



Международный журнал информационных технологий и энергоэффективности

Сайт журнала:

<http://www.openaccessscience.ru/index.php/ijcse/>



УДК 004.02

СПОСОБ И АЛГОРИТМЫ РЕАЛИЗАЦИИ ГЕНЕТИЧЕСКОГО ОПЕРАТОРА РЕКОМБИНАЦИИ НА ОСНОВЕ ТУРНИРНОЙ СТРАТЕГИИ СО СЛУЧАЙНЫМ ИСХОДОМ

Зернов М.М., ¹Картавенков А.А.

*Филиал ФГБОУ ВО Национальный исследовательский университет «МЭИ» в г. Смоленске, Россия, (214013, г. Смоленск, Энергетический проезд, 1),
e-mail: ¹ kartawenkov.aleksandr@yandex.ru*

В статье предлагается два варианта способа отбора особей популяции в новое поколение, отличающиеся применением модифицированных вариантов стратегии турнирной селекции. Модификации турнирной стратегии состоят в случайном определении исхода турнира как на основе равной вероятности выбора особей, так и с применением принципа рулетки.

Ключевые слова: генетический алгоритм, оператор селекции, оператор рекомбинации, отбор методом колеса рулетки, турнирная селекция, равновероятностная турнирная стратегия, турнирная стратегия с использованием колеса рулетки.

METHOD AND ALGORITHMS FOR IMPLEMENTING THE GENETIC RECOMBINATION OPERATOR BASED ON A TOURNAMENT STRATEGY WITH A RANDOM OUTCOME

Zernov M.M., Kartavenkov A. A.

*Smolensk Branch of the National Research University "Moscow Power Engineering Institute", Smolensk, Russia (214013, Smolensk, Energeticheskyy proezd, 1),
e-mail: kartawenkov.aleksandr@yandex.ru*

New method of survival selection, using two variants of tournament selection with random outcome, proposed. Variants differ by using equal-probability individuals selection or limited (by tournament group) roulette wheel selection.

Keywords: genetic algorithm, selection operator, recombination operator, selection by the roulette wheel method, tournament selection, tournament strategy, equal-probability tournament selection, different-probability tournament strategy, tournament selection using the roulette wheel.

Введение

Генетический алгоритм до сих пор за счет своей простоты и универсальности является одним из популярных методов оптимизации [2]. Несмотря на глубокую проработку многих вопросов, постоянно ведутся поиски эффективных способов конструирования решений и управления динамикой эволюционного процесса. Одним из направлений является раскрытие потенциала турнирной стратегии, в том числе и за счет внесения в неё элементов случайности. В работе рассматривается вариант стационарного генетического алгоритма (алгоритма с постоянной численностью популяции) с двумя видами отбора – отбором родительских особей в промежуточную популяцию (классическая селекция) и отбором особей из объединённой популяции «родители+потомки» в новое поколение. Последний вид отбора в некоторых источниках называют оператором рекомбинации [5]. В других же [7-8] называют стратегией выживания (Survivor Selection Policy или Survival Selection). Рассмотренные положения имеют смысл и для случая поколенческих генетических алгоритмов (с ростом численности популяции от поколения к поколению) с частичным отсевом особей при переходе к новому поколению.

Вообще говоря, стратегии отбора, применимые в случае селекции родительских особей, могут быть с успехом применены и в операторе рекомбинации [5,7-8]. Выбор вида стратегии позволяет настроить баланс между направленной и случайной составляющей поиска решения в эволюционном процессе. Применение турнирной стратегии, как правило, делает генетический алгоритм более «эксплуатирующим», т.е. направленным. Стратегия же на основе колеса рулетки позволяет повысить выживаемость менее приспособленных особей и, таким образом, увеличить долю случайности (что важно при оптимизации целевых функций со многими локальными оптимумами).

1. Постановка задачи

В данной работе предлагаются модифицированные варианты реализации оператора рекомбинации (стратегии выживания) на основе турнирной селекции путём введения в результат турнира элемента случайности. Предлагается использовать турниры с равновероятностными исходами (вырожденная стратегия, практически эквивалентная случайному выбору особей в новое поколение) и разноразностными исходами (с вероятностями, определёнными в соответствии с принципом колеса рулетки, но только для конкурирующих в пределах турнира особей).

