

УДК 681.324

А.Ю. Новосад, студент

Севастопольский национальный технический университет

ул. Университетская 33, г. Севастополь, Украина, 99053

E-mail: kvt@sevgtu.sebastopol.ua

ЗАДАЧА УПРАВЛЕНИЯ НАГРУЗКОЙ УЗЛА СЕТИ

Рассматривается задача оптимизации распределения нагрузки узла сети с применением аппарата имитационного моделирования, исследуются средства автоматизации обработки результатов моделирования с целью повышения эффективности исследования сложных систем. Приводится численный пример.

Протоколы канального уровня, применяемые в существующих локальных сетях, используют метод доступа к среде, основанный на ее совместном использовании несколькими узлами за счет разделения во времени. В случае разделения ресурсов случайным потоком запросов могут возникать очереди. Для описания этого процесса наиболее часто используется теория массового обслуживания.

Задача проектирования оптимальной сети сводится к выбору наилучшей топологии, эффективного управления нагрузкой узла сети, а также определению рабочих характеристик сети и этапов будущего развития. Перечисленные подзадачи наиболее эффективно решаются методами имитационного моделирования, зарекомендовавшими себя как высокоэффективные средства исследования сложных систем [1]. Однако управление нагрузкой узла сети будет оптимальным только тогда, когда результаты моделирования будут обработаны с максимальной точностью. Методика обработки результатов известна, но к числу ее недостатков относятся: громоздкость формул, многогранность экспериментальных работ и, как следствие, неоправданные временные затраты на ее реализацию.

Целью данной статьи является составление детальной методологии обработки результатов имитационного моделирования, необходимой для разработки автоматизированного комплекса управления нагрузкой узла сети. Данная стратегия позволит уменьшить временные затраты на обработку данных, а также сократить взаимодействие пользователя со средой моделирования, что повысит эффективность управления в целом.

Общая схема исследования нагрузки узла сети выглядит так, как показано на рисунке 1. В задачи аппарата имитационного моделирования входит многократный прогон модели узла сети с целью получения статистических выборок значений его характеристик.

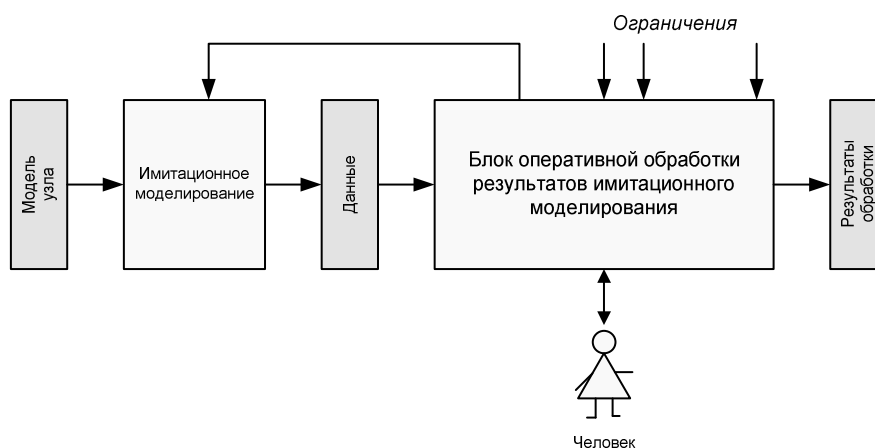


Рисунок 1 — Структура исследования узла сети

Рассмотрим модуль обработки результатов имитационного моделирования более детально. В его функции необходимо включить следующее:

- исследование модели сложной системы;
- первичный и вторичный анализы результатов моделирования;
- планирование и проведение экспериментов с моделью;
- возможность оптимизации системы.

С целью синтеза наиболее адекватной структуры помимо функций модуля априорно задаются ограничения на его параметры и критерии качества результата проектирования. Качество разработанной структуры можно оценить, пользуясь следующими показателями:

- точностью получаемых результатов;
- скоростью обработки данных;
- объемом памяти, необходимым для представления внутренних структур данных;
- уровнем автоматизации.

Описанные выше функции разрабатываемого объекта представим с большей степенью детализации на рисунке 2.

