

УДК 004.855.5

МНОГОЦЕЛЕВАЯ ОПТИМИЗАЦИЯ ДЛЯ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ СТРОГО БИНАРНЫХ ДЕРЕВЬЕВ

Л. А. Демидова, д.т.н., профессор кафедры ВПМ РГРТУ; liliya.demidova@ Rambler.ru
Н. Н. Астахова, аспирант РГРТУ; asnadya@yandex.ru

Рассматривается задача многоцелевой оптимизации, возникающая при разработке моделей прогнозирования на основе строго бинарных деревьев. Целью работы является улучшение поисковых характеристик модифицированного алгоритма клонального отбора, применяемого при разработке моделей прогнозирования на основе строго бинарных деревьев, посредством вовлечения в процесс отбора лучших моделей прогнозирования двух показателей качества моделей: показателя аффинитета и показателя несовпадения тенденций. Учет двух показателей качества модели прогнозирования реализован с использованием понятия «Парето-множество», применяемого в процессе формирования новых популяций моделей прогнозирования в модифицированном алгоритме клонального отбора. При формировании новой популяции моделей прогнозирования для поддержания ее высокого разнообразия предложено учитывать значения расстояния скученности моделей прогнозирования. Приведены результаты экспериментальных исследований, подтверждающие эффективность предложенного подхода к улучшению поисковых характеристик модифицированного алгоритма клонального отбора.

Ключевые слова: временной ряд, модель прогнозирования, строго бинарное дерево, модифицированный алгоритм клонального отбора, генетический алгоритм, многоцелевая оптимизация, Парето-доминирование, расстояние скученности.

Введение

При решении задачи прогнозирования временных рядов (ВР) ключевым вопросом является вопрос выбора адекватной модели прогнозирования. В случае применения тех или иных технологий искусственного интеллекта для разработки моделей прогнозирования, в частности эволюционных алгоритмов оптимизации, поиск адекватной модели прогнозирования сводится к выбору «лучшей» модели с точки зрения какого-либо показателя качества модели из популяции моделей, полученных с использованием общих принципов формирования моделей прогнозирования. Таким показателем качества может служить, например, средняя относительная ошибка прогнозирования, которая должна быть минимизирована.

При разработке моделей прогнозирования на основе строго бинарных деревьев (СБД) и модифицированного алгоритма клонального отбора (МАКО) [1 – 3], показавших свою эффективность при решении задач краткосрочного прогнозирования ВР, общий принцип формирования моделей прогнозирования заключается

в представлении их в виде закодированных с помощью случайно выбранных символов последовательностей – антител, которые преобразуются в аналитические зависимости, применяемые для получения прогнозных значений ВР.

Поиск искомой модели прогнозирования осуществляется в процессе итерационных вычислений. При этом на каждом шаге этого процесса выявляются «лучшие» антитела (и соответствующие им «лучшие» модели прогнозирования), которые на следующем шаге становятся родителями для нового поколения антител (и соответствующих им моделей прогнозирования) [1 – 11]. Очевидно, что правильный выбор «лучших» антител является залогом эффективности применения МАКО при разработке моделей прогнозирования, а также его сходимости.

Традиционный подход, обычно используемый при разработке моделей краткосрочного прогнозирования ВР (в том числе моделей на основе СБД и МАКО), заключается в оценивании моделей с учетом одного показателя качества, например показателя качества, основанного на вычислении средней относительной

ошибки прогнозирования (или средней квадратической ошибки прогнозирования) на обучающей выборке данных [1 – 3, 5 – 11]. При этом в качестве лучшей выбирается модель с минимальным значением такого показателя качества.

Однако использование только одного показателя качества, основанного, например, на вычислении средней относительной ошибки прогнозирования, далеко не всегда является достаточным для получения действительно лучшей модели прогнозирования. Зачастую целесообразно учитывать дополнительные показатели качества модели прогнозирования, позволяющие учесть соответствие сезонным тенденциям или тренду временного ряда, отсутствие выбросов, сложность модели прогнозирования и т.п. [6]. Чаще всего не представляется возможным выбрать единственно необходимый показатель качества модели прогнозирования, поэтому актуальной является проблема разработки подходов к созданию моделей прогнозирования, при оценке качества которых одновременно учитывается несколько показателей качества.

Так, при оценке качества моделей прогнозирования на основе СБД и МАКО с целью повышения их эффективности при решении задачи краткосрочного прогнозирования и обеспечения возможности их применения при решении задачи среднесрочного прогнозирования целесообразно дополнительно (наряду с показателем качества, основанным на вычислении средней относительной ошибки прогнозирования [6]) использовать показатель качества, который позволит оценить общую тенденцию изменения значений известных элементов ВР (например, показатель несовпадения тенденций ВР) [3].

Теоретическая часть

В настоящее время известны и широко используются различные подходы к разработке моделей прогнозирования ВР. Модели прогнозирования, основанные на эволюционных алгоритмах оптимизации, в частности на генетических алгоритмах (ГА) [12 – 13] и алгоритмах клонального отбора (АКО) [7, 8, 12, 13], обеспечивают высокую точность прогнозных решений при приемлемых временных затратах. Данный факт объясняется тем, что такие модели базируются на принципах естественного отбора, использование которых позволяет в итоге осуществлять построение моделей прогнозирования, наилучшим образом описывающих известные значения элементов ВР.

В частности, МАКО, моделирующий законы

функционирования естественной иммунной системы, обеспечивает формирование сложных аналитических зависимостей при разработке моделей прогнозирования на основе СБД [5, 7].