2. Применяемая схема генетического алгоритма

В дальнейшем, будем рассматривать эволюционный процесс, соответствующий схеме, представленной на рисунке 1. Так, формирование начальной популяции происходит в начале работы алгоритма и представляет собой случайно сгенерированную последовательность неповторяющихся решений. Далее следует процесс выбора родительских пар и их скрещивание. Полученные потомки подвергаются мутации и добавляются в объединённую популяцию родителей и потомков. После этого происходит отбор особей в новое поколение. Так происходит до тех пор, пока не будет достигнут критерий останова генетического алгоритма. Критерием окончания процесса работы алгоритма может служить заданное количество поколений или схождение популяции. Схождением называется такое состояние

популяции, когда все элементы популяции почти одинаковые и находятся в области некоторого экстремума. В такой ситуации кроссинговер почти никак не может изменить популяцию, так как создаваемые при нем потомки представляют собой копии родителей с переменными участками хромосом. Вышедшие из этой области за счет мутации хромосомы склонны к вымиранию, так как чаще всего имеют меньшую приспособленность, особенно если данный экстремум является глобальным. Таким образом, схождение популяции обычно означает, что найдено лучшее или близкое к нему решение [1].

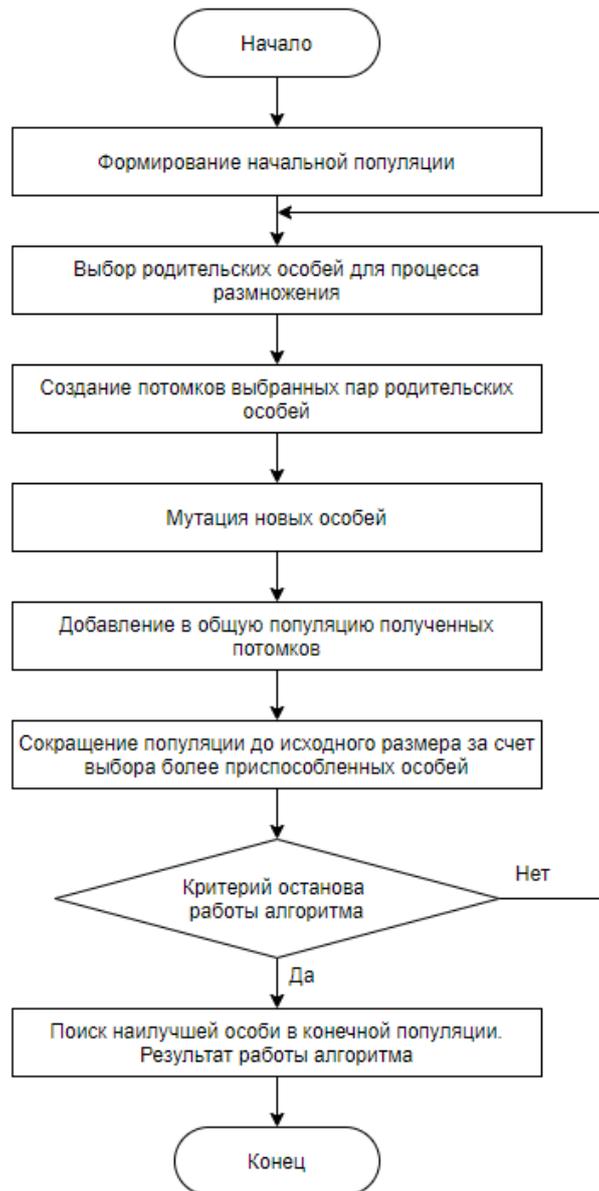


Рисунок 1 – Схема алгоритма генетического алгоритма

3. Отбор методом колеса рулетки

В основе данной стратегии лежит сопоставление каждой особи определенного сектора, размер которого пропорционален значению ее приспособленности [6], как показано на рисунке 2. Далее происходит выбор фиксированной точки и «вращение» колеса рулетки.

Особь, напротив которой останавливается точка, попадает в промежуточную популяцию. Так продолжается до тех пор, пока не будет отобрано определенное количество особей.

