

ИССЛЕДОВАНИЕ СЛОЖНЫХ СИСТЕМ НА ОСНОВЕ КОМБИНИРОВАННОГО ПОДХОДА

Т. И. Алиев (Санкт-Петербург)

При исследовании сложных технических систем с использованием методов математического моделирования возникает ряд проблем, обусловленных большой размерностью задачи и необходимостью учета множества факторов, присущих системе и влияющих на качество решения задачи. Примерами таких систем могут служить информационно-вычислительные системы, построенные на основе локальных и территориально-распределенных вычислительных сетей [1, 2]. Случайный характер процессов формирования, обработки и передачи данных обуславливает необходимость применения стохастических моделей, в качестве которых широко используются модели массового обслуживания, представляющие собой системы и сети массового обслуживания различных классов [1–3]. Для расчета характеристик функционирования систем с использованием указанных моделей могут применяться различные методы моделирования: аналитические, численные и имитационные.

Современные телекоммуникационные системы и компьютерные сети характеризуются наличием приоритетного обслуживания передаваемых и обрабатываемых данных. Аналитические методы исследования приоритетных дисциплин обслуживания заявок разработаны, в основном, для дисциплин с одним классом приоритетов, а большинство результатов получены при различных допущениях и предположениях, ограничивающих их применение на практике. Детальное исследование приоритетных систем обслуживания предполагает применение комбинированного подхода к моделированию, позволяющего, как показывает практика, получить результаты, значительно изменяющие наше представление о влиянии приоритетов на качество функционирования приоритетных систем обслуживания.

Принципы моделирования сложных систем

При исследовании сложных технических систем с дискретным характером функционирования наиболее широкое применение получили аналитические и имитационные методы моделирования [4–6].

Одним из основных требований, предъявляемых к модели, является ее адекватность реальной системе, которая достигается за счет использования моделей с различным уровнем детализации, зависящим от особенностей структурно-функциональной организации системы и целей исследования. Процессы функционирования реальных систем невозможно описать полно и детально, что обусловлено существенной сложностью таких систем. Основная проблема при разработке модели состоит в нахождении компромисса между простотой ее описания и необходимостью учета многочисленных особенностей, присущих реальным системам. Попытка построить единую универсальную модель обречена на неудачу, ввиду ее необозримости и невозможности расчета.

Математическое моделирование сложных технических систем должно базироваться на ряде принципов, обеспечивающих корректность и достоверность результатов моделирования и, в конечном счете, качественное проектирование систем. Среди этих принципов можно выделить три основных принципа:

- 1) системный подход при решении задач анализа и синтеза;
- 2) принцип иерархического многоуровневого моделирования;
- 3) принцип множественности моделей.

В основе исследования сложных технических систем с использованием математического моделирования лежит системный подход, конечной целью которого является

системотехническое проектирование, направленное на построение системы с заданным качеством. Для решения задач проектирования необходимо располагать знаниями о том, как влияют различные способы структурно-функциональной организации на характеристики функционирования системы, т. е. решать задачи системного анализа.

Принцип иерархического многоуровневого моделирования базируется на иерархическом описании исследуемой системы и процессов, протекающих в ней. При этом система и протекающие в ней процессы представляются семейством моделей, каждая из которых описывает поведение системы с точки зрения различных уровней абстрагирования, отличающихся рядом характерных особенностей и параметров, с помощью которых и описывается поведение системы.

Применительно к моделям сложных технических систем с дискретным характером функционирования предлагается выделить два направления иерархии:

- 1) иерархия по вертикали, в которой деление моделей по уровням осуществляется в зависимости от структурно-функциональных особенностей системы;
- 2) иерархия по горизонтали, в которой деление моделей по уровням осуществляется в зависимости от методов их исследования.

В иерархии по вертикали, в общем случае, можно выделить три уровня моделей:

- уровень базовых моделей, содержащий простейшие модели, на основе которых строятся и могут быть рассчитаны другие более сложные модели второго и третьего уровней;
- уровень локальных моделей, отображающих отдельные особенности структурно-функциональной организации систем и позволяющих решать частные задачи анализа и синтеза;
- уровень глобальных моделей, наиболее полно отображающих структурные и функциональные особенности организации исследуемых систем и представляющих собой модели с высокой степенью детализации.

Глобальные модели строятся на основе базовых и локальных моделей.