В [3, 5, 7] приведено подробное описание МАКО, реализующего построение и отбор моделей прогнозирования на основе СБД. При этом каждой модели прогнозирования соответствует некоторое антитело, которое интерпретируется в аналитическую зависимость, применяемую для получения прогнозных значений ВР. В процессе реализации МАКО должна быть определена модель прогнозирования, наилучшим образом описывающая известные значения ВР и обеспечивающая получение минимального значения средней относительной ошибки прогнозирования (Average Forecasting Error Rate) *AFER* :

$$AFER = \frac{\sum_{j=k+1}^n |(f^j - d^j)/d^j|}{n - k} 100\%, \quad (1)$$

где d^j и f^j – соответственно реальное и предсказанное значения для j -го элемента ВР; n – количество элементов ВР.

В контексте работы с МАКО средняя относительная ошибка прогнозирования *AFER* [1 – 3] выступает в роли такого показателя качества модели прогнозирования, как показатель аффинитета (affinity) *Aff* .

Модели прогнозирования на основе СБД и МАКО показали свою высокую эффективность при решении задач краткосрочного прогнозирования ВР.

В МАКО в простейшем виде реализованы принципы распознавания антигенов с помощью антител. При этом для представления антигена *Ag* используются известные значения элементов прогнозируемого ВР, а в роли антитела *Ab* выступает некоторая аналитическая зависимость, определяющая модель прогнозирования. В качестве «лучшего» антитела *Ab* выбирается антитело, обеспечивающее минимальное значение показателя аффинитета *Aff* [3, 5, 7, 11].

Кодирование антитела *Ab* осуществляется посредством записи в строку символов, выбираемых из трёх символьных алфавитов [3, 5]: алфавита арифметических операций (операций сложения, вычитания, умножения и деления) – *Operation* = {'+', '-', '*', '/'}; алфавита функционалов *Functional* = {'S', 'C', 'Q', 'L', 'E', '_'} , в котором символы 'S', 'C', 'Q', 'L', 'E' определяют математические функции «синус», «косинус», «квадратный корень», «натуральный логарифм», «экспонента», а символ '_' – отсутствие математической

функции; алфавита терминалов $Terminal = \{ 'a', 'b', \dots, 'z', '?' \}$, в котором символы $'a', 'b', \dots, 'z'$ определяют аргументы аналитической зависимости, а символ '?' – константу. При использовании указанных выше трёх символьных алфавитов удается обеспечить корректное преобразование случайным образом формируемых антителей, структура которых может быть описана с помощью СБД, в аналитические зависимости [3, 5, 7, 11].

Количество символов из алфавита терминалов $Terminal$ в антители Ab определяет максимально возможный порядок K модели прогнозирования ($K \geq k$, где k – реальный порядок модели) [3, 5]. Если максимально возможный порядок модели прогнозирования равен K , то это означает, что на момент времени j при прогнозировании значения элемента d^j ВР могут использоваться K значений элементов ВР: $d^{j-K}, \dots, d^{j-2}, d^{j-1}$.

Для обеспечения высокой точности прогнозирования ВР со сложно формализуемым математическим законом при формировании антителей используются СБД, структура которых аналогична структуре СБД, приведенного на рисунке 1 [3, 5, 7]. СБД, соответствующее модели прогнозирования наперед заданного порядка, строится в результате композиции одного «левого» поддерева максимально возможного порядка $K = 3$ и некоторого количества «правых» поддеревьев максимально возможного порядка $K = 2$. В этом случае термин «левое» поддерево («правое» поддерево) используется для указания того, в какую ветвь (в левую или в правую) некоторого уровня СБД следует включать новое поддерево [3, 5].

При формировании антителя целесообразно сначала реализовать разбиение СБД на поддеревья, далее выполнить обход вершин каждого поддерева с формированием упорядоченных списков символов, находящихся в его вершинах, а затем – последовательное объединение этих списков [5]. При формировании упорядоченного списка символов на основе поддерева осуществляется последовательный двухкратный обход его вершин. При этом сначала при движении по поддереву снизу вверх слева направо попарно обходятся вершины, содержащие символы из алфавита терминалов $Terminal$, и соответствующие им расположенные сверху вершины, содержащие символы из алфавита функционалов $Functional$. Затем при движении в том же направлении (снизу вверх слева направо) попарно обходятся вершины, содержащие символы из алфавита арифмети-

ческих операций $Operation$, и соответствующие им расположенные сверху вершины, содержащие символы из алфавита функционалов $Functional$ [5].

Антитело (при движении по нему слева направо) в качестве первых двух символов содержит пару символов нулевого уровня СБД из алфавита функционалов $Functional$ и алфавита арифметических операций $Operation$. Затем в антители располагаются списки символов, соответствующие «правым» поддеревьям максимально возможного порядка $K = 2$ (при движении по СБД сверху вниз), и, наконец, список символов, входящих в «левое» поддерево максимально возможного порядка $K = 3$ [5]. При использовании такого способа формирования антителя удается обеспечить наглядность представления структуры СБД в виде объединения поддеревьев. При этом само антители легко интерпретируется в аналитическую зависимость.

