

ПЛАНИРОВАНИЕ МАШИННЫХ ЭКСПЕРИМЕНТОВ ПРИ ИМИТАЦИОННОМ МОДЕЛИРОВАНИИ

В. Н. Прищенко, Ю. Н. Саенко, А. Н. Татаринев (Санкт-Петербург)

При имитационном моделировании так же, как и при любом другом методе анализа и синтеза системы, весьма существенен вопрос его эффективности. Эффективность имитационного моделирования может оцениваться рядом критериев, в том числе точностью и достоверностью результатов моделирования, временем построения и работы с моделью, затратами машинных ресурсов.

Как правило, имитационные модели (ИМ) содержат случайные события, величины, процессы. Поэтому моделирование представляет собой выборочные эксперименты, анализ результатов которых имеет статистические аспекты. Последние связаны с этапами моделирования:

– на этапе постановки задачи необходимо определить цель (сравнение вариантов на основе точечной оценки показателей, определение функциональной зависимости показателя и факторов, определение оптимальной комбинации факторов);

– на этапе сбора и обработки данных об объекте следует решить основные задачи математической статистики (определение законов распределения случайных величин, проверка правдоподобия гипотез или определение параметров распределения по статистическим данным) и выбирать начальные условия для исследования переходного или устойчивого состояния;

– на этапе разработки модели необходимо выбрать факторы и комбинации их уровней для экспериментирования в соответствии с целями (планирование экспериментов);

– на этапе реализации модели и интерпретации результатов осуществляется манипулирование выборкой для уменьшения дисперсии ошибок и определяется объем выборки, исходя из заданной статистической надежности.

После того, как имитационная модель реализована на ЭВМ, необходимо выполнить испытания модели, исследование ее свойств и составить план проведения имитационных экспериментов на ЭВМ.

При исследовании свойств модели необходимо установить диапазон изменения отклика модели \vec{Y} при изменении каждой компоненты вектора параметров \vec{X} . В зависимости от диапазона изменения откликов \vec{Y} определяется стратегия планирования экспериментов на ИМ. Если при значительной амплитуде изменений некоторой компоненты вектора параметров модели \vec{X} отклик \vec{Y} меняется незначительно, то это означает, что точность представления этой компоненты в ИМ не играет существенной роли. Кроме того, в планировании имитационных экспериментов эта компонента не будет использоваться как основная. Если же отклик модели \vec{Y} окажется высокочувствительным к изменению некоторой компоненты вектора \vec{X} , то это указывает на необходимость представления ее в модели с максимально возможной точностью.

Оценка чувствительности модели по i -й компоненте вектора параметров \vec{X} определяется парой значений $(\partial X_i^o, \partial Y_i^o)$. Эта пара чисел показывает, на сколько процентов может измениться отклик модели при увеличении i -й компоненты параметров на ∂X_i^o процентов. В результате получается множество пар значений $\{\partial X_i^o, \partial Y_i^o\}$, $i = \overline{1, h}$, где h – размерность вектора параметров модели X .

Каждая i -я компонента вектора \vec{X} отклоняется от значения его в центральной точке в обе стороны на длину выбранного интервала его изменения ($\min X_i, \max X_i$). Приращение i -й компоненты вектора параметров модели, которая является приращением вектора параметров \vec{X} при изменении только одной компоненты i , определяется по следующему выражению

$$\partial X_i^o = \frac{(\max X_i - \min X_i) \cdot 2}{(\max X_i + \min X_i)} \cdot 100\% . \quad (1)$$

Изменение вектора отклика \vec{Y} определяется максимальным значением из ∂Y_n ; $\partial Y_i^o = \max\{\partial Y_n\}$, $n = \overline{1, L}$; L – размерность вектора отклика. При этом ∂Y_n определяется по формуле

$$\partial Y_n = \frac{|\max Y_n - \min Y_n| \cdot 2}{(\max Y_n + \min Y_n)} \cdot 100\% . \quad (2)$$

Эффективность машинных экспериментов с имитационными моделями, которые формализованы в виде схем массового обслуживания, существенно зависит от выбора плана эксперимента, так как именно план определяет объем и порядок проведения вычислений на ЭВМ, приемы накопления и статистической обработки результатов моделирования системы в целом, что влияет на эффективность использования ресурсов ЭВМ при моделировании.

При реализации на ЭВМ статистических моделей процессов эксплуатации (функционирования) производится многократный розыгрыш случайных ситуаций, которые могут иметь место при известных законах распределения случайных величин. Накопленные реализации могут рассматриваться как выборка конечного объема N (N – общее число розыгрышей) и обрабатываются теми же методами, что и выборка, полученная в ходе натуральных наблюдений.

Очевидно, что для того чтобы иметь возможность поставить на ЭВМ эксперимент, необходимо:

- программным образом получить случайные величины с заданным законом распределения;
- описать процессы порождения запросов на обслуживание, а также процессы формирования очередей и собственно обслуживания;
- накапливать статистическую информацию в ходе машинного эксперимента.

Математические методы планирования экспериментов основаны на кибернетическом представлении процесса проведения эксперимента, наиболее подходящей моделью которого является абстрактная схема типа «черного ящика» вида $Y = \varphi(X)$, где $\vec{Y} = (\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n)$ – множество векторов зависимых выходных переменных, называемых реакциями.

При проведении машинного эксперимента с моделью для оценки характеристик процесса функционирования исследуемого НС необходимо создать такие условия, которые способствовали бы выявлению влияния существенных факторов, находящихся в функциональной зависимости с искомой характеристикой.

