

УДК 004.81

ГЕОИНФОРМАТИКА: ТЕОРИЯ БИОИНСПИРИРОВАННОГО ПОИСКА ОПТИМАЛЬНЫХ РЕШЕНИЙ И ЕЕ ПРИЛОЖЕНИЯ ДЛЯ ОБРАБОТКИ ПРОБЛЕМНО-ОРИЕНТИРОВАННЫХ ДАННЫХ И ЗНАНИЙ

Сергей Иванович Родзин, канд. техн. наук, проф.

E-mail: srodzin@yandex.ru

Лада Сергеевна Родзина, аспирантка

E-mail: raisin25@yandex.ru

Южный федеральный университет

http://www.sfedu.ru

В статье обсуждаются основные элементы теории биоинспирированного поиска оптимальных решений. Представлены оригинальные меметические алгоритмы, сочетающие локальный поиск, кооперацию и соревнование. Эксперименты для NP-сложных задач оптимизации показали, что с помощью разработанной теории эти задачи решаются быстро, надежно и точно.

Ключевые слова: геоинформатика, биоинспирированный поиск, меметический алгоритм, обработка проблемно-ориентированных знаний.

Исследование выполнено за счет гранта Российского научного фонда (проект № 14-11-00242) в Южном федеральном университете.

Введение

Многие задачи в геоинформатике сводятся к задачам поиска оптимальных решений. Особенности таких задач являются нелинейность, многоэкстремальность, недифференцируемость, отсутствие аналитического выражения, высокая размерность пространства поиска и т.д. Для эффективного решения указанных задач интенсивно разрабатываются биоинспирированные алгоритмы поиска оптимальных решений. Возрастает интерес исследователей к решению математических проблем геоинформатики, особенно в области разработки алгоритмов и компьютерных программ, использующих простые механизмы из-



С.И. Родзин



Л.С. Родзина

менчивости и отбора по аналогии с эволюцией в природе.

Биоинспирированные алгоритмы весьма разнообразны. Их разновидностями, наряду с генетическими алгоритмами, являются алгоритмы генетического программирования, эволюционных стратегий, эволюционного программирования, обучающие классификаторы, алгоритмы Монте-Карло, роевого интеллекта, меметики, гармоничного поиска и др. Одной из основных проблем, с которой сталкиваются их разработчики, является проблема обеспечения баланса между скоростью сходимости алгоритма и диверсификацией поиска. Быстрая сходимость алгоритма означает снижение разнообразия популяции. Напротив, диверсификация поиска расширяет пространство поиска и увеличивает вероятность локализации глобального экстремума задачи. Это – фундаментальные проблемы, ввиду их теоретической и практической важности. Биоинспирированные алгоритмы являются математическими преобразованиями. Они трансформируют входной поток информации в выходной и основаны на правилах имитации механизмов эволюционного синтеза, а также на статистическом подходе к исследованию ситуаций и итерационном приближении к искомому решению. Неполнота знаний о внешнем мире, непредсказуемость реальных ситуаций заставляют исследователей, проектирующих геоинформационные системы, ис-

пользовать модели, инспирированные природными системами, способные подстраиваться к изменению внешней среды и самостоятельно ориентироваться в сложных условиях. Используя компьютерное моделирование на основе биоинспирированных алгоритмов, можно создать сложные системы, для которых не существует аналитического описания.

Для целенаправленного синтеза биоинспирированных алгоритмов с развитыми механизмами решения проблемы баланса, адаптации и самоадаптации, необходима общая теория биоинспирированного поиска оптимальных решений. В статье именно эта задача является основной.

Постановка задачи поиска оптимальных решений

Поиск оптимальных решений является задачей нахождения экстремума (минимума или максимума) целевой функции в некоторой области конечномерного векторного пространства, ограниченной набором линейных и/или нелинейных равенств и/или неравенств. Одним из наиболее трудных классов задач поиска оптимальных решений являются задачи комбинаторной оптимизации. Постановка этого класса задач состоит в следующем.

Пусть задан отрезок натурального ряда $M = \{1, 2, \dots, m\}$; D – дискретное пространство, в котором все точки изолированы друг от друга в некотором смысле; гомоморфизм $f: M \rightarrow D$, представляющий собой такое соотношение между M и D , что: а) каждому элементу M соответствуют один элемент D (но не наоборот); б) когда для ряда элементов M выполняется некоторое отношение, то и для соответствующих элементов D выполняется соответствующее отношение, причем f удовлетворяет некоторой системе ограничений Z . Тройка $\langle f, D, Z \rangle$ является комбинаторным объектом. Тогда задача поиска оптимального решения заключается в поиске такого решения $x_0 \in X$, что

$$x_0 = \arg \min_{x \in X_{\text{доп}} \subseteq X} F(x),$$

где X – пространство решений задачи, элементами которого являются комбинаторные объекты;
 $D \subseteq X$ – подпространство допустимых решений, определяемое некоторым предикатом;
 $F: X \rightarrow R$ – целевая функция задачи (R – множество вещественных чисел).

