

ТЕХНИЧЕСКИЕ НАУКИ

УДК 681.3.681.5

Р.А. НЕЙДОРФ, В.Г. КОБАК, Д.В. ТИТОВ

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ЭФФЕКТИВНОСТИ ВАРИАНТОВ ТУРНИРНОГО ОТБОРА ГЕНЕТИЧЕСКОГО АЛГОРИТМА РЕШЕ- НИЯ ОДНОРОДНЫХ РАСПРЕДЕЛИТЕЛЬНЫХ ЗАДАЧ

В данной работе дана оценка турнирного отбора, применяемого в генетическом алгоритме для однородных систем обработки информации, путем сравнения стандартной схемы турнирного отбора с её модификациями при помощи вычислительного эксперимента.

Ключевые слова: теория расписаний, задача планирования, генетический алгоритм, турнирный отбор, вычислительный эксперимент, вектор загрузки.

Введение. В настоящее время для решения задач из различных областей науки и производства интенсивней стали применяться многопроцессорные вычислительные системы обработки данных, которые резко увеличивают скорость нахождения результатов. Но для эффективного выполнения множества независимых задач, имеющих конкретные временные характеристики, на данных системах необходимо оптимально спланировать распределение заданий по процессорам. Для этого применяют различные алгоритмы распределения задач по параллельно работающим устройствам. При этом в зависимости от алгоритма они могут возвращать как оптимальные решения (точные), так и близкие к оптимальным (списочные, эволюционные). В случаях, когда необходимо достаточно быстро получить результаты, а при этом допускается решение близкое к оптимальному, применяют приближенные алгоритмы. Это обусловлено большими вычислительными и временными затратами применения точных алгоритмов, а при решении задач большой размерности получение точного решения крайне проблематично.

Постановка задачи. Пусть вычислительная система состоит из n идентичных параллельных процессоров $P = \{p_1, \dots, p_n\}$. Множество независимых заданий $Q = \{q_1, \dots, q_m\}$ образует параллельную программу, причем известно время выполнения j -го задания t_j на любом из процессоров вычислительной системы, где $j = \overline{1, m}$. Требуется найти такое распределение заданий по процессорам, при котором время завершения T_{\max} параллельной программы минимально, т.е.

$$T_{\max} = \max_{1 \leq i \leq n} \{T_i\} \rightarrow \min, \quad (1)$$

где $T_i = \sum_{t_j \in Q_i} t_j$ – загрузка i -го процессора (время окончания выполнения множества заданий $Q_i \subseteq Q$, назначенных на процессор p_i , где $i = \overline{1, n}$).

С увеличением числа заданий и количества используемых приборов сложность построения оптимальных расписаний экспоненциально возрастает. Поэтому большое значение имеют простые в построении и удобные в

реализации приближенные расписания параллельных вычислительных процессов, близкие к оптимальным распределениям с точки зрения обеспечиваемого времени выполнения параллельных программ [1].

Генетические алгоритмы. Генетические алгоритмы (ГА) относятся к эвристическим методам и представляют собой алгоритмы поиска лучшего решения, а не оптимального, используя при этом принципы, сходные с принципами естественного отбора и генетики.

Генетические алгоритмы имеют вероятностную природу, поэтому полученные результаты отличаются друг от друга при каждом запуске алгоритма и определяются случайной последовательностью, переданной в схему алгоритма. Точность алгоритма зависит не только от входной последовательности случайных чисел, но и от условий задачи, таких как размерность задачи и конкретное распределение весов.

Общая схема работы генетического алгоритма следующая:

Ш.1 Формирование начального поколения, состоящего из заданного числа особей.

Ш.2 Отбор особей и применение операторов кроссовера и мутации ГА с заданной вероятностью, формирование нового поколения.

Ш.3 Проверка условия останова, если не выполнилось, то переход на Ш.2.

Ш.4 Лучшая особь выбирается как найденное решение.

Такая схема является наиболее общим алгоритмом для решения разнотипных задач. В данном случае решения задачи теории расписания минимаксный критерий будет являться оптимизационной функцией, а условием останова будет неизменность лучшего решения в течение заданного числа поколений. При отборе родительских пар для оператора кроссовера берется текущая особь и случайно выбранная из исходного вектора особей. Для формирования нового поколения используется турнирный отбор, при котором из заданного числа особей (в данном случае две особи) выбирается лучшая, которая перейдет в новое поколение. Лучшей особью считается особь, для которой значение оптимизационной функции минимально.

