

Программная система PARETO RATING для оценки качества Парето-аппроксимации в задаче многокритериальной оптимизации

07, июль 2014

DOI: 10.7463/0714.0720253

Грошев С. В.¹, профессор, д.ф.-м.н. Карпенко А. П.¹, Сабитов Д. Р.², Шибитов И. А.²

УДК 519.6

¹Россия, МГТУ им. Н.Э. Баумана,

²Россия, ООО "Крок"

^aapkarpenko@mail.ru

Рассматриваем задачу оценки качества численной аппроксимации множества (фронта) Парето в задаче многокритериальной оптимизации (МКО-задаче). Имеем в виду, что Парето-аппроксимация получена с помощью того или иного популяционного, например, генетического алгоритма. В конечном счете, целью работы является сравнительная оценка эффективности популяционных алгоритмов Парето-аппроксимации. Разработано большое число характеристик (индикаторов) качества Парето-аппроксимации. Поэтому задачу оценки качества Парето-аппроксимации рассматриваем также как многокритериальную (многоиндикаторную). Известен ряд программных систем, которые в разной степени решают задачу оценки качества Парето-аппроксимации. Общим недостатком этих систем является отсутствие WEB-интерфейса, а также отсутствие поддержки многоиндикаторной оценки качества Парето-аппроксимации (хотя поддержка вычисления значений большого числа этих индикаторов имеется). Программная система PARETO RATING призвана устранить указанные недостатки известных систем.

Ключевые слова: задача многокритериальной оптимизации; Парето-аппроксимация; индикаторы качества; методы оценки качества Парето-аппроксимации

Введение

Постановка задачи многокритериальной оптимизации (МКО-задачи) фиксирует лишь множество допустимых значений вектора варьируемых параметров и вектор целевых функций. Как правило, этой информации недостаточно для однозначного решения задачи - данная информация позволяет лишь выделить соответствующее множество Парето (а, тем самым и фронт Парето). Поэтому часто говорят, что решением МКО-задачи является множество Парето этой задачи, точнее – аппроксимация этого множества (Парето-аппроксимация).

Выделяют не популяционные и популяционные алгоритмы построения Парето-аппроксимации. В отличие от не популяционных алгоритмов, в популяционных алгоритмах реализуется одновременный поиск не одной, а большого числа недоминируемых точек, близких к множеству Парето. В настоящее время в вычислительной практике наиболее широкое применение находят популяционные алгоритмы. Ограничиваемся поэтому рассмотрением этих алгоритмов.

В силу высокой практической значимости МКО-задач разработано большое число популяционных алгоритмов Парето-аппроксимации [1]. Поэтому актуальной является задача сравнительной оценки эффективности этих алгоритмов.

В практически значимых МКО-задачах, как правило, используется более двух целевых функций, и поэтому визуальный анализ качества соответствующих Парето-аппроксимаций затруднён. В содержательных терминах качество Парето-аппроксимации может быть оценено с помощью следующих характеристик: близость найденных решений к точному множеству Парето рассматриваемой МКО-задачи; равномерность распределения решений в полученной Парето-аппроксимации; мощность найденного множества решений. Разработано большое число характеристик (индикаторов) качества Парето-аппроксимации, формализующих указанные характеристики [2, 3]. Таким образом, задача оценки качества Парето-аппроксимации сама является многокритериальной (многоиндикаторной).

Известен ряд программных систем [4], которые в разной степени решают задачу оценки качества Парето-аппроксимации. Общим недостатком этих систем является отсутствие *WEB*-интерфейса, а также отсутствие поддержки многоиндикаторной оценки качества Парето-аппроксимации (хотя поддержка вычисления значений большого числа этих индикаторов имеется). Представляемая в данной работе программная система *PARETO RATING* призвана устранить указанные недостатки известных систем.

Популяционные алгоритмы Парето-аппроксимации являются, как правило, стохастическими. Поэтому в случае многократного запуска такого алгоритма для одной и той же МКО-задачи, каждый раз получается, вообще говоря, новая аппроксимация. Известны три основных класса статистических методов оценки качества двух и более Парето-аппроксимаций: методы на основе ранжирования указанных аппроксимаций; методы на основе индикаторов качества; методы на основе так называемых эмпирических функций достижимости [5].

В первом разделе работы приводим формальную постановку МКО-задачи и общую схему популяционных алгоритмов её решения. Второй раздел содержит обзор известных индикаторов качества Парето-аппроксимации. В третьем разделе представлен обзор статистических методов оценки качества Парето-аппроксимаций. Четвёртый и пятый разделы посвящены описанию архитектуры системы *PARETO RATING* и основным особенностям её программной реализации соответственно. Шестой раздел иллюстрирует эффективность принятых алгоритмических и программных решений. В заключении формулируем основные результаты работы и перспективы её развития.

1. Постановка МКО-задачи и общая схема алгоритмов её решения

Запись вида $|A|$, где A – счётное множество, далее означает мощность (число элементов) этого множества; запись вида $|B|$, где B – вектор, означает его размерность.

Полагаем, что множеством допустимых значений вектора варьируемых параметров X является множество

$$D_X = \{X \mid G(X) \geq 0\} \subset \{X\} = R^{|X|},$$

где $G(X)$ - ограничивающая вектор-функция.

Критериальная вектор-функция $F(X) = (f_1(X), f_2(X), \dots, f_{|F|}(X))$ со значениями в критериальном пространстве $\{F\} = R^{|F|}$ определена в области D_X . Лицо, принимающее решения (ЛПР), стремится минимизировать в этой области каждую из частных критериальных функций $f_1(X), f_2(X), \dots, f_{|F|}(X)$.

Множество достижимости МКО-задачи обозначаем $D_F \subset \{F\}$, фронт Парето - $D_F^* \subset D_F$, а множество Парето - $D_X^* \subset D_X$ [6].

Пусть A^F, A^X - так называемые, архивные множества, содержащие не доминируемые точки F_j^A и соответствующие им точки X_j^A ; $j \in [1:|A|]$. Идея популяционных алгоритмов Парето-аппроксимации состоит в итерационном уточнении множеств точек в архивах A^F, A^X . Если при этом на данной итерации появляется новая точка F_i , доминирующая некоторые точки из архива A^F , то все доминируемые точки, а также соответствующие им точки из архива A^X удаляем. При удовлетворении некоторого критерия останова, текущее содержимое архивов A^F, A^X полагаем искомой аппроксимацией фронта D_F^* и множества Парето D_X^* соответственно.