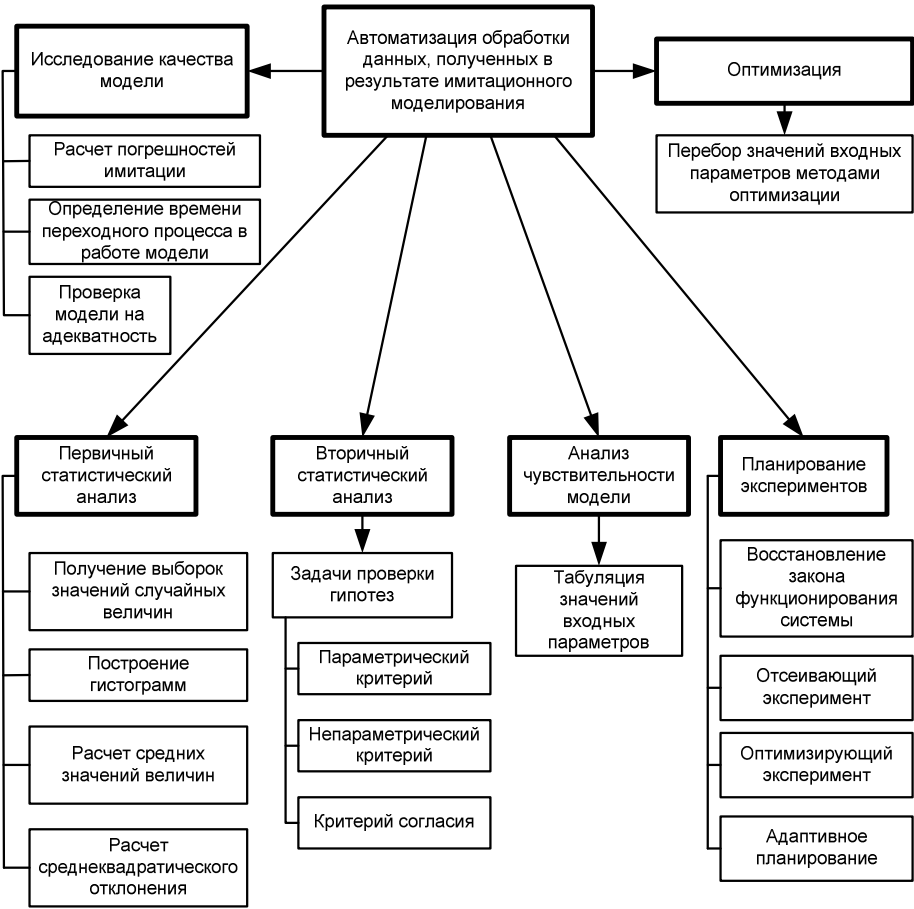


Рисунок 2 — Схема обработки результатов имитационного моделирования

Из шести составляющих процесса анализа результатов имитации (рисунок 2) некоторые частично автоматизированы в существующих разработках, но в статье говорится о комплексном подходе в данной области. Это позволит перенести процесс исследования сложных систем на качественно новый уровень. Для иллюстрации подхода сопроводим каждый этап исследования соответствующим примером, что позволит перенести разработку модуля на алгоритмический уровень.

Имитационная модель должна пройти многоэтапную проверку, что позволит в дальнейшем гарантировать достоверность результатов имитации. Качество модели определяется преимущественно временем переходного процесса и точностью результатов. Способность принятой структуры модели наиболее точно описывать динамику реальной системы определяет ее *адекватность*. В случае если статистическая совместимость выборок моделируемых и реальных значений не выполняется, производится калибровка модели.

В качестве объекта исследований выбрана модель разомкнутой многопроцессорной системы с отказами смешанного типа. В данной системе ограничены емкость буфера и время существования заявки. Критерий оптимальности системы зависит от коэффициента потерь и коэффициента простоя системы за время моделирования. В соответствии со структурой исследования системы (рисунок 2) выполним расчет погрешностей имитации, проводимой с использованием аппарата GPSS World.

Ранее говорилось, что одной из составляющих имитационной модели является генератор случайных чисел. В действительности, все генераторы, используемые в современных системах моделирования, являются генераторами псевдослучайных чисел. Невысокое качество генераторов является источником погрешности результатов имитационного моделирования и, как следствие, снижает адекватность модели. Данная особенность отражается в таблице 1. В GPSS World может быть любое число генераторов [2], из которых рассмотрим десять. Номер генератора случайных чисел совпадает с его начальным числом.