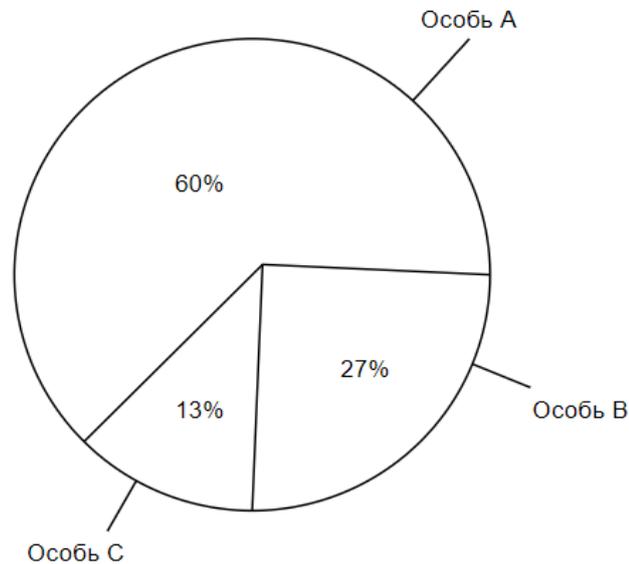


Рисунок 2 – Сопоставление особям секторов

Величина сектора, то есть вероятность выбора особи выполняется по следующей формуле:

$$P(i) = \frac{f(i)}{\sum_{i=1}^N f(i)}, \quad (1)$$

где $P(i)$ – вероятность выбора i особи

$f(i)$ – значение функции приспособленности для i особи,

N – количество анализируемых особей (в популяции в целом или турнирной группе, как будет предложено далее).

Также у данного способа возможна модификация, в которой колесо рулетки отбирает особи без повторений. Для этого после каждой итерации отбора следует пересчет процентного соотношения приспособленности для оставшихся особей. Эта модификация более ресурсоемкая, но зато она позволяет решить проблему преждевременной сходимости.

4. Турнирная селекция

Другим популярным методом отбора особей в родительские пары является турнирная селекция, которая представляет собой метод, в котором из популяции, содержащей N особей, случайно отбираются несколько особей, после чего лучшая из них записывается в промежуточный массив. Эта операция повторяется N раз (для получения $N/2$ родительских пар, но можно выбирать и меньшее количество). В дальнейшем особи, записанные в этот массив, будут участвовать в процессе скрещивания [6]. На рисунке 3 изображена схема турнирной селекции.

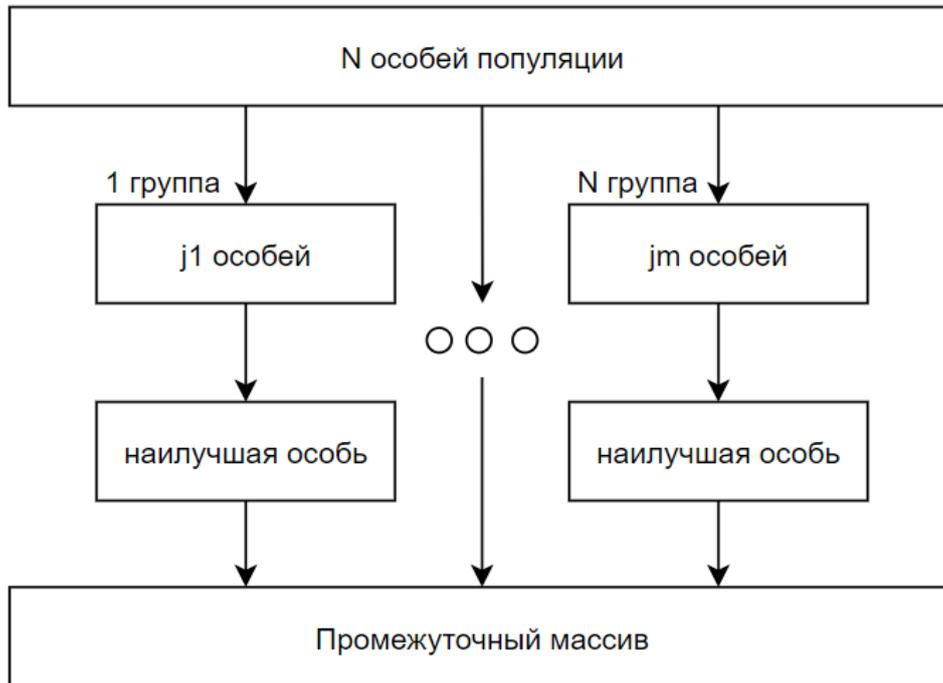


Рисунок 3 – Схема турнирной селекции

Турнирную стратегию помимо отбора особей в родительские пары, также можно использовать в операторе рекомбинации. На рисунке 4 изображена схема турнирной селекции для выбора особей, переходящих в новую популяцию. Возможен и обратный подход – выбор особей, удаляемых из популяции.