Иерархия по горизонтали включает четыре уровня моделей в зависимости от методов их исследования:

- модели, поддающиеся точному расчету, позволяющему получить результаты либо аналитически в явном виде, либо численно с использованием численных методов анализа;
- модели, поддающиеся приближенному аналитическому расчету с приемлемой для инженерных применений точностью, причем результаты могут быть получены либо в явном виде, либо в виде границ (верхней и нижней);
- модели, требующие применения статистических методов расчета, основанных на имитационном моделировании;
- модели, использующие аналитико-имитационные методы расчета.

Базовые модели допускают применение точных и приближенных аналитических методов и позволяют получить результат в явном виде. Локальные модели обычно предполагают применение имитационных методов, а глобальные – наряду с перечисленными методами моделирования могут использовать аналитико-имитационные методы. На практике при исследовании сложных систем наиболее эффективным является комбинированный подход к моделированию, основанный на применении на различных этапах исследования разных моделей и методов моделирования. Так, например, на этапах анализа свойств системы и синтеза в соответствии с заданным критерием эффективности оптимальной системы целесообразно использовать модели, поддающиеся точному или приближенному аналитическому расчету. Имитационное моделирование обычно используется для аттестации приближенных методов и детального анализа

свойств и потенциальных возможностей спроектированной системы на моделях большой сложности, а также с целью разработки на основе полученных результатов приближенных и эвристических методов расчета.

Основное достоинство имитационного моделирования заключается в универсальности, т. е. в возможности исследования систем практически любой сложности с любой степенью детализации. Применительно к моделированию приоритетных систем эта универсальность проявляется в возможности исследования свойств систем при любых законах распределения случайных величин, описывающих, в частности, интервалы времени между поступающими в систему заявками и длительности обслуживания заявок. Кроме того, имитационное моделирование предоставляет возможность анализа различных экзотических дисциплин буферизации и обслуживания заявок, не поддающихся аналитическому описанию. Однако на практике оказывается, что и имитационное моделирование имеет определенные ограничения, обусловленные как возможностями средств вычислительной техники, с помощью которых реализуется имитационная модель, так и присущими имитационному моделированию недостатками. К таким недостаткам при моделировании приоритетных систем, в первую очередь, относятся следующие.

1. Длительность моделирования должна быть такой, чтобы результаты моделирования были достоверными для каждого класса заявок, включая класс заявок, создающий наименьшую нагрузку. Это означает, что длительность моделирования должна определяться исходя из условия: для класса заявок, создающего самую маленькую нагрузку, количество заявок, прошедших через имитационную модель, должно быть достаточно большим для получения статистически устойчивых результатов, что в некоторых случаях может привести к неоправданно длительному процессу моделирования.

2. При большой нагрузке получить достоверные результаты для низкоприоритетных заявок становится просто невозможно, поскольку в области больших нагрузок начинает сказываться свойство защиты от перегрузок [1], приводящее к резкому увеличению времени ожидания низкоприоритетных заявок.

3. Имитационное моделирование приоритетных систем массового обслуживания не позволяет, в общем случае, решать задачи оптимального синтеза (проектирования), в то время как применение аналитического моделирования даже с использованием приближенных методов позволяет получить, пусть и не точное, но решение в явном виде, которое в дальнейшем может уточняться на основе имитационного моделирования.

Взаимодействие моделей разных уровней иерархии осуществляется путем пересчета характеристик, полученных на одном уровне, в параметры модели, используемой на другом (соседнем) уровне. На каждом уровне может использоваться множество различных моделей. Состав моделей каждого уровня зависит от структурно-функциональной организации системы и целей исследования. Последнее также определяет степень детализации моделей одного и того же уровня.

Реализация принципа иерархического многоуровневого моделирования базируется на структурно-функциональной декомпозиции исследуемой системы, направленной на выделение и исследование наиболее существенных аспектов структурно-функциональной организации. Структурно-функциональная декомпозиция систем позволяет на разных этапах исследования использовать модели разных уровней: на этапе функционального проектирования – базовые модели, на этапе структурного проектирования – локальные модели и на завершающем этапе структурно-функционального проектирования – глобальные модели. Такой подход позволяет существенно упростить решение задачи системотехнического проектирования реальных систем, характеризу-

ющейся значительной сложностью ввиду ее большой размерности и громоздкости результатов.

Одним из основополагающих принципов моделирования сложных систем является принцип множественности моделей, заключающийся, с одной стороны, в возможности отображения многих различных систем и процессов с помощью одной и той же модели и, с другой стороны, в возможности представления одной и той же системы множеством различных моделей в зависимости от целей исследования. Использование этого принципа позволяет отказаться от подхода, когда для каждой исследуемой системы разрабатывается своя модель, и предложить новый подход, при котором разрабатываются абстрактные математические модели разного уровня (в основном базовые и локальные), используемые для исследования систем различных классов. При этом задача моделирования сводится к грамотной параметризации моделей и интерпретации полученных результатов.

