электронный журнал
МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". Эл No. ФС77-51038.

УДК 004.942

Методы оптимизации имитационных моделей стохастических систем

А.В. Трубкина

Студент, кафедра «Программное обеспечение ЭВМ и информационные технологии»
МГТУ им. Н.Э. Баумана, г. Москва, Россия

Научный руководитель: Рудаков И.В., к.т.н.,
доцент кафедры «Программное обеспечение ЭВМ и информационные технологии»
МГТУ им. Н.Э. Баумана, г. Москва, Россия

МГТУ им. Н.Э. Баумана
trubkina.a@gmail.com

Введение

В настоящее время разработан целый ряд методов оптимизации имитационных моделей, часть этих методов включена в состав программных средств для компьютерного моделирования. Однако в связи с существенными различиями в рассматриваемых имитационных моделях реальных систем не удается сконструировать универсальный метод, способный одинаково эффективно определить глобальный минимум выходной функции для любой модели стохастической системы с различными ограничениями.

Классификация методов оптимизации в контексте моделирования проводится в ряде работ ([1,2,3]). Целью данной работы является обзор и классификация методов оптимизации, используемых для имитационных моделей стохастических систем с непрерывной областью определения параметров для последующего выбора метода оптимизации модели размещения баз машин скорой помощи. В работе рассматриваются общие требования, предъявляемые к методу оптимизации. Дается обзор методов на основе метамоделей, эвристических и вероятностных методов. Приводятся требования к этапам гибридного метода.

Задача размещения баз машин скорой помощи

Рассматривается модель для задачи размещения баз машин скорой помощи. Необходимо определить местоположение каждой из $p(i), i = 1, 2, \ldots, d$, баз машин скорой помощи таким образом, чтобы время прихода машины к месту вызова было минимальным (d – общее число баз, рассматриваются $d < 15$). Вызовы (звонки) поступают по Пуассоновскому закону с параметром λ_a. Все звонки поступают из области $[0,1]^2$. Предполагается, что каждая машина скорой помощи передвигается с постоянной скоростью v, а время работы бригады в точке вызова подчиняется экспоненциальному распределению с параметром λ_s. На поступивший звонок отвечает ближайшая свободная машина скорой помощи. Если свободных машин нет, то звонок ставится в очередь типа FIFO. Критерием останова служит общее время моделирования.

Система из описанной выше задачи обладает следующими особенностями:

- непрерывная область определения параметров (координат x и y каждой базы), по которым необходимо провести оптимизацию;
- выходная функция (время ожидания прихода машины к месту вызова) имеет скалярный вид;
- система является стохастической.

Необходимо выбрать метод оптимизации модели, наиболее полно учитывающий указанные особенности.

Формальная постановка задачи оптимизации в контексте моделирования

Несмотря на длительные разработки в области оптимизации имитационных моделей, эффективные и достоверные алгоритмы до сих пор не созданы и, вполне вероятно, в принципе не существуют в связи с широтой постановки задачи оптимизации [3], которая может быть сформулирована следующим образом:

$$\mu^* = \min_{x \in W} f(x),$$

где функция $f(\cdot)$ определяется моделью с вектором параметров длины m.

Вектор значений x называется конфигурацией. Выходом при моделировании является случайная переменная μ^*. Задача оптимизации сводится к поиску минимума ожидаемого значения (т. е. $f(\cdot) = E[g(x)]$, где $g(x)$ – выходное значение для модели, называемое откликом). Множество возможных параметров является подмножеством N^m (далее дискретные области определения параметров не рассматриваются), R^m или определятся набором неравенств вида $g_i(x) \leq 0$ $(i = 1, 2, \ldots, k)$, т. е. является ограниченным. Предполагается, что W задано и не требует дополнительных модельных экспериментов. Т.к. $f(\cdot)$ определяется стохастической моделью, то отклик может быть получен лишь путем проведения эксперимента.

Для заданной выборочной функции θ, которая для моделирования определяется последовательностью случайных чисел, результат моделирования строго определен. Пусть $f(x,\theta)$ – отклик для выборочной функции θ. В этом случае математическое ожидание и дисперсия для выходного значения будут определяться следующими формулами:

$$\hat{\mu}_x = \sum_{i=1}^{n} f(x,\theta_i),$$

$$\hat{s}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n} (f(x,\theta_i) - \hat{\mu}_x)^2,$$

где n - число экспериментов для независимой выборочной функции θ_i.