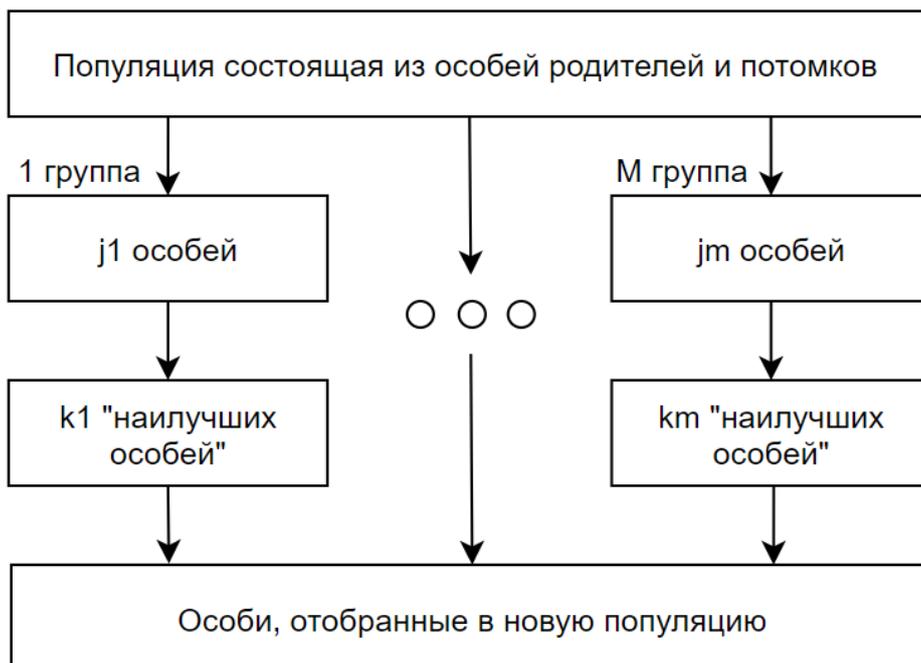


Рисунок 4 – Схема турнирной селекции

Предлагаемые способы реализации оператора рекомбинации, основанные на турнирной стратегии со случайным исходом, имеют ряд общих этапов и отличаются от «классического» способа непосредственно принципом определения победителей.

Общими являются следующие этапы: группировка особей, расчета количества особей, которые должны выйти из группы.

4.1 Разбиение особей на группы

В процессе разбиения особей на группы участвует объединенная популяция родительских и дочерних особей. Все особи этой популяции должны оказаться в турнирных группах, но без дублирования, то есть не должно быть такой ситуации, когда одна и та же особь присутствует в нескольких группах. Вначале требуется рассчитать количество групп. Исходя из одинакового размера, количество групп рассчитывается как:

$$Group\ Amount = \frac{JoinPopulationSize}{Group\ Size}, \quad (2)$$

где *Group Amount* – количество групп,
JoinPopulationSize – размер объединённой популяции родителей и потомков,
Group Size – размер группы.

Если количество групп получается дробным, то производится округление в большую сторону. Последняя группа может оказаться неполной, что не является пробой.

После расчета количества групп следует распределение особей по группам. Данный процесс протекает в нескольких циклах. Внешний имеет размер, равный рассчитанному выше количеству групп, внутренний же равен размеру группы. Внутри последнего и происходит отбор особей. Данный процесс является абсолютно случайным.

Однако при распределении особей по группам возникает проблема поочередного случайного выбора всех элементов в последовательности, так как генератор случайных чисел имеет возможность спустя какое-то количество итераций повторить уже отобранное число. Поэтому простая проверка на наличие выбранной особи в группах не является корректной из-за того, что цикл становится практически бесконечным при случайном выборе последних элементов последовательности. Таким образом, набор особей в группы идёт из дополнительного списка особей, которые еще не были отобраны. Процесс заполнения продолжается до тех пор, пока доступных особей не останется. На рисунке 5 проиллюстрирован процесс группировки особей.