Общий принцип работы базового генетического алгоритма на Ш.2 следующий: на вход алгоритма поступает исходный вектор особей Ch_i ; по порядку, начиная с первой, выбирается очередная особь j ; затем, случайно, выбирается особь l , и с заданной вероятностью P_{cross} выполняется оператор кроссовера; далее выполняется оператор мутации с заданной вероятностью P_{mut} ; потом выполняется турнирный отбор для результирующей особи и случайно выбранной особи l ; если обработаны не все особи в векторе Ch_i , то процесс повторяется, иначе формируется новое поколение Ch_{i+1} . Схематически работа генетического алгоритма на Ш.2 представлена на рис.1.

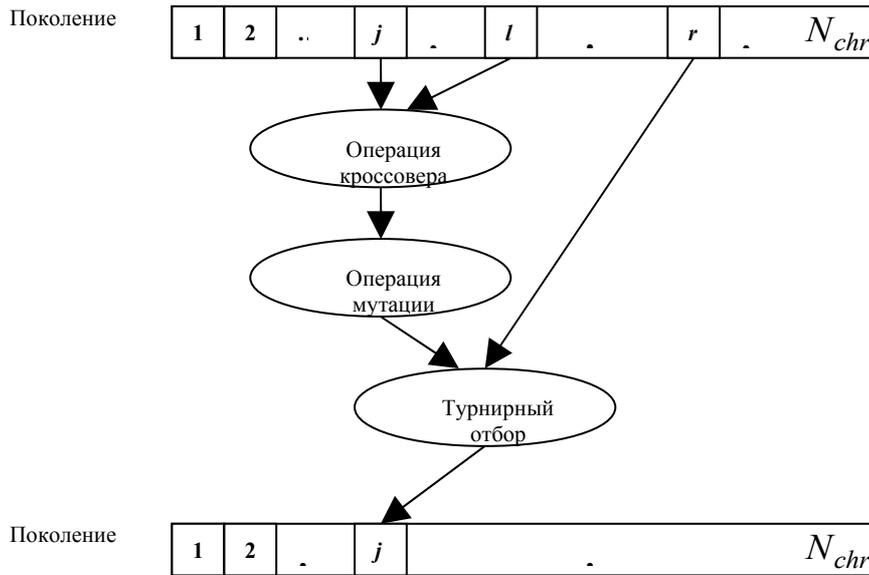


Рис. 1. Схема работы базового ГА на Ш.2

Модификация выбора родительских пар. В данной работе рассмотрена модификация (по сравнению с базовой схемой ГА) выбора родительских пар для оператора кроссовера, которая заключается в использовании в качестве оператора выбора турнирного отбора. В турнирном отборе будут участвовать две (бинарный турнирный отбор) и четыре случайно выбранные особи [2], из которых лучшая особь будет скрещиваться с очередной особью из вектора особей.

При данной модификации алгоритм работы генетического алгоритма на Ш.2 с использованием бинарного турнирного отбора выглядит следующим образом: на вход алгоритма поступает исходный вектор особей Ch_i ; затем, по порядку, начиная с первой, выбирается очередная особь j ; если сработает оператор кроссовера с заданной вероятностью P_{cross} , то выполняется турнирный отбор для случайно выбранных двух особей l и k , затем выполняется операция кроссовера для очередной и лучшей особей; потом выполняется оператор мутации с заданной вероятностью P_{mut} ; затем выполняется турнирный отбор для результирующей особи и случайно выбранной особи r ; если обработаны не все особи в векторе Ch_i , то процесс повторяется, иначе формируется новое поколение Ch_{i+1} . Схематически работа генетического алгоритма на Ш.2 при использовании данной модификации приведена на рис.2. При использовании турнирного отбора для четырех случайно выбранных особей алгоритм выглядит аналогично.