Таблица 1 — Значения критерия оптимальности при различных генераторах случайных чисел

<i>i</i>	<i>Число прогонов, n_i</i>	<i>Номер генератора случайных чисел</i>									
		<i>1</i>	<i>10</i>	<i>20</i>	<i>50</i>	<i>100</i>	<i>120</i>	<i>150</i>	<i>200</i>	<i>300</i>	<i>350</i>

1	10	14,276	14,281	14,175	14,373	14,352	14,25	14,415	14,262	14,314	14,376
2	100	14,256	14,282	14,273	14,294	14,251	14,294	14,28	14,293	14,278	14,287

На основе выборок рассчитаем числовые характеристики случайной величины: математическое ожидание (W_{1cp} для $n_1=10$, W_{2cp} для $n_2=100$) и дисперсию:

$$W_{1cp} = 14,307; D_1 = 0,0052;$$

$$W_{2cp} = 14,279; D_2 = 0,0002.$$

Вычислим доверительные интервалы при вероятности $p=0,5$ и числе степеней свободы $N=10$ (коэффициент Стьюдента равен $t_{0,95} = 2,228$):

$$W_{cp} - t_{0,95} \sqrt{\frac{D}{N-1}} \leq W \leq W_{cp} + t_{0,95} \sqrt{\frac{D}{N-1}}. \quad (1)$$

В результате получим значения критерия эффективности узла сети при количестве прогонов $n_1=10$ и $n_2=100$ соответственно:

$$W_1 = 14,307 \pm 0,054;$$

$$W_2 = 14,279 \pm 0,010.$$

Результаты анализа относительно простой системы демонстрируют наличие погрешности, вносимой генераторами. В реальных условиях для более сложных систем, когда проведение прогонов ограничено, влияние погрешности будет более масштабным. В таком случае возникает задача управления качеством генератора, сводимая к проектированию управляемых генераторов псевдослучайных чисел.

Число прогонов во многом определяет точность результатов. Для множества моделей систем эта величина варьируется в зависимости от структуры модели, и, как показывает опыт исследований в данной области, в среднем необходимо выполнить $10^2 \dots 10^4$ прогонов.

Стационарная работа модели в значительной степени определяется выбором *начальных условий* испытаний сложных систем, на основе которых был разработан алгоритм функционирования модели [3]. Удачно подобранные условия позволяют получить наиболее точное математическое описание исследуемого процесса. Таким образом, имитационная модель будет адекватно отражать влияния внешних факторов на выходную характеристику системы.

Для повышения точности спроектированной модели важно провести коррекцию результатов имитационного моделирования. Процедура корректировки заключается в исключении величины смещения в точечных оценках выходных показателей. Величину смещения можно оценить по формуле

$$\Delta_R(c) = \frac{1}{2} \left(\int \frac{\partial^2}{\partial c^2} R dc \right) \Big|_c \{D(\mathcal{E} - c)\}, \quad (2)$$

где c — истинный вектор параметров выходной характеристики системы, $\mathcal{E}$ — оценка вектора c , D — дисперсия, $\frac{\partial^2}{\partial c^2} R$ — вторая производная выходной характеристики.

Отклонение результатов моделирования можно оценить более простой, однако, менее точной формулой *среднеквадратической ошибки калибровки* R_k . Формула использует значения выходных характеристик системы (y_i) и ее модели (v_i), образующих векторы Y и V

$$R_k = \sqrt{\frac{\sum_{i=1}^n (y_i - v_i)^2}{n}}, \quad (3)$$

где y_i — очередное значение выходной характеристики системы, v_i — значение выходной характеристики модели, n — число прогонов модели.

Сформулируем последовательность действий автоматизированной системы на этапе исследования качества модели узла сети.

1. Многократный прогон модели с сопутствующей сменой генераторов случайных чисел (для этого в системе GPSS World предусмотрена команда RMULT).

2. Расчет числовых характеристик на основе полученной выборки значений.

3. Оценка смещения результата моделирования.

4. Вычисление среднеквадратической ошибки калибровки.

5. Сравнение с величиной ошибки, заданной пользователем. Если ошибка незначительна, то этап закончен.

6. Калибровка модели и переход в п.1

Реализация данного алгоритма позволит получить отлаженную и в наибольшей степени адекватную исследуемой системе модель, пригодную для дальнейших исследований. Калибровка модели может не дать положительного результата после множества ее модификаций. В таком случае необходимо сделать заключение о непригодности составленной модели и сообщить об этом пользователю.