Для этого необходимо:

- отобрать факторы x_i , $i = 1, k$, влияющие на искомую характеристику, и описать функциональную зависимость;
- установить диапазон изменения факторов $x_{\min} \dots x_{\max}$;

– определить координаты точек факторного пространства $\{x_1, x_2, \dots, x_k\}$, в котором следует проводить эксперимент;

– оценить необходимое число реализаций и их порядок в эксперименте.

Если число факторов равно k , то при их варьировании на двух уровнях (-1; +1) число сочетаний (комбинаций) уровней факторов или, что то же самое, число опытов N эксперимента определится выражением $N = 2^k$.

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом (ПФЭ). Начальный этап планирования эксперимента для получения коэффициентов линейной модели основан на варьировании факторов на двух уровнях: нижнем X_n и верхнем X_v .

Планирование ПФЭ с любым числом факторов k сводится к записи в матрицу всех неповторяющихся сочетаний уровней этих факторов.

Любой эксперимент сопровождается погрешностями (ошибками воспроизводимости). Для оценки воспроизводимости необходимо осуществить параллельные опыты, то есть каждый i -й опыт матрицы планирования выполняется в конечном итоге несколько раз. Число серий n характеризует параллельность опытов матрицы планирования. Каждая серия должна включать N неповторяющихся опытов матрицы планирования.

Число параллельных опытов, а следовательно, и число серий опытов n рекомендуется выбирать из условия $n \geq 2 \div 5$.

Необходимо отметить, что оценка воспроизводимости опытов по сути сводится к расчету так называемой дисперсии воспроизводимости.

С целью уменьшения детерминированных факторов при реализации плана ПФЭ необходимо провести так называемую рандомизацию, то есть опыты каждой серии выполняются не по порядку, как они записаны в матрице планирования, а в случайной последовательности. Случайная последовательность опытов каждой серии выбирается по таблицам случайных чисел.

Выполнение ПФЭ состоит в последовательном проведении опытов каждой серии. В пределах серии опыты выполняются в очередности, полученной в результате рандомизации.

При проведении опытов для всех факторов принудительно устанавливаются также их натуральные (в своей размерности) значения, которые соответствуют кодированным значениям этих факторов, указанном в матрице планирования для выполняемого опыта.

Последовательность выполнения статистической обработки и проверки пригодности для практики построенной модели такова. Необходимо определить средние и дисперсии отклика в i -м опыте по формулам:

$$M(y_i) = \frac{\sum_{u=1}^n (y_i)_u}{n}; \quad (3)$$

$$D(y_i) = \frac{\sum_{u=1}^n [(y_i)_u - M(y_i)]^2}{n-1}, \quad (4)$$

где $(y_i)_u$ – значение отклика в i -м опыте (строке) u -й серии опытов; n – число параллельных опытов (серий опытов).

По критерию Кохрена необходимо проверить дисперсию $D(y_i)$ на однородность $G_{расч} < G_{кр}$, где $G_{расч}, G_{кр}$ – расчетное и критическое (табличное) значения критерия Кохрена, найденное для заданной доверительной вероятности $\gamma = 0,95$.

Проверка значимости коэффициентов регрессии выполняется по критерию Стьюдента $t_{расч} > t_{кр}$, где $t_{расч}, t_{кр}$ – расчетное и критическое значения критерия Стьюдента, найденное для заданной доверительной вероятности $\gamma = 0,95$.

Проверка адекватности модели и заключение о ее пригодности для практики проводится по критерию Фишера $F_{расч} < F_{кр}$, где $F_{расч}, F_{кр}$ – расчетное и критическое (табличное) значения критерия Фишера, найденное для заданной доверительной вероятности $\gamma = 0,95$.

Процесс статистической обработки результатов активных факторных экспериментов целесообразно проводить на ЭВМ, например, с помощью пакетов прикладных программ STADIA, STATGRAPHICS, STATISTICA.

В связи с тем, что количество испытаний в ПФЭ значительно превосходит число определяемых коэффициентов линейной модели плана эксперимента, т. е. ПФЭ обладает большой избыточностью и поэтому возникает проблема сокращения их количества. С увеличением k число опытов N растет очень быстро. Так, например, при $k=10$ число опытов $N = 2^{10} = 1024$. Поэтому проблема сокращения числа опытов стоит достаточно остро. При этом задача заключается в том, чтобы не только уменьшить число опытов, но и получить достаточное количество информации об объекте исследования.

Учитывая эти причины, предлагается применять дробный факторный эксперимент (ДФЭ). Правило проведения ДФЭ следующее: для сокращения числа испытаний новому фактору присваивается значение вектор-столбца матрицы, принадлежащего взаимодействию, которым можно пренебречь.

Применение ДФЭ предусматривает проведение меньшего числа опытов (в два и более раза) по сравнению с ПФЭ, что является экономически более целесообразным. Правильно построенный план ДФЭ точно так же, как и план ПФЭ, обладает свойствами симметричности, нормировки и ортогональности.

Литература

1. **Иванов А.Ю., Полковников С.П., Ходасевич Г.Б.** Военно-технические основы построения и математическое моделирование перспективных средств и комплексов автоматизации. – СПб.: ВАС, 1997. – 419 с.
2. **Коваленко И.А., Кузнецов Н.Ю.** Статистическое моделирование. Моделирование высоконадежных систем / Надежность технических систем. Справочник. – М.: Радио и связь, 1985. – 456 с.
3. **Максимей И.В.** Имитационное моделирование на ЭВМ. – М.: Радио и связь, 1988. – 232 с.
4. **Тюрин Ю.Н., Макаров А.А.** Статистический анализ данных на компьютере/под ред. В.Э.Фигурнова. – М.: ИНФРА-М, 1988. – 528 с.