К ускорению имитационного моделирования

О. И. Кутузов
Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

Т. М. Татарникова
Санкт-Петербургский государственный университет аэрокосмического приборостроения

***Abstract*. The problem solved by the machine simulation method in the study of technical systems with a distributed structure is discussed. The complex application of a stratified sample and balanced modeling in the evaluation of rare events makes it possible to significantly reduce the tests number without losing the accuracy of the analyzed system characteristics under study.**

Keywords: Monte Carlo method; balanced layered modeling; rare events estimation; experiments number; efficiency by experiments number; simulation modeling acceleration

I. Введение

В настоящее время имитационное моделирование переживает второе рождение [1, 2]. В практической реализации оно опирается на четыре объектно-ориентированные доктрины, составляющие основу библиотеки методов и их реализаций [3]. Объединенная логика имитационного моделирования (ИМ) открывает путь к фабрикам имитационных моделей, первые образцы которых уже существуют [4].

Технология ИМ стохастических систем реализуется с привлечением метода статистических испытаний Монте-Карло [5],[6], известный недостаток которого заключается в его медленной сходимости.

Проблема точности моделирования стохастических систем является центральной в ИМ. Общей стратегией снижения цены точности является ускорение сходимости вычисляемых оценок. Ускорение может достигаться как за счет соответствующего аналитического преобразования решаемой задачи [7], так и системотехническими методами ускорения расчетов, в частности, путем организации параллельных вычислений и распределенного моделирования [8]. Наибольшего эффекта удается достичь тогда, когда методы ускорения учитывают специфику моделируемых объектов, решаемых задач и алгоритмов их решения. Так, например, сочетание расслоенной выборки [7] и равновзвешенного моделирования [9] при анализе сетевых моделей дает весьма ощутимый эффект.

Покажем это на двух примерах.

II. Метод равновзвешенного расслоенного моделирования

Во многих приложениях представление показателя качества моделируемой системы используется в виде среднего риска:

$$Q = M\left\{ f\left(\bar{\alpha}, \bar{\beta}\right)\right\},$$

где $\bar{\alpha} = \left(\alpha_1, ..., \alpha_n\right)$ – случайный вектор, описывающий процессы в моделируемой системе; $\bar{\beta} = \left(\beta_1, ..., \beta_r\right)$ – вектор, задающий параметры модели; $f(.)$ – функция, значение которой определяется в ходе вероятностного моделирования при различных реализациях x_i, $i = \overline{1, N}$ вектора $\bar{\alpha}$.

В частности, если $f(.)=\xi=1$ при выполнении моделируемой системой заданных требований и $f(.)=\xi=0$ в противном случае, то $Q\left(\bar{\beta}\right)$ есть вероятность выполнения заданных требований. Решение задач, связанных с оценкой редких событий, может быть сведено к оцениванию математического ожидания $M\xi$ двоичной случайной величины (СВ) ξ, заданной в виде функции $\xi = f\left(\bar{\alpha}\right)$, причем СВ $\bar{\alpha} = \left(\alpha_1, ..., \alpha_n\right)$ имеет закон распределения $p\left(\alpha \sim p\right)$, который известен.

Расчет оценки для $M\xi$ может выполняться с помощью различных аналитико-статистических алгоритмов. Их эффективность можно характеризовать дисперсией оценки при фиксированном числе N испытаний. Чем меньше дисперсия оценки, тем точнее алгоритм и тем он эффективнее, если, конечно, сложность расчетов в сравниваемых алгоритмах различается незначительно.

При прямом статистическом моделировании («чистый» метод Монте-Карло) генерируется N независимых реализаций $\alpha^1, \dots, \alpha^N$ случайной величины $\bar{\alpha}$ и вычисляются $\xi^i = f\left(\alpha^i\right)$, $i = \overline{1, N}$. Оценка $\hat{m}$ определяется как $\hat{m} = \left(\xi^1 + ... + \xi^N\right)/N$ и дисперсия $D\hat{m}$ оценки $\hat{m}$ составляет величину $D\xi/N$.

При методе расслоения [6] искомая оценка имеет вид $\hat{m}_p = \sum_i^L p_i \hat{m}_i$, где L – число слоев; $\hat{m}_i$ и p_i – оценка и вероятность i-го слоя, $\hat{m}_i = \frac{1}{N_i}\sum_1^{N_i} \xi_i^j$ для $M\xi_i$, $i = \overline{1, L}$, $\sum_i N_i = N$.