Рассмотрим две некоторые Парето-аппроксимации A^1, A^2 . Далее нам понадобятся для этих аппроксимаций отношения предпочтения, указанные в таблице 1

Таблица 1. Отношения предпочтения для Парето-аппроксимаций A^1, A^2

Отношение	Интерпретация в пространстве критериев	
Строго доминирует	$A^1 \prec\prec A^2$	Каждый $F^2 \in A^2$ строго доминируется хотя бы одним $F^1 \in A^1$
Доминирует	$A^1 \prec A^2$	Каждый $F^2 \in A^2$ доминируется хотя бы одним $F^1 \in A^1$

Лучше	$A^1 \triangleleft A^2$	Каждый $F^2 \in A^2$ слабо доминируется хотя бы одним $F^1 \in A^1$ и $A^1 \not\sim A^2$
Слабо доминирует	$A^1 \preceq A^2$	Каждый $F^2 \in A^2$ слабо доминируется хотя бы одним $F^1 \in A^1$
Несравнимы	$A^1 \parallel A^2$	Ни $A^1 \preceq A^2$, ни $A^2 \preceq A^1$
Безразличны	$A^1 \sim A^2$	$A^1 \preceq A^2$ и $A^2 \preceq A^1$

2. Индикаторы качества Парето-аппроксимации

Унарные индикаторы. Среднее расстояние до точного фронта (*Generational Distance*) определяет формула

$$I_{GD}(A) = \frac{\left(\sum_{j=1}^{|A|} \|F_j, F_j^*\|_E \right)^{1/|F|}}{|A|} \rightarrow \min,$$

где $\|\cdot\|_E$ - символ евклидовой нормы; F_j^* - ближайшая к точке F_j точка множества D_F^* .

Индикатор характеризует близость найденных решений к точному фронту Парето D_F^* .

Среднее рассеяние (*Spacing*) $I_S(A)$ представляет собой меру равномерности распределения решений, полученных в результате Парето-аппроксимации. Индикатор задаёт формула

$$I_S(A) = \sqrt{\frac{1}{|A|-1} \sum_{j=1}^{|A|} \text{abs}(\bar{d} - d_j)} \rightarrow \min,$$

где $d_j = \min_{k \in [1:|A|], k \neq j} \|F(X_j), F(X_k)\|_M$ - минимальное манхеттоновское расстояние между архивным решением $F(X_j)$ и остальными архивными решениями [2]; $\bar{d}$ - среднее

всех этих величин.

Максимальное рассеяние (*Maximum Spread*) $I_{MS}(A)$ определяется аналогично индикатору $I_S(A)$:

$$I_{MS}(A) = \sqrt{\sum_{j=1}^{|A|} \max_{k \in [1:|A|]} \|F_j, F_k\|_M} \rightarrow \min.$$

Отклонение от равномерного распределения (*Deviation from Uniform distribution*) $I_{DU}(A)$, аналогично индикаторам $I_S(A)$, $I_{MS}(A)$, определяет равномерность распреде-

ления архивных не доминируемых решений A в целевом пространстве. Индикатор задаёт формула

$$I_{DU}(A) = \sum_{j=1}^{|A|} \frac{abs(d_j - \bar{d})}{|A|},$$

где d_j - евклидово расстояние решения F_j до ближайшего из решений множества A ; $\bar{d}$ - среднее этих величин.

Мощность множества решений (*Overall Nondominated Vector Generation*) $I_{ONVG}(A)$ есть не что иное, как число элементов аппроксимации A , то есть

$$I_{ONVG}(A) = |A| \rightarrow \max.$$

Объем объемлющего гиперкуба (*Hypercube enclosing indicator*) $I_{HC}(A)$ определяют величины [2]

$$I_{HC_1}(A) = \max_{P \in R^{|F|}} (P \succ F_j, \forall j \in [1:|A|]) = p, \quad P = (p, p, \dots, p),$$

$$I_{HC_2}(A) = \min_{Q \in R^{|F|}} (Q \succ F_j, \forall j \in [1:|A|]) = q, \quad Q = (q, q, \dots, q).$$

Относительное число ошибочных решений (*Error Ratio*) $I_{ER}(A)$ - относительное число решений аппроксимации A , не принадлежащих точному фронту Парето D_F^* рассматриваемой МКО-задачи:

$$I_{ER}(A) = \frac{1}{|A|} \sum_{j=1}^{|A|} e_j \rightarrow \min, \quad e_j = \begin{cases} 1, & F_j \notin D_F^*, \\ 0, & F_j \in D_F^*. \end{cases}$$

Максимальная ошибка аппроксимации (*Maximum Pareto front Error*) $I_{MPE}(A)$ также для своего вычисления требует знания точного фронта Парето. Индикатор определяет формула

$$I_{MPE}(A) = \max_{j \in [1:|A|]} \|F_j, F_j^*\|_M^{|F|} \rightarrow \min,$$

где, как и ранее, F_j^* - ближайшая к точке F_j точка множества D_F^* .

Минимальная ошибка аппроксимации (*Spacing Distribution in ε -area*)

$$I_{SD(\varepsilon)}(A) = I_{\varepsilon}(D_F^*, A)$$

представляет собой минимальную величину ε , для которой справедливо отношение $F_j \succ_{\varepsilon} F_k^*$ для любых $F_j \in A$ и любых $F_k^* \in D_F^*$ [2]. Для определения значения индикатора $I_{SD(\varepsilon)}(A)$ необходимо знание точного фронта Парето D_F^* .

Индикатор протяжённости (*Dimensions Extent*) $I_{DE}(A)$ позволяет оценить суммарную протяжённость Парето-аппроксимации в критериальном пространстве во всем $|F|$ измерениям. Значения индикатора вычисляются по формуле

$$I_{DE}(A) = \sqrt{\sum_{j=1}^{|F|} \max_{k,l \in [1:|A|]} |f_j(X_k^A) - f_j(X_l^A)|} \rightarrow \max.$$

Бинарные индикаторы. Положим, что аппроксимация $A_*^F = A_*$ недоминируемых решений в критериальном пространстве D_F является «эталонной», а аппроксимация $A^F = A$ - аналогичной оцениваемой аппроксимацией.

ε -индикатор $I_\varepsilon(A_*, A)$ строят на основе унарного критерия $I_\varepsilon(A_1, A_2)$. Величина $I_\varepsilon(A_*, A)$ формализует близость аппроксимации A к аппроксимации A_* как минимальную величину ε , при которой для любого вектора $F_j \in A$ существует такой вектор $F_k \in A_*$, что имеет место отношение $F_j \succ_\varepsilon F_k$.

ε_+ -индикатор $I_{\varepsilon_+}(A_*, A)$ определяем аналогичным образом.

Индикатор $I_C(A_*, A)$, называемый покрытием (*Coverage*), может быть использован в том случае, когда аппроксимация A_* является счётной. Величина $I_C(A_*, A)$ имеет смысл относительного числа решений, принадлежащих аппроксимации A_* , которые доминируются некоторыми из решений аппроксимации A :

$$I_C(A_*, A) = \frac{|\{F_k \in A_* \mid \exists F_j \in A: F_j \succ F_k\}|}{|A_*|} \rightarrow \max.$$

Индикатор гиперобъёма разности покрытий (*HyperVolume of coverage difference*) $I_{HV}(A_*, A)$ строят на основе величины, которая называется разностью покрытий (*coverage differences*) $\delta(A_1, A_2)$. В качестве меры $\zeta(\cdot)$ индикатор использует объем соответствующей части множества достижимости [2].

Ряд индикаторов качества Парето-аппроксимации построен Хансеном (*M. P. Hansen*) и Жажкевичем (*A. Jaszkiwicz*) в 1998 г. на основе функции предпочтений ЛПП (*decision-maker's utility function*) [7].

