

УДК 004.89+004.021

ОПТИМИЗАЦИЯ МНОГОЭКСТРЕМАЛЬНЫХ ФУНКЦИЙ НА ОСНОВЕ КЛАСТЕРНОЙ МОДИФИКАЦИИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

П.В. Казаков¹

В работе представлена модификация стандартного генетического алгоритма на основе использования кластеров решений. Данная версия генетического алгоритма расширяет функциональность простого генетического алгоритма при решении задач оптимизации многоэкстремальных функций. Представлены результаты экспериментов, демонстрирующие возможности предлагаемой модификации генетического алгоритма.

Введение

Стандартный генетический алгоритм (СГА) является высокоэффективным средством оптимизации, но отсутствие в нем эффективных средств поддержания разнообразия в популяции на протяжении всего процесса поиска приводит к ее преждевременному вырождению. Следствием этого является возможность СГА локализовать за один запуск лишь одно экстремальное решение, что существенно снижает эффективность его использования при решении многоэкстремальных задач. Однако, обладая свойством параллелизма, генетический алгоритм (ГА) как универсальный метод оптимизации имеет полноценные возможности к нахождению всех оптимальных решений, распределенных в поисковом пространстве. Такие свойства у ГА появляются при наделении его дополнительными средствами поддержания разнообразия в популяции на протяжении всего генетического поиска. Практическая реализация этого позволит всесторонне исследовать ГА пространство решений, что является ключевым моментом в многоэкстремальной оптимизации, свойственной множеству практических задач из различных прикладных областей.

¹ 241035, Брянск, бул. им. 50-летия Октября, 7, БГТУ, pyk_mail@list.ru

1. Основные подходы к расширению возможностей СГА для решения многоэкстремальных задач оптимизации

Как правило, связаны с интеграцией в СГА средств распределенного поиска, а также выделения и группировки по специальным признакам хромосом в популяции. Анализ литературы [Гладков и др., 2002], [Пупков и др., 2004], [Bäck, 2000], [Chambers, 1995], [Chambers, 1999] позволил выделить три основных подхода к разработке таких генетических алгоритмов:

- имитационный;
- на основе обработки локальных подпопуляций;
- на основе обработки распределенных популяций.

Имитационный подход предполагает многократные (по предварительно подготовленному плану) запуски СГА с варьированием значений управляющих параметров и позволяет получить «хорошие» результаты лишь при малом числе экстремумов и «предсказуемом» пространстве решений.

Второй подход подразумевает выделение в основной популяции подпопуляций хромосом, обладающих определенными свойствами. Как правило, для этого изменяются, реже добавляются новые генетические операторы. Представителями таких алгоритмов являются ГА с логистической моделью [Гладков и др., 2002], дифференциальной эволюции [Пупков и др., 2004], а также алгоритмы, где в оператор отбора добавлено использование так называемых нишевых механизмов [Bäck, 2000].

Последний подход объединяет алгоритмы, в которых используются разделенные популяции, эволюционирующие изолированно, либо взаимодействуя друг с другом, например для передачи хромосом, кодирующих перспективные решения [Chambers, 1999].

Представленные модификации СГА позволяют поддерживать разнообразие в популяции для предотвращения преждевременной сходимости к единственному решению. Однако использование для своего функционирования дополнительных управляющих параметров оказывает существенное влияние на эффективность поиска в силу определенных трудностей с их настройкой.

Данный недостаток устраняется в предлагаемой кластерной модификации генетического алгоритма (КГА). В ней используется единственная подпопуляция, содержащая кластеры хромосом, формируемых по принципу фенотипического отличия друг от друга. Это отличает КГА от других модификаций ГА, в частности кластерно-ориентированного генетического алгоритма [Гладков и др., 2002], где

пространство поиска разбивается с помощью нишевых механизмов на участки в зависимости от границ значений fitness-функции.

2. Особенности организации КГА

Под кластером хромосом понимается группа решений, имеющих похожие свойства, то есть кодирующие их хромосомы с похожим фенотипом. Для определения меры близости могут быть использованы как вещественная (Евклидова), так и бинарная (Хемминга) метрика d . Число кластеров зависит от задаваемого в качестве дополнительного управляющего параметра ГА радиуса гиперсферы кластера R_c . Хромосомы, располагающиеся в пределах R_c до центроида кластера Z_i рассматриваются как похожие и принадлежащие Z_i .