В работе [9] предложен оригинальный алгоритм взвешенного моделирования. Вместо случайной величины α разыгрываются равновероятные ее реализации $\gamma \sim p_0(x),\ x\in X,\ p_0(x_j)=1/n,\ j=\overline{1,n}$ и показатель $M\xi$ оценивается как математическое ожидание случайной величины $\zeta = nf(\gamma)p(\gamma)$ на том основании, что $M\xi = M\zeta$:

$$M\zeta = \sum_{x\in X} nf(x)\,p(x)\,p_0(x) = \sum_{x\in X} f(x)\,p(x) = M\xi.$$

Этот метод равновзвешенного моделирования (РВМ) по существу реализует случайный выбор с равными вероятностями для возможных состояний системы.

В общем случае РВМ проигрывает в числе испытаний $N_{\text{РВМ}}$ числу испытаний $N_{\text{п}}$ при прямом моделировании методом Монте-Карло. В [10] получена оценка этого проигрыша в виде

$$t = \frac{N_{\text{прм}}}{N_{\text{п}}} = 1+\rho^2, \tag{1}$$

где ρ – коэффициент вариации множества вероятностей для единичных значений функции $f(\overline{\alpha})$.

Из (1) следует, что при РВМ требуется провести в $(1+\rho^2)$ раз большее число прогонов модели по сравнению с прямым моделированием, чтобы получить одинаковую точность оценок.

Уменьшить значение показателя t можно за счет совместного применения расслоения и равновзвешенного моделирования. Оценим такую возможность.

Эффективность равновзвешенного моделирования оценим по соотношению числа испытаний (опытов), необходимых для получения оценки с одинаковой точностью как при РВМ, так и при прямом моделировании методом Монте-Карло.

В общем случае при прямом моделировании полное количество испытаний $N_{\text{п}}$, необходимых для получения оценки с заданной точностью, условно представим в виде двух подмножеств $N'_{\text{п}}$ и $N''_{\text{п}}$ таких, что

$$N_{\text{п}} = N'_{\text{п}} + N''_{\text{п}},$$

где $N'_{\text{п}}$ – число испытаний, реализации которых могут содержать положительный исход, то есть в каждой из этих реализаций возникает состояние, для которого с вероятностью $0<p^+<1$ возможно значение $\xi = f(x_i)=1,\ i\in N'_{\text{п}}$. Назовем такие реализации содержательными; $N''_{\text{п}}$ – число испытаний, реализации которых не могут содержать положительного исхода, т.е. в каждой из этих реализаций возникает состояние, для которого с вероятностью 1 значение $\xi = f(x_i)=0,\ i\in N''_{\text{п}}$. Назовем такие реализации пустыми.

Тогда, чтобы при прямом моделировании получить $N'_{\text{п}}$ содержательных реализаций, необходимо провести в среднем $N_{\text{п}} = N'_{\text{п}}/p^+$ испытаний. При этом получаем оценку, точность которой характеризуется дисперсией $D\hat{\xi}$.

При совместном применении расслоения и РВМ проводим только содержательные испытания, число которых пусть будет равно $N_{\text{РВРМ}} = N'_{\text{п}}(1+\rho^2)$ и получаем оценку с дисперсией $D\hat{\zeta} = D\hat{\xi}$. Соответственно, соотношение числа испытаний при условии $D\hat{\zeta} = D\hat{\xi}$ равно

$$t = \frac{N_{\text{п}}}{N_{\text{РВРМ}}} = \frac{1}{p^+(1+\rho^2)}. \tag{2}$$

Таким образом, выигрыш в числе испытаний при переходе от прямого моделирования к совместному применению расслоения и равновзвешенного моделирования имеет место при условии, что

$$p^+(1+\rho^2) < 1.$$

Покажем, что применение расслоения уменьшает значение t.

III. Тестовый пример

Оценим этот возможный выигрыш при моделировании маловероятных событий на примере расчета надежности системы, структурная схема которой изображена в виде случайного графа на рис. 1 [11].

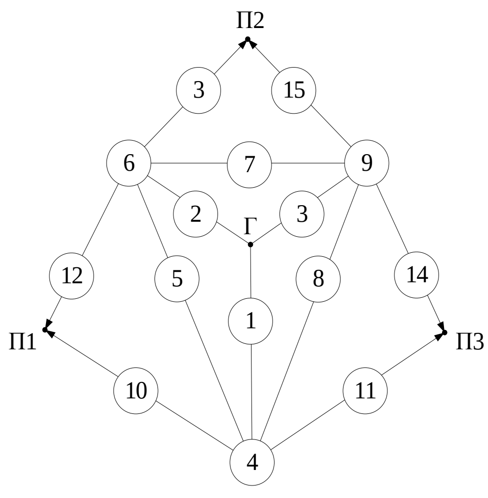

Рис. 1. Структурная схема рассматриваемой системы

Вершины (элементы системы) 1,...,15 имеются в графе с вероятностями $q_1,\ldots,q_{15}$ соответственно, дуги абсолютно надежны. Система работоспособна, если из полюса Г есть пути во все полюса П1, П2, П3 (движение против ориентированных дуг запрещено). Требуется найти вероятность отказа системы при следующих вариантах:

