УДК 519.6, 519.8

В.А. ПЕПЕЛЯЕВ, Ю.М.ЧЕРНЫЙ

О ВОЗМОЖНОСТЯХ ПРИМЕНЕНИЯ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ В ОПТИМИЗАЦИОННО-ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТАХ

Рассматриваются потенциальные возможности использования генетического алгоритма для глобальной оптимизации произвольных мультимодальных функций, зависящих от многих переменных, заданных на множествах дискретных и/или непрерывных значений.

Введение. При исследовании сложных систем с использованием методов и средств имитационного моделирования необходимо иметь в наличии инструментарий, позволяющий автоматизировать процесс поиска оптимальных проектов и решений [1]. Критерий (критерии) оптимальности, согласно которому сравниваются исследуемые проекты, задается исследователем – лицом, принимающим решения (ЛПР). Такой критерий принято называть целевой функцией (ЦФ) или функцией цели.

ЦФ зависит от значений входных параметров (или *факторов*) модели, которые ЛПР задает перед началом сеанса имитации, и от значений ограниченного множества выходных результатов моделирования (откликов модели на набор входных параметров). Входные переменные определяют конфигурацию сложной системы, влияют на характер функционирования модели в целом и ее отдельных частей и объектов, и, следовательно, отклики модели также зависят от входных параметров. Однако явную функциональную зависимость значений выходных данных от значений входных переменных в задачах моделирования задать невозможно из-за сложных связей и зависимостей между составными частями сложной системы. Также выход-ные данные, как правило, являются случайными величинами, так как при моделировании учитывается возможное влияние на реальную систему внешних случайных воз-действий.

Множество допустимых значений для каждого фактора определяется структурой и назначением моделируемой системы и задается ЛПР. Среди факторов могут присутствовать параметры, задающие как количественные характеристики системы, так и качественные (условия обмена информацией между частями и объектами модели, приоритеты, переключатели режимов и т. п.). Область определения допустимых значений факторов может быть задана как конечным множеством дискретных значений, так и на ограниченном непрерывном отрезке.

Неявная зависимость ЦФ в моделировании от входных переменных, наличие одновременно дискретных и непрерывных переменных существенно ограничивают выбор методов оптимизации при проведении имитационно-оптимизационных экспериментов. Необходимо также учитывать, что ЦФ при исследовании сложных систем являются мультимодальными, т. е. имеют более одного глобального и множество локальных экстремумов. Поэтому использование градиентных, субградиентных и подобных им методов в рассматриваемых задачах практически невозможно. Одним из возможиых вариантов является применение различных метаэвристических подходов [1, 2], которые используют в работе предлагаемых алгоритмов информацию только о типе переменных (непрерывный или дискретный), значениях переменных, множестве допустимых значений для каждой переменной и значении ЦФ, которое достигается при заданном входном векторе-решении. К настоящему времени предложено и исследовано значительное количество разнообразных эвристических методов оптимизации, одним из которых является генетический алгоритм (ГА) [3, 4].

В предлагаемой работе представлен анализ различных аспектов применения ГА в качестве метода безусловной оптимизации мультимодальных функций многих переменных.

Задача оптимизации. Пусть имеется набор (вектор) переменных $\{ x_1, \ldots, x_i, \ldots, x_n \}$, $i = 1,n$. Каждая переменная может принимать значения из своего множества определения D_i. Множества допустимых значений могут быть заданы одним из трех способов [5]:

1) непрерывным отрезком $[a_i, b_i] \in R^1$, $a_i < b_i$;

2) конечным отрезком $[a_i, b_i] \in R^1$ и шагом $s_i < b_i - a_i$ для выбора дискретных значений из отрезка;

3) конечным множеством, состоящим из k целых или действительных чисел $\{a_1, \ldots, a_i, \ldots, a_k\}$.