3. Методы оценки эффективности алгоритмов Парето-аппроксимации

Положим, что речь идёт о сравнении эффективности $n \geq 2$ стохастических алгоритмов Парето-аппроксимации. Пусть выполнено $r_i \geq 1$ запусков i -го алгоритма, в ре-

результате чего получена коллекция аппроксимаций

$$C = \{A_1^1, A_2^1, \dots, A_{r_1}^1, \dots, A_1^n, \dots, A_{r_n}^n, i \notin [1:n]\}.$$

Методы оценки на основе ранжирования. Отношения, представленные в таблице 1, могут быть использованы для установления частичного порядка в коллекции аппроксимаций C . На основе этого порядка можно определить показатели качества или ранги для каждой из аппроксимаций, входящих в коллекцию.

Известно несколько подходов к определению ранга (*rank*) аппроксимации на основе отношения доминирования. Например, ранг можно принять равным числу аппроксимаций, которые доминирует данная аппроксимация [8]. Можно использовать недоминируемую сортировку коллекции C [9]. Последний подход в сочетании с использованием отношения «Лучше» (таблица 1) определяет формула

$$\text{rank}(A^i) = 1 + |\{A^j \in C : A^j \triangleleft A^i\}|.$$

Здесь, напомним, символ $|C|$ означает мощность множества C . Наименьший ранг соответствует наилучшей аппроксимации в коллекции.

Поскольку речь идёт о стохастических алгоритмах Парето-аппроксимации, к полученным рангам алгоритмов следует применить статистический ранговый критерий [10].

Методы на основе индикаторов качества. Представленные выше индикаторы ставят в соответствие алгоритму Парето-аппроксимации A^i совокупность r_i наборов вещественных чисел $\{I_j(A_k^i), j \in [1:m], k \in [1:r_i]\}$ - значений используемых индикаторов $I_1, I_2, \dots, I_m$. Рассматриваемые методы предполагают применение к указанным наборам стандартных процедур статистического тестирования.

В противовес методам, основанным на ранжировании, с помощью данного подхода можно получить количественные оценки качества алгоритмов даже для несравнимых наборов аппроксимаций. Использование нескольких Парето совместимых индикаторов [2] может в данном случае дать более полноценную оценку эффективности сравниваемых алгоритмов. Так, если два Парето совместимых индикатора оказываются по разным индикаторам противоречащими друг другу, то это означает, что рассматриваемые аппроксимации несравнимы.

Если с помощью индикаторов качества сравниваются два алгоритма Парето-аппроксимации, то может быть использован ранговый статистический критерий Манна-Уитни [10]. Если сравниваются более двух алгоритмов, то используют ранговый статистический критерий Крускала-Уоллиса [10]. В качестве альтернативы ранговым тестам может быть использован пермутационный статистический критерий Фишера [11].

Эмпирические функции достижимости. Функция достижимости алгоритма Парето-аппроксимации A может быть оценена на основе Парето-аппроксимаций, полученных в результате r независимых запусков этого алгоритма, с помощью эмпирической функцией достижимости (*Empirical Attainment Function, EAF*)

$$\alpha_r(A, F) = \frac{1}{r} \sum_{j=1}^r I(A_j \preceq F), \quad F \in D_F,$$

где A_j - аппроксимация, полученная на j -том запуске алгоритма; $I(.)$ - функция-индикатор, равная 1, если её аргумент имеет значение *true*, и равная 0 в противном случае. Для каждого вектора $F \in D_F$ функция EAF имеет смысл относительной частоты, с которой данный вектор достижим в наборе аппроксимаций $\{A_j, j \in [1:r]\}$, то есть слабо доминируется элементами этого набора [5].

Сравнение эмпирических функций достижимости различных алгоритмов Парето-аппроксимации может быть выполнено путём сравнения соответствующих статистических критериев.

Функция EAF алгоритма является обобщением одномерной эмпирической кумулятивной функции распределения (*Empirical cumulative distribution function, ECDF*) [5]. Для проверки того, что две $ECDF$ различны, применяют статистический критерий Колмогорова-Смирнова (*KS*), который «измеряет» максимальное различие между двумя $ECDF$ и оценивает статистическую значимость этих различий [12].

EAF могут быть использованы также для визуализации результатов запусков алгоритмов Парето-аппроксимации. Визуальное представление EAF является ценными по следующим причинам:

- ЛПР может иметь предпочтения относительно определённых частей фронта Парето, которые он может сформировать на основе этого представления;
- некоторые индикаторы качества могут неадекватно выражать количественное превосходство одной аппроксимации над другой;
- визуализация может обеспечить понимание сильных и слабых мест алгоритма;
- визуализация может обеспечить возможность быстрой проверки используемых индикаторов качества.

Для визуализации EAF может быть использован специальный метод, предложенный в работе [13].

Методы на основе эмпирических функций достижимости отличаются от двух предыдущих классов методов тем, что могут показать более тонкие различия между сравниваемыми алгоритмами. В то же время, методы требуют достаточно высоких вычислительных затрат, и применимы только для задач с относительно небольшим числом критериев.

4. Архитектура программной системы

Общая структурная схема программной системы представлена на рисунке 1.

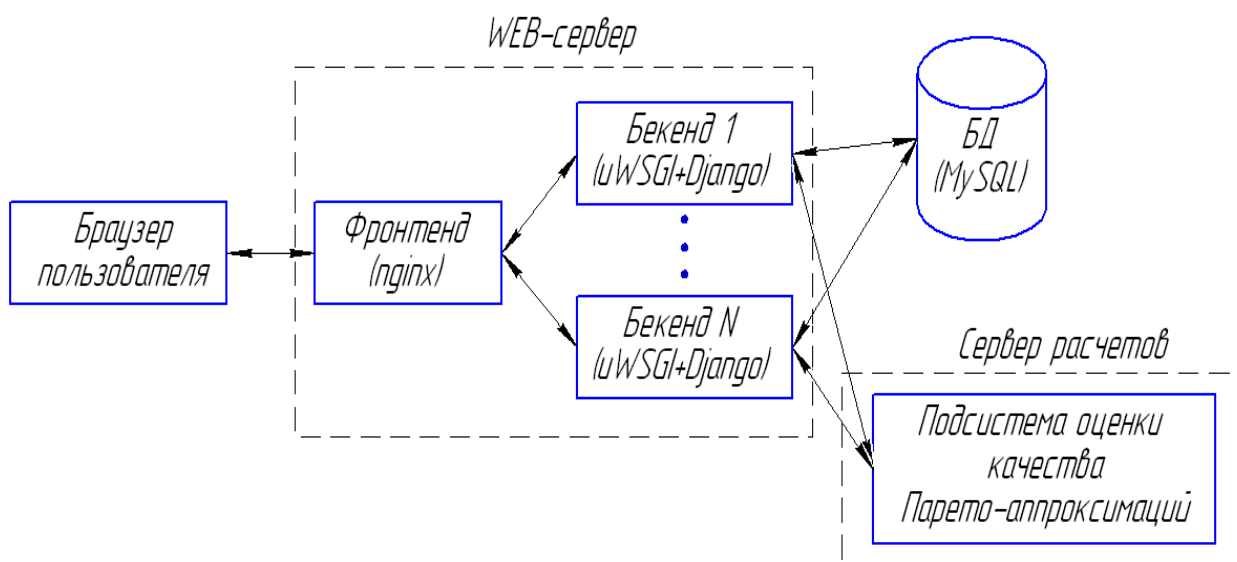


Рисунок 1. Структурная схема системы PARETO RATING

Структуру подсистемы оценки качества Парето-аппроксимации иллюстрирует рисунок 2.