I. $q_1=q_2=q_3=10^{-5}$,

II. $q_4=q_6=q_9=2\cdot10^{-5}$,

III. $q_{10}=\ldots=q_{15}=5\cdot10^{-5}$.

Для вариантов I и II известны точные решения [11] (табл. 1, графа *A*). В [9] приведены результаты расчетов методом РВМ без расслоения, содержащие значения оценок $\hat{Q}_\xi$ для $M\xi$ и оценки $\hat{\vartheta}$ для коэффициентов вариации СВ $\hat{Q}_\xi$, (табл. 1, графа *B*). Видно, что при хорошей точности оценок $\hat{Q}_\xi$ в целом, при появлении некоторого разброса в исходных данных (вариант III) значение $\hat{\vartheta}$ заметно возрастает.

ТАБЛИЦА I

Вариант	A	B		C	
		N=1000		N=1020	
	$M\xi\cdot10^8$	$\hat{Q}_\xi\cdot10^8$	$\hat{\vartheta}$	$\hat{m}_0\cdot10^8$	$\hat{\vartheta}$
I	12,00	12,40	0,086	11,82	0,069
II	0,120	0,1271	0,085	0,1183	0,069
III	–	1,549	0,354	1,505	0,068

Эта же задача решена методом равновзвешенного моделирования с расслоением. Система, изображенная на рис. 1, имеет связность γ=2. Анализ показывает, что для данных значений отказов элементов при расслоенном моделировании достаточно рассмотреть слои с кратностью отказов $i=\overline{2,3}$.

Разделение элементов системы на группы состояло в выделении для i=2 и i=3 одних и тех же 4 групп элементов с номерами 1,2,3; 4,6,9; 5,7,8 и 10,...,15 соответственно. При этом получены результаты графа *C*, приведенные в таблице. Можно видеть, что для варианта III показатель $\hat{\vartheta}$ оказался примерно таким же, как и для первых двух вариантов.

Для рассматриваемого примера $p^+=2,7\cdot10^{-4}$ и $\rho_\gamma^2=0,303$. Согласно (2) выигрыш в числе испытаний при расслоении и равновзвешенном моделировании по сравнению с прямым моделированием по методу Монте-Карло составит $t_1=2,84\cdot10^3$ раз.

IV. Заключение

Комбинированное применение аналитико-статистических методов моделирования позволяет ускорить численный расчет и анализ характеристик исследуемой системы методом машинной имитации. Выполнена оценка эффективности совместного применения аналитико-статистических методов моделирования в сравнении с непосредственным моделированием по числу испытаний.

Приведенный тестовый пример показывает насколько само по себе эффективно применение расслоения при моделировании редких событий. Кроме этого пример демонстрирует, каким образом применение РВМ упрощает процесс моделирования.

Список литературы

[1] The Big Book of Simulation Modeling. Multimethod modeling with AnyLogic 6. AnyLogic North America Publ., 2013. 614 p.

[2] Clark J.S., Gelfand A.E. Hierarchical Modelling for the Environmental Sciences. Statistical Methods and Applications. Oxford University Press, 2006. 205 p.

[3] Кутузов О.И., Татарникова Т.М. К анализу парадигм имитационного моделирования // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 3. С. 552–558. doi: 10.17586/2226-1494-2017-17-3-552-558.

[4] Карпов Ю.Г. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5. СПб.: БХВ-Петербург, 2005. 400 с.

[5] Соболь И. М. Численные методы Монте-Карло. М.: Наука, 1973. 212 с.

[6] Поляк Ю.Г. Вероятностное моделирование на ЭВМ. М.: Сов. Радио, 1971. 400 с.

[7] Kleijnen J.P.C. Design and analysis of simulation experiments. New York; London: Springer, 2008. 216 p.

[8] Олзоева С.И. Распределенное моделирование в задачах разработки АСУ. Улан-Удэ: ВСГТУ, 2005. 219 с.

[9] Плакс Б.И. Расчет надежности систем со сложной структурой ускоренным методом Монте-Карло// Изв. АН СССР. Техн. кибернетика. 1983. №6. С. 158-162.

[10] Кутузов О.И., Задорожный В.Н. Аналитико-статистический метод расчета высоконадежных систем связи // Техника средств связи, серия Техника проводной связи. 1990. Вып. 1. С. 121-130.

[11] Рябинин И.А., Черкесов Г.Н. Логико-вероятностные методы исследования структурно-сложных систем. М.: Радио и связь, 1981. 264 с.