Сложность решения задач данного класса связана с тем, что многие из них относятся к переборным или NP-сложным задачам. Для них отсутствуют эффективные точные методы решения. Проблема усугубляется дополнительными обстоятельствами. Во-первых, многие практические задачи являются многокритериальными, что существенно усложняет поиск оптимального решения, т.к. требуется определить некоторый компромиссный результат, который в общем случае не является оптимальным ни по одному критерию. Во-вторых, необходимо принимать во внимание нелинейный характер ряда параметров решаемых задач. В-третьих, нестационарность задач, трудность идентификации основных параметров приводят к необходимости поиска решений на основе приближительных и «нечетких» исходных данных.

Полный перебор (например, поиск в ширину или в глубину) в задачах комбинаторной оптимизации пригоден только для самых простейших задач. Для практических задач полный перебор нереален. К тому же исходные данные в задаче часто имеют определенные погрешности, что окончательно делает значительные вычислительные затраты для нахождения точного решения неоправданными.

Развитие методов поиска оптимальных решений вышло за рамки первичного накопления знаний. Ведется анализ и разработка стохастических и эвристических поисковых процедур, позволяющих достичь повышения эффективности алгоритмов оптимизации. К таким процедурам относятся биоинспирированные методы поиска оптимальных решений. Разновидностями биоинспирированных методов являются эволюционные алгоритмы (*genetic algorithms* [1], *genetic programming* [2], *evolution strategies* [3], *evolutionary programming* [4]), *algorithms swarm intelligence (ant colony* [5], *swarm of bees* [6], *flock of birds* [7], *school of fish* [8], *swarm of salmonella bacterium* [9], *behavior of cuckoos during the forced*

nest parasitism [10], *glowworm swarm* [11], *weed farmland colonization* [12], *behaviour of frog groups during the food searching process* [13], *swarm of flies* [14], *behaviour of frog groups during the food searching process* [15], *flock of bats* [16]) и др. Необходима теория, обобщающая существующие биоинспирированные алгоритмы поиска оптимальных решений.

Теория биоинспирированного поиска

Теория биоинспирированного поиска оптимальных решений основывается на следующих гипотезах и закономерностях.

Алгоритмы биоинспирированного поиска представляют собой не просто случайный поиск, а последовательное преобразование множества решений с использованием информации, накопленной в процессе эволюции. Они характеризуются существенно меньшей зависимостью от особенностей приложения, являются более универсальными и обеспечивают лучшую степень приближения к оптимальному решению. Алгоритмы биоинспирированного поиска применимы к решению задач, у которых фазовое пространство переменных не обязательно является метрическим. Применение алгоритмов биоинспирированного поиска эффективно в задачах, где необходимо получить всю историю поведения системы, ее эволюцию в целом. Ключевым элементом формализованных моделей биоинспирированного поиска является построение начальной модели, правил, по которым она эволюционирует, а также подходов к представлению (кодированию) решений. Модели биоинспирированного поиска могут быть лингвистическими и не иметь количественного выражения. Алгоритмы биоинспирированного поиска моделируют процесс поиска оптимальных решений посредством эволюционных операторов селекции, репродукции, мутации и др., поддерживают популяцию структур, которая эволюционирует в окружающей среде. Селекция фокусирует внимание на отборе индивидуумов с более высокими значениями целевой функции, а репродукция, мутация и другие операторы генерируют новых индивидуумов. Все алгоритмы биоинспирированного поиска представляют собой итерационные эвристические процедуры, не имеют ограничений на вид целевой функции. Популяция в алгоритмах биоинспирированного поиска играет роль памяти о структуре пространства поиска решений. Память не обязательно ограничивается лишь лучшими решениями. Различия в алгоритмах биоинспирированного поиска не носят методологический характер и не затрагивают фундаментальные принципы, присущие эволюции независимо от формы и уровня абстракции используемой модели.

Основные положения теории биоинспирированного поиска оптимальных решений представляют собой следующую систему идей. Моделирование эволюции популяции происходит посредством процессов циклического вычисления и оценки функции качества решений с последующей их селекцией и репродукцией. Модели биоинспирированного поиска поддерживают популяции структур, которые эволюционируют в соответствии с правилами отбора и операторами поиска (рекомбинация, мутация). Эволюционировать могут только популяции, которые содержат в себе процессы репродукции и подвержены изменчивости. Процессы репродукции и отбора присущи самоорганизующимся системам. Например, эти процессы автоматически действуют в социальных системах: люди, машины, технологии, системы управления всегда сравниваются и отбираются в этих системах. Невозможно найти хотя бы один факт, когда эволюционировали бы системы, которые не имеют этого алгоритма. Изменчивость элементов популяции не одинакова.

Суть теории биоинспирированного поиска сводится к следующему.

Фиксируется множество объектов X (популяция решений). Среди объектов необходимо выбрать наилучшие по критерию оптимальности F . Критерий оптимальности формируется на основе свойств объектов, представляет отображение вида $F: X \rightarrow R$, которое каждому объекту $x \in X$ из множества X сопоставляет значение $F(x)$.

Фенотипическая природа множества объектов произвольна, поэтому строится кодированное представление множества объектов в конечном векторном множестве S (генотип). Отображение вида $\varphi: X \rightarrow S$ описывает связь между исследуемыми объектами, кото-

рые являются решениями задачи поиска, и объектами, управление и манипулирование которыми осуществляет алгоритм биоинспирированного поиска.