4.2 Расчет количества особей, которые должны выйти из группы

После разбиения популяции на группы появляется новая проблема, а именно, сколько особей должно выйти из каждой группы и попасть в новое поколение с учётом постоянной численности нового поколения.

Прежде всего требуется рассчитать среднее число особей, которые могут выйти из группы. Расчет происходит по формуле 4.2., которая также была найдена эмпирическим путем.

$$AvgPassCandidate = \frac{PopulationSize}{GroupCount}, \quad (4.2)$$

где *AvgPassCandidate* – среднее число особей, которые могут выйти из группы;
PopulationSize – заданный размер популяции;

GroupCount – количество турнирных групп, которое рассчитывается на этапе группировки особей;

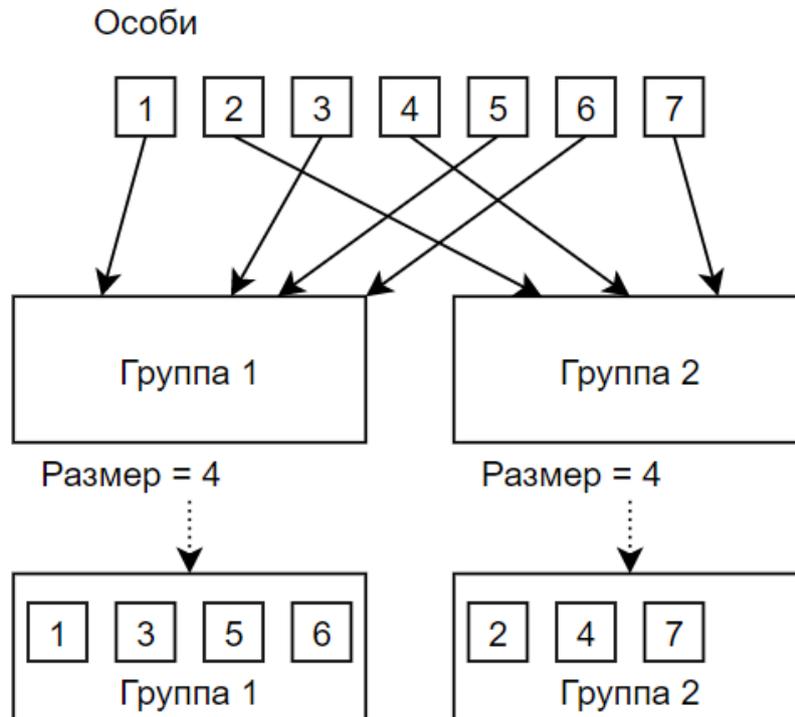


Рисунок 5 – Схема группировки особей

Вообще говоря, рекомендуется задавать размер объединённой популяции (*JoinPopulationSize*) кратным размеру группы, а размер основной популяции кратным количеству групп. Тогда из групп будет отбираться одинаковое количество особей. Однако это условие может и не выполняться. Предположим, что из каждой группы должно выйти *AvgPassCandidate* особей. Далее берется последняя группа, и если она является неполной, а её размер меньше или равен *AvgPassCandidate*, то количество выходящих из нее особей становится равным размеру группы. После этого идет суммирование значений количества выходящих особей из групп. Если это значение оказывается меньше, чем заданный размер популяции, то далее следует процесс дозаполнения, то есть поочередное увеличение для каждой группы значения выходящих в новую популяцию особей. Процесс дозаполнения продолжается до тех пор, пока разница между суммой особей выходящих из каждой группы и заданным размером популяции не станет равна нулю. Таким образом, на выходе будет получен список, который будет содержать количество выходящих особей из каждой группы. На рисунке 6 изображены группы, где красным цветом отмечены особи, которые пройдут отбор.

4.3 Отбор в турнирной селекции

После того, как все особи были разбиты на группы и для каждой этих групп было рассчитано количество выходящих из нее особей, следует непосредственно отбор в группах.

В классической турнирной стратегии отбор представляет собой сортировку особей по их приспособленности. После чего, в зависимости от значения количества выходящих особей, отбираются первые особи текущей группы.