Установив, что имитационная модель наиболее достоверно отражает процессы в реальной системе, приступаем к этапу *первичного статистического анализа* результатов имитации.

Проведено 10 независимых прогонов модели и получена выборка значений критерия качества узла сети. Результаты отображены в таблице 2.

Таблица 2 — Выборка значений критерия

n	1	2	3	4	5	6	7	8	9	10
W, усл.ед.	14,620	14,641	14,489	14,565	14,412	15,093	14,412	14,228	14,668	14,998

Вычислим абсолютные и относительные частоты попадания выборочных значений в соответствующие интервалы. Результаты расчета приведены в таблице 3.

Таблица 3 — Абсолютные и относительные частоты

Интервал	1	2	3	4	5
	14,228 ... 14,401	14,401... 14,574	14,574... 14,747	14,747... 14,92	14,92... 15,093
Абсолютные частоты	1	4	3	0	2
Относительные частоты	0,1	0,4	0,3	0	0,2

Гистограмма выходной характеристики системы изображена на рисунке 3.



Рисунок 3 — Гистограмма относительных частот f попадания значения критерия эффективности в интервал с заданным номером m

Математическое ожидание и дисперсия равны соответственно $m_x=14,613$ и $D_x=0,070$.

Реализация данных расчетов на ЭВМ возможна средствами среды Matlab, оснащенной мощной базой статистической обработки данных. Эта особенность не исключает возможности программирования расчетов в среде Delphi. Исследователями уже выведены алгоритмы, позволяющие сэкономить память ЭВМ, на которой проводятся расчеты. Так, математическое ожидание и дисперсия вычисляются по рекуррентным формулам путем наращивания итогов при появлении очередного измерения случайной величины [4].

Результаты первичного статистического анализа важны для дальнейших исследований. Автоматизация требует аналитического описания происходящих процессов. Поведение случайной величины можно описать одним из законов распределения в случае идентичности его графика гистограмме относительных частот случайной величины (рисунок 3). Данная процедура проводится в рамках *вторичного статистического анализа*.

В качестве примера была реализована проверка гипотезы H_0 о согласии статистического распределения с показательным законом. Значение критерия Хи-квадрат вычислялось по формуле

$$\chi^2 = n \sum_{i=1}^k \frac{(p_i^* - p_i)^2}{p_i}, \quad (4)$$

где p_i — теоретическая частота, рассчитанная по предполагаемому закону; p_i^* — статистическая частота, определенная по выборочным данным; k — число разрядов; n — объем выборки.

$$\text{Параметр распределения } \lambda^* = \frac{1}{m_x^*} = \frac{1}{14,613} = 0,68.$$

Рассчитаны статистические и теоретические частоты попадания значения в соответствующий интервал. Результаты расчетов отражены в таблице 4.

Таблица 4 — Теоретические и статистические относительные частоты

Теоретическая (р)	0,572	0,397	0,167	0,558	0,478
Статистическая (р*)	0,1	0,4	0,3	0	0,2

По формуле (4) получено статистическое значение критерия, равное $\chi^2 = 32,14$.

Рассчитанное значение близко к табличному, тем не менее, гипотеза о нормальном распределении отвергается.

Выполнение вторичного статистического анализа функционирования узла сети связано с перечисленной ниже последовательностью действий.

1. Получение выборки значений критерия качества узла сети.
2. Построение гистограммы.
3. Расчет числовых характеристик.
4. Проверка гипотез.
5. Составление отчета.

Эксперимент должен быть не только эффективным, но и оправданным, в связи с чем применяется теория планирования. Определим чувствительность имитационной модели к входным параметрам. Итоги этой процедуры определяют целесообразность прогонов модели. Результаты анализа приведены в таблице 5.

Таблица 5 — Значения критерия оптимальности системы при табулировании указанных параметров

Емкость буфера	1	2	3	4	5	6	7
W, усл.ед.	14,282	14,279	14,276	14,272	14,269	14,266	14,262
Количество процессоров	1	2	3	4	5	6	7
W, усл.ед.	14,282	21,958	55,464	66,596	73,276	77,729	80,909
Интенсивность входного потока заявок, с ⁻¹	1	2	3	4	5	6	7
W, усл.ед.	14,282	2,397	52,565	64,143	71,195	75,894	79,286
Время обслуживания, с	1	2	3	4	5	6	7
W, усл.ед.	14,282	19,358	21,955	23,533	24,592	25,355	25,929

В данном случае критерий качества работы узла наиболее чувствителен к количеству обслуживающих каналов и параметру входного потока заявок.