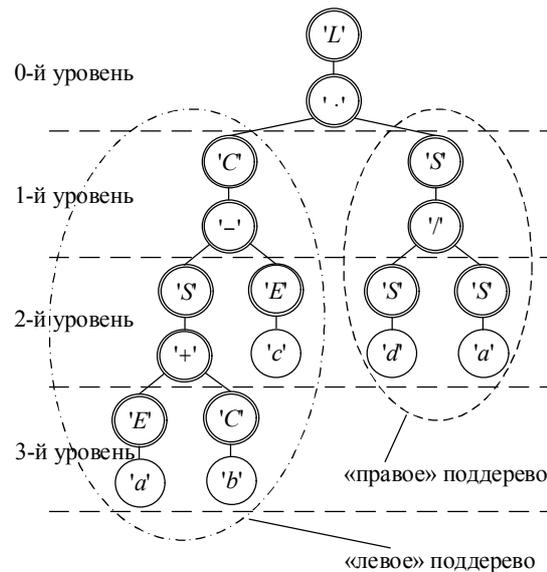


Рисунок 1 – Пример строго бинарного дерева

Так, антители, сформированное на основе СБД, представленного на рисунке 1, кодируется строкой символов:

$$L \cdot S / S e S d C - S + E a C b E a . \tag{2}$$

При разработке модели прогнозирования k -го порядка при $k = 4$ код антителя (2) с учетом упорядоченности символов a, b, c, d в алфавите терминалов $Terminal$ может быть записан как:

$$f(d^{j-1}, d^{j-2}, d^{j-3}, d^{j-4}) = \ln(\cos(\sin(\exp(d^{j-1}) + \cos(d^{j-2})) - \exp(d^{j-3})) \cdot \sin(\sin(d^{j-4}) / \sin(d^{j-1}))). \tag{3}$$

МАКО, используемый для выбора «лучшего» антителя, определяющего «лучшую» аналитическую зависимость, включает в себя подготовительную часть, которая отвечает за

формирование начальной популяции антител, и итерационную часть, в которой последовательно реализуются: упорядочение антител по возрастанию значений аффинитета Aff ; отбор и клонирование части «лучших» антител (то есть антител с наименьшими значениями аффинитета Aff); гипермутация клонов антител; самоуничтожение клонов антител, «похожих» на другие клоны и антитела текущей популяции; вычисление аффинитета клонов антител и формирование новой популяции антител (которая становится текущей популяцией антител); супрессия текущей популяции антител; генерация новых антител и добавление их к текущей популяции до получения ее исходного размера; проверка условия завершения работы МАКО. В работе [5] приведено подробное описание МАКО.

Как уже было отмечено выше, качество модели прогнозирования на основе СБД и МАКО оценивается с помощью показателя аффинитета Aff по формуле (1). Для повышения эффективности данных моделей прогнозирования при решении задачи краткосрочного прогнозирования и обеспечения возможности их применения при решении задачи среднесрочного прогнозирования при оценке качества моделей наряду с показателем аффинитета предлагается использовать еще один показатель качества – показатель несовпадения тенденций ВР.

Показатель несовпадения тенденций $Tendency$ ВР может быть вычислен как:

$$Tendency = \frac{h}{n-k-1}, \quad (4)$$

где h – количество отрицательных произведений $(f^{j-1} - f^j) \cdot (d^{j-1} - d^j)$ при $j = \overline{r+2, n}$; f^j и d^j – предсказанное и реальное значения ВР для j -го отсчета времени; n – количество элементов ВР; k – порядок модели; $n-k-1$ – общее количество произведений $(f^{j-1} - f^j) \cdot (d^{j-1} - d^j)$.

Оба рассматриваемых показателя качества модели прогнозирования [Aff (1) и $Tendency$ (4)] позволяют оценить сходство прогнозных значений известных элементов анализируемого ВР с реальными, но с применением разных принципов оценивания. Если показатель аффинитета Aff используется при реализации МАКО для оценки «приспособляемости» антитела Ab (и соответствующей ему модели прогнозирования) к антигену Ag (и соответствующему ему ВР), то показатель несовпадения тенденций $Tendency$ позволяет оценить качество антитела Ab (и

соответствующей ему модели прогнозирования) с учетом совпадения с трендом анализируемого ВР, соответствующего антигену Ag . При этом оба показателя качества должны быть минимизированы. То есть показатель аффинитета Aff оценивает, прежде всего, сходство и различие прогнозных и реальных значений известных элементов анализируемого ВР, а показатель несовпадения тенденций $Tendency$ оценивает сходство и различие направлений изменения прогнозных и реальных значений известных элементов анализируемого ВР, позволяя проанализировать тенденции изменения ВР, а также наличие сезонных колебаний. Таким образом, при разработке моделей прогнозирования на основе СБД и МАКО очевидна целесообразность реализации одновременного учета при оценке качества моделей прогнозирования наряду со значением показателя аффинитета Aff (1) и значения показателя несовпадения тенденций $Tendency$ (4).

Для решения задачи одновременного учета двух показателей качества при разработке модели прогнозирования могут быть применены различные хорошо зарекомендовавшие себя подходы [12]. При этом особо следует выделить подход, предполагающий использование тех или иных алгоритмов многоцелевой оптимизации, в том числе эволюционных алгоритмов, обеспечивающих решение проблемы учета нескольких целевых функций (критериев, показателей качества) при решении различных прикладных задач.

Наибольшее применение среди эволюционных алгоритмов многоцелевой оптимизации в настоящее время находят ГА. Следует также отметить наличие и АКО, предназначенных для решения задач многоцелевой оптимизации. Однако данные алгоритмы менее проработаны и, в своем большинстве, заимствуют принципы многоцелевой оптимизации, заложенные в генетические алгоритмы. Возможность данного заимствования может быть объяснена многими сходными механизмами реализации эволюционного процесса в ГА и АКО.