Данные значения позволяют определять доверительные интервалы стандартными образом [4].

Требования к методу оптимизации

Метод оптимизации в идеале должен удовлетворять следующим требованиям [3]:

1. Если алгоритм работает бесконечно долгое время, то вероятность найти точку x, такую что $|\mu^* - f(x)| < \varepsilon$ должна стремиться к нулю при $\varepsilon > 0$. Данная формулировка соответствует т.н. *сходимости почти наверное*.

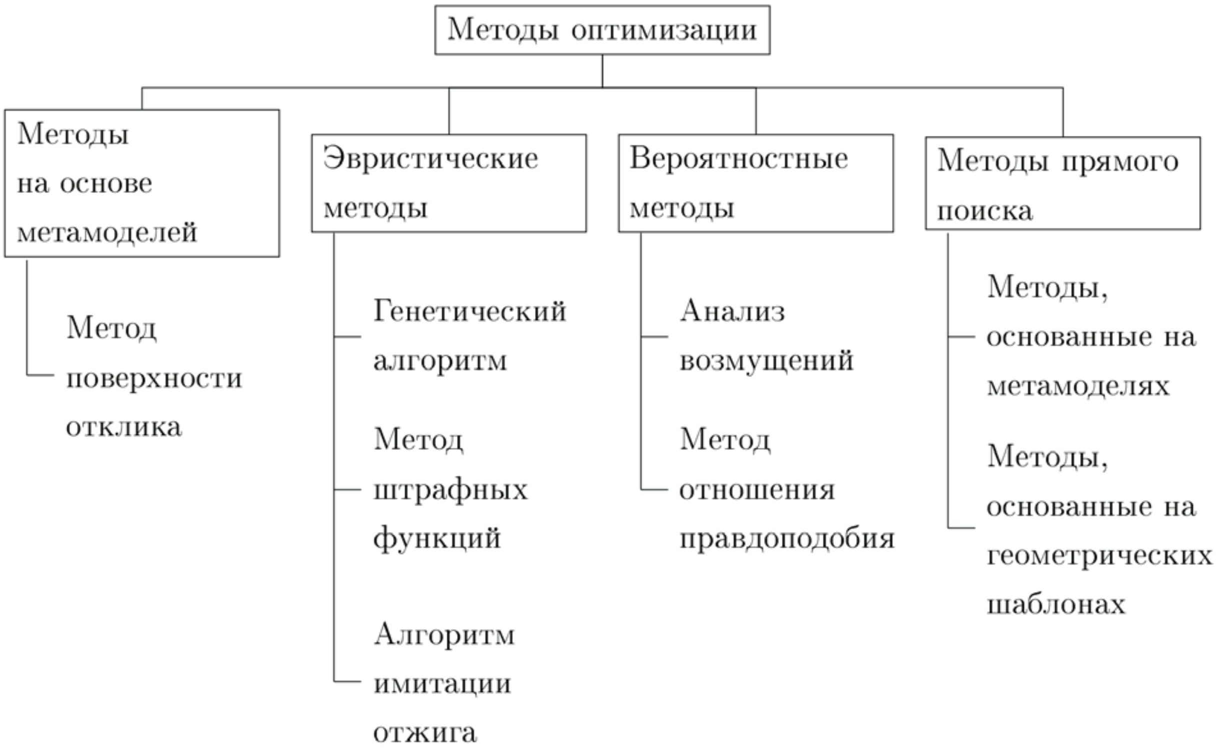
2. Алгоритм должен быстро находить точки с малым значением отклика (эмпирический критерий).

3. Если алгоритм определяет точку x как точку с минимальным значением выходной функции, тогда доверительный интервал для $|\mu^* - f(x)|$ должен быть вычислим, если W конечно.

Следует отметить, что требования 1 и 2 требуют кардинально различной реализации алгоритма оптимизации.

Обзор методов оптимизации

Рассмотрим методы оптимизации, подходящие для моделей с непрерывной областью определения параметров и скалярным откликом, применяемые в том числе в ряде коммерческих программ моделирования. Их можно разделить на несколько категорий, которые представлены на рисунке.



