

УДК 004.421.2

Ж.С. ЕСЕНГАЛИЕВА

*Казахский национальный университет им. аль-Фараби, Алматы, Казахстан;
e-mail: Zhanna.Yessengaliyeva@kaznu.kz*

Разработка генетического алгоритма в контексте оценки научно-инновационных проектов

Предложен генетический алгоритм, позволяющий производить отбор научно-инновационных проектов. Представлена функция приспособленности, которая численно характеризует научно-инновационный проект. Введено понятие индекса привлекательности проекта. Описаны результаты лабораторных экспериментов, направленных на принятие решений при оценке научно-инновационных проектов. Установлено, что рассмотрение научно-инновационных проектов как особей существенно способствует улучшению результата их отбора, а также оптимизирует процесс оценки и выбора проектов.

Ключевые слова: генетический алгоритм, научно-инновационный проект, отбор, функция приспособленности, популяция.

Zh. Yessengaliyeva

Development of genetic algorithm in the context of the assessment of scientific and innovative projects

The genetic algorithm, allowing making selection of scientific and innovative projects is offered. Fitness function which numeric characterizes the scientific and innovative project is presented. The concept of an index of attractiveness of the project is entered. Results of the laboratory experiments directed on decision-making at an assessment of scientific and innovative projects are described. It is established that consideration of scientific and innovative projects as individuals significantly promotes improvement of result of selection and optimizes process of an assessment and a choice of projects.

Key words: genetic algorithm, scientific and innovative project, selection, fitness function, population.

Ж. Есенгалиева

Ғылыми-инновациялық жобаларының сұрыптауындағы генетикалық алгоритмды әзірлеу

Ғылыми-инновациялық жобаның сұрыпын жасау үшін генетикалық алгоритм қолданылған. Ғылыми-инновациялық жобаны сандық түрінде сипаттайтын икемнің функциясы ұсынылған. Жобаның тартымдылық индексі деген ұғымы енгізілген. Ғылыми-инновациялық жобаның сарапшылығында лабораториялық эксперименттің нәтижелері баяндалған. Ғылыми-инновациялық жобаның дарақ ретінде қарастыруы таңдаудың нәтижесін жақсартудың айтарлықтай мүмкіндік туғызатыны және жобаның бағалау мен іріктеу процессінің оптимизациялауы анықталды.

Түйін сөздер: генетикалық алгоритм, ғылыми-инновациялық жоба, сұрыптау, икемнің функциясы, популяция.

Введение

В качестве одного из методов снижения негативного воздействия рисков, неопределенностей, под воздействием которых реализуются научно-инновационные проекты, рассмотрим метод оптимизации - генетический алгоритм, где отличительной особенностью является возможность одновременного манипулирования многими параметрами и генерация новых решений посредством использования генетических операторов. Генетические алгоритмы (ГА) - это эвристические поисковые алгоритмы, используемые для решения задач моделирования путем последовательного подбора, комбинирования и вариации искомым параметров с применением механизмов, напоминающих биологическую эволюцию. Они нашли широкое применение в оптимизации, искусственном интеллекте, инженерии и других областях [1].

Принципы биологии и генетики являются основой данного метода. Концепция ГА состоит в создании популяции особей, представленных в виде хромосомы. Искомая хромосома есть возможное решение моделируемой задачи. Для поиска прогнозных решений необходимо только значение целевой функции, или функции приспособленности. В процессе решения задачи значение функции приспособленности особи отражает соответствие полученных значений ожидаемым результатам. Хромосома представляет собой генотип исследуемого объекта и состоит из конечного числа генов, так называемых наследственных признаков. На уровне генотипа в популяции ведется эвристический эволюционный поисковый процесс, и применяются биологические операторы: скрещивание, мутация, инверсия, селекция и др. Решением ГА является основополагающий механизм эволюции - естественный отбор, где по известному принципу "выживает сильнейший". Популяция постоянно обновляется при помощи генерации новых особей и уничтожения старых, и каждая новая популяция становится лучше и зависит только от предыдущей [1]. Джон Холланд является основоположником теории генетических алгоритмов, его научный труд "Адаптация в естественных и искусственных системах" (1992) лег в первоначальную основу исследований в данной области.

Идею ГА подсказала сама природа и работы Дарвина. В 1975 г. Дж.Х. Холланд предложил схему генетического алгоритма [2]. Эти труды стали основой главных направлений разработки эволюционных генетических алгоритмов.