На данный момент широко известны следующие ГА многоцелевой оптимизации:

- VEGA (Vector Evaluated Genetic Algorithm);
- FFGA (Fonseca and Fleming's Genetic Algorithm);
- NPGA (Niche Pareto Genetic Algorithm);
- NSGA (Non-dominated Sorting Genetic Algorithm);
- NSGA-II (Non-dominated Sorting Genetic Algorithm-II);

– NSGA-III (Non-dominated Sorting Genetic Algorithm-III);

– SPEA (Strength Pareto Evolutionary Algorithm).

Самые ранние многоцелевые ГА – алгоритм VEGA (автор Schaffer D.; предложен в 1985 году) [13] и алгоритм FFGA (авторы Fonseca C.M., Fleming P.J.; предложен в 1993 году) [14] – позволяют обеспечить высокую сходимость и высокую скорость вычислений. При этом алгоритм FFGA, использующий в отличие от алгоритма VEGA принципы Парето-доминирования [15], реализует формирование множества Парето-оптимальных решений (хромосом). Недостатком данных алгоритмов является отсутствие механизмов защиты от попадания в локальные экстремумы.

Алгоритм NPGA (авторы Horn J., Nafpliotis N., Goldberg D.E.; предложен в 1994 году) [16] и NSGA (авторы Srinivas N., Deb K.; предложен в 1994 году) [17] значительно превосходят своих предшественников. Так, они эффективно борются с проблемой попадания в локальные экстремумы. Алгоритм NPGA применяет для этого принципы турнирной селекции, а алгоритм NSGA вводит в использование понятие «фенотипическая дистанция», предполагающее анализ расстояний до решений-соседей. Если решение находится слишком близко к другим решениям с лучшими значениями ранга (то есть «фенотипическая дистанция» между решениями меньше установленного порогового значения), то оно не будет выбрано для переноса в следующее поколение. При этом под рангом понимается количество решений, доминирующих над данным решением (чем меньше ранг, тем лучше решение). Данные алгоритмы обеспечивают поддержание высокого разнообразия как в популяции решений, так и в итоговом множестве Парето-оптимальных решений. Однако использование этих алгоритмов сопряжено со значительными вычислительными затратами.

В настоящее время известны модификации алгоритма NSGA: NSGA II (авторы Deb K., Agrawal S., Pratap A., Meyarivan T.; предложен в 2002 году) [18] и NSGA III (автор Deb K.; предложен в 2012 году) [19]. Алгоритм NSGA-II использует улучшенный алгоритм сортировки, который значительно снижает вычислительную сложность расчетов, а для поддержания высокого разнообразия в популяции решений алгоритм использует понятие «расстояние скученности» (crowding-distance), позволяющее оценить, насколько некоторое решение близко к решениям-соседям. Некоторое решение может быть выбрано для переноса в следующее

поколение, если у него ранг меньше или его расстояние скученности больше, чем у другого решения с тем же рангом. Алгоритм NSGA-II использует в своей реализации принципы элитизма и является одним из лучших на данный момент алгоритмов многоцелевой оптимизации, поскольку реализует механизмы защиты от попадания в локальные экстремумы и характеризуется сравнительно небольшими вычислительными затратами [18]. Алгоритм NSGA-III решает проблему сходимости к локальному экстремуму с помощью априори задаваемых опорных векторов, задающих направления поиска решений, что приводит к некоторому увеличению сложности расчетов (относительно аналогов).

Алгоритм SPEA (авторы Zitzler E., Thiele L.; предложен в 1998 году) [20] использует в своей реализации принципы элитизма, осуществляя хранение решений, недоминируемых другими решениями, входящими в популяцию, в специальном внешнем массиве. Однако так как количество недоминируемых решений может быть велико, для уменьшения количества решений, хранящихся во внешнем массиве, используется процедура кластеризации решений. Для защиты от преждевременной сходимости, характерной для элитизма, в алгоритме используется механизм образования ниш, формируемых не по расстоянию между решениями, а на основе принципов Парето-доминирования. В результате таких дополнительных манипуляций происходит увеличение вычислительных затрат на реализацию алгоритма SPEA, поэтому его применение целесообразно именно в задачах, не допускающих потерю лучшего решения.

Наибольшее применение при решении широкого спектра прикладных задач многоцелевой оптимизации в настоящее время находят алгоритмы SPEA, NSGA и NSGA-II. Они успешно применяются при решении задачи календарного планирования, задачи составления расписаний, задачи коммивояжера и т.п. [21]. При этом алгоритм NSGA-II существенно превосходит алгоритмы SPEA, NSGA и NSGA-III с точки зрения минимизации вычислительных затрат.

В качестве наиболее известных АКО многоцелевой оптимизации следует указать:

– MISA (Multiobjective Immune System Algorithm);

– MOIA (Multiobjective Immune Algorithm);

– MOCSA (Multiobjective Clonal Selection Algorithm);

– IDCMA (Immune Dominance Clonal Multiobjective Algorithm);

- ACSAMO (Adaptive Clonal Selection Algorithm for Multiobjective Optimization);
- CNMOIA (Constrained Nonlinear Multiobjective Optimization Immune Algorithm);
- NNIA (Nondominated Neighbor Immune Algorithm).

Все перечисленные АКО основаны на принципах Парето-доминирования.

Алгоритм MISA (авторы Coello Coello C.A., Cruz Cortés N.; предложен в 2002 году) [21] является фактически первым АКО многоцелевой оптимизации и использует в своей реализации принципы элитизма. В этом алгоритме все решения (антитела) разбиваются на классы: допустимые и недоминируемые; недопустимые и недоминируемые, недопустимые и доминируемые. Количество решений, используемых для переноса в следующее поколение, определяется пропорционально априори заданным для каждого класса коэффициентам.