Методы на основе метамоделей

Данный класс методов был изначально предложен для оптимизации систем физических экспериментов [4]. Суть методов состоит в построении одной или нескольких математических моделей A, которые называются метамоделями, для аппроксимации функции f с целью простого и быстрого (по сравнению с модельными экспериментами) вычисления выходной функции в каждой точке x пространства параметров. На основе метамодели осуществляется поиск точки, в которой исходная система потенциально имеет минимальное значение.

В качестве метамоделей используются регрессионные модели и соответствующий им метод поверхности отклика. Модель $A(\cdot)$ определяется путем минимизации разности $A(\cdot)$, $A(\cdot) \in L$ и функции $f(\cdot)$ в точках выборки S, где L – класс функциональных форм. Наиболее часто используемыми L являются линейные и квадратичные модели [5]. Полиномиальные модели обладают требуемой простотой вычисления, но не являются пригодными для сложных, нетривиальных функций. Возможно применение моделей Крайгинга для задач с большой (более 15) размерностью параметрического пространства, но стоит учитывать квадратичное увеличение сложности построения модели Крайгинга в зависимости от числа точек в S.

Недостатками методов этого класса применительно к задаче размещения баз скорой помощи являются:

- значительный рост сложности построения метамоделей с ростом размерности задачи;
- отсутствие встроенных средств учета стохастичности.

Эвристические методы

Эвристические методы доказали свою эффективность во множестве практических задач. При решении задач оптимизации имитационных моделей наиболее часто применяются следующие из них [5]:

- генетический алгоритм;
- метод штрафных функций (tabu search);
- алгоритм имитации отжига (simulated annealing).

Генетический алгоритм является итерационным алгоритмом оптимизации, осуществляющий поиск решения путём случайного подбора, комбинирования и вариации искомых параметров с использованием механизмов, напоминающих биологическую эволюцию.

Метод штрафных функций основан на методе локального поиска, который на каждой итерации перемещается к соседнему решению до тех пор, пока не будут удовлетворены определенные критерии останова. Для предотвращения зацикливания (возврата к уже проверенным решениям) сохраняется список запрещенных шагов. Данный список обновляется на каждом шаге.

Алгоритм имитации отжига выбирает одну из соседних точек из списка случайным образом. Если значение в данной точке лучше предыдущего, то переход происходит с вероятностью 1, в противном случае – с вероятностью, строго меньшей 1, которая определяется разницей значений функции в текущей и рассматриваемой точках. Для сходимости алгоритма необходимо, чтобы вероятность перехода в точку с худшим значением функции уменьшалась с увеличением числа итераций.

Каждый из указанных трех алгоритмов относится к алгоритмам поиска глобального оптимума, т.к. все они способны выходить из областей локального минимума.

Преимуществом методов этого класса является отсутствие ограничений по размерности задачи. *Недостатком* является отсутствие встроенных средств учета стохастичности.

Вероятностные методы

Вероятностные методы относятся к методам, использующим производные, и обычно осуществляют итеративный поиск решения со смещением в направлении градиента:

$$x_{k+1} = x_k + a_k \nabla \tilde{f}(x_k),$$

где a_k – длина шага.

Используемые методы:

- анализ возмущений;
- метод отношения правдоподобия.

Различие данных методов состоит в способе вычисления градиента. Применение подобных методов в стохастических моделях является неэффективным в связи с отсутствием возможности вычисления градиента с высокой достоверностью, особенно в областях, близких к минимуму, где значение градиента стремится к 0.

Методы прямого поиска

Методы данного класса не используют градиент, что позволяет избежать проблем в области, близкой к локальному решению.

Можно выделить две группы методов:

- методы, основанные на метамоделях;
- методы, основанные на геометрических шаблонах.

Первая группа методов составляет последовательность локальных метамоделей, аппроксимирующих целевую функцию, и алгоритм переходит к следующему шагу, используя предсказание метамодели. Альтернативой служит метод, основанный на геометрических шаблонах, который напрямую вычисляет значения целевой функции в точках, заранее определенных геометрическим шаблоном.