Рисунок 2. Структура подсистемы оценки качества Парето-аппроксимации

В системе реализованы следующие основные режимы: оценка; расчёт; анализ.

Режим «оценка» предоставляет ЛПР возможность передать на сервер свои аппроксимации фронтов Парето и в качестве ответа системы получить его визуализации и значения следующих индикаторов оценки качества [2].

1) Унарные индикаторы: расстояние до точного фронта $I_{GD}(\Theta)$; среднее рассеяние $I_S(\Theta)$; максимальное рассеяние $I_{MS}(\Theta)$; отклонение от равномерного распределения $I_{DU}(\Theta)$; мощность множества решений $I_{ONVG}(\Theta)$; объем объемлющего куба $I_{HC}(\Theta)$; относительное число ошибочных решений $I_{ER}(\Theta)$; максимальная ошибка аппроксимации $I_{MPE}(\Theta)$; минимальная ошибка аппроксимации $I_{SD(\varepsilon)}(\Theta)$; критерий протяжённости $I_{DE}(\Theta)$; критерий гиперобъема разности покрытий $I_H^-(\Theta)$; ε -критерий $I_\varepsilon(\Theta)$; ε_+ -критерий $I_{\varepsilon_+}(\Theta)$; критерии на основе функции предпочтения $I_{R_1}(\Theta), I_{R_2}(\Theta), I_{R_3}(\Theta)$.

2) Бинарные индикаторы: критерий гиперобъема разности покрытий $I_{HV}(\Theta_*, \Theta)$; ε -критерий $I_\varepsilon(\Theta_*, \Theta)$; ε_+ -критерий $I_{\varepsilon_+}(\Theta_*, \Theta)$; критерии на основе функции предпочтения $I_{R_1}(\Theta_*, \Theta), I_{R_2}(\Theta_*, \Theta), I_{R_3}(\Theta_*, \Theta)$.

Режим «расчёт» даёт ЛПП возможности выбора алгоритмов Парето-аппроксимации и тестовых задач. Также имеется возможность использовать собственные реализации алгоритмов Парето-аппроксимации и тестовых задач. Ответами системы в этом режиме являются Парето-аппроксимация выбранной тестовой задачи (в виде текстового файла), визуализация указанной аппроксимации, а так же соответствующие значения унарных индикаторов качества.

В системе реализованы широко известные алгоритмы Парето-аппроксимации *NSGA-II*, *SPEA-2*, *IBEA*, *FEMO*, *SIBEA*, *ECEA*, *EPSMOEA*, а также «стандартные» тестовые задачи *ZDT1-ZDT6*, *DTLZ1-DTLZ7*.

В **режиме «анализ»** ЛПП предоставляется возможность выбора алгоритмов Парето-аппроксимации и одной или нескольких тестовых задач оптимизации, с помощью которых выполняется анализ эффективности этих алгоритмов. В качестве ответа системы ЛПП получает результаты статистического анализа данных, полученных при многократном запуске исследуемых алгоритмов Парето-аппроксимации.

Реализованы следующие статистические методы анализа эффективности алгоритмов Парето-аппроксимации: доминантное ранжирование с последующим использованием статистического рангового критерия; методы на основе критериев Манна-Уитни, Крускала-Уоллиса, а также на основе пермутационного критерия Фишера и Т-критерий Вилкоксона.

5. Особенности программной реализации

Основные задачи, решаемые системой *PARETO RATING*, имеют высокую вычислительную сложность. Поэтому в силу кросс-платформенности и высокой эффективности в качестве основного языка программирования используем C++. В качестве библиотеки, предоставляющей доступ к обобщённым алгоритмам, стандартным контейнерам и структурам данных, применяем стандартную библиотеку шаблонов (*Standard Template Library, STL*). Для компиляции используется набор компиляторов *GCC (GNU Compiler Collection)*.

Система использует следующие сторонние библиотеки:

- *GNU Scientific Library (GSL)* – библиотека численного анализа;
- *PISA. A Platform and Programming Language Independent Interface for Search Algorithms*;
- *cppMetal. Object-oriented C++-based framework for multi-objective optimization with metaheuristics*.

В качестве основной операционной системы выбрана *Linux Ubuntu 12.04 LTS*, которая обеспечивает хорошую переносимость функционирующего под её управлением программного обеспечения.

Для высокоуровневых операций с файлами, а также для объединения подпрограмм, написанных на языке программирования *C++* в логические и функциональные модули использован язык командной оболочки *bash*. Построение ящичковых диаграмм осуществляем с помощью языка программирования *R*.

Все перечисленное программное обеспечение является свободным и распространяется под лицензией *GPL (GNU Public License)*.

Каждая тестовая МКО-задача представлена в виде класса с соответствующим названием. Аналогичную структуру имеют подпрограммы, реализующие алгоритмы Парето-аппроксимации. Для создания объекта языка программирования *C++*, соответствующего данной МКО-задаче, используется класс-фабрика.

ЛПП предоставляет возможность гибкой настройки параметров алгоритмов Парето-аппроксимации и тестовых задач посредством изменения конфигурационных файлов. Формат конфигурационных файлов представляет собой упрощённый вариант общеизвестного формата *.ini*.

За счёт модульности программная система *PARETO RATING* предоставляет широкие возможности для параллельного выполнения, как одной МКО-задачи (в рамках данной ЭВМ), так и распределённого выполнения нескольких МКО-задач (в рамках одного проекта) на разных ЭВМ.

Тестирование системы выполнено на «стандартных» тестовых задачах *ZDT1-ZDT6*, *DTLZ1-DTLZ7*.

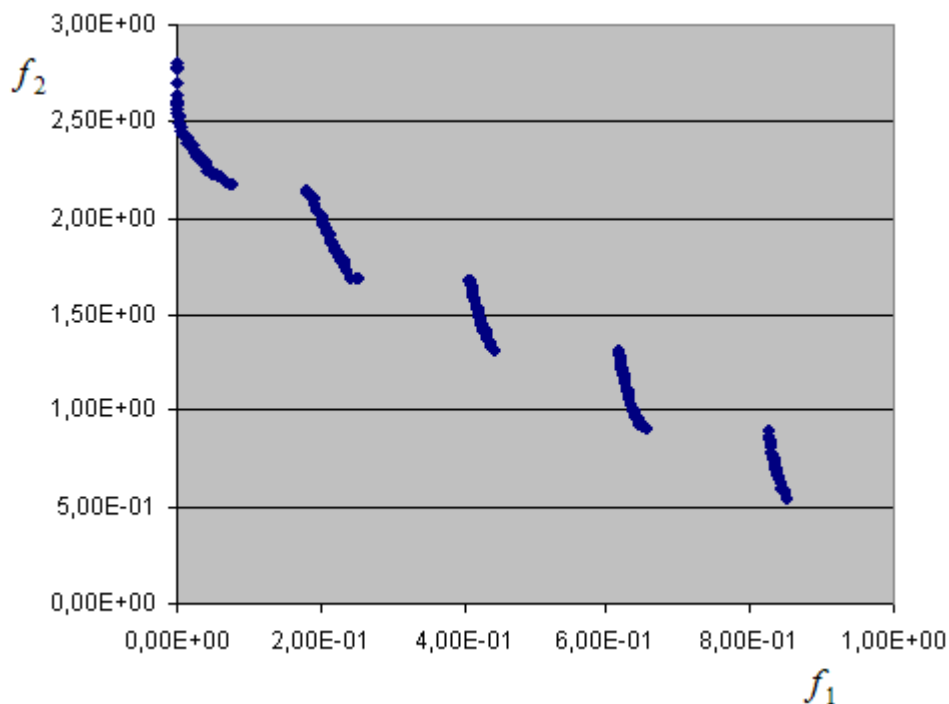
6. Вычислительные эксперименты

Исследование эффективности разработанного алгоритмического и программного обеспечения выполнено на примере широко известных алгоритмов Парето-аппроксимации *NSGA-II*, *SPEA-2* и относительно нового алгоритма *IBEA*. В качестве тестовых задач использованы двухкритериальные *ZDT1 - ZDT5* и трёхкритериальные *DTLZ1 - DTLZ3* задачи.