На множестве переменных неявно задана некоторая ЦФ $F(\vec{x})$, которая для вектора значений $\vec{x} = (x_1,\ldots,x_n)$ выдает число – значение функции. Задача глобальной оптимизации, в частности, минимизации, заключается в поиске такого (таких) вектора значений $\vec{x}^*$, для которого:

$$F(\vec{x}) \geq F(\vec{x}^*) = F^*, \forall \vec{x} \in D = \otimes D_i, i = 1,n.$$

Генетический алгоритм использует в своей работе заимствованную из природы живых организмов концепцию эволюции популяции решений [3]. Каждое решение представляется в виде вектора значений переменных-факторов, прямо или косвенно (через отклики модели) влияющих на значение ЦФ, которое используется в качестве оценки степени «приспособленности» отдельной особи-решения, что влияет на «выживаемость» особи в процессе эволюции. В задачах минимизации более приспособленной считается особь с меньшим значением ЦФ, а в задачах максимизации – с большим значением. Поэтому при применении ГА оптимизации ЦФ принято называть *функцией приспособленности* (*fitness function* – ФП).

Вектор входных переменных задачи принято называть *хромосомой* или *особью*, поскольку в работе алгоритм использует одновременно не одно решение, а целую *популяцию* из *M* векторов. Факторы, которые формируют вектор решения, называются *генами*, а набор их текущих (на очередной итерации) значений – *генотипом*. Текущее значение ФП особи также называют *фенотипом*.

Работа ГА заключается в искусственной эволюции популяции из *M* особей-решений. Начальная популяция может быть задана ЛПР или генерируется алгоритмом случайным образом из множества допустимых значений каждого гена. Для начальной популяции и на очередной итерации алгоритма производится вычисление функций приспособленности особей-решений из нового поколения. Эволюция популяции и генерация новых поколений осуществляется с помощью специальных операторов – скрещивания (crossover, иногда – crossingover), мутации (mutation) и отбора (селекции – selection).

Оператор отбора, используя значения ФП и выбранную заранее стратегию, на очередной итерации алгоритма определяет, какие из особей-решений «выживут», т. е. останутся в популяции на следующую итерацию, а остальные решения заменяются новым поколением. Особи в новое поколение формируются или путем скрещивания двух или более решений (родителей) из текущей популяции с возможной мутацией новой особи, или непосредственной мутацией одного родителя без скрещивания.

Операция скрещивание заключается в том, что новая особь-потомок случайным образом получает от каждого из своих родителей значения части генотипа, наследуя информацию, накопленную от «выживших» на предыдущей итерации особей. Оператор мутации с случайным образом выбирает от одного до нескольких генов в потомке и также случайным образом меняет их текущие значения на другие допустимые для каждого гена значения.

За несколько десятилетий исследования и применения ГА для решения задач разного типа было предложено множество возможных стратегий отбора, скрещивания и мутаций [4].

Базовый принцип, положенный в основу ГА, заключается в репродукции частей генотипа (значений только нескольких переменных) от «лучших» найденных на очередной итерации решений-родителей в новые генерируемые решения-потомки. С помощью мутаций предполагается разнообразить поиск по всей области допустимых решений, не ограничиваясь накопленным в популяции «генофондом». Таким способом алгоритм пытается собрать оптимальное решение как бы «по частям». Одинаковые значения части генотипа у нескольких особей называются шаблоном.

Дж. Холланд [3], один из пионеров в разработке теории ГА, описавший репродуктивный план эволюции решений, сформулировал и доказал теорему шаблонов о возможности сходимости алгоритма путем отбора особей с лучшими ФП и дальнейшего наследования их «генетического материала».

Формат представления и использования данных в оптимизационно-имитационных экспериментах. Теоретически ГА может работать с переменными любого типа – дискретными и непрерывными. На начальном этапе развития теории генетических алгоритмов последние работали, как правило, с битовыми строками, в которые кодировались и декодировались значения переменных других типов. В работе [3] такому преобразованию посвящен целый раздел. В дальнейшем генетическими стали называть и алгоритмы, работающие и непосредственно со значениями переменных.