Существует обратное отображение вида $\varphi^{-1}: S \rightarrow X$, где каждому элементу популяции $s \in S$ соответствует элемент во множестве X . Тогда процесс оптимизации с помощью алгоритма биоинспирированного поиска состоит в построении множества объектов-решений $X_{opt} \in X$, для которых выполняются следующие условия:

$$X_{opt} = \operatorname{argmax} F[\varphi^{-1}(s)], s \in S.$$

В процессе оптимизации множество X эволюционирует к оптимальному состоянию, изменяя свой состав и параметры входящих в него объектов. Способ построения множества объектов $s \in S$ определяется алгоритмом биоинспирированного поиска. Эволюция множества X определяется эволюцией популяции S . На множестве S определяется подмножество P_0 – случайная начальная популяция. Решение на каждом шаге эволюции определяется разностной вычислительной схемой $P_{t+1} = A(P_t)$, где A – композиция различных эволюционных операторов. Критерий оптимальности вычисляется на каждом шаге в процессе отбора решений по критерию, реализуемому в композиции операторов эволюции A .

Другая идея теории относится к реализации системы биоинспирированного поиска. Она состоит в следующем. Алгоритмы биоинспирированного поиска работают не в исходном множестве решений, а в его представлении в кодовом пространстве. Замысел состоит в том, что модель и алгоритм биоинспирированного поиска *разделены*, а система биоинспирированного поиска представляет собой систему с обратной связью.

Структура системы биоинспирированного поиска оптимальных решений приведена на рисунке 1.

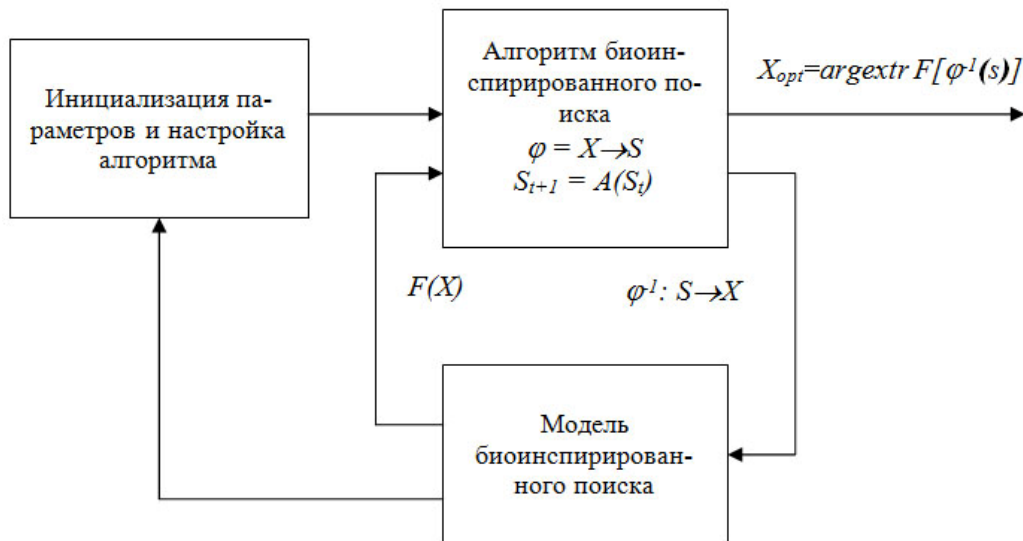


Рисунок 1 – Структура системы биоинспирированного поиска оптимальных решений

Модель биоинспирированного поиска получает от алгоритма очередной набор значений параметров, характеризующих решения $X = (x_1, x_2, \dots, x_n)$ и выдает соответствующее значение функции $F(X)$. Данное значение используется биоинспирированным алгоритмом при отборе и формировании новых решений. Процесс останавливается, когда текущий набор значений решений удовлетворяет заданному критерию, то есть найдено оптимальное решение $X_{opt} = (x_{1opt}, x_{2opt}, \dots, x_{nopt})$.

Структура системы биоинспирированного поиска является адаптивной за счет использования композиции различных эволюционных операторов A в разностной вычислительной схеме $P_{t+1} = A(P_t)$.

С математической точки зрения процесс эволюции представляет собой движение по ландшафту целевой функции в направлении поиска ее экстремума.

Пусть x_j – переменные, описывающие состояние природной системы и ее реакцию на внешние воздействия M_k . Обозначим через b_m параметры структуры, определяющие функционирование природной системы. Тогда ее функционирование определяется системой из уравнений вида

$$U(x_j, M_k, b_m) = 0.$$

При заданных значениях M_k и b_m можно определить состояние системы и ее реакцию на внешние воздействия. В процессе эволюции производится отбор наиболее приспособленных особей. Этому препятствуют ограничения – реальные параметры всегда ограничены по величине. Процесс эволюции моделирует поиск экстремума некоторой функции при наличии ограничений. В качестве аналога процесса эволюции можно использовать функцию следующего вида:

$$Y(X, \Lambda) = F(x_j) + \sum_i \lambda_i U_i(x_j, M_k, b_m) .$$

Здесь U_i – ограничения в задаче, взятые с коэффициентами λ_i , $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$, $i = 1, 2, \dots, n$, где n – число ограничений.

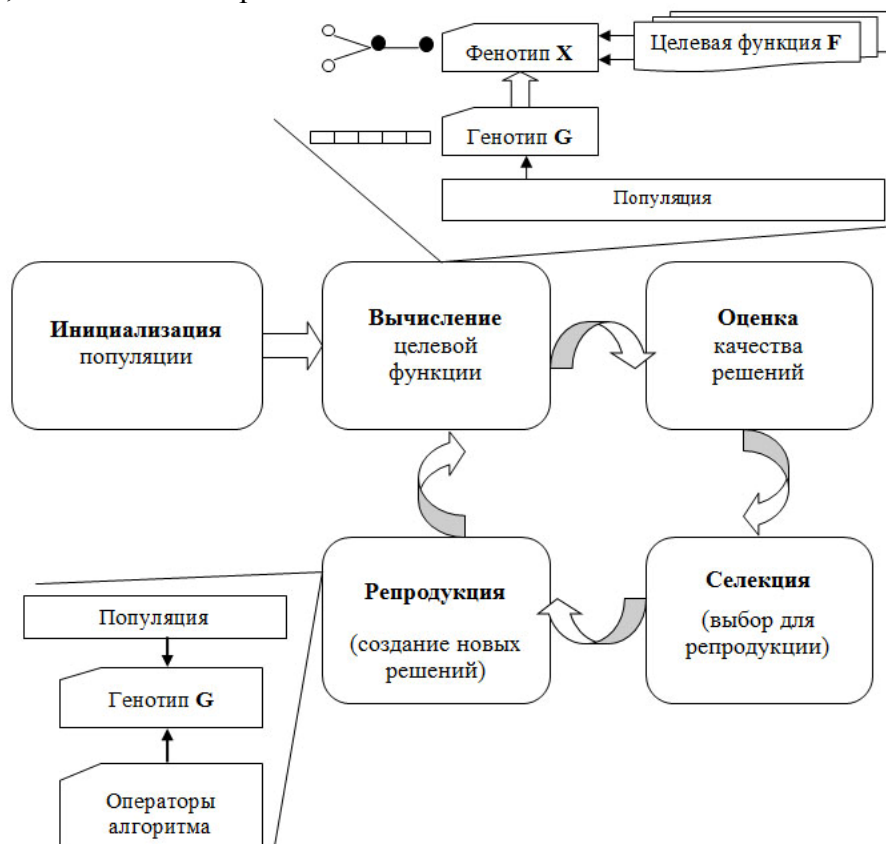


Рисунок 2 – Базовый цикл в эволюционном алгоритме

Для алгоритмов биоинспирированного поиска справедлива *No Free Lunch (NFL)* теорема Вольперта–Макрида [17]. Согласно этой теореме все эвристические алгоритмы «в среднем» являются идентичными, т.е. их производительность одинакова, если усреднить результаты по всевозможным целевым функциям. Теорема устанавливает, что в любом возможном ландшафте целевой функции лучшей стратегией будет хаотичное движение. Чтобы выработать более перспективную стратегию поиска оптимума, необходимо иметь некоторую информацию о характере исследуемого ландшафта. Эвристический алгоритм, хорошо зарекомендовавший себя для некоторого класса задач, может потерпеть неудачу для другого класса задач с другими целевыми функциями. Согласно следствию теоремы, единственным способом сокращения времени и повышения качества биоинспирированного поиска является специализация схемы поиска. Это выражается в использовании проблемно-ориентированных знаний о задаче. Эти знания влияют на эволюцию особей в популяции, на

настройку параметров алгоритма биоинспирированного поиска. Покажем это на примерах конкретных алгоритмов и задач.

Следствием общей теории биоинспирированного поиска оптимальных решений является наличие базового цикла в эволюционном алгоритме, графическое изображение которого приводится на рисунке 2.

Из рисунка 2 видно, что базовый цикл включает следующую последовательность шагов: вычисление целевой функции, оценку качества решений, селективный отбор решений для репродукции и репродукцию, т.е. создание новых решений. Базовый цикл соответствует основным положениям общей теории биоинспирированного поиска оптимальных решений, и является неотъемлемым элементом структуры системы биоинспирированного поиска.

Алгоритм организации вычислений в базовом цикле имеет следующий вид:

Input: Функция для оценки качества решений cmp_F

Input: ps – размер популяции решений

Data: t – текущий номер поколения

Data: $P(t=0)$ – исходная популяция решений

Data: параметры алгоритма, включая целевую функцию

Output: X^* – найденное оптимизированное решение

begin

$t := 0$

Pop: = *init Pop(ps)* /*функция *init* выполняет первоначальную случайную инициализацию популяции */

while (критерий останова) *do*

$v := F(Pop, cmp_F)$

$P(t) := selection(Pop, v)$

$t := t + 1$

Pop: = $P(t)$ /*репродукция потомков из отобранных родительских решений с использованием композиции операторов эволюции $A: P(t+1) = A(P(t))$ */

return /*восстановление фенотипа Pop */

end

Базовый цикл не исключает существования различных стратегий организации популяции в эволюционном алгоритме. Например, стратегия вымирания предполагает, что популяция на следующем шаге эволюции состоит только из потомков предыдущей популяции, в которых очередная популяция не содержит родительских хромосом, а включает лишь хромосомы сгенерированных потомков.

Другой стратегией является способ, когда популяция на следующем шаге эволюции $P(t+1)$ формируется как комбинация текущей популяции $P(t)$ и образованных на ее основе потомков. В этом случае продолжительность жизни отдельных особей в популяции может превышать одно поколение, следовательно, родители и их потомки конкурируют друг с другом за выживание. Например, в такой разновидности эволюционных алгоритмов, как эволюционные стратегии для описания перехода от одного поколения к другому используются следующие обозначения: λ – число потомков, μ – число родителей. Тогда запись вида (μ, λ) при $\lambda \geq \mu$ означает, что из созданных λ потомков от μ родителей $(\lambda - \mu)$ худших потомков будут исключены из следующего поколения. Стратегия $(\lambda + \mu)$ при $\lambda > \mu$ будет обозначать, что в следующее поколение будут отобраны по определенной стратегии μ особей из родителей и их потомков.

Еще одной стратегией организации популяции в эволюционном алгоритме является элитизм: в следующем поколении обеспечивается сохранение, по крайней мере, одного лучшего решения из текущего поколения. Следствием этой стратегии является то, что если найден глобальный оптимум, то эволюционный алгоритм гарантирует сходимость процесса, хотя возрастает риск попадания в «локальную яму».

Алгоритм организации элитизма в процессе эволюционных вычислений, представленный на псевдокоде, имеет следующий вид:

Input: Функция для оценки качества решений cmp_F , представляющая собой компаратор, сравнивающий качество решений

Input: ps – размер популяции решений

Input: as – размер архива элитных решений

Data: t – текущий номер поколения

Data: $P(t=0)$ – исходная популяция решений

Data: Arc – архив элитных решений

Data: параметры алгоритма, включая целевую функцию

Output: X^* – найденное оптимизированное решение

begin

$t := 0$

$Arc := \emptyset$

$Pop := \text{init Pop}(ps)$

while (критерий останова) *do*

$Arc := \text{updateOptimalSetN}(Arc, Pop)$ /*обновленное оптимальное множество*/

$Arc := \text{pruneOptimalSet}(Arc, as)$ (сокращение оптимального множества решений до размера ps)

$v :=$ функция качества (Pop, Arc, cmp_F)

$P(t) :=$ селекция (Pop, Arc, v, ps)

$t := t + 1$

$Pop :=$ репродукция $P(t)$

return /*восстановление фенотипа оптимального множества решений

$(Pop \cup Arc)$ */

end

Уточним различия в представленных выше на псевдокоде алгоритмах, реализующих базовый цикл и стратегию элитизма.

Во-первых, создается архив Arc элитных решений, который первоначально является пустым множеством, а затем обновляется функцией « updateOptimalSetN », которая сохраняет и обновляет полученные элитные решения (алгоритмы реализации этой функции будут подробно рассматриваться в последующих разделах). Во-вторых, если множество элитных решений становится слишком большим, то функция « pruneOptimalSet » сокращает его до величины ps . В принципе, в алгоритмах, построенных по принципам, отличным от элитизма, такой архив можно сделать пустым множеством ($Arc := \emptyset$).

Конечно, для поддержки теории биоинспирированного поиска оптимальных решений с помощью эволюционных алгоритмов необходимы не только адекватные модели и эффективные алгоритмы. Нужны инструментальные средства и соответствующая среда, обеспечивающая возможность рационального выбора конкретных моделей и алгоритмов эволюции, их подбора под решаемую задачу, позволяющей осуществлять быстрое построение и динамическую настройку параметров алгоритмов. В настоящее время предложено большое количество различных реализаций биоинспирированных алгоритмов, некоторые из них опробованы на тестовых задачах, однако их применение для решения практических задач геоинформатики затруднено ввиду отсутствия или недостаточной гибкости соответствующих программных систем. Существующие на текущий момент программные средства, реализующие генетический поиск, обладают рядом существенных недостатков:

- разработанное программное обеспечение жестко привязано к решаемой задаче на этапах кодирования и декодирования ее решений;
- программирование биоинспирированного алгоритма начинается практически с нуля;
- программное обеспечение закрыто для доработки или для интеграции с другими программами;

- результатов поиска и промежуточных состояния популяции решений не сохраняются, их невозможно анализировать в дальнейшем;
- когнитивные возможности эволюционных операторов являются ограниченными.

В этой связи в дальнейшем предстоит разработать архитектуру среды поддержки биоинспирированных алгоритмов. Она должна включать общесистемную среду и среду поддержки действий пользователя. Общесистемная среда должна обеспечивать общее управление процессами обработки геоинформации, содержать базу данных и знаний моделей, алгоритмов и эволюционных операторов, а также инструменты для сборки рабочих программ. Среда поддержки действий пользователя должна включать интерфейсную часть, поддерживать действия пользователя по выбору вариантов алгоритма, по настройке его параметров с помощью соответствующих диалоговых окон. Среда должна помогать подготавливать определенные форматы исходных данных для решения конкретных задач, а также обеспечивать визуализацию результатов.

С теоретической точки зрения подобного рода структура позволит устранить практически все описанные выше недостатки, но сложность ее практической реализации, возможно, окажется значительной, так как потребуются поддержка различных моделей и биоинспирированных алгоритмов для решения задач обработки геоинформации.

Более перспективной представляется инструментальная среда поддержки биоинспирированных алгоритмов в геоинформатике, которая бы удовлетворяла следующим требованиям:

- возможность встраивания инструментальной среды в различные существующие геоинформационные системы;
- простота настройки биоинспирированного поиска, доступная для неспециалиста в области эволюционного моделирования и программирования;
- наличие блока подбора и адаптации операторов и параметров алгоритмов под решаемую задачу;
- возможность использования внешних функций, реализующих когнитивные возможности эволюционных операторов, способы кодирования и вычисления целевых функций;
- наличие встроенных библиотек алгоритмов и операторов, типов кодирования и тестовых функций, а также возможностей гибкой настройки базового цикла алгоритма и использования внешних функций;
- возможность запуска алгоритма с созданной внешними средствами или ранее сохраненной начальной популяцией;
- наличие блока анализа исходных данных, сохранения, обработки и анализа результатов с расчетом основных статистических данных;
- простота разработки инструментальных средств.

Система должна быть принципиально открытой и процесс ее обучения никогда не завершается созданием окончательной формализованной модели. В этом случае инструментальные средства исполняют роль «информационной машины», которая способна действовать в непрерывном режиме усвоения и реструктуризации информации. Роль обратной связи в такой системе смогут выполнять экспертные знания и правила работы с ними.

При разработке инструментальных средств и среды поддержки биоинспирированного поиска необходимо следовать следующим рекомендациям.

Создать программы, реализующие биоинспирированный поиск, позволяющие настраивать и управлять всеми его параметрами. Речь идет о параметрах, устанавливаемых для популяций и не изменяемых во время работы алгоритма (количество запусков алгоритма, функция качества), а также о параметрах, устанавливаемых для популяций и которые могут изменяться.

База знаний инструментальной среды поддержки биоинспирированного поиска в геоинформатике должна включать базу фактов и базу правил, предназначенную для под-

бора параметров алгоритма и их коррекции. Это могут быть продукционные правила с нечетким заданием переменных или фреймовые структуры.

Логичным выглядит также требование о возможности вычисления целевых функций и функции качества в инструментальной среде путем выбора из библиотеки тестовых функций или путем задания их пользователем или с помощью внешней процедуры.

Немаловажным представляется наличие в инструментальной среде средств для сохранения результатов моделирования, сбора статистики, графического представления и анализа результатов.

С целью получения достоверной информации о качестве работы алгоритма биоинспирированного поиска инструментальная среда должна позволять производить несколько запусков с одинаковыми исходными параметрами, использовать методы математической статистики.

Подобного рода структура инструментальной среды позволит применять её как при разработке и исследовании новых моделей и алгоритмов биоинспирированного поиска, так и в интересах совершенствования и практического применения уже существующих геоинформационных систем.

4 Приложения теории биоинспирированного поиска

В процессе развития общества люди накапливают информацию об окружающем мире. Эту информацию можно считать базой знаний общества, которую его члены могут эксплуатировать в целях оптимизации своего поведения.

В алгоритме биоинспирированного поиска эту базу знаний формализуют в виде продукционных правил и используют в качестве элемента воздействия на популяцию. Основными компонентами алгоритма являются популяция особей-решений и база знаний. Они взаимодействуют посредством двух функций: принятия и влияния. Функция принятия определяет множество наилучших решений, корректирующих базу знаний. Функция влияния задает правила, по которым знания влияют на эволюцию особей в популяции. При этом возможна модификация любого из существующих операторов, обеспечивающих эволюционные изменения в популяции.

Псевдокод данного алгоритма биоинспирированного поиска (*BIK*) имеет вид:

Begin:

$t = 0;$

Инициализация популяции $P(t)$;

Инициализация базы знаний $KB(t)$;

Оценка популяции $P(t)$;

Repeat

Взаимодействие ($P(t)$, $KB(t)$);

Коррекция базы знаний $KB(t)$;

Оценка популяции $P(t)$;

Эволюция популяции $P(t)$ на текущей итерации t ;

$t = t + 1;$

Until (*проверка условий останова алгоритма*)

End

Эксперименты на бенчмарках известных задач комбинаторной оптимизации (*Travelling salesman problem* [18], задача диспетчирования [19], задача о расписании [20] и др.) показывают, что адаптация и эволюция среди особей популяции происходит быстрее с использованием базы знаний. Идет накопление знаний, которые передается другим поколениям.

Другой меметический алгоритм (*BMA*) использует концепцию мемов и принцип эволюции [21]. Мем – это информационная единица, способная к самовоспроизводству при помощи людей. Люди хранят и передают эту информацию другим людям посредством коммуникации [22]. В *BMA* мем является реализацией какого-либо алгоритма локальной

оптимизации, уточняющего текущие решения на каждой итерации, либо через несколько итераций. По сути меметический алгоритм – это комбинация глобального и локального поиска путем интеграции эволюционных вычислений с обучением отдельных особей.

Общая схема ВМА:

Шаг 1 Создание начальной популяции решений $P_0 = \{p_1^0, \dots, p_n^0\}$. Формирование множества мемов. В качестве мемов используется множество различных алгоритмов $M = \{M_1, \dots, M_k\}$, например, *Simulated annealing*, *Tabu search*, *gradient method* и др. [23].

Шаг 2 Для каждого решения вычисляем фитнес-функцию F_i .

Шаг 3 Кооперация путем обмена информацией между решениями (аналогично кроссинговеру в генетических алгоритмах), применение эволюционных операторов.

Шаг 4 Локальный поиск. Лучший мем M_i выбирается из роя мемов на основе h запусков локального поиска путем вычисления значений целевой функции оптимизации.