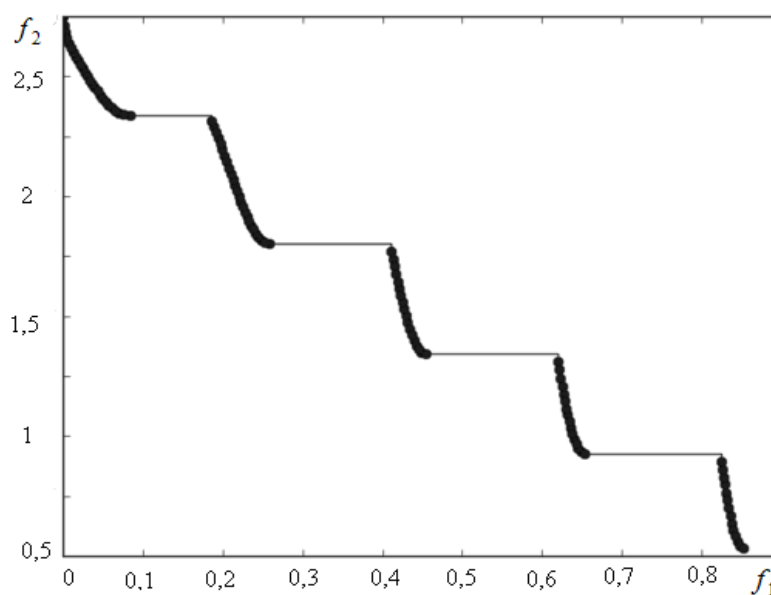
Исследование проводилось при 150 итерациях каждого из указанных алгоритмов Парето-аппроксимации. Использовано нормирование аппроксимаций фронта Парето – приведение всех значений частных целевых функций к интервалу [1, 2]. В качестве опорной точки (*reference point*) во всех случаях выбиралась точка с координатами (2,1; 2,1),

которая в силу указанной нормализации целевых функций удовлетворяет условию доминируемости всеми точками аппроксимаций фронта Парето. Для каждой пары «алгоритм-задача» выполнено по 30 запусков. Уровень значимости статистических критериев принят равным $\alpha = 0,05$.

В качестве примера приведём результаты, полученные системой для алгоритма *SPEA-2* и тестовой задачи *ZDT3* (рисунок 3, таблицы 2, 3).



а) аппроксимация фронта Парето: алгоритм *SPEA-2*; тестовая задача *ZDT3*



б) точный фронт Парето: тестовая задача *ZDT3*

Рисунок 3. Аппроксимация (а) и точный фронт Парето (б) тестовой задачи *ZDT3*

Таблица 2. Значения индикаторов качества для Парето-аппроксимации и точного фронта Парето: тестовая задача ZDT3

Индикатор	I_H^1	I_{R1}^1	I_{R2}^1	I_{R3}^1	$I_{\varepsilon+}^1$	I_{ε}^1
Фронт, полученный алгоритмом SPEA-2	0,78	1,0	-0,62	-4e27	-0,61	0,71
Точный фронт	0,79	1,0	-0,61	-4e27	-0,61	0,72

Доминантное ранжирование. Для тестовых задачах ZDT1, ZDT3, ZDT4, ZDT6 не было обнаружено аппроксимационных наборов, которые бы доминировали остальные наборы (все аппроксимации имели ранг, равный 1). Другими словами, доминантное ранжирование не выявило различий между алгоритмами NSGA-II, SPEA-2, IBEA. Результаты статистической обработки результатов доминантного ранжирования для задачи ZDT2 представлены в таблице 3.

Таблица 3. Результаты (p -значения) применения критерия Манна-Уитни: задача ZDT1

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,72	0,99
SPEA-2	0,27	-	0,99
IBEA	0,001	0,005	-

Из данных, представленных в таблице 3, следует, что по критерию Манна-Уитни алгоритм IBEA превосходит алгоритмы NSGA-II, SPEA-2 (соответствующие p -значения меньше выбранного уровня значимости $\alpha = 0,05$).

На тестовых задачах DTLZ1 - DTLZ3 различий между алгоритмами не обнаружено.

Унарные индикаторы качества. К выборкам значений индикаторов качества применён непараметрический статистический критерий Крускала-Уоллиса. Для задачи ZDT1 этот критерий показал отсутствие различий между алгоритмами NSGA-II, SPEA-2, IBEA с точки зрения индикатора $I_{\varepsilon+}^1$, а для задачи ZDT3 – с точки зрения индикаторов I_H^1 , I_{R2}^1 . Подчеркнём, что речь идёт об уровне значимости $\alpha = 0,05$.

Некоторые результаты исследования представлены в таблицах 4 – 9 и на иллюстрирующих их рисунках 4 – 6.

Таблица 4. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $ZDT1$; индикатор I_H^-

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,63	0,99
SPEA-2	0,37	-	0,99
IBEA	0,002	0,001	-

Таблица 5. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $ZDT1$; индикатор I_{R2}^1

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,63	0,99
SPEA-2	0,37	-	0,99
IBEA	0,0003	0,0005	-

Таблица 6. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $ZDT3$; индикатор $I_{\varepsilon+}^1$

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,51	0,99
SPEA-2	0,49	-	0,99
IBEA	8.1e-4	2.1e-5	-

Таблица 7. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $DTLZ2$; индикатор

$$I_{\varepsilon+}^1$$

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,75	0,99
SPEA-2	0,25	-	0,99
IBEA	3,9e-4	0,006	-

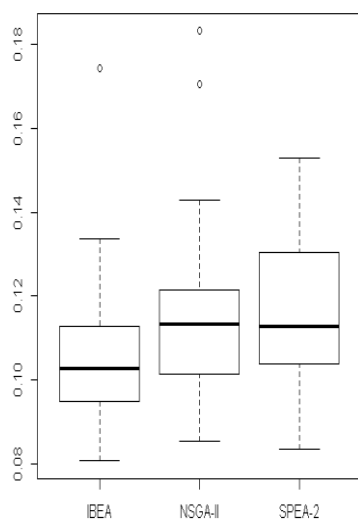
Таблица 8. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $DTLZ2$; индикатор I_H^-

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	0,99	0,99
SPEA-2	2,7e-4	-	0,99
IBEA	1,7e-6	1,5e-7	-