Однако, при больших размерностях задачи и большом количестве допустимых значений переменных пространство поиска для ГА возрастает экспоненциально, что значительно усложняет и замедляет работу алгоритма. При оптимизации функций от непрерывных переменных решение ищется с некоторой допустимой точностью, что позволяет заменять непрерывный отрезок значений его аналогом с применением некоторого шага дискретизации (квантования).

В задачах моделирования, особенно на начальных этапах исследования систем, допустимые интервалы значений, если позволяет тип данных, разбиваются на конечное количество уровней квантования, каждый из которых может содержать не одно значение, а множество значений. При квантовании данных предполагается, что значения ФП не очень чувствительны к вариации переменной в пределах одного уровня. Таким образом, в оптимизационно-имитационных экспериментах ГА работает в первую очередь с индексами уровней квантования, а лишь затем непосредственно со значениями переменных. В работе [5] рассматриваются структуры шаблонов представления входных данных для оптимизационно-имитационных экспериментов и возможность использования предлагаемых шаблонов при применении ГА в качестве метода оптимизации.

Генерация тестовых функций приспособленности (целевых функций) для сравнительного анализа эффективности применения ГА в задачах моделирования осуществлялась с использованием предложенной в работе [6] методике. ФП $F(\vec{x})$ для задачи глобальной оптимизации строится с помощью m специальных функций $g_i(\vec{x})$ следующим образом:

$$F(\vec{x}) = \min_i g_i(\vec{x}), \tag{1}$$

где

$$g_i(\vec{x}) = H - (H - C_i) * \exp(-u_i * Q_i), \quad u_i > 0, \; C_i < H, \; C_i, H \in R^1, \tag{2}$$

$$Q_i(\vec{x}) = \sum_{j=1}^{n} k_j * | x_j - a_{ij} |^{p_j}, \quad k_j > 0, \; p_j > 0,$$

$$x_j \in D_j,\ a_{ij} \in D_j,\quad i = 1,m,\ j = 1,n. \qquad (3)$$

Каждая из функций $g_i(\vec{x})$ имеет один минимум со значением C_i в точке $\vec{x} = (a_{i1},\ldots,a_{in})$, а функция $F(\vec{x})$ имеет m минимумов, локальных или глобальных в зависимости от значений C_i .

Параметры k_j и p_j позволяют при тестировании алгоритмов изменять чувствительность ЦФ к отдельной переменной, а с помощью коэффициента u_i имеется возможность увеличивать и уменьшать вероятность попадания пробных решений в окрестности точек экстремумов.

Функцию, заданную по формулам (1) – (2), принято называть «полем для гольфа», так как она имеет практически постоянное значение H почти на всей области определения, за исключением точек минимумов функций $g_i(\vec{x})$ и их окрестностей, образующих как бы «лунки» на ровной поверхности. Чтобы усложнить для различных алгоритмов, в том числе и ГА, задачу поиска оптимальных решений, для задания ФП (1) кроме функций типа (2) также возможно использовать периодическую функцию в виде

$$g_0(\vec{x}) = K - (K - L)*\cos(\varphi(\vec{x})),\ \max_i C_i < L < K \le H, \qquad (4)$$

где $\varphi(\vec{x})$ - некоторая функция, зависящая от всех или нескольких входных переменных. Функция (4) имеет множество локальных минимумов со значением L в точках, в которых $\varphi(\vec{x}) = 2\pi s$ ($s = 0, \pm1, \pm2,\ldots$), что позволяет создавать «взрыхленную» поверхность для ФП. Условия, заданные при определении функции $g_0(\vec{x})$ гарантируют, что все значения C_i будут меньше любого значения $g_0(\vec{x})$.

Если исследователя при проведении экспериментов для оценивания эффективности алгоритмов оптимизации не интересует расположение точек минимумов функции (4), то тип функции $\varphi(\vec{x})$ не имеет особого значения. Это может быть линейная, полиномиальная, тригонометрическая или иная функция от переменных задачи, как, например, предлагается в [7].