Шаг 5 Создание новой популяции путем соревнования/селекции.

Шаг 6 Проверка критерия окончания поиска (число итераций, оценка улучшения результата), переход на шаг 2, иначе выбор лучшего решения из популяции.

Общая схема меметических алгоритмов открывает широкие возможности для разработки многих его вариантов:

- локальный поиск использовать не на каждой итерации, а через некоторое число итераций по определенному закону;
- для всех элементов популяции решений на всех итерациях может быть использован один и тот же алгоритм-мем или разные мемы для разных элементов популяции и/или для различных итераций;
- основная задача конструирования ВМА – выбор целесообразной стратегии (гиперэвристики) использования того или иного мема из роя доступных мемов M ;
- построение самоадаптирующихся алгоритмов, когда каждое решение кодируется мультихромосомой, первая часть которой – решение, а вторая часть – мем, с помощью которого получено данное решение.

ВМА описывается следующей синтаксической моделью:

$$MA = \{P_0, t, \mu, \delta_0, \lambda, N, F, q, L\},$$

где $P_0 = (P_0^1, P_0^2, \dots, P_0^m)$ – исходная популяция мемов из конечномерного дискретного пространства;

t – шаг эволюции $(0, 1, \dots, t_{max})$;

m – размер популяции мемов;

δ_0 – начальный набор параметров для операторов эволюции;

λ – размер множества мемов-потомков, сгенерированных на очередном шаге эволюции;

N – размерность задачи;

F – целевая функция;

q – отображение, описывающее связь между мемами и кодируемыми объектами;

$L = \{L_1, L_2, \dots, L_m\}$ – операторы генерации мемов, используемые алгоритмом в процессе эволюции.

В рамках синтаксической модели можно описать все разнообразие алгоритмов ВМА, используя различные сочетания эволюционных операторов.

Для характеристики отдельных классов ВМА введем индекс

$$I(BMA) = (i_3, i_2, i_1, i_0).$$

Индекс представляет собой 4-битовое двоичное число. Младший бит указывает на использование унарного оператора эволюции, аргумент которого включает один мем, следующий бит указывает на использование не более m мемов, следующий бит указывает на использование не более $(m + \lambda)$ мемов, старший разряд – на использование $2^{P_0 \cup \dots \cup P_{i-1}}$ мемов. Если операторы генерации мемов $L_j(t)$ изменяют свои параметры, то го-

ворят об адаптации мемов. Если изменяется сам оператор $L_j(t)$, то говорят о самоадаптации мемов.

Экспериментальные исследования *ВМА* проводились для задачи построения расписания [20]. Постановка этой задачи состоит в следующем.

Дано множество требований к расписанию (длительность, стоимость обработки требования, время поступления требования, директивный срок окончания обслуживания требования и т.п.). Известно множество объектов (машин, приборов, аудиторий и т.п.), на которых требования должны по порядку обслуживаться. Имеются ограничения в задаче: «жесткие» (должны быть удовлетворены и не могут нарушаться) и «мягкие». Задача была формализована на примере составления расписания экзаменационной сессии.

Согласно алгоритму, полученные в ходе эволюции решения в пределах текущей популяции, улучшались в течение одного поколения за счет использования алгоритма локального меметического поиска.

Алгоритм был запрограммирован в *Matlab*, выполнен на компьютере *Intel Pentium 4* (2,33 ГГц) и проверен на реальных наборах тестовых данных.

Результаты экспериментов показали, что качество найденных решений лучше, по сравнению с использованием известных генетических алгоритмов [23].

Алгоритм *ВМА* обеспечивает неплохие результаты для решения целого ряда *NP*-сложных задач: быстро, надежно и точно. Алгоритму присущ параллелизм. Он проще генетических и роевых алгоритмов.

Заключение

Авторы считают, что в данной работе новыми являются следующие результаты: разработаны основные элементы общей теории биоинспирированного поиска оптимальных решений, оригинальные алгоритмы биомететики, сочетающие такие компоненты, как локальный поиск, кооперация и соревнование. Проведенные эксперименты для ряда *NP*-сложных задач оптимизации показали, что с помощью разработанной теории биоинспирированного поиска эти задачи решаются быстро, надежно и точно. Алгоритмы *ВИК* и *ВМА* проще генетических и роевых алгоритмов.

Биоинспирированные алгоритмы – активно развивающаяся область методов оптимизации и принятия решений. На данный момент наиболее перспективным следует считать создание адаптивных версий биоинспирированных алгоритмов поиска, учитывающих предысторию поиска, а также проблемно-ориентированную информацию об области поиска оптимальных решений. Основным инструментом конструирования новых высокоэффективных биоинспирированных алгоритмов является гибридизация. Объем статьи позволил включить лишь очень небольшую часть примеров гибридизации. По этой причине не рассмотрены вопросы параллелизма алгоритмов *ВИК* и *ВМА*, а также задачи многокритериальной оптимизации.

Особенностью биоинспирированных алгоритмов является наличие значительного числа свободных параметров. От них может сильно зависеть эффективность алгоритмов. Имеются лишь экспериментальные рекомендации по выбору значений этих параметров. Необходимы дальнейшие исследования по адаптации и самоадаптации этих параметров.

