная модель системы массового обслуживания типа $(M/M/c):(GD/N/\infty)$ с потерями.

В данную систему заявки на обслуживание поступают в соответствии с распределением Пуассона в среднем 20 заявок в час. Длительность обслуживания заявки является экспоненциально распределенной случайной величиной с математическим ожиданием 11,5 мин. Количество параллельно работающих каналов, которые могут обслужить заявку, равно 4-м. Список заявок, которые могут пребывать в очереди, ожидая обслуживание, ограничивается 16-ю заявками.

Основная суть данной модели заключается в том, что, если в момент поступления очередного требования в системе уже находится одно и более требование и при проверке наличия свободного сервиса обнаружится, что все сервисы заняты, то заявка поступит в очередь и будет обслужена, в случае, если очередь еще не достигла своего максимума, иначе заявка получит отказ.

В соответсвии с обозначениями, принятыми в разделе (3), основные параметры системы можно описать следующим образом:

— $\lambda = 20$ - интенсивность поступления клиентов в систему человек в час;
— $\mu = \dfrac{60}{11.5} = 5.217391$ - интенсивность выходного потока обслуженных клиентов человек в час;
— $c = 4$ - количество параллельно работающих сервисов;
— $N = c + 16 = 20$ - емкость системы.

Если в момент поступления очередного требования в системе уже находится одно и более требование, то проверяется наличие свободного сервиса, если все сервисы заняты - заявка поступает в очередь, пока очередь не достигнет своего максимума - иначе заявка получит отказ.

Поведение системы изменяется в заданные моменты времени – моменты поступления требований в систему и в моменты, когда обслуженное требование покидает систему - эти события и являются главными.

Входными данными программы являются следующие:

— λ - интенсивность входящего потока требований (число требований в час);
— μ - интенсивность обслуживания требований;
— n - общее число требований.

При этом время между поступлениями заявок и время осблуживания заявок генерируется показательным законом распределения.

Выходными данными программы являются следующие:

— L_q - среднее число требований в очереди;
— W_q – среднее время ожидания в очереди;
— L_s - среднее число требований в очереди;
— W_s - среднее время пребывания требования в системе.

Данные характеристики при имитационном моделировании расчитываются по формулам:

$$L_q = \frac{\text{суммарное время ожидания заявки в очереди}}{\text{время работы системы}},$$
$$W_q = \frac{\text{суммарное время ожидания заявки в очереди}}{\text{количество смоделированных заявок}},$$
$$L_s = \frac{\text{суммарное время ожидания заявки в системе}}{\text{время работы системы}},$$
$$W_s = \frac{\text{суммарное время ожидания заявки в системе}}{\text{количество смоделированных заявок}}.$$

Поскольку конечной целью является построение имитационной модели для анализа полученной модели, то характеристики, полученные в ходе построения имитационной модели системы массового обслуживания типа $(M/M/c) : (GD/N/\infty)$ с потерями сравниваются с теоретическими характеристиками:

Command Window

Функциональная характеристика	Аналитич. знач.	Иммитац. мод. (среднее значение)	Нижняя граница дов. инт.	Верхняя граница дов. инт.
Lq	5.85	3.56	2.52	4.59
Ls	9.55	7.00	5.90	8.09
Wq	0.30	0.18	0.13	0.23
Ws	0.49	0.35	0.30	0.40
cc	3.70	3.44	3.33	3.55

Рисунок 1. Характеристики, построенные для n = 100

Command Window

Функциональная характеристика	Аналитич. знач.	Иммитац. мод. (среднее значение)	Нижняя граница дов. инт.	Верхняя граница дов. инт.
Lq	5.85	5.83	5.63	6.02
Ls	9.55	9.53	9.31	9.74
Wq	0.30	0.29	0.28	0.30
Ws	0.49	0.48	0.47	0.49
cc	3.70	3.70	3.68	3.72

Рисунок 2. Характеристики, построенные для n = 10000

Можно заметить, что с увеличением числа требований n доверительный интервал сужается и значение, полученное в ходе имитационного моделирования, приближается к аналитическому. Таким образом можно сделать вывод о том, что построенную имитационную модель системы $(M/M/4)$: $(GD/20/\infty)$ с потерями можно использовать для моделирования реальных систем, меняя интенсивности входных потоков и интенсивности обслуживания требований.

В заключении описаны результаты проделанной работы о том, что:

— была построена математическая модель изученной системы массового обслуживания;

— была азработана программа, моделирующая работу такой системы, и позволяющая вычислять основные характеристики системы на основании дискретной модели и по теоретическим формулам;

— был проведен сравнительный анализ теоретических характеристик и характеристик в ходе построения имитационной модели СМО с потерями для изучения эффективности функционирования системы. В ходе данного анализа было выявлено, что имитационную модель можно эффективно использовать для моделирования реальной системы.