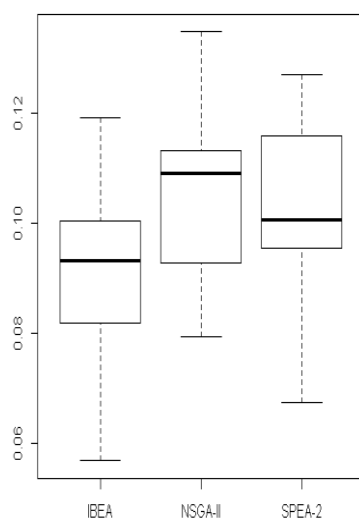
Таблица 9. Результаты (p -значения) применения критерия Крускала-Уоллиса: задача $DTLZ2$; индикатор

$$I_{R2}^1$$

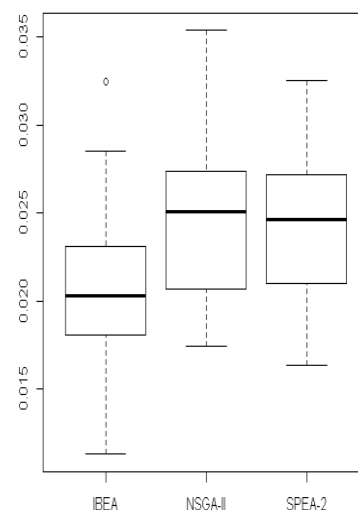
	<i>NSGA-II</i>	<i>SPEA-2</i>	<i>IBEA</i>
<i>NSGA-II</i>	-	2,4e-4	0,03
<i>SPEA-2</i>	0,99	-	0,96
<i>IBEA</i>	0,93	0,04	-



а) индикатор $I_{\epsilon+}^1$

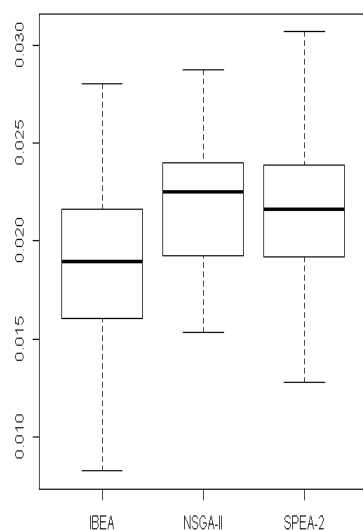


б) индикатор I_H^-

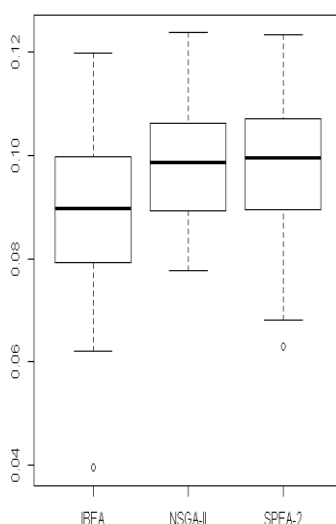


в) индикатор I_{R2}^1

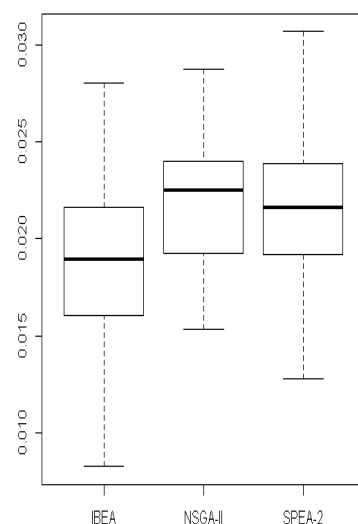
Рисунок 4. Ящичковые диаграммы для задачи $ZDT1$



а) индикатор $I_{\epsilon+}^1$



б) индикатор I_H^-



в) индикатор I_{R2}^1

Рисунок 5. Ящичковые диаграммы для задачи $ZDT3$

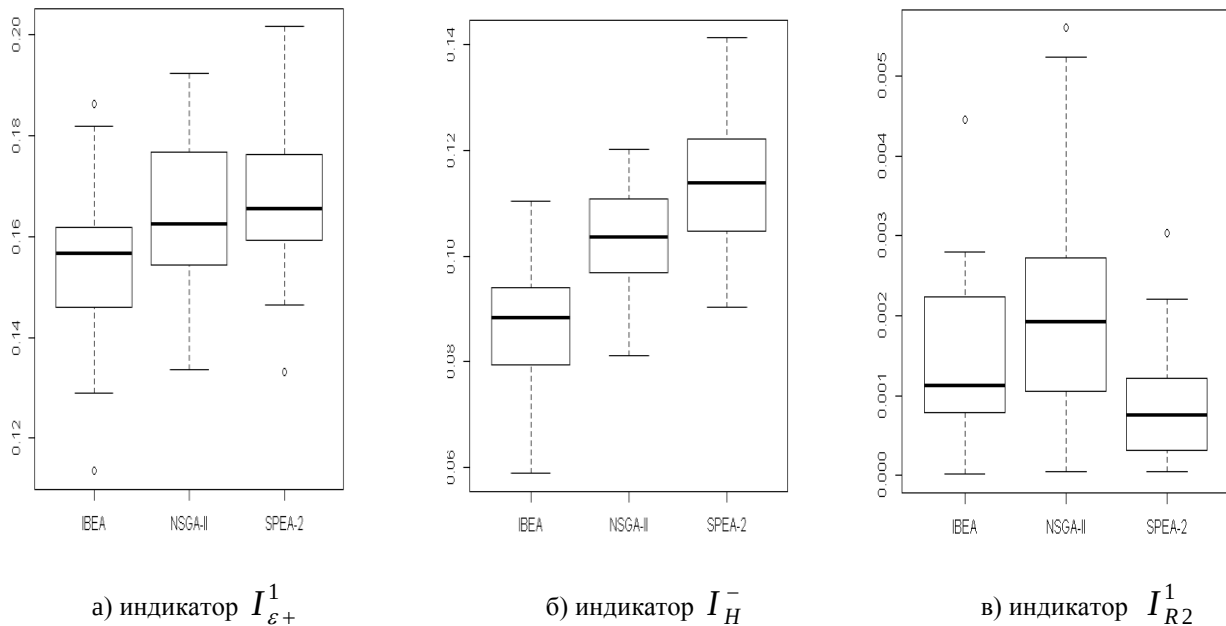


Рисунок 6. Ящичковые диаграммы для задачи DTLZ2

Из представленных результатов вытекают следующие выводы.

- 1) Алгоритм *IBEA* превосходит алгоритмы *NSGA-II*, *SPEA-2* для тестовой задачи *ZDT1* по индикаторам I_H^- , I_{R2}^1 ; по индикатору $I_{\epsilon+}^1$ существенные различия в эффективности этих алгоритмов отсутствуют.
- 2) Алгоритм *IBEA* превосходит алгоритмы *NSGA-II*, *SPEA-2* для тестовой задачи *ZDT3* по индикатору $I_{\epsilon+}^1$; по индикаторам I_H^- , I_{R2}^1 существенные различия в эффективности этих алгоритмов не выявлены.
- 3) Алгоритм *NSGA-II* превосходит алгоритм *SPEA-2* для тестовой задачи *DTLZ2* по индикатору I_{R2}^1 .