Литература:

1. *Holland J.H.*: Adaptation in Natural & Artificial Systems. Ann Arbor: Uni of Michigan Press, 1975.
2. *Kureichik V.M. and Rodzin S.I.* Evolutionary algorithms: Genetic Programming. J. Computer System Science Int. 2002. Vol. 41. No. 1. P. 123–132.
3. *Rodzin S.I.* Schemes of Evolution Strategies. Proc. of IEEE Int. Conf. on Art. Intel. Systems (ICAIS, sept. 2002): Los Alamos, California. 2002. P. 375–380.
4. *Fogel D.B.* Evolutionary Computation: The Fossil Record. Piscataway, NJ: IEEE Press, 1998.
5. *Dorigo M., Maniezzo. V., Colorni. A.* The Ant System: Optimization by a colony of cooperating objects. IEEE / Trans. on Systems, Man, and Cybernetics. Part B. 1996. No. 26(1). P. 29–41.

6. *Pham D.T., Ghanbarzadeh A., Koc E., Otri S., Rahim S., Zaidi M.* The Bees Algorithm. Tech. Note, Manufacturing Eng. Centre, Cardiff University, UK, 2005.
7. *Bastos-Filho C.J.A., Lima-Neto F.B., Lins A., Nascimento A., Lima M.* Fish School Search. Nature-inspired Algorithms for Optimization (NISCO'2010). Springer; Heidelberg, 2009. Vol. 193. P. 261–277 (2009).
8. *Reynolds C.* Flocks, Herds, and Schools: A Distributed Behavioral Model. Computer Graphics, 1987. No. 4(21). P. 25–34.
9. *Das S., Biswas A., Dasgupta S., Abraham A.* Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis and Applications. Foundations of Computational Intelligence. Publisher: Springer, 2009. Vol. 203. P. 23–55.
10. *Yang X.-S., Deb S.* Cuckoo Search Via Levy Flights. Proc. of the World Congress on Nature & Biologically Inspired Computing (NaBIC'2009, India). IEEE Publ., USA, 2009. P. 210–214.
11. *Yang Xin-She.* Firefly Algorithm, Stochastic Test Functions and Design Optimization // Int. Jour. of Bioinspired Computation. 2010. Vol. 2. No. 2. P. 78–84.
12. *Mehrabiana A.R., Lucas C.* A Novel Numerical Optimization Algorithm Inspired from Weed Colonization. Ecological informatics. 2006. No. 1. P. 355–366.
13. *Mucherino A., Seref O.* Monkey Search: a Novel Meta-heuristic Search for Global Optimization / Proc of AIP Conf. «Data Mining, System Analysis and Optimization in Biomedicine». P. 162–173 (2007).
14. *Abidin Z.Z., Arshad M.R., Ngah U.K.* A Simulation Based Fly Optimization Algorithm for Swarms of Mini Autonomous Surface Vehicles Application. Indian Jour. of Geo-marine Sciences. 2011. Vol. 40 (2). P. 250–266.
15. *Eusuff M.M., Lansey K., Pasha F.* Shuffled Frog Leaping Algorithm: a Memetic Meta-heuristic for Discrete Optimization. Engineering Optimization. 2006. Vol. 38. No. 2. P. 129–154.
16. *Yang X.-S.* A New Metaheuristic Sat-inspired Algorithm. Nature Inspired Cooperative Strategies for Optimization (NISCO'2010). Berlin: Springer, 2010. Vol. 284. P. 65–74.
17. *Wolpert D.H., Macready W.G.* The no Free Lunch Theorems for Optimization. IEEE Trans. Evol. Comp. 1997. Vol. 1. No. 1. P. 67–82.
18. *Cormen T.H., Leiserson C.E., Rivest R.L., Stein C.* Introduction to Algorithms. MIT Press, 2009.
19. *Rodzin S.I.* Smart Dispatching and Metaheuristic Swarm Flow Algorithm. J. of Comp. and Syst. Sc. Inter. 2014. Vol. 53. No. 1. P. 109–115.
20. *Conway R.W., Maxwell W.L., Miller L.W.* Theory of Scheduling. Courier Dover Publ., 2003.
21. *Moscato P.* Memetic algorithms. Handbook of Applied Optimization. Oxford: Uni Press, 2002.
22. *Rodzina L., Kristofferson S.* Context-dependent Car Navigation as Kind of Human-Machine Collaborative Interaction. Proc. of the 2013 Int. Conf. on Collaboration Technologies & Systems (CTS'2013, may 20–24, 2013, San Diego, California, USA.). Publ. of the IEEE, 2013. P. 253–259.
23. *Kureichik V.V., Kureichik V.M., Rodzin S.I.* Theory of Evolutionary Computation. Fizmatlit: Moscow, 2012.

Geoinformatics: theory of bioinspired search for optimal solutions and its application for the processing of problem-oriented data and knowledge

Sergey Ivanovich Rodzin, professor, Southern Federal University

Lada Sergeevna Rodzina, postgraduate, Southern Federal University

The paper discusses the main elements of the theory of bioinspired search for optimal solutions. An original memetic algorithms combining local search, cooperation and competition. Experiments for NP-hard optimization problems have shown that using the developed theory, these problems are solved quickly, reliably and accurately.

Keywords: geoinformatics, bioinspired search, memetic algorithm, processing of problem-oriented knowledge.