Алгоритм MOIA (авторы Luh G.-C., Chueh C.-H., Liu W.-W.; предложен в 2004 году) [22] позволяет задать для различных целевых функций индивидуальные коэффициенты значимости, используемые на этапе мутации. При этом большей мутации подвергаются те решения (антитела), которые имеют лучшие значения по менее приоритетным целевым функциям (критериям, показателям качества) и худшие значения по более приоритетным. В результате удается задать правильное направление развития решений при одновременном сохранении решений с лучшими значениями по самым приоритетным целевым функциям. Применение данного алгоритма целесообразно в задачах с возможной экспертной оценкой важности целевых функций.

Алгоритм MOCSA (авторы Campelo F., Guimarães F.G., Saldanha R.R., Igarashi H., Noguchi S., Lowther D.A., Ramirez J.A.; предложен в 2004 году) [23] использует обратно пропорциональную зависимость между рангом решения и количеством созданных для этого решения клонов, что, с одной стороны, обеспечивает быструю сходимости алгоритма, а с другой – создает опасность попадания в локальные экстремумы. В связи с этим алгоритм MOCSA целесообразно использовать в задачах, в которых скорость вычислений является главным приоритетом.

Алгоритм IDCMA (авторы Jiao L., Gong M., Shang R., Du H., Lu B.; предложен в 2005 году) [24] использует в своей реализации принципы элитизма решений и вводит понятие «средства антител» (решений), которое применяется как мера, определяющая, насколько рассматриваемое

решение сходно с решениями-соседями.

Алгоритм ACSAMO (авторы Wang X.L., Mahfouf M.; предложен в 2006 году) [25] использует в своей реализации принципы элитизма решений и для увеличения скорости сходимости в каждом поколении антител (решений) в качестве антигенов лучшие решения предыдущего поколения. При этом выбор родителей при формировании следующего поколения определяется с учетом принципов недоминируемой сортировки.

Алгоритм CNMOIA (автор Zhang Z.H.; предложен в 2006 году) [26] оценивает каждое антитело (решение) с учетом значений аффинитета и «плотности» антитела, которая определяется на основе количества решений, находящихся в непосредственной близости к рассматриваемому решению. В результате такой оценки для каждого решения вычисляется количество создаваемых клонов-решений.

Алгоритм NNIA (авторы Jiao L., Gong M., Du H., Wo L.; предложен в 2008) [27] является модификацией алгоритма IDCMA. Одно из принципиальных отличий данного алгоритма – использование расстояния скученности (как в алгоритме NPGA-II) в процессе отбора решений.

В результате анализа работ, посвященных вопросам многоцелевой оптимизации на основе эволюционных алгоритмов, в частности ГА и АКО, было выявлено, что преобладающее большинство алгоритмов используют в своей реализации принципы Парето-доминирования. В связи с этим был сделан вывод о перспективности использования принципов Парето-доминирования при усовершенствовании МАКО, применяемого для отбора моделей прогнозирования на основе СБД. С учетом особенностей решаемой задачи и вычислительных характеристик ГА и АКО многоцелевой оптимизации было принято решение о целесообразности адаптации идей, заложенных в алгоритм NSGA-II.

В контексте решения задачи разработки моделей прогнозирования на основе СБД и МАКО под решением следует понимать модель прогнозирования (и соответствующее ей антитело), а под целевой функцией – показатель качества модели. Используя понятие «Парето-доминирование», все модели прогнозирования можно разделить на доминируемые и недоминируемые.

Пусть $Q_{s,v}$ – значение v -го показателя качества для s -й модели прогнозирования; $v = \overline{1, V}$; $s = \overline{1, S}$; V – количество показателей качества модели прогнозирования; S – количество моделей прогнозирования (например, в популяции, то есть размер популяции).

В соответствии с понятием Парето-доминирования будем говорить, что z -я модель прогнозирования доминирует над s -й моделью прогнозирования ($s = \overline{1, S}$; $z = \overline{1, S}$), если z -я модель не хуже s -й модели по всем v -м ($v = \overline{1, V}$) показателям качества прогнозирования и хотя бы по одному показателю качества превосходит s -ю модель, то есть для всех значений v ($v = \overline{1, V}$) выполняется условие:

$$Q_{s,v} \geq Q_{z,v} \quad (5)$$

и существует хотя бы один v^* -й ($1 \leq v^* \leq V$) показатель, для которого выполняется условие:

$$Q_{s,v^*} > Q_{z,v^*}. \quad (6)$$

При этом все v -е ($v = \overline{1, V}$) показатели качества прогнозирования должны быть минимизированы.

Таким образом, s -я модель прогнозирования доминируема z -й моделью, если z -я модель не хуже s -й модели прогнозирования по всем показателям качества и хотя бы по одному показателю качества прогнозирования лучше s -й модели. В противном случае, если такие z -е ($z = \overline{1, S}$) модели прогнозирования не существуют, s -я модель прогнозирования недоминируема.

Следует отметить, что в случае когда какой-либо показатель качества должен быть максимизирован, он легко может быть преобразован к показателю, который необходимо минимизировать.

Для каждой s -й модели прогнозирования ($s = \overline{1, S}$) вычисляется ранг R_s . Ранг s -й модели прогнозирования равен количеству моделей в популяции, которые доминируют над s -й моделью прогнозирования. Так, если для некоторой модели прогнозирования ее значения показателей качества больше или равны соответствующим значениям показателей качества трех других моделей прогнозирования, то ранг такой модели равен 3. Ранг R_s недоминируемой s -й модели прогнозирования полагается равным нулю.