Эмпирические функции достижимости. Поверхности 50%-ой достижимости для тестовой задачи *ZDT1* представлены на рисунке 7.

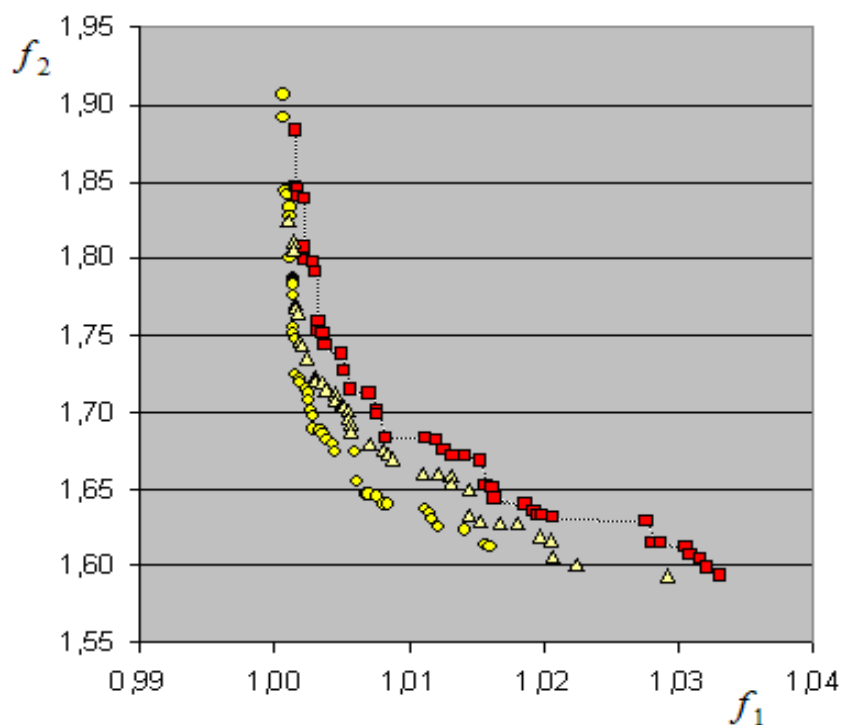


Рисунок 7. Поверхности 50% достижимости для тестовой задачи ZDT1: Δ - алгоритм NSGA-II; $\blacksquare$ - SPEA-2; $\bullet$ - IBEA

Результаты применения статистического критерия Колмогорова-Смирнова к наборам EAF представлены в таблицах 10, 11.

Таблица 10. Результаты (p -значения) применения критерия Колмогорова-Смирнова к наборам EAF: задача ZDT1

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	более 0,05	0,039
SPEA-2	более 0,05	-	более 0,05
IBEA	0,039	более 0,05	-

Таблица 11. Результаты (p -значения) применения критерия Колмогорова-Смирнова к наборам EAF: задача ZDT3

	NSGA-II	SPEA-2	IBEA
NSGA-II	-	более 0,05	более 0,05
SPEA-2	более 0,05	-	0.016
IBEA	более 0,05	0.016	-

Для задач ZDT2, ZDT4, ZDT6 критерий Колмогорова-Смирнова показал отсутствие значительных различий в EAF.

Обсуждение. В результате доминантного ранжирования с последующим применением критерия Манна-Уитни было выявлено значительное превосходство алгоритма IBEA

над алгоритмами *NSGA-II*, *SPEA-2* с точки зрения качества Парето-аппроксимаций, полученных с помощью этих алгоритмов.

Сравнение распределений значений индикаторов качества $I_{\varepsilon+}^1$, I_H^- , I_{R2}^1 не позволило обнаружить превосходство какого-либо из исследуемых алгоритмов над остальными. Другими словами, алгоритмы *NSGA-II*, *SPEA-2*, *IBEA* оказались несравнимыми по данным индикаторам качества.

Сравнение эмпирических функций достижимости с применением критерия Колмогорова-Смирнова показало наличие существенных различий только в качестве Парето-аппроксимаций, полученных с помощью алгоритмов *SPEA-2* и *IBEA*.

Заключение

Программную систему *PARETO RATING* отличает от подобных известных систем комплектность подхода и *WEB*-ориентированность. В работе представлена архитектура системы в целом, а также архитектура подсистемы, реализующей оценки качества Парето-аппроксимации. Представлены основные особенности принятых программных решений.

Функциональное тестирование системы, выполненное на «стандартных» тестовых задачах *ZDT1-ZDT6*, *DTLZ1-DTLZ7*, показало корректность работы системы. Широкое исследование эффективности системы позволяет сделать вывод об адекватности принятых алгоритмических и программных решений.

Поскольку практические МКО-задачи имеют высокую вычислительную сложность, актуальным является распараллеливание алгоритмов Парето-аппроксимации для различных классов параллельных вычислительных систем. Архитектурные и программные решения, использованные в системе *PARETO RATING*, позволяют с небольшими затратами реализовать параллельный вариант системы.

В развитие системы авторы планируют также расширить функции системы как системы поддержки принятия многокритериальных (многоиндикаторных) решений.

Список литературы

1. Карпенко А.П., Семенихин А.С., Митина Е.В. Популяционные методы аппроксимации множества Парето в задаче многокритериальной оптимизации. Обзор // Наука и образование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2012. № 4. Режим доступа: <http://www.technomag.edu.ru/doc/363023.html> (дата обращения 01.06.2014).
2. Zitzler E., Deb K., Thiele L. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results // Evolutionary Computation. 2000. Vol. 8, no. 2. P. 173-195. DOI: [10.1162/106365600568202](https://doi.org/10.1162/106365600568202)
3. Белоус В.В., Грабик А.В., Грошев С.В., Шибитов И.А. Качество Парето-аппроксимации в задаче многокритериальной оптимизации // XVIII Байкальская Всероссийская конференция «Информационные и математические технологии в науке и управлении»: материалы. Ч. 1. Иркутск: ИСЭМ СО РАН, 2013. С. 6-12.