Описание генетического алгоритма

Концепция генетического естественного отбора предполагает скрещивание двух жизнеспособных особей, с целью получения новой особи, с высокой вероятностью ее жизнеспособности и приспособленности к условиям существования. В ГА хромосомы представлены в виде бинарных векторов, в которых закодированы значения аргументов заданной функции приспособленности. Значение целевой функции выражает жизнеспособность хромосомы, точнее ее пригодность. В рассматриваемом алгоритме исходными решениями являются родительские хромосомы. Потомками называются новые решения, получаемые из исходных решений.

Для реализации генетического алгоритма используют моделирование эволюции. Борьба за выживание в генетических алгоритмах представляет собой процесс замещения в популяции хромосом с худшей пригодностью на хромосомы с лучшей пригодностью, получаемые в результате скрещивания и мутации. Генетический алгоритм, являясь искусственной эволюцией, включает несколько этапов (см.рис 1): инициализацию популяции случайными хромосомами, отбор и группирование хромосом в пары, скре-

щивание и мутацию родительских хромосом, оценку пригодности потомков, проверку условий завершения поиска. При этом хромосомы начальной популяции в процессе эволюции могут замещаться другими, более приспособленными хромосомами.



Рисунок 1. Блок-схема генетического алгоритма.

ГА позволяет решать задачи прогнозирования, классификации, поиска оптимальных вариантов и совершенно незаменим в тех случаях, когда в обычных условиях решение задачи основано на интуиции или опыте, а не на строгом (в математическом смысле) ее описании [3]. Достоинство ГА заключается в его способности манипулировать одновременно многими параметрами [3].

Численные эксперименты

В исследуемом процессе оптимизации научно-инновационного развития какого-либо объекта следует найти приемлемое, наиболее подходящее решение из предложенных научно-инновационных проектов (НИП). Генетический алгоритм случайным образом генерирует начальную популяцию. Работа генетического алгоритма представляет итерационный процесс, который зависит от заданного числа поколений.

Для организации работы генетического алгоритма при оценке научно-инновационных проектов введем понятие индекса привлекательности НИП в качестве характеристики инвестиционной привлекательности элитных особей, выживших в процессе естественного отбора.

Индекс привлекательности - это фактор, позволяющий учитывать степень рискованности инвестирования НИП, базирующийся на имеющихся знаниях о "выживших" проектах и на данных предлагаемых новых проектов. Необходимо определить используемые генетические операторы:

1. *Оператор селекции (selection)*. В рассматриваемом случае используется стратегия элитизма, когда в популяцию выбираются элитные элементы на основе сравнения значений целевой функции. Данное значение рассчитывается для каждой хромосомы на основе закодированных данных в генотипе с целью выбора наиболее приспособленных особей, которые получают возможность скрещиваться и давать потомство. Наихудшие особи удаляются из популяции и не дают потомства. В классических реализациях алгоритма размер популяции постоянен.

Здесь целевая функция численно характеризует научно-инновационный проект, базируясь на методологический подход, основанный на экспертных оценках показателей инновационности и конкурентоспособности с применением аддитивно-мультипликативного метода расчета. Подробно данная методология расчета опубликована в работе [4].

Таблица 1: Инициализация начальной популяции.

Особи	Нормированные оценки вектора приоритета критерия инновационности и конкурентоспособности	Индекс привлекательности	Примечание	Функция приспособленности
Проект № 1	6,26; 5,38	0,5	Новый	5,82
Проект № 2	4,62; 4,32	0,5	Новый	4,47
Проект № 3	3,2; 2,8	0,5	Новый	3,00
Проект № 4	2,51; 5,54	0,8	Элитный	6,44
Проект № 5	5,54; 6,65	0,9	Элитный	10,97
Проект № 6	7,8; 2,39	0,6	Элитный	6,11
Проект № 7	5,8; 3,39	0,6	Элитный	5,51
Проект № 8	3,1; 2,7	0,4	Реализованный	2,32
Проект № 9	5,1; 5,7	0,4	Реализованный	4,32
Проект № 10	4,02; 4,0	0,4	Реализованный	3,21

В представленной целевой функции индекс привлекательности выражается в процентном соотношении:

$$f(x) = \left(\sum_{i=1}^n I_j + \sum_{k=1}^m K_j \right) * IA_j \rightarrow \max,$$

$$0 \leq IA_j \leq 0.4 \quad (\text{худшие проекты})$$

$$IA_j = 0, 5 \quad (\text{предлагаемые новые проекты})$$

$$0, 6 \leq IA_j \leq 1 \quad (\text{элитные проекты})$$

где I – инновационность j -го проекта, K – конкурентоспособность j -го проекта, IA_j – индекс привлекательности j -го проекта, характеризующий процент успешного внедрения элитного элемента популяции, n – количество показателей инновационности, m – количество показателей конкурентоспособности. В таблице 1 на основе предварительно обработанных нормированных оценок векторов инновационности и конкурентоспособности [4] получена начальная популяция генетического алгоритма.

В данном операторе, как отмечено выше, используется стратегия элитизма, которая с помощью различных преобразований позволяет выбирать элитные элементы. Процесс продолжается до тех пор, пока появляются элитные элементы. В новое поколение включается заданное количество лучших особей предыдущего поколения (часто одна лучшая особь). Использование элитизма не допускает потери лучших решений. К примеру, если популяция сошла в локальном максимуме, а мутация вывела одну из строк в область глобального, то при замещении родителей весьма вероятно, что эта особь в результате скрещивания будет потеряна, и решение задачи не будет получено. Если же используется элитизм, то полученное хорошее решение будет оставаться в популяции до тех пор, пока не будет найдено лучшее.

2. *Оператор кроссинговера* (crossing-over). Далее для реализации работы исследуемого генетического алгоритма применяется двухточечный кроссинговер. При этом две особи, выбранные оператором селекции или отбора, становятся "родителями" и обмениваются генами для воспроизведения новых двух особей - потомков. Нормированные оценки векторов инновационности и конкурентоспособности складываются и полученное значение переводится в двоичную систему счисления, то есть получаем генотип проекта. При двухточечном скрещивании особи обмениваются генами: выбираются случайные две точки раздела, и родители обмениваются промежутками между ними. В таблице 2 приведены фенотипы и генотипы исследуемых особей.

Таблица 2: Фенотипы и генотипы хромосом.

Особь	Фенотип	Генотип
Особь 1 (Проект № 1)	5,82	101,1101000111
Особь 2 (Проект № 2)	4,47	100,011110000
Особь 3 (Проект № 3)	3,00	011,000000000
Особь 4 (Проект № 4)	6,44	110,011100001
Особь 5 (Проект № 5)	10,97	1010,11111000
Особь 6 (Проект № 6)	6,11	110,000111000
Особь 7 (Проект № 7)	5,51	101,100000101
Особь 8 (Проект № 8)	2,32	010,010100011
Особь 9 (Проект № 9)	4,32	100,010100011
Особь 10 (Проект № 10)	3,21	011,001101011

Рассмотрим применение оператора кроссинговера, при этом "родителями" являются особь 1 (Проект № 1) и особь 6 (Проект № 6):

$$\begin{array}{l} 101,0.1010.0011 \rightarrow 101,1.0011.0011 \\ 110,0.0011.1000 \quad 110,0.1010.1000 \end{array}$$

Аналогичным образом, скрещиваем Особь 2 и Особь 7, затем Особь 3 и Особь 8.

3. *Оператор мутации* (mutation). В исследуемом алгоритме используется точечная мутация и инверсия. Ген у двух особей, выбранных оператором селекции, заменяется случайным значением с учетом его известной предельной величины. При этом происходит переход мутированных особей в следующее поколение, здесь же применяется инверсия:

$$\begin{aligned} 111, 1.0010.0010 \\ 100, 0.1011.1001 \end{aligned}$$

4. *Оператор редукции* (reduction). Оператор редукции выполняется после оператора мутации и позволяет выбрать две лучшие особи из пар родителей и потомков, имеющих максимальное значение целевой функции, и помещает их в популяцию. При этом две оставшиеся особи в популяцию не попадают. По итогам отбора остаются 10 особей, которые войдут в итоговую популяцию. Остальные особи погибают.

На данном этапе используем селекцию по методу рулетки. Для скрещивания особи отбираются с помощью запусков рулетки N раз. Колесо рулетки содержит по одному сектору для каждой особи из популяции. Размер g -го сектора пропорционален вероятности попадания особи p_g , которая рассчитывается по формуле:

$$p_g = f_g / \sum f_g$$

Очевидно, что для всей популяции выполняется условие $\sum p_g = 1$. При таком отборе члены популяции с более высокой приспособленностью с большей вероятностью будут чаще выбираться, чем особи с низкой приспособленностью (табл. 3 и рис. 2).

Таблица 3: Метод рулетки. Размер популяции = 10. Суммарная пригодность популяции = 6,34.