Автоматизированная система должна проводить на порядок больше экспериментов с моделью для сбора более точных данных в случае, если при небольших изменениях параметра значительно меняются результаты.

Наиболее результативным с точки зрения исследования систем является блок проведения экспериментов с моделью системы и оптимизации ее функционирования. Теория планирования экспериментов охватывает все вопросы организации и проведения исследований объектов. В современных пакетах программ часть упомянутых на рисунке 2 модулей обработки данных автоматизированы, однако носят внутренний характер и обособлены от неквалифицированных пользователей [2]. Среди таких подпрограмм справедливо отметить модуль проведения дисперсионного анализа на базе системы моделирования GPSS World. В качестве исходных данных используется перечень факторов, подлежащих исследованию, а также уровни их влияния, отображенные в таблице 6. Результаты проведения эксперимента показаны в таблице 7.

Таблица 6 — Факторное пространство исследуемой модели

Фактор		Минимальный уровень	Максимальный уровень
1. Интервал времени между поступлением заявок на обслуживание, с	dT	5	15
2. Время обслуживания, с	Tobs	5	15
3. Емкость буфера	D	5	10
4. Время существования заявки в системе, с	T	10	30
5. Штраф за потерю требования, усл.ед.	StPo	5	10

6. Штраф за простой процессора, усл.ед.	StPr	5	10
---	------	---	----

Таблица 7 — Результаты дисперсионного анализа

Фактор		F-статистика	Критическое значение F
1. Интервал времени между поступлением заявок на обслуживание	dT	11,496	5,12
2. Время обслуживания	Tobs	5,944	5,12
3. Емкость буфера	D	0,006	5,12
4. Время существования заявки в системе	T	0,022	5,12
5. Штраф за потерю требования	StPo	0,001	5,12
6. Штраф за простой процессора	StPr	0,006	5,12
7. Критерий эффективности, усл.ед.	W	3,883	

Таким образом, по результатам моделирования можно сделать вывод, что при данных потоках и характеристиках узла значение критерия эффективности равно 3,883. Для уменьшения материальных затрат автоматизированная система должна продолжить исследование каждого значимого фактора и в первую очередь фактора 1. Факторы, имеющие F-статистику менее критического значения (факторы 3...6) являются незначимыми, и их влияние на переменную отклика в дальнейшем не рассматривается.

Недостатком проведения эксперимента в среде GPSS World является его трудоемкость. Необходимо сократить взаимодействие пользователя со средой моделирования в процессе проведения экспериментов, разработав интерфейс, требующий от него лишь введения набора исходных данных.

Предлагаемая методика позволяет систематизировать исследование сложной системы, в данном случае узла сети, и перенести этапы обработки результатов имитационного моделирования на алгоритмический уровень. Разработка программного комплекса на базе описанной методики позволит сократить влияние пользователя на процесс обработки данных в три раза за счет полной автоматизации 60 % выполняемых операций. Средства, затраченные на разработку автоматического управляющего устройства, восполняются выигрышем во времени и сокращением материальных расходов. Полная автоматизация процесса управления на данный момент невозможна по причине множества факторов, среди которых следует отметить необходимость поддержки принятия решения системой на этапах калибровки модели и проведения отсеивающих и оптимизирующих экспериментов. Необходима разработка средств сбора наиболее полной информации о модели. Задачей дальнейших исследований является поиск алгоритма введения точек принятия решений в управлении нагрузкой узла сети, что позволит повысить уровень автоматизации процесса управления на 20...30 %.

Библиографический список

1. Бусленко В.Н. Автоматизация имитационного моделирования сложных систем / В.Н. Бусленко. — М.: Наука, 1977. — 240 с.
2. Боев В.Д. Моделирование систем. Инструментальные средства GPSS World: учеб. пособие / В.Д. Боев. — СПб.: БХВ-Петербург, 2004. — 368 с.
3. Шаракшанэ А.С. Сложные системы: учеб. пособие / А.С. Шаракшанэ. — М.: Высшая школа, 1977. — 247 с.
4. Советов Б.Я. Моделирование систем: учебник для вузов / Б.Я. Советов, С.А. Яковлев. — М.: Высш. шк., 1985. — 271 с.

Поступила в редакцию 22.04.2008 г.