Организация экспериментов. Исследование возможностей применения ГА в качестве метода оптимизации в задачах имитационного моделирования проводилось на четырех целевых функциях, которые отличались типом входных переменных и имели схожие параметры для конфигурации поверхности отклика (значений ЦФ).

В частности, все функции имели по 50 входных переменных:

– функция-1 F_1: все переменные булевого типа, могли принимать только два значения - 0 или 1;

– функция-2 F_2: все переменные дискретного типа, количество допустимых значений - от 2 до 1000 для разных переменных в одной задаче;

– функция-3 F_3: все переменные заданы как непрерывные с допустимыми интервалами значений (в одной задаче) от 10 до 10000, но с дискретизацией до 0,1 % от ширины интервала значений каждой переменной;

– функция-4 F_4: переменные смешанного типа: присутствуют данные всех типов – от булевых (дискретных с двумя значениями) до непрерывных.

Для всех четырех функций случайным образом на множествах их определения были сгенерированы 10 точек минимумов с известными (контролируемыми) для экспериментатора значениями переменных – 1 глобальный минимум со значением ЦФ равным 400 и 9 локальных минимумов с различными значениями ЦФ в интервале от 450 до 750. Различные значения глобального и локальных минимумов позволяли легко определять по ходу процесса оптимизации какие решения находит алгоритм.

Параметры k и p в функциях (3) подбирались таким образом, чтобы для переменных с малым количеством допустимых значений (до 5) каждое значение было бы относительно «изолировано» от остальных (например, при $p < 1$), а для переменных с большим количеством допустимых значений при $p > 1$ выполнялось бы, например, условие $k^*|x-y|^p < 0{,}03$ (x и y – соседние допустимые значения переменной). Начальные значения параметров u_i для всех функций $g_i(\vec{x})$ были заданы равными, что обеспечивало приблизительно равные вероятности попадания алгоритма в окрестности всех заданных экстремальных точек.

На некоторых этапах экспериментов использовалась также функция $g_0(\vec{x})$ вида (4) с параметром $L = 1000$.

Сложность задачи оптимизации в рассматриваемых примерах можно оценить такими основными параметрами:

– количество всех возможных комбинаций значений переменных (альтернативных вариантов для алгоритма) – от $1{,}12^*10^{15}$ для функции-1 до более, чем 10^{80} (в зависимости от уровня дискретизации) для функции-3;

– максимальное количество значений для одной переменной (с учетом дискретизации непрерывных) – 1001;

- количество переменных – 50.

Общее количество всех возможных комбинаций значений переменных позволяет оценить потенциальное количество возможных альтернатив в случае прямого перебора вариантов, а количество допустимых значений для одной переменной является важной характеристикой для оценки вероятности выбора конкретного значения при мутации гена.

Эксперименты по оптимизации заданных функций проводились для исследования поведения ГА с такими параметрами:

– вероятность скрещивания – 0,8;

– тип скрещивания – равномерное скрещивание;

– вероятность мутации после скрещивания – 0,2;

– максимальное количество одновременно мутирующих генов – до 10 (20 % от общего количества);

– размер популяции – 100 (для функций $F_{1,2,3,4}$) или 300 (для функций $F_{3,4}$);

– часть нового поколения в популяции – 50 % от размера популяции;

– рассматривались две стратегии отбора: элитный или равномерный отбор.

Равномерный тип скрещивания был выбран потому, что все переменные рассматривались как независимые одна от другой, и их индексация (порядок расположения в векторе-решении) осуществлялась произвольным образом.

В качестве стратегий отбора исследовались:

1) элитный отбор – особи с наилучшими значениями ФП принимают участие в создании нового поколения и остаются на следующий этап эволюции, а особи с худшими значениями ФП удаляются из популяции, не оставляя потомков и освобождая место для особей нового поколения;

2) равномерный отбор – из популяции удаляются «близкие» по генотипу особи независимо от значений ФП, но перед удалением все особи популяции с одинаковой вероятностью могут принять участие в создании нового поколения. «Близость» особей в алгоритме определялась по одной из двух метрик – расстоянию Хэмминга или расстояние городских кварталов (манхэттенское расстояние) [8]. Если одна из метрик имела значение меньше 5 (10 % от длины хромосомы), то особи считались «близкими».

Проведение экспериментов для анализа поведения ГА на задачах разной сложности осуществлялось в три этапа.

Этап 1. Все четыре целевые функции $F_{1,2,3,4}$ имели вид «поля для гольфа» с ровной поверхностью отклика и в одном варианте экспериментов с тремя известными точками экстремума (один глобальный), а затем в другом варианте – с десятью точками экстремума. При этом коэффициенты u_i при построении каждой из ЦФ имели равные значения, и, следовательно, точки экстремумов имели примерно равные вероятности нахождения их алгоритмом.

Этап 2. Коэффициент u изменялся (уменьшался) у одной или нескольких точек, чтобы увеличить вероятность попадания алгоритма в окрестность данной точки.

Этап 3. При построении по формулам (1) – (3) функций $F_{1,2,3,4}$ к известным контролируемым экстремальным точкам добавлялась функция вида (4), что коренным образом изменяло ландшафт поверхности отклика ЦФ.

На всех этапах отдельно исследовалось применение обеих стратегий отбора.

Для обнаружения устойчивых тенденций и характерных свойств в работе ГА для каждой конфигурации ЦФ $F_{1,2,3,4}$ или изменении параметров настройки ГА (стратегии отбора и размера популяции) выполнялось по 80 –120 запусков алгоритма для набора минимальных статистических данных.

При проведении экспериментов не ставилась задача обязательно найти с помощью ГА глобальный минимум ЦФ. Основной целью было получить информацию об эффективности поиска и обнаружения новых лучших решений при различной геометрии поверхности отклика ЦФ и разных размерах областей допустимых решений. Поэтому в качестве критерия остановки работы алгоритма задавалось предельное значение рассмотренных алгоритмом в ходе эволюции вариантов (количество вычислений ЦФ), которое не превышало 300000, а в некоторых экспериментах было менее 50000. Выбор таких порядков чисел для количества рассмотренных вариантов был обусловлен тем обстоятельством, что в имитационных экспериментах для получения значения ЦФ для одного варианта решения требуется выполнять прогон модели, что часто требует значительных вычислительных и временн*ы*х затрат. Поэтому алгоритмы, требующие для нахождения оптимального или хотя бы приближения к оптимальному решению более 10^4 – $5*10^5$ вычислений ЦФ независимо от сложности решаемой задачи, не могут рассматриваться для использования в оптимизационно-имитационных экспериментах.

Результаты экспериментов и их анализ. На этапе 1, как и ожидалось, для всех функций алгоритм с примерно равной вероятностью обнаруживал одну из экстремальных точек. Скорость нахождения такой точки существенно зависела от типа данных, на которых была задана тестовая ЦФ, т. е. от общего количества допустимых вариантов. Если обнаруженная точка оказывалась локальным экстремумом, то алгоритму в дальнейшей работе до момента останова практически не удавалось перейти к глобальному минимуму или к лучшему локальному. Алгоритм продолжал порождать похожие по генотипу решения, которые не могли быть существенно изменены с помощью оператора мутации, который в проводимых экспериментах мог изменить не более 20 % генов в особи.

На этапе 2 алгоритм чаще находил ту точку, для которой повышалась вероятность ее обнаружения. При этом участились случаи перехода в эту точку из других локальных экстремумов с худшими значениями ЦФ. Однако, если для повышения вероятности обнаружения выбиралась точка локального экстремума, то переход из нее к глобальному экстремуму оставался практически неосуществимым, если вероятность обнаружения глобального экстремума была ниже вероятности попадания в этот локальный минимум.

Добавление на этапе 3 функции вида (4) существенно осложнило и замедлило работу алгоритма даже для относительно простой задачи оптимизации ЦФ F_1. Алгоритм подолгу «застревал» в локальных минимумах функции (4), особенно, если вероятности обнаружить контролируемые экстремальные точки были относительно малыми.