Множество центроидов кластеров Z_c образуют отдельную подпопуляцию хромосом. Для кластеризации элементов популяции используется принцип доминирования. Пусть Z_1, Z_2, \dots, Z_k – k фрагментов популяции P_n , представляющих собой кластеры. Хромосома $C^* \in Z_i$ является доминирующей в кластере, если для $\forall C' \in Z_i : f(C^*) \leq f(C')$. Здесь предполагается, что решается задача минимизации. Знак нестрогого неравенства означает, что в кластере может быть более одного доминирующего решения. Поэтому хромосома C^* является центроидом кластера Z_i тогда и только тогда, если для $\forall C' \in Z_i : d(C^*, C') \leq R_c$. Важно, что отдельные хромосомы, соответствующие этому неравенству могут принадлежать одновременно и другим кластерам, а также доминировать сразу в нескольких кластерах (рис. 1)

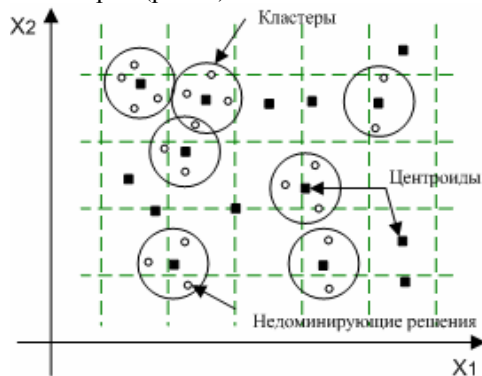


Рис.1. Выделение кластеров для двумерного пространства решений

Подпопуляция найденных центроидов кластеров представляет собой механизм поддержания разнообразия популяции для параллельного исследования всех областей поискового пространства.

3. Принципы функционирования КГА

Расширение функциональных возможностей СГА для локализации множества экстремумов реализуется в КГА с помощью процедур выделения и копирования кластеров. Они определяются в текущей популяции, а затем сохраняются в отдельной подпопуляции. Представляя, в определенном смысле, лучшие среди «непохожих» между собой хромосом, центроиды кластеров копируются в новую популяцию, моделируя этим распределенную стратегию элитизма. Число кластеров может быть произвольным и зависит от свойств задачи или потребностей пользователя КГА.

Процедура выделения кластеров состоит в определении N_z числа кластеров, определяемых координатами центроидов, каждый из которых соответствует хромосоме, которая доминирует над другими хромосомами-членами кластера, находящиеся на расстоянии не превышающем R_c от центроида.

Центроиды найденных кластеров сохраняются в отдельной подпопуляции. Основная популяция подвергается применению генетических операторов и следовательно претерпевает изменения, поэтому найденные кластеры теряются. Для предотвращения этого предварительно сохраненные центроиды по специальному алгоритму копируются в новую популяцию, тем самым восстанавливая «присутствие» ГА в соответствующих участках поискового пространства.

Копирование кластеров в новую популяцию происходит за счет замены «худшей» хромосомы из некоторого кластера его центроидом. При отсутствии хромосом, принадлежащих данному кластеру его центроид заменяет «худшую» хромосому (не принадлежащих не одному из кластеров) новой популяции.

С практическим использованием КГА связан ряд особенностей. В частности, было исследовано влияние двух дополнительных процедур, связанных с определением и копированием кластеров на временную эффективность КГА. Дополнительные вычисления связаны с расчетом меры близости между хромосомами. Общая сложность вычислений варьировалась от линейной до квадратичной (худший случай). Радиус гиперболы кластера играет ключевую роль при поиске решений и как правило определяется экспериментально. Уменьшение значения этого параметра приводит к увеличению числа кластеров и наоборот. Частный случай, когда $R_c \rightarrow 1$, соответствует стратегии элитизма, когда из популяции в популяцию копируется единственный кластер, соответствующий лучшему решению.

После достижения заданного числа поколений подпопуляция кластеров содержит разнообразные решения с различной оптимальностью.

Для выделения их них хромосом, кодирующих оптимальные решения, используется зависимость $|f(Zc_i) - f(C^*)| \leq e$,

где $f(Zc_i)$ - оптимальность i - го центроида кластера;

$f(C^*)$ – оптимальность лучшей хромосомы последней популяции;

$e > 0$ – параметр, определяющий верхнюю границу «глобального» оптимума.

Задавая значение параметра e , можно изменять уровень оптимальности решений от глобального до рационального. Последнее особенно важно для многих практических задач, когда важно обеспечить выбор среди рациональных решений, найденных за короткое время с минимальными затратами на настройку КГА.

4. Практическое применение КГА для решения многоэкстремальных задач оптимизации

Для исследования эффективности КГА решались тестовые задачи мультимодальной оптимизации (табл. 1).

Табл. 1

Функция	Число глобальных оптимумов	Параметры КГА
$f(x_1, x_2) = (4 - 2.1 \cdot x_1^2 + \frac{1}{3} \cdot x_1^4) + x_1 \cdot x_2 + (-4 + 4 \cdot x_2^2) \cdot x_2^2$ $-3 \leq x_1 \leq 3, -2 \leq x_2 \leq 2$	2	$N_p = 100, P_c = 0,9,$ $P_m = 0,03, R_c = 0,03$
$f(x_1, x_2) = \prod_{i=1}^2 \sum_{j=1}^5 j \cos[(j+1)x_i + j]$ $-10 \leq x_i \leq 10$	18	$N_p = 300, P_c = 0,9,$ $P_m = 0,03,$ $R_c = 0,025$

КГА для каждой из функций запускался по несколько раз и для каждого случая он определял все экстремумы. Наряду с КГА для оптимизации применялись другие ГА (СГА – для первой функции и ГА с дифференциальной эволюцией для второй). В сравнении с ними КГА сумел локализовать все экстремумы за минимальное время и с минимальным числом управляющих параметров. На рис. 2, 3 показаны соответственные графики функции и плотности исследования пространства решений КГА для второй тестовой функции.