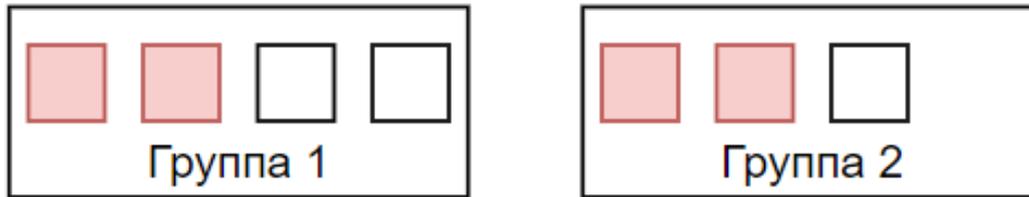


Рисунок 6 – Схема группы с отмеченными особями, которые пройдут отбор

5. Равновероятностная турнирная селекция

В качестве крайнего варианта турнирной стратегии при рекомбинации было предложено равновероятностный способ, при котором особи, вне зависимости от их приспособленности, будут с одинаковой вероятностью попадать в следующее поколение. Отличие от простого случайного отбора особей в новую популяцию состоит исключительно в структуре процесса, полностью соответствующей последовательности турнирной стратегии, что облегчает сравнение с другими вариантами (классическим и равновероятностным). Этапы способа представлены ниже:

Этап №1: Объединение родительской и дочерней популяции. Удаление из объединенной популяции особей, выходящих за границу допустимых решений. Таким образом, в процессе отбора будут участвовать заведомо корректные особи.

Этап №2: Разбиение особей популяции на турнирные группы. Таким образом отбор происходит непосредственно в группах.

Этап №3: Расчет числа выходящих особей. Таким образом для каждой группы известно количество особей, которые должны выйти из неё.

Этап №4: Отбор особей в новую популяцию, используя турнирную стратегию, где особи имеют одинаковую вероятность выхода из группы.

«Равновероятностная турнирная селекция» имеет 2 вида реализации оператора рекомбинации – с выходом дублирующих особей в следующее поколение и без. Оба этих варианта подразумевают использование генератора случайных чисел на этапе отбора. Для варианта без дубликатов используется дополнительный список, который будет содержать оставшиеся в группе особи.

В равновероятностном турнире с выходом дублирующих особей при помощи генератора случайных чисел на диапазоне от 0 до индекса последнего элемента группы будет получено число, которое соответствует индексу особи в этой группе. После чего по этому индексу выбирается особь, которая добавляется в следующее поколение. Процесс турнира продолжается пока не будет отобрано требуемое количество особей. После чего турнир проводится для следующей группы.

В равновероятностном турнире без дублирующих особей вначале происходит инициализация списка доступных особей в группе, после чего он принимает в качестве

значения турнирную группу. Далее начинается отбор. Используя генератор случайных чисел на диапазоне от 0 до индекса последнего элемента списка доступных особей группы, будет получено число, которое соответствует индексу особи в данном списке. По этому индексу выбирается особь, которая добавляется в следующее поколение, а также удаляется из списка доступных элементов. Процесс продолжается до тех пор, пока не будет отобрано необходимое число особей. После этого турнир проводится для других групп.

6. Разновероятностная турнирная селекция

Вторым (и основным) вариантом реализации оператора рекомбинации на основе турнирной стратегии со случайным исходом является способ на основе разновероятностной турнирной селекции. В данном способе каждая особь в зависимости от ее приспособленности имеет свою вероятность выхода из группы в соответствии с принципом рулетки (в пределах группы).

От предыдущего, рассматриваемый способ отличается этапом №4.

Этап отбора является комбинацией использования турнирной селекции и колеса рулетки. Таким образом, в каждой турнирной группе происходит вращение колеса рулетки, и особь, значение сектора которой соответствует выбранному при помощи генератора случайных чисел значению, отбирается в следующее поколение. Количество вращений колеса соответствует требуемому количеству особей, которые должны выйти из группы.

«Разновероятностная турнирная селекция» также имеет 2 вида реализации – с выходом из группы дублирующих особей и без.