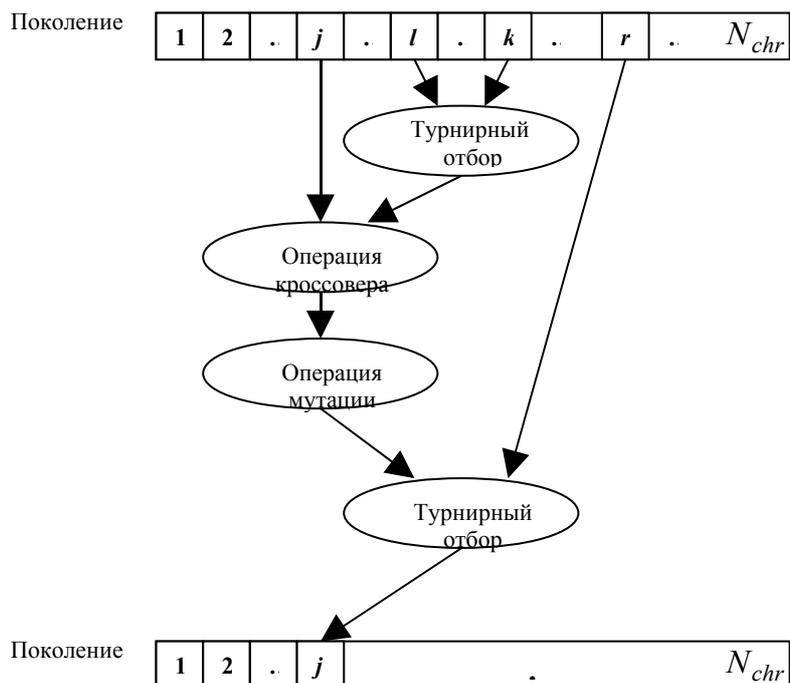


Рис. 2. Схема работы ГА на Ш.2 с использованием турнирного отбора для выбора родительских пар

Модификация формирования нового поколения. В целях повышения производительности работы генетического алгоритма предлагается изменить существующий алгоритм турнирного отбора для формирования нового поколения. Изменение заключается в том, что в турнирном отборе будут участвовать очередная особь и результирующая, полученная в ходе выполнения операторов кроссовера и мутации.

При использовании данной модификации генетический алгоритм на Ш.2 примет следующий вид: на вход алгоритма поступает исходный вектор особей Ch_i ; по порядку, начиная с первой, выбирается очередная особь j ; затем выполняется оператор кроссовера с заданной вероятностью P_{cross} для очередной особи j и случайно выбранной особи l ; далее выполняется оператор мутации с заданной вероятностью P_{mut} ; потом выполняется турнирный отбор для очередной особи j и результирующей особи; если обработаны не все особи в векторе Ch_i , то процесс повторяется, иначе формируется новое поколение Ch_{i+1} . Схематически работа генетического алгоритма при данной модификации на Ш.2 представлена на рис. 3.

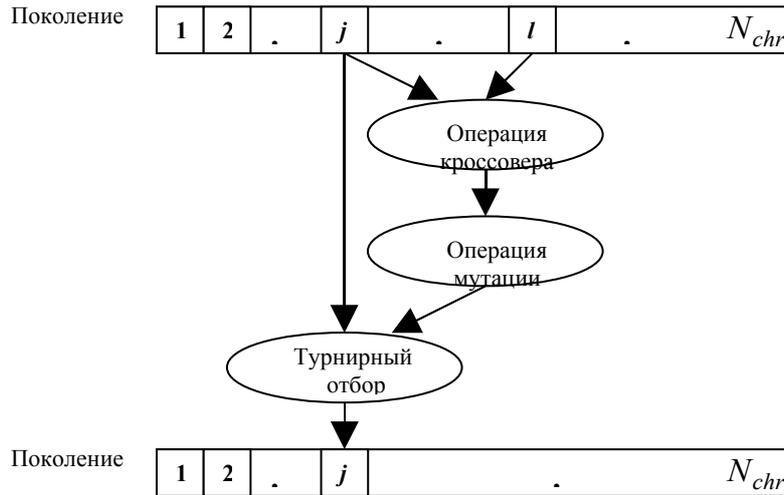


Рис.3. Схема работы ГА на Ш.2 с использованием турнирного отбора для формирования нового поколения

В данный алгоритм было внесено изменение, заключающееся в применении двоичного турнирного отбора для выбора одной из особей для оператора кроссовера вместо случайного выбора особи l из вектора Ch_i .

Методы испытаний. Для сравнительного исследования алгоритмов был проведен ряд вычислительных экспериментов для 2-4-приборных систем обработки информации, которые являются наиболее распространенными [3,4]. Количество работ, распределяемых между приборами, увеличивалось с 25 до 30.