В случае когда количество показателей качества модели прогнозирования равно 2 ($V = 2$), а $Q_{s,1} = Aff_s$, $Q_{s,2} = Tendency_s$, где Aff_s и $Tendency_s$ – соответственно значения показателя аффинитета (1) и показателя несовпадения тенденций (4) для s -й модели прогнозирования ($s = \overline{1, S}$), s -я модель прогнозирования доминируема z -й моделью прогнозирования

($s = \overline{1, S}$; $z = \overline{1, S}$), если выполняются условия:

$$Q_{s,1} \geq Q_{z,1} \text{ и } Q_{s,2} > Q_{z,2}$$

или

$$Q_{s,1} > Q_{z,1} \text{ и } Q_{s,2} \geq Q_{z,2},$$

то есть

$$Aff_s \geq Aff_z \text{ и } Tendency_s > Tendency_z$$

или

$$Aff_s > Aff_z \text{ и } Tendency_s \geq Tendency_z.$$

Для вычисления расстояний скученности τ_s ($s = \overline{1, S}$) для всех моделей прогнозирования может быть использован следующий алгоритм [18].

1. Вычислить ранг всех моделей прогнозирования в популяции. Модели, имеющие одинаковые значения ранга, объединить в одну группу.

2. Для каждой группы моделей:

– отсортировать модели прогнозирования по возрастанию значений каждого показателя качества;

– принять в качестве значений расстояний скученности τ_s для самой первой и самой последней в списке сортировки (то есть при $s = 1$ и $s = G_w$) максимально возможное наперед заданное значение η_{max} ; все остальные расстояния скученности τ_s при $s = \overline{2, G_w - 1}$ положить равными нулю ($w = \overline{1, W}$; W – количество групп; G_w – количество моделей в w -й группе);

– вычислить расстояние скученности τ_s как:

$$\tau_s = \sum_{v=1}^V \frac{Q_{s-1,v} - Q_{s+1,v}}{Q_v^{\max} - Q_v^{\min}}, \quad (7)$$

где $Q_{s-1,v}$ и $Q_{s+1,v}$ – значения v -го показателя качества ($v = \overline{1, V}$) для моделей с номерами $(s-1)$ и $(s+1)$, являющихся «соседями» для s -й модели; Q_v^{\min} и Q_v^{\max} – минимальное и максимальное значения v -го показателя качества ($v = \overline{1, V}$) соответственно.

На рисунке 2 представлена геометрическая интерпретация, применяемая при расчете расстояния скученности при наличии двух показателей качества. Закрашенные «кружочки» соответствуют моделям с минимальным (нулевым) значением ранга. Для определения расстояния скученности для s -й модели требуется определить значения обоих показателей качества для $(s-1)$ -й и $(s+1)$ -й моделей, являющихся

ближайшими к s -й модели и имеющих тот же ранг, а также лучшее и худшее значения каждого показателя качества (вне зависимости от ранга модели). Расстояние скученности для s -й модели с учетом двух показателей качества может быть вычислено как:

$$\tau_s = \frac{(Q_{s-1,1} - Q_{s+1,1})}{(Q_1^{\max} - Q_1^{\min})} + \frac{(Q_{s-1,2} - Q_{s+1,2})}{(Q_2^{\max} - Q_2^{\min})}. \quad (8)$$

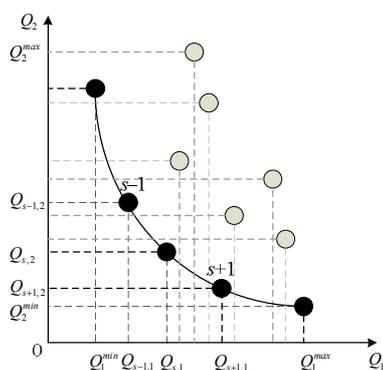


Рисунок 2 – Геометрическая интерпретация, применяемая при расчете расстояния скученности

При реализации расчетов s -я модель прогнозирования будет считаться лучшей, чем z -я модель, если:

- $R_s < R_z$;
- $R_s = R_z$ и $\tau_s > \tau_z$.

Если s -я модель прогнозирования лучше, чем z -я модель, то s -я модель является кандидатом для переноса в новое поколение.

Следует отметить, что модели (и соответствующие антитела) в процессе реализации многоцелевого МАКО будут дважды подвергаться процедуре отбора, обеспечивающей поддержание разнообразия в популяции: сначала при выполнении супрессии антител (на основе сходства их кодировок), а затем – при одновременном анализе значений рангов и значений расстояний скученности.

Многоцелевой МАКО может быть описан следующей последовательностью шагов.

1. Сгенерировать популяцию антител, каждое из которых кодируется на основе СБД и определяет некоторую модель прогнозирования.

2. Отсортировать популяцию антител по «недоминированию» (то есть по возрастанию рангов антител) на основе значений двух показателей качества модели прогнозирования – значения показателя аффинитета Aff (1) и значения показателя несовпадения тенденций $Tendency$ (4).

3. Осуществить выбор антител-родителей для формирования следующего поколения антител-клонов на основе значения рангов,

вычисляемых для каждого антитела (каждой модели прогнозирования) с учетом количества доминирующих над этим антителом (этой моделью) антител (моделей), и «расстояний скученности», характеризующих сходство каждого антитела (каждой модели прогнозирования) с соседними (ближайшими) антителами (моделями). При этом при равенстве рангов предпочтение отдается антителам (моделям) с большим значением расстояния скученности. Использование и сохранение таких антител (моделей) способствуют поддержанию разнообразия в популяции и, как следствие, помогают избежать сходимости алгоритма к локальному минимуму. Клонирование антител, гипермутация клонов антител, самоуничтожение клонов антител, «похожих» на другие клоны и антитела текущей популяции, супрессия текущей популяции антител и пр. осуществляются в соответствии с принципами, заложенными в МАКО [1, 2].