Популяция	Пригодность	Вероятность выбора	Доля
Особь1 (Проект№1+Проект№6)	7,57	0,13	13%
Особь2(Проект№1+Проект№6)	4,36	0,07	7%
Особь3(Проект№2+Проект№7)	4,12	0,07	7%
Особь4(Проект№2+Проект№7)	7,76	0,13	13%
Особь5(Проект№3+Проект№8)	7,0	0,12	12%
Особь6(Проект№3+Проект№8)	1,82	0,05	5%
Особь7(Проект № 5)	10,97	0,18	18%
Особь8(Проект № 6)	6,11	0,10	10%
Особь9(Проект № 7)	5,51	0,09	9%
Особь10(Проект № 9)	4,32	0,07	7%
Суммарная пригодность	60,542		100%

Таким образом, предлагаемый метод позволяет учитывать опыт предыдущих реализованных научно-инновационных проектов, входящих в состав элитных особей, что приводит к снижению степени риска инвестирования. Экспериментально-теоретические исследования генетического алгоритма при отборе НИП показывают, что поведение нейтральных, по экспертному мнению, проектов может привести к высоким результатам.

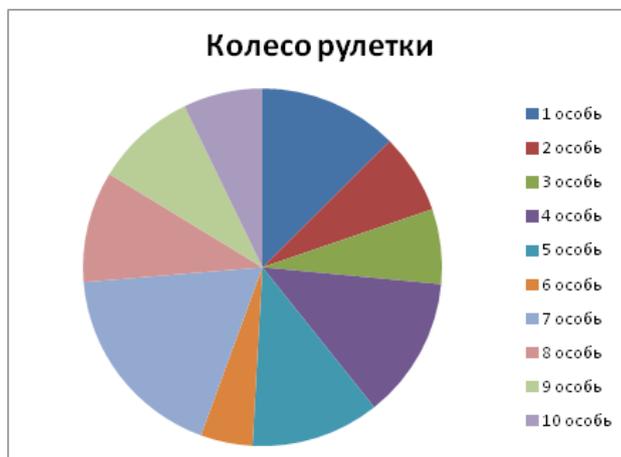


Рисунок 2. Оператор селекции типа колеса рулетки с секторами пропорциональными функции приспособленности.

Заключение

Данная статья посвящена искусственному интеллекту, где рассматривается некоторое пространство гипотез, из которых мы должны выбрать лучшее решение. Эволюционное мутирование одного из них может привести к неожиданному улучшению результата. Использование предлагаемого аппарата позволяет улучшить качество отбора научно-инновационных проектов. Итак, на основании вышеизложенного следует отметить, что генетические алгоритмы являются универсальным методом оптимизации функций и позволяют решать задачи научно-технологического спектра.

Список литературы

- [1] *Паклин Н.* Непрерывные генетические алгоритмы - математический аппарат. – 2013. – (http://www.basegroup.ru/library/optimization/real_coded_ga/)
- [2] *Мясников А.С.* Островной генетический алгоритм с динамическим распределением вероятностей выбора генетических операторов. – 2013. – (<http://technomag.bmstu.ru/doc/136503.html>)
- [3] *Ковшов Е.Е., Горяева О.В.* Применение генетического алгоритма при оценке рисков инновационных проектов. // Российское предпринимательство. – 2010. – №11. – В.3(172). – С. 85-91.
- [4] *Есенгалиева Ж.С.* Исследование метода и графической модели оценки инновационности и конкурентоспособности инновационных проектов. // Инновации в образовании: Материалы 42-й международной научно-практической конференции. – Алматы: Казак университеті. – 2012. – Кн.2. – С. 261–267.

References

- [1] *Paklin N.* Nепрерывные генетические алгоритмы - математический аппарат. – 2013. – (http://www.basegroup.ru/library/optimization/real_coded_ga/)
- [2] *Myasnikov A.C.* Ostrovnoi генетический алгоритм с динамическим распределением вероятностей выбора генетических операторов. – 2013. – (<http://technomag.bmstu.ru/doc/136503.html>)
- [3] *Kovshov E.E., Goryaeva O.V.* Primenenie генетического алгоритма при отсенке рисков инновационных проектов. // Rossiyskoe predprinimatel'stvo. – 2010. – №11. – V.3(172). – S. 85-91.
- [4] *Yessengaliyeva Zh.S.* Issledovanie metoda i modeli otsenki innovatsionnosti i konkurentosposobnosti innovatsionnykh projektov. // Innovatsii v obrazovanii: Materialy 42-й mezhdunarodnoi nauchno-prakticheskoi konferentsii. – Almaty: Qazaq universiteti. – 2012. – Kn.2. – S. 261–267.

Поступила в редакцию 3 мая 2013 года