Заключение

В статье предложено 2 варианта способа реализации оператора отбора особей в новое поколение генетического алгоритма (оператора рекомбинации) на основе турнирной стратегии со случайным исходом. Варианты способа имеют общую структуру и отличаются применяемыми алгоритмами отбора особей из турнирных групп. Первый вариант, на основе равновероятностного отбора, формирует вырожденную стратегию эквивалентную случайному отбору. Второй вариант, на основе разновероятностного отбора позволяет сочетать свойства турнирной стратегии и стратегии на основе колеса рулетки. В дальнейшем предстоит провести практический анализ результатов применения рассмотренных вариантов, а также рассмотреть способы на основе обратного турнира – в котором выбираются выбывающие особи.

Список литературы

1. Батищев Д.И., Неймарк Е.А., Старостин Н.В. Применение генетических алгоритмов к решению задач дискретной оптимизации. Учебно-методический материал по программе повышения квалификации «Информационные технологии и компьютерное моделирование в прикладной математике». – Нижний Новгород: Нижегородский государственный университет им. Н.И. Лобачевского, 2007. – 85 с.
2. Букатова И. Л., Михасев Ю. И., Шаров А. М. Эвоинформатика: Теория и практика эволюционного моделирования. – М.: Наука, 1991. – 206 с.

3. Бураков М. В. Генетический алгоритм: теория и практика: учеб. пособие – СПб.: ГУАП, 2008. – 164 с.
4. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы / под ред. В.М. Курейчика. – 2-е изд., испр. и доп. – М.: Физматлит, 2006. – 320 с.
5. Емельянов В. В., Курейчик В. В., Курейчик В. М. Теория и практика эволюционного моделирования. – М.: Физматлит, 2003. – 432 с.
6. Панченко Т.В. Генетические алгоритмы: учебно-методическое пособие / под ред. Ю. Ю. Тарасевича. – Астрахань: Издательский дом «Астраханский университет», 2007. – 87 с
7. Genetic Algorithms - Survivor Selection// Материалы сайта www.tutorialspoint.com [Электронный ресурс]: - Режим доступа: https://www.tutorialspoint.com/genetic_algorithms/genetic_algorithms_survivor_selection.htm, свободный. – Загл. с экрана. – Яз. англ. Дата доступа: 20 сент. 2020
8. Luke S. Essentials of Metaheuristics. A Set of Undergraduate Lecture Notes. Zeroth Edition. Online Version 2.3. February, 2016 [Электронный документ]: - Режим доступа: <https://cs.gmu.edu/~sean/book/metaheuristics/Essentials.pdf>, свободный. – Загл. с экрана. – Яз. англ. Дата доступа: 20 сент. 2020

References

1. Batishchev D.I., Neymark E.A., Starostin N.V. Application of genetic algorithms to solving discrete optimization problems. Educational and methodical material according to the development program "Information technologies and computer modeling in applied mathematics". – Nizhny Novgorod: N.I. Lobachevsky State University of Nizhni Novgorod, 2007. – 85 p.
 2. Bukatova I. L., Mikhasev Yu. I., Sharov A. M. Evoinformatics: Theory and practice of evolutionary modeling. – М.: Science, 1991. - 206 p.
 3. Burakov M.V. Genetic algorithm: theory and practice: textbook – St. Petersburg: GUAP, 2008. – 164 p.
 4. Gladkov L.A., Kureichik V.V., Kureichik V.M. Genetic algorithms/ed. V.M. Kureichik. – 2nd ed., Correct and additional – М.: Fizmatlit, 2006. – 320 p.
 5. Emelyanov V.V., Kureichik V.V., Kureichik V.M. Theory and practice of evolutionary modeling. – М.: Fizmatlit, 2003. – 432 p.
 6. Panchenko T.V. Genetic algorithms: educational manual/ed. Yu. Yu. Tarasevich. – Astrakhan: Publishing House "Astrakhan University," 2007. – 87 p.
 7. Genetic Algorithms - Survivor Selection. Available at: https://www.tutorialspoint.com/genetic_algorithms/genetic_algorithms_survivor_selection.htm (accessed 20 September 2020)
 8. Luke S. Essentials of Metaheuristics. A Set of Undergraduate Lecture Notes. Zeroth Edition. Online Version 2.3. February, 2016. Available at: <https://cs.gmu.edu/~sean/book/metaheuristics/Essentials.pdf> (accessed 20 September 2020)
-