4. Перейти к п. 5, если достигнуты желаемые значения показателей качества моделей прогнозирования [показателя аффинитета Aff (1) и показателя несовпадения тенденций $Tendency$ (4)] или исчерпано предварительно заданное количество поколений в МАКО. В противном случае перейти к п. 2.

5. Выбрать в последней популяции антител, отсортированной по «недоминированию» на основе значений двух показателей качества модели прогнозирования – значения показателя аффинитета Aff (1) и значения показателя несовпадения тенденций $Tendency$ (4), антитело с минимальным (нулевым) рангом и минимальным значением показателя аффинитета Aff (1) и считать его «лучшим» (искомым) антителом. Использовать «лучшее» антитело для построения искомой модели прогнозирования с целью выполнения краткосрочных и среднесрочных прогнозов.

При реализации данного алгоритма в п. 5 значение показателя несовпадения тенденций $Tendency$ (4) в явном виде не учитывается [оно учтено только при сортировке антител по «недоминированию», а определяющим при выборе «лучшего» антитела является значение показателя аффинитета Aff (1)].

Следует отметить, что в результате реализации предлагаемого алгоритма будет получено Парето-множество недоминируемых (Парето-оптимальных) моделей прогнозирования, обеспечивающих лучшие сочетания значений используемых показателей качества модели прогнозирования для анализируемого ВР. Данные модели прогнозирования могут быть

использованы в дальнейшем, например при создании ансамбля моделей прогнозирования.

Полученную в результате реализации многоцелевого МАКО модель прогнозирования планируется использовать при решении задач среднесрочного прогнозирования, что значительно расширит сферу применения моделей прогнозирования на основе СБД и МАКО, а также для прогнозирования групп ВР [5].

Экспериментальные исследования

Для подтверждения выдвинутого предположения о перспективности предлагаемого алгоритма были проведены экспериментальные исследования по прогнозированию 5 ВР:

T_1 (ВР «Стоимость нефти марки Brent»; период наблюдения 15.12.2015 – 02.01.2016; единица измерения – рубль) = [46.04; 45.11; 44.46; 44.38; 44.30; 43.97; 44.70; 44.70; 44.08; 45.92; 47.08; 47.32; 47.46; 48.05; 48.84; 50.25; 48.90; 49.50];

T_2 (ВР «Общее число произведенных энергоресурсов»; период наблюдения 1992 – 2012 гг.; единица измерения – квадриллион БТЕ) = [47.99; 44.69; 42.30; 41.42; 39.35; 38.74; 39.07; 40.81; 41.70; 42.63; 44.16; 47.16; 49.86; 51.05; 52.06; 52.52; 52.52; 50.01; 53.74; 54.63; 55.30];

T_3 (ВР «Число интернет-пользователей на 100 человек»; период наблюдения 1990 – 2014 гг.; единица измерения – процент) = [0.00000; 0.00067; 0.01345; 0.05382; 0.14815; 0.26975; 0.47296; 0.81274; 1.01899; 1.97723; 2.94437; 4.12827; 8.29886; 12.85939; 15.22667; 17.02328; 24.66000; 26.83000; 29.00000; 43.00000; 49.00000; 63.80000; 67.97000; 70.52000];

T_4 (ВР «Пассажиропоток»; период наблюдения 1989 – 2012 гг.; единица измерения – млн пассажиро-км.) = [3669.202; 3635.441; 3305.227; 3145.413; 3087.016; 3251.615; 2714.314; 3268.434; 3743.247; 4249.406; 4468.123; 4489.569; 4651.419; 4726.832; 4986.468; 5054.200; 5482.314; 5655.296; 5129.103; 5987.268; 5933.041; 5973.308];

T_5 (ВР «Добавленная стоимость в сельском хозяйстве»; период наблюдения 1992 – 2013 гг.; единица измерения – доллар) = [270112; 274000; 255000; 253200; 272475; 227102; 192117; 181200; 170300; 152900; 141042; 167100; 158000; 152900; 155573; 164272; 164262; 173699; 173411; 175800; 153500; 139028; 139842; 144612].

Выбранные ВР характеризуются различными количеством значений, уровнем шкалы значений, точками дискретизации, амплитудой и периодом, что позволяет считать дальнейшие выводы универсальными.

Значения ошибок прогнозирования

Вид ошибки	Ошибка прогнозирования на основе:	
	<i>Aff</i>	<i>Aff</i> и <i>Tendency</i>
ВР №1 (стоимость нефти марки Brent)		
Ошибка обучения, %	3,12	1,10
На 1 шаг вперед, %	1,43	0,22
На 2 шага вперед, %	0,85	0,01
На 3 шага вперед, %	0,67	0,38
На 4 шага вперед, %	1,06	0,22
На 5 шагов вперед, %	0,45	0,57
Средняя ошибка, %	0,89	0,28
ВР №2 (количество произведенных энергоресурсов)		
Ошибка обучения, %	5,23	2,89
На 1 шаг вперед, %	1,14	0,62
На 2 шага вперед, %	0,87	0,20
На 3 шага вперед, %	1,53	0,73
На 4 шага вперед, %	0,29	0,29
На 5 шагов вперед, %	0,97	0,68
Средняя ошибка, %	0,96	0,50
ВР №3 (количество интернет-пользователей)		
Ошибка обучения, %	5,49	3,29
На 1 шаг вперед, %	1,48	0,57
На 2 шага вперед, %	3,75	0,13
На 3 шага вперед, %	7,41	0,65
На 4 шага вперед, %	5,86	0,12
На 5 шагов вперед, %	9,14	0,42
Средняя ошибка, %	5,53	0,38
ВР №4 (пассажиропоток)		
Ошибка обучения, %	5,47	2,54
На 1 шаг вперед, %	1,58	1,75
На 2 шага вперед, %	2,84	1,42
На 3 шага вперед, %	3,54	1,98
На 4 шага вперед, %	3,17	1,10
На 5 шагов вперед, %	3,69	1,38
Средняя ошибка, %	2,96	1,52
ВР №5 (добавленная стоимость в сельском хозяйстве)		
Ошибка обучения, %	3,98	2,75
На 1 шаг вперед, %	0,82	0,82
На 2 шага вперед, %	1,37	0,54
На 3 шага вперед, %	1,89	0,86
На 4 шага вперед, %	2,29	1,09
На 5 шагов вперед, %	3,97	0,22
Средняя ошибка, %	2,27	0,71