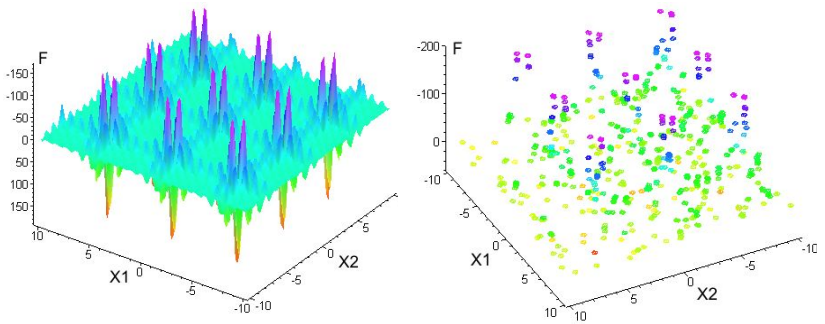


Рис.2. Вторая тестовая функции и плотность ее исследования КГА

Из рисунка видно, что оптимизация выполнялась не за счет хаотичного исследования поискового пространства, а благодаря направленному синтезу решений (кластеров) в наиболее перспективных его участках. На рис. 3 представлен результат работы программы оптимизации функции с 18 экстремумами с помощью КГА.

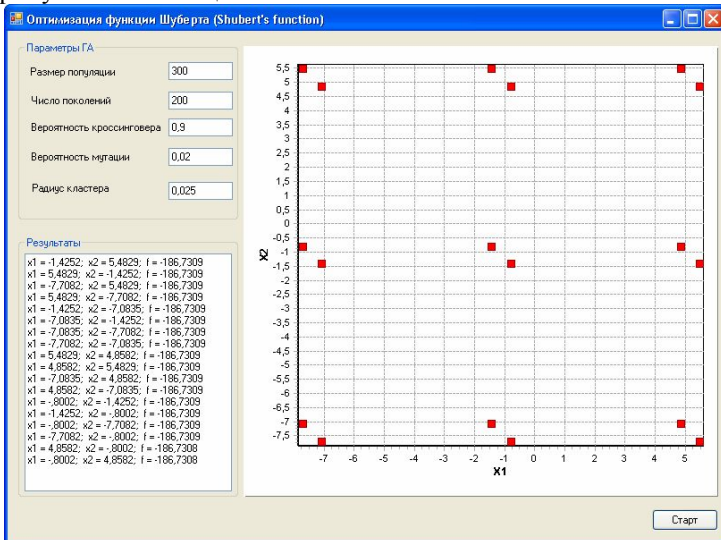


Рис.3. Программная реализация КГА для оптимизации многоэкстремальной функции

Также анализировалось влияние значений параметров размер популяции и радиус кластера на результативность поиска. Было выявлено, что при несложной настройке можно определить такие их значения, при которых КГА находит все оптимумы за минимальное число поколений. Однако такая настройка должна быть индивидуальной для каждой задачи.

В частности, для второй тестовой функции результаты показали, что увеличение Rc приводит к уменьшению числа глобальных оптимумов. Очевидно, что при больших Rc некоторые решения, соответствующие глобальным оптимумам могут размещаться в одном кластере и тем самым теряться. Граничные значения Rc , при котором определялись все 18 глобальных оптимумов за установленное время составило 0,025, при большем значении их число уменьшалось. При $Rc > 0,4$ определялся только один глобальный оптимум, что соответствовало классической стратегии элитизма.

Заключение

Таким образом, использование в КГА принципов кластеризации позволяет всесторонне исследовать пространство решений, благодаря сохранению разнообразия в популяции. Варьирование значения параметра радиус кластера позволяет настраивать КГА на локализацию как группы глобальных, так и множества различных субоптимальных решений. Последнее особенно важно для многих практических задач, когда важно обеспечить выбор среди рациональных решений, найденных за короткое время с минимальными затратами на настройку генетического алгоритма.

Список литературы

- [Гладков и др., 2002] Гладков Л.А. Методы генетического поиска. - Таганрог: Изд-во ТРТУ, 2002.
- [Пупков и др., 2004] Пупков К.А. Алгоритм дифференциальной эволюции для задач технического проектирования // Информационные технологии. 2004. №8.
- [Bäck, 2000] Bäck T, Fogel D., Michalewicz Z. Evolutionary computations 2. Advanced algorithms and operators. - IOP Publishing, 2000.
- [Chambers, 1995] Chambers L. Practical handbook of genetic algorithms : new frontiers. – CRC Press, Inc., 1995.
- [Chambers, 1999] Chambers L. Practical handbook of genetic algorithms : complex coding systems. – CRC Press, Inc., 1999.