Для проведения эксперимента были случайным образом сгенерированы 100 векторов загрузки, содержащие работы в диапазоне 25, 30; каждый из векторов решался базовым генетическим алгоритмом и предложенными модификациями данного алгоритма. Для всех используемых генетических алгоритмов были выбраны следующие фиксированные параметры: число особей составляло 50; условием останова являлось появление в процессе решения более 500 поколений с лучшей одинаковой особью; вероятность кроссовера составляла 90%, а вероятность мутации 10%.

Полученные результаты усреднялись по количеству экспериментов, т.е.

$$\bar{T}_{\max} = \frac{\sum_{i=1}^{100} T_{\max i}}{100}, \quad (2)$$

где T_{\max} определялось по (1), а также аналогично усреднялись по времени решения задач T_{cp} .

Основные результаты вычислительного эксперимента отображены на рис. 4-9.

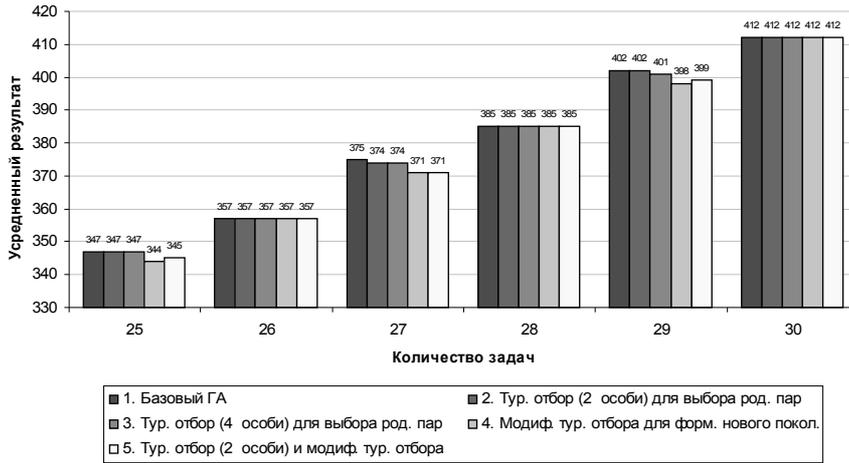


Рис.4. Усредненный результат работы алгоритмов при n=2

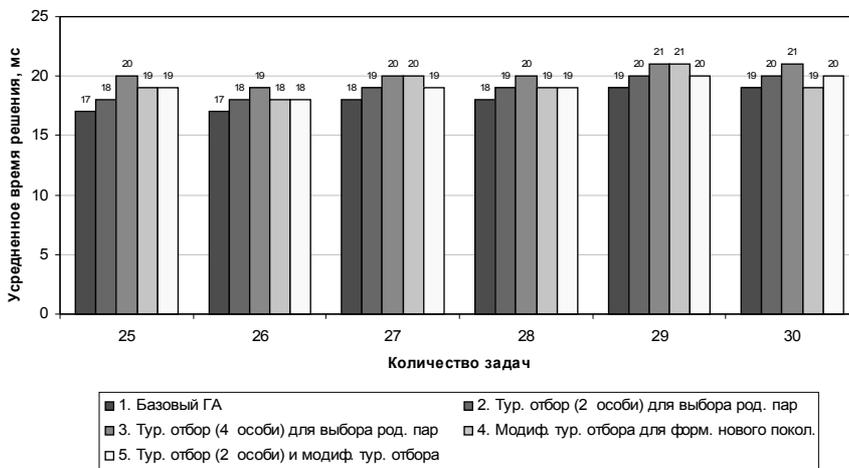


Рис.5. Усредненное время решения алгоритмов при n=2

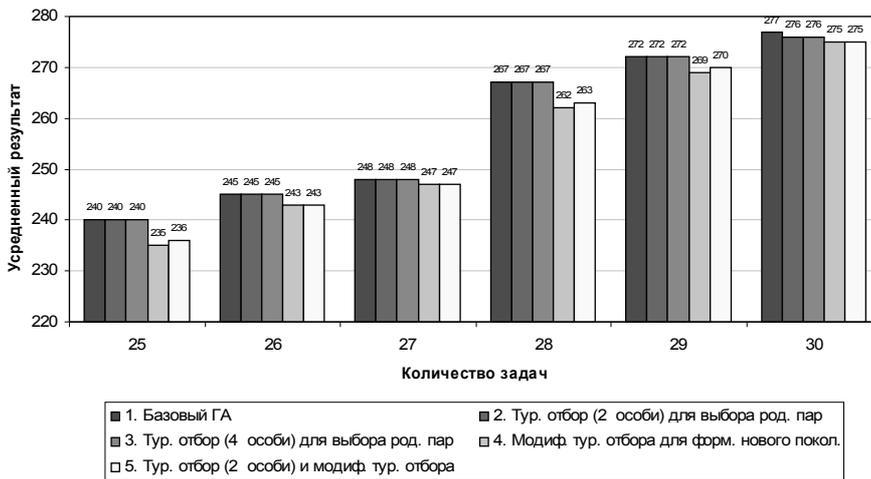


Рис.6. Усредненный результат работы алгоритмов при n=3

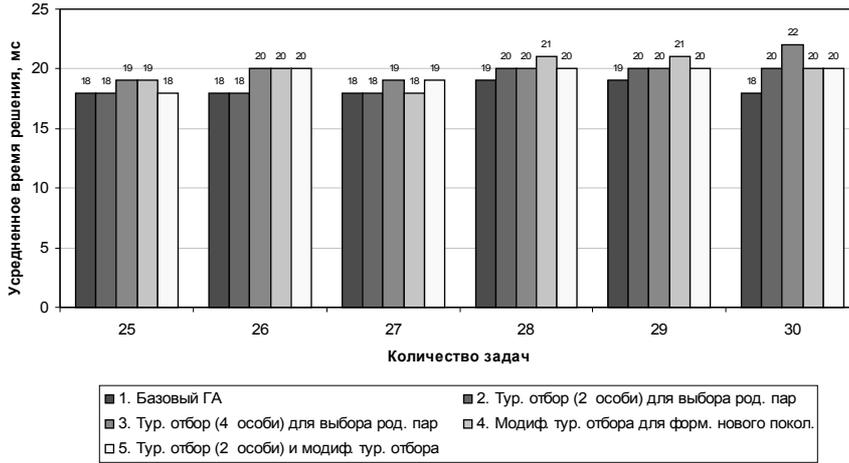


Рис. 7. Среднее время решения алгоритмов при n=3

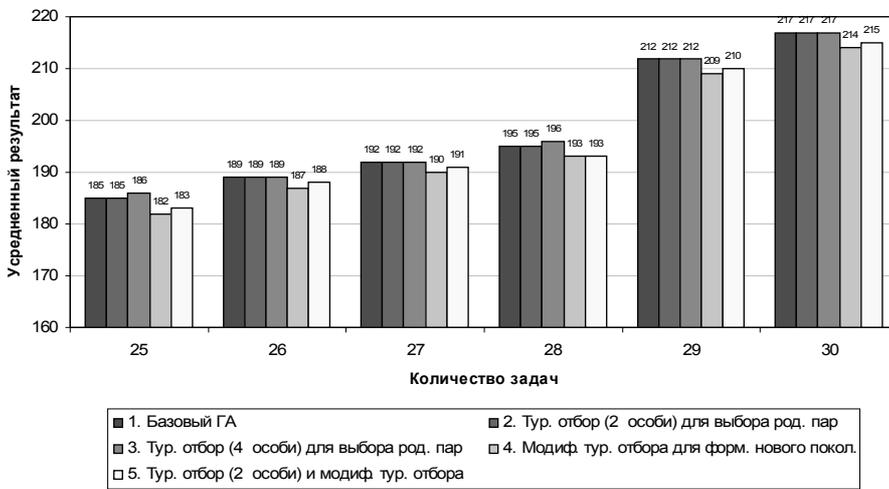


Рис. 8. Средний результат работы алгоритмов при n=4

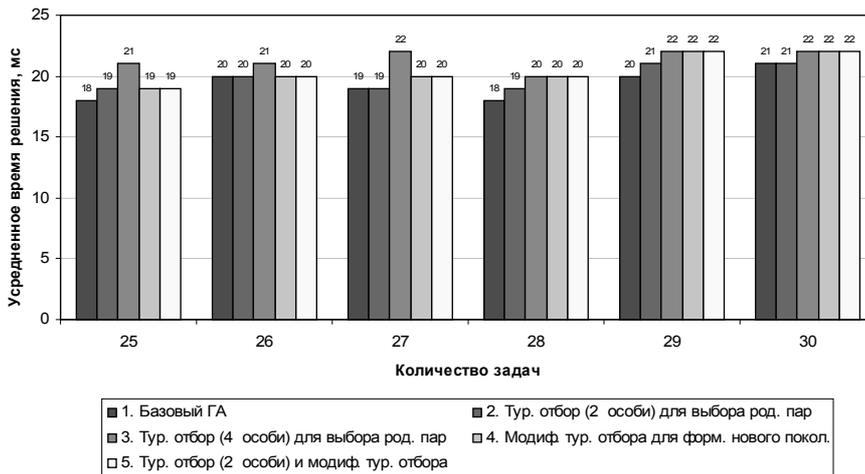


Рис. 9. Среднее время решения алгоритмов при n=4

Проанализировав результаты вычислительного эксперимента, можно прийти к следующим выводам, основанным на совокупном сопоставительном анализе диаграмм рис.4-9.