Ниже применение комбинированного подхода к моделированию иллюстрируется на примере приоритетной системы массового обслуживания.

Модели приоритетных систем

В качестве базовых моделей дискретных систем с приоритетами обычно используются одноканальные или многоканальные системы массового обслуживания с однородным или неоднородным потоком заявок.

Основой для аналитического исследования сложных систем на базовых моделях с одним обслуживающим прибором могут служить аналитические зависимости для расчета характеристик обслуживания заявок, полученные в [4] в предположении об обслуживании неоднородного потока заявок в соответствии с дисциплиной со смешанными приоритетами. В частности, среднее время ожидания заявок класса k определяется по формуле:

$$w_k^{СП} = \frac{\sum_{i=1}^H (2 - q_{ki})(1 + q_{ki})\lambda_i b_i^2 (1 + v_i^2)}{(2 - \sum_{i=1}^H q_{ik}(3 - q_{ik})\rho_i)(2 - \sum_{i=1}^H (1 - q_{ki})(2 - q_{ki})\rho_i)} + \frac{b_k \sum_{i=1}^H q_{ik}(q_{ik} - 1)\rho_i}{2 - \sum_{i=1}^H q_{ik}(q_{ik} - 1)\rho_i},$$

где $\rho_i = \lambda_i b_i$ – нагрузка, создаваемая заявками класса i ; q_{ik} – элементы матрицы приоритетов ($i, k = 1, \dots, H$), принимающие следующие значения: 0, если заявки класса i не имеют приоритета по отношению к заявкам класса k ; 1, если приоритет относительный; 2, если приоритет абсолютный, причем $q_{ii} = 0$.

При этом предполагается, что в систему поступает H простейших потоков заявок с интенсивностями $\lambda_1, \dots, \lambda_H$, длительности обслуживания заявок распределены по произвольному закону со средними значениями b_1, \dots, b_H и коэффициентами вариации v_1, \dots, v_H , и емкости накопителей – неограниченные.

На основе среднего времени ожидания w_k ($k = 1, \dots, H$) с использованием фундаментальных соотношений теории массового обслуживания, справедливых для широкого класса моделей, могут быть рассчитаны другие характеристики, в частности, среднее время пребывания заявки (ожидания и обработки): $u_k = w_k + b_k$; среднее число запросов, находящихся в состоянии ожидания l_k и в системе m_k (в состоянии ожидания и на обработке): $l_k = \lambda_k w_k$; $m_k = \lambda_k u_k$.

Полученные для базовой модели математические зависимости характеристик обслуживания заявок от параметров системы при использовании приоритетных дисциплин

плин обслуживания (ДО) заявок позволяют выполнить достаточно полный анализ свойств исследуемой системы. Так, например, анализ влияния суммарной загрузки системы на характеристики обслуживания заявок показывает, что среднее время пребывания в системе заявок всех классов растет с увеличением суммарной загрузки R , причем более резко в области больших значений загрузки, особенно для заявок низкоприоритетных классов. В области перегрузок, когда $R \geq 1$, проявляется свойство защиты от перегрузок высокоприоритетных заявок за счет отказа в обслуживании низкоприоритетным заявкам. При этом время пребывания низкоприоритетных заявок возрастает неограниченно и стремится к бесконечности, в то время как для высокоприоритетных заявок время пребывания имеет конечное значение.

Анализ влияния числовых моментов более высокого порядка интервалов времени между поступающими в систему заявками и длительности обслуживания на характеристики обслуживания заявок приводит к одному важному результату, а именно: для определения первых l моментов характеристик обслуживания заявок, необходимо задать $(l+1)$ моментов длительности обслуживания ($l = 1, 2, \dots$), т. е. на один момент больше. В частности, для определения средних характеристик достаточно задать кроме среднего значения лишь второй момент длительности обслуживания. В то же время, на средние значения характеристик обслуживания заявок существенное влияние кроме двух первых моментов интервалов между заявками в потоке оказывают моменты более высоких порядков, в частности – третий момент распределения.

В случае произвольных потоков заявок, отличных от простейших, для которых коэффициент вариации интервалов между заявками $v_a \neq 1$, значительно усложняется аналитический расчет характеристик обслуживания заявок, что предопределяет применение имитационного моделирования. С использованием разработанных имитационных моделей в среде GPSS была проведена серия экспериментов с целью выявления зависимости среднего времени ожидания заявок разных классов от характера поступающих в систему потоков заявок, задаваемых средними значениями a_i и коэффициентами вариации v_{a_i} интервалов времени между поступающими заявками.

Результаты многочисленных экспериментов, выполненных для различных дисциплин обслуживания заявок в широком диапазоне значений параметров исследуемой системы, позволили сформулировать следующие важные для практики выводы, существенно изменяющие общепринятые представления об эффектах, достигаемых за счет применения приоритетных ДО.

Известные в теории [1] утверждения о равенстве средних значений времени ожидания заявок различных классов при бесприоритетной ДО справедливо лишь для систем с простейшими потоками заявок. При $v_a \neq 1$ средние времена ожидания заявок разных классов для бесприоритетной ДО могут различаться, если создаваемые ими нагрузки не одинаковы. Введение приоритетных ДО в случае $v_a \neq 1$ не всегда обеспечивает меньшее время ожидания для заявок высокоприоритетного класса.

Указанные эффекты особенно сильно проявляются в области малых значений загрузки системы R и при больших различиях между нагрузками ρ_i , создаваемыми заявками разных классов.

Выполненные исследования на основе комбинированного подхода с использованием точных аналитических методов и имитационного моделирования [6] свидетельствуют о существенном влиянии третьего момента $a^{(3)}$ интервалов времени между заявками в потоке, задаваемого в виде коэффициента асимметрии $\gamma = a^{(3)} / [a^{(2)}]^{3/2}$, где $a^{(2)}$, $a^{(3)}$ – второй и третий центральные моменты, на средние значения характеристик

обслуживания заявок. С увеличением γ среднее время ожидания заявок уменьшается, причем эта зависимость особенно сильно проявляется при малых загрузках R системы и уменьшается с ее увеличением. Так, при значениях $R = 0,1$ времена ожидания при разных значениях γ отличаются в несколько раз, а при $R = 0,99$ эта разница составляет несколько процентов, что обуславливает возможность получения качественных (робастных) аналитических зависимостей для расчета верхних границ характеристик обслуживания заявок при большой загрузке системы [1]. В области значений R от 0,3 до 0,7, наиболее характерной для реальных систем, эта разница значительна и может составлять десятки и сотни процентов, причем она растет с увеличением второго момента $a^{(2)}$ интервалов времени между поступающими в систему заявками.

Локальные модели дискретной системы представляют собой более сложные модели и используются для оценки эффективности отдельных подсистем и, возможно, всей системы в целом. В качестве локальных моделей выступают разомкнутые и замкнутые сети массового обслуживания, представляющие собой совокупность взаимосвязанных систем массового обслуживания разных классов [1, 3].

Заключение

Базовые и локальные модели служат основой для построения более сложных глобальных моделей реальных систем. С использованием аналитических методов расчета этих моделей могут решаться задачи проектирования, связанные с формированием требований к структурным и функциональным параметрам, обеспечивающим заданное качество функционирования проектируемой системы, однако получаемые при этом результаты могут иметь значительную погрешность. Для повышения достоверности результатов следует применять комбинированный подход, основанный на сочетании аналитических и имитационных методов моделирования. При этом аналитические методы должны применяться на этапах анализа свойств и синтеза оптимальной системы. Имитационные методы, основанные на использовании специализированных языков моделирования, таких как GPSS [5], позволяют выполнять исследование систем практически любой сложности с любой степенью детализации и должны применяться на заключительном этапе детального анализа спроектированной системы.

Литература

1. Клейнрок Л. Вычислительные системы с очередями. – М.: Мир, 1979. – 600 с.
2. Шварц М. Сети ЭВМ. Анализ и проектирование/Под ред. В.А. Жожикашвили. – М.: Радио и связь, 1981. – 336 с.: ил.
3. Алиев Т.И., Новиков Г.И. Метрическая теория и мониторинг компьютерных систем: состояние и проблемы//Изв. вузов.- Приборостроение, 2000.- № 3. – С. 40–44.
4. Алиев Т.И. Характеристики дисциплин обслуживания заявок с несколькими классами приоритетов//Известия АН СССР – Техническая кибернетика, 1987.- № 6.- С.188-191.
5. Учебное пособие по GPSS World/Перевод с английского/Казань: Изд-во «Мастер Лайн», 2002. – 272 с.
6. Алиев Т.И., Гореславко И.Э., Муравьева Л.А. Характеристики обслуживания заявок в системах с неоднородными потоками произвольного вида//Распределенные автоматизированные системы массового обслуживания. Тезисы докладов 4-й Всесоюзной школы-семинара. – Кутаиси, 1987. – С.135–136.