Проведенные эксперименты показали низкую эффективность ГА при глобальном поиске новых вариантов решений в задачах оптимизации произвольных функций. Но при локальном поиске среди решений с близкими генотипами и значительно ограниченном количестве доступных значений для мутаций ГА показал хорошую производительность и точность нахождения решений. Для повышения эффективности глобального поиска при использовании ГА за многие годы были предложены и продолжают предлагаться различные стратегии повышения популяционного разнообразия [9], однако специфические свойства работы ГА таковы, что применение таких стратегий несколько повышает возможности глобального поиска, но существенно не влияет на его эффективность.

Выводы. Применение генетического алгоритма в качестве основного метода глобальной оптимизации при проведении имитационно-имитационных экспериментов не может рассматриваться как эффективный инструмент глобальной оптимизации. ГА стремится воспроизводить значения переменных из «лучших» на очередном этапе решений при генерации новых поколений, однако такая стратегия не позволяет вести активный поиск по всему пространству решений. ГА может применяться как вспомогательный инструмент уточнения найденных другими алгоритмами глобального поиска «оптимальных» решений.

В.А. Пепеляєв, Ю.М. Чорний

ПРО МОЖЛИВІСТЬ ВИКОРИСТАННЯ ГЕНЕТИЧНИХ АЛГОРИТМІВ В ОПТИМІЗАЦІЙНО-ІМІТАЦІЙНИХ ЕКСПЕРИМЕНТАХ

Розглядаються потенційні можливості використання генетичних алгоритмів для глобальної оптимізації довільних мультимодальних функцій, що залежать від багатьох змінних, визначених на множинах дискретних та/або неперервних значень.

V.A. Pepeliayev, Yu.M. Czornyy

THE POSSIBILITIES OF THE USING GENETIC ALGORITHMS IN SIMULATION OPTIMIZATION

The potential possibilities of using a genetic algorithm for the global optimization of arbitrary multimodal functions depending on many variables defined on sets of discrete and/or continuous values are considered.

Список литературы

1. Пепеляев В.А. К вопросу об интеграции методов оптимизации и имитационного моделирования. *Теория оптимальных решений.* К.: Ин-т кибернетики имени В.М. Глушкова НАН Украины. 2003. № 2. С. 51 – 60.
2. Пепеляев В.А., Сахнюк М.А., Черный Ю.М., Шваб Н.Д. К вопросу о реализации мета-эвристических стратегий оптимизации моделирования. *Компьютерная математика*. К.: Ин-т кибернетики имени В.М. Глушкова НАН Украины. 2005. № 2. С. 26–33.
3. Holland J.H. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. *Univ. Michigan Press*. 1975. 226 p.
4. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы. М.: ФИЗМАТЛИТ, 2010. 368 с.
5. Бигдан В.Б., Криковлюк А.А., Пепеляев В.А. Унификация структур входных данных для оптимизационных алгоритмов в имитационных экспериментах. *Компьютерная математика.* К.: Ин-т кибернетики имени В.М. Глушкова НАН Украины. 2017. № 1. С. 45 – 54.
6. Пепеляев В.А., Черный Ю.М. Принципы построения целевых функций для тестирования алгоритмов глобальной оптимизации. *Компьютерная математика.* К.: Ин-т кибернетики имени В.М. Глушкова НАН Украины. 2017. № 2. С. 62 – 71.
7. Rönkkönen J., Li X., Kyrki V., Lampinen J. A generator for multimodal test functions with multiple global optima. *Proceedings of the 7th International Conference on Simulated Evolution and Learning (SEAL '08)*, Berlin, Heidelberg. Springer-Verlag. 2008. P. 239 – 248.
8. Деза Е.И., Деза М.-М. Энциклопедический словарь расстояний. М.: Наука, 2008. 446 с.
9. Gupta D., Ghafir Sh. An Overview of methods maintaining Diversity in Genetic Algorithms. *International Journal of Emerging Technology and Advanced Engineering.* 2012. Vol. 2, Issue 5. P. 56 – 60.

Получено 12.03.2019