1. Предложенная модификация турнирного отбора для формирования нового поколения дает лучшие результаты, не увеличивая среднее время решения.

2. С увеличением числа особей, участвующих в турнирном отборе для выбора родительской пары с 2 до 4, результат не улучшается, а среднее время решения алгоритма увеличивается, поэтому следует ограничиться одной родительской парой.

3. Из приведенной диаграммы на рис.4 видно, что при количестве процессоров равных двум и при кратности количества задач количеству процессоров усредненные результаты работы предложенных алгоритмов совпадают. Однако из диаграмм, изображенных на рис. 6 и 8, видно, что при количестве процессоров больше двух эта тенденция значительно ослабляется.

4. Среднее время работы предложенных алгоритмов мало и практически соизмеримо между собой.

Выводы. Таким образом, предлагаемая модификация турнирного отбора для формирования нового поколения, заключающаяся в использовании в нём не только «потомков», но и «родителей», дает лучшие результаты при практически одинаковом времени решения задачи, поэтому с вычислительной точки зрения она является более эффективной и её можно рекомендовать к использованию.

Библиографический список

1. Головкин Б.А. Расчет характеристик и планирование параллельных вычислительных процессов. / Б.А. Головкин. – М.: Радио и связь, 1983.

2. Норенков И.П. Генетические методы структурного синтеза проектных решений / И.П. Норенков // Информационные технологии. – 1998. – № 1. – С. 9-13.

3. Ларионов А.М. Вычислительные комплексы, системы и сети. / А.М. Ларионов, С.А.Майоров, Г.И. Новиков. – Л.: Энергоатомиздат, 1987.