Гибридные методы

Для выполнения требований 1 и 2 к методу оптимизации он должен состоять из 2-х этапов. На первом этапе (т.н. этапе исследования) производится поиск в области допустимых значений областей, потенциально содержащих минимум функции. На следующем этапе проводится поиск минимумов в отобранных на первом этапе участках, результаты из областей сравниваются для определения глобального минимума. Сравнение может быть реализовано путем применения методов ранжирования и выбора [6].

Помимо рассмотренных требований к методу оптимизации в случае оптимизации стохастических моделей выбор двухфазного метода позволяет нивелировать составляющие ошибки выходного значения. На этапе построения модели ошибка в получаемом выходном значении состоит из двух частей: случайной ошибки и ошибки модели [5]. Случайная ошибка появляется непосредственно из случайных величин, влияющих на выходное значение. Ошибка модели появляется из-за некорректного выбора функциональных форм для описания данных (например, аппроксимация кубической поверхности с использованием квадратичной функции приведет к большому расхождению, если границы интересующей области определены нечетко). В глобальном масштабе модельная ошибка превосходит случайную, но в то же время для локальных областей уменьшение составляющей случайной ошибки становится более критичным. Метод двухэтапной оптимизации позволяет снизить влияние ключевой на данном этапе составляющей ошибки.

На первом этапе гибридного метода могут быть использованы различные алгоритмы. Основной целью данного этапа является поиск областей, потенциально содержащих глобальный минимум в интересующем пространстве решений. В случае использования на первом этапе методов оптимизации, к ним предъявляется требование сходимости к глобальному минимуму. Однако, возможно использование упрощенных алгоритмов исследования, которые лишь определяют наличие в рассматриваемой области потенциального минимума (алгоритмы классификации, поиск разбиением). На втором этапе гибридного метода используются алгоритмы поиска локального минимума.

Переход от этапа к этапу включает в себя определение границ областей потенциального минимума (области не должны пересекаться), формирование множества начальных точек поиска. Эти данные подаются на вход второго этапа оптимизации, на котором ввиду решения нескольких независимых однотипных задач вычисления могут осуществляться параллельно.

Таким образом, гибридные методы имеют преимущество перед другими методами для рассматриваемой модели по причине наличия у нее свойства стохастичности.

Заключение

Проведенная классификация методов оптимизации моделей стохастических систем с непрерывной областью определения параметров позволила выделить преимущества и недостатки различных методов применительно к задаче размещения баз машин скорой помощи. На основе анализа для решения поставленной задачи выбран гибридный метод, позволяющий учесть стохастичность системы и использовать достоинства метода локальной оптимизации после определения областей потенциального минимума на первом этапе работы метода.

Для первого этапа выбран алгоритм разбиения прямоугольниками, его достоинствами являются исследование всего пространства параметров на предмет наличия потенциального минимума, а также отсутствие необходимости обучения. Для второго этапа выбран алгоритм Нелдера-Мида (представитель группы методов поиска по шаблону), не использующий критичное для стохастических систем вычисление градиента и не требующий построения метамоделей, а также успешно зарекомендовавший себя в математических пакетах.

Список литературы

1. Fu M. C. A tutorial review of techniques for simulation optimization // Winter Simulation Conference. 1994. С. 149–156.
2. Andradottir S. A review of Simulation Optimization Techniques // Winter Simulation Conference. 1998. С. 151–158.
3. Buchholz . P. Optimization of Stochastic Discrete Event Simulation Models // Dagstuhl, Germany: Models and Algorithms for Optimization in Logistics. 2009.
4. Law A. M., Kelton W. D. Simulation modeling and analysis. New York City, USA: Wiley, 2000. 972 с.
5. Deng G. Simulation–based optimization. дис. . . . Doctor of Philosophy. Wisconsin, 2007. 229 с.
6. Swisher J., Jacobson S. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey // ACM Trans. Model. Comput. Simul. 2003. №. 13. С. 134–154.
7. Santner T. J, Wlliams D. J., Notz W. I. The design and analysis of computer experiments. New York: Springer, 2003. 548 с.
8. Hong L., Nelson B. Speeding up COMPASS for high-dimensional discrete optimization via simulation // Oper. Res. Lett. 2010. №. 6. С. 550–555.
9. Ranking and selection for steady state simulation / Goldsman D., Marshall W., Seong-Hee K., Nelson B. // Ranking Winter Simulation Conference. 2000. С. 544–553.