4. Белоус В.В., Грошев С.В., Карпенко А.П., Шибитов И.А. Программные системы для оценки качества Парето-аппроксимации в задаче многокритериальной оптимизации. Обзор // Наука и образование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2014. № 4. DOI: [10.7463/0414.0709198](https://doi.org/10.7463/0414.0709198)
5. Fonseca C.M., Grunert da Fonseca V., Hall A.O. Inferential performance assessment of stochastic optimizers and the attainment functions // In: Evolutionary Multi-Criterion Optimization. Proc. First International Conference, EMO 2001 / Zitzler E., Deb K., Thiele L., Coello C.A.C., Corne D., eds. Springer Berlin Heidelberg, 2001. P. 213-225. DOI: 10.1007/3-540-44719-9_15
6. Лотов А. В., Поспелова И. И. Многокритериальные задачи принятия решений: учеб. пособие. М.: МАКС Пресс, 2008. 197 с.
7. Density and approximations of μ -distribution for different testproblems // System Optimization: сайт. Режим доступа: <http://www.tik.ee.ethz.ch/sop/download/supplementary/testproblems/> (дата обращения 01.06.2014).
8. Fonseca C. M., Fleming P. J. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization // Proc. of the 5th International Conference on Genetic Algorithms, San Mateo, California, 1993. P. 416-423.
9. Goldberg D.E. Genetic Algorithms in Search, Optimization and Machine Learning. Boston, MA, USA: Addison-Wesley Longman Publishing Co., 1989. 201 p.
10. Conover W.J. Practical Nonparametric Statistic. 3rd ed. New York: John Wiley and Sons, 1999. 583 p.
11. Efron B., Tibshirani R. An introduction to the bootstrap. London: Chapman and Hall, 1993. 436 p.
12. Shaw K. J., Nortcliff A. L., Thompson M., Love J., Fleming P.J., Fonseca C.M. Assessing the Performance of Multiobjective Genetic Algorithms for Optimization of a Batch Process Scheduling Problem // Proc. of the 1999 Congress on Evolutionary Computation. CEC 99. Vol. 1. IEEE Service Center, 1999. P. 37-45. DOI: [10.1109/CEC.1999.781905](https://doi.org/10.1109/CEC.1999.781905)
13. Knowles J. A summary-attainment-surface plotting method for visualizing the performance of stochastic multiobjective optimizers // Proc. of the 5th International Conference on Intelligent Systems Design and Applications, 2005. ISDA '05. IEEE, 2005. P.552-557. DOI: [10.1109/ISDA.2005.15](https://doi.org/10.1109/ISDA.2005.15)

The PARETO RATING Software System for the Pareto-approximation Quality Assessment in Multi-criteria Optimization Problem

07, July 2014

DOI: 10.7463/0714.0720253

S. V. Groshev¹, A.P. Karpenko^{1,a}, I. A. Sabitov², D. R. Shbitov²¹Bauman Moscow State Technical University,
105005, Moscow, Russian Federation²LLC "Croc", Moscow, 111033, Russia^aapkarpenko@mail.ru

Keywords: [multiobjective optimization](#), [Pareto set](#)

We consider the task to assess the quality of Pareto set (front) numerical approximation in a multi-criteria optimization (MOC) problem. We mean that Pareto-approximation is obtained by means of this or that population e.g. genetic algorithm.

Eventually, the purpose of work is a comparative assessment of the efficiency of population algorithms of Pareto-approximation. The great number of characteristics (indicators) of the Pareto-approximation quality is developed. Therefore an assessment problem of the Pareto-approximation quality is also considered as multi-criteria (multi-indicator). There are a number of well-known software systems to solve an assessment problem of the Pareto-approximation quality in different degree. Common drawback of these systems is a lack of both the WEB INTERFACE and the support of a multi-indicator assessment of Pareto-approximation quality (though there is a support to calculate the values of a large number of these indicators). The PARETO RATING software system is urged to eliminate the specified shortcomings of known systems. As population algorithms of Pareto-approximation are, as a rule, stochastic, we consider statistical methods to assess the quality of two and more Pareto-approximations (and thereby the estimates of algorithms used to obtain these approximations as well) as follows: methods based on the ranging of the specified approximations; methods based on the quality indicators; methods based on the so-called empirical functions of approachability. We give formal statement of the MOC-problem and general scheme of the population algorithms of its solution, present reviews of known indicators of Pareto-approximation quality and statistical methods for assessment of

Pareto-approximation quality. We describe the system architecture and main features of its software implementation and illustrate efficiency of made algorithmic and software solutions.

References

1. Karpenko A.P., Semenikhin A.S., Mitina E.V. Review: population methods of Pareto set approximation in multi-objective optimization problem. *Nauka i obrazovanie MGTU im. N.E. Bauman* = *Science and Education of the Bauman MSTU*, 2012, no. 4. Available at: <http://technomag.edu.ru/en/doc/363023.html>, accessed 01.06.2014. (in Russian).
2. Zitzler E., Deb K., Thiele L. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evolutionary Computation*, 2000, vol. 8, no. 2, pp. 173-195. DOI: [10.1162/106365600568202](https://doi.org/10.1162/106365600568202)
3. Belous V.V., Grabik A.V., Groshev S.V., Shibitov I.A. [Quality of Pareto-approximation in multicriteria optimization problem]. 18 *Baykal'skaya Vserossiyskaya konferentsiya "Informatsionnye i matematicheskie tekhnologii v nauke i upravlenii": materialy* [Proc. of the 18th Baikal all-Russian conference "Information and Mathematical Technologies in Science and Management"]. Pt. 1. Irkutsk, Publ. of Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences (ESI SB RAS), 2013, pp. 6-12. (in Russian).
4. Belous V.V., Groshev S.V., Karpenko A.P., Shibitov I.A. Programme Systems to Estimate the Pareto-Approximation Quality in the Problem of Multi-Criteria Optimization. A Review. *Nauka i obrazovanie MGTU im. N.E. Bauman* = *Science and Education of the Bauman MSTU*, 2014, no. 4. DOI: [10.7463/0414.0709198](https://doi.org/10.7463/0414.0709198) (in Russian).
5. Fonseca C.M., Grunert da Fonseca V., Hall A.O. Inferential performance assessment of stochastic optimisers and the attainment functions. In: Zitzler E., Deb K., Thiele L., Coello C.A.C., Corne D., eds. *Evolutionary Multi-Criterion Optimization. Proc. First International Conference, EMO 2001*. Springer Berlin Heidelberg, 2001, pp. 213-225. DOI: [10.1007/3-540-44719-9_15](https://doi.org/10.1007/3-540-44719-9_15)
6. Lotov A.V., Pospelova I.I. *Mnogokriterial'nye zadachi prinyatiya resheniy* [Multicriteria decision making problems]. Moscow, MAKS Press, 2008. 197 p. (in Russian).
7. System Optimization. Density and approximations of μ -distribution for different testproblems. System Optimization: website. Available at: <http://www.tik.ee.ethz.ch/sop/download/supplementary/testproblems/>, accessed 01.06.2014.
8. Fonseca C. M., Fleming P. J. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. *Proc. of the 5th International Conference on Genetic Algorithms*, San Mateo, California, 1993, pp. 416-423.

9. Goldberg D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA, USA, Addison-Wesley Longman Publishing Co., 1989. 201 p.
10. Conover W.J. *Practical Nonparametric Statistic*. 3rd ed. New York, John Wiley and Sons, 1999. 583 p.
11. Efron B., Tibshirani R. *An introduction to the bootstrap*. London, Chapman and Hall, 1993. 436 p.
12. Shaw K. J., Nortcliff A. L., Thompson M., Love J., Fonseca C.M., Fleming P.J. Assessing the Performance of Multiobjective Genetic Algorithms for Optimization of a Batch Process Scheduling Problem. *Proc. of the 1999 Congress on Evolutionary Computation. CEC 99*. Vol. 1. IEEE, 1999, pp. 37-45. DOI: [10.1109/CEC.1999.781905](https://doi.org/10.1109/CEC.1999.781905)
13. Knowles J. A summary-attainment-surface plotting method for visualizing the performance of stochastic multiobjective optimizers. *Proc. of the 5th International Conference on Intelligent Systems Design and Applications, 2005. ISDA '05*. IEEE, 2005. P.552-557. DOI: [10.1109/ISDA.2005.15](https://doi.org/10.1109/ISDA.2005.15)