4. Панфилов И.В., Вычислительные системы. / И.В. Панфилов, А.М. Половко. – М.: Советское радио, 1980.

Материал поступил в редакцию 10.03.09.

R.A.NEYDORF, V.G.KOBAK, D.V.TITOV

**COMPARATIVE ANALYSIS OF EFFICIENCY VARIANTS
OF TOURNAMENT SELECTION OF GENETIC ALGORITHM
OF THE DECISION OF HOMOGENEOUS DISTRIBUTION PROBLEMS**

This paper gives the estimation of the tournament selection, used in the genetic algorithm, for the uniform instruments of information processing, via the comparison of the standard diagram of tournament selection with its modifications, with the aid of the computational experiment.

НЕЙДОРФ Рудольф Анатольевич (р. 1944), заведующий кафедрой «ПОВТ и АС» ДГТУ, доктор технических наук (1988), профессор (1994). Окончил НПИ (1967).

Основные направления научной деятельности: синтез и структурно-параметрическая оптимизация алгоритмов управления, процессы обработки информации и управления в технических системах.

Автор более 200 научных работ, 11 учебных пособий. Имеет 34 авторских свидетельства на изобретения, 9 свидетельств Роспатента о регистрации программ.

КОБАК Валерий Григорьевич (р. 1961), докторант ДГТУ, кандидат технических наук. Окончил НПИ (1983).

Научные интересы: методы решения экстремальных задач в однородных вычислительных системах.

Автор более 70 научных работ.

ТИТОВ Дмитрий Вячеславович (р. 1981), аспирант ДГТУ. Окончил магистратуру в ДГТУ (2005).

Научные интересы: методы решения минимаксной задачи в однородных вычислительных системах точными и приближенными алгоритмами.

Автор 9 научных работ.

neyruan@yandex.ru