Для каждого ВР было выполнено прогнозирование с использованием одного (*Aff*) и двух (*Aff* и *Tendency*) показателей качества модели прогнозирования. Результаты расчетов значений ошибок прогнозирования представлены в таблице.

Представленные значения ошибок прогнозирования – это средние значения, определенные по результатам 1000 экспериментов по поиску лучшей модели прогнозирования за 100 итераций при размере популяции в 20 антител.

Значения относительных ошибок прогнози-

рования на 1 – 5 шагов вперед свидетельствуют о том, что предлагаемый подход к отбору моделей прогнозирования является более эффективным как для решения задач краткосрочного прогнозирования (на 1 – 3 шага вперед), так и для решения задач среднесрочного прогнозирования (на 4, 5 шагов вперед).

Использование дополнительного показателя качества модели прогнозирования (показателя *Tendency*) позволило осуществить поиск модели прогнозирования в нужном (правильном) направлении. В результате этого для всех рассмотренных примеров ВР за небольшое количество поколений МАКО удалось получить меньшие значения показателя аффинитета (*Aff*) [ошибки обучения (1)] и в большинстве случаев – ошибок прогнозирования на 1 – 5 шагов вперед (таблица).

На рисунках 3 и 4 представлены результаты прогнозирования значений стоимости барреля нефти марки Brent, полученные при использовании моделей прогнозирования на основе одного и двух показателей качества моделей соответственно. Из представленных графиков отчетливо видно, что вторая из моделей лучше повторяет математический закон исходного ВР на обучающей выборке данных. Более того, эта особенность также остается верна на тестовой выборке (при прогнозировании на 5 шагов вперед). При этом модель прогнозирования, полученная при использовании двух показателей качества (*Aff* и *Tendency*), не только демонстрирует лучшую выживаемость для удлиненного горизонта прогнозирования, но и оказывается эффективнее и для краткосрочного прогнозирования.

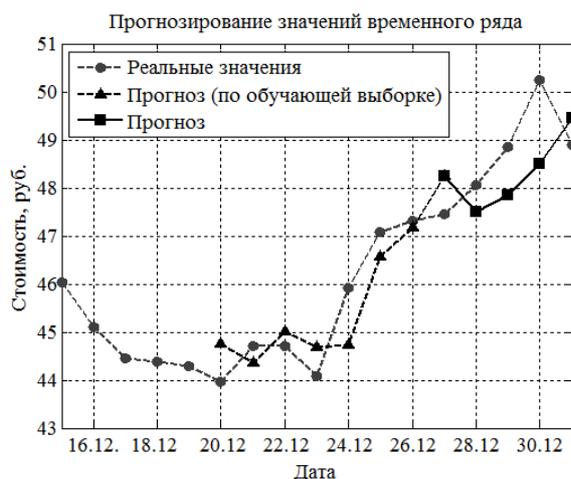


Рисунок 3 – Прогнозирование стоимости марки Brent (один показатель качества *Aff*)

На основе представленных расчетов можно сделать вывод о том, что использование второго показателя качества модели прогнозирования позволяет увеличить срок жизни модели прогно-

зирования и при этом сохранить ее эффективность для выполнения краткосрочных прогнозов, что в целом подтверждает успешность предлагаемого подхода.

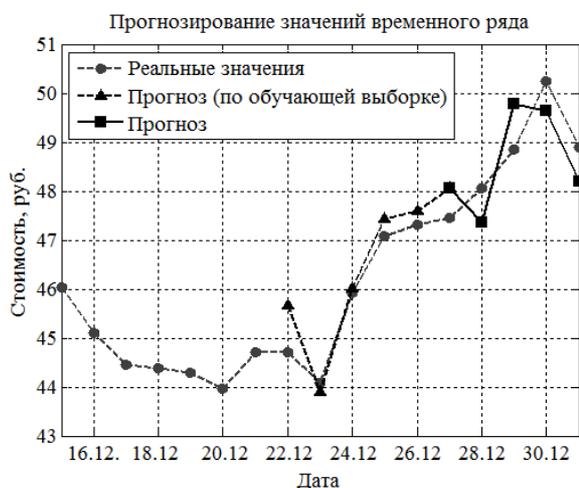


Рисунок 4 – Прогнозирование стоимости марки Brent (два показателя качества *Aff* и *Tendency*)

Заключение

МАКО, первоначально разработанный для решения задач краткосрочного прогнозирования, по результатам полученных исследований может быть использован и при решении задач среднесрочного прогнозирования.

Применение к моделям прогнозирования на основе СБД принципов Парето-доминирования позволяет получить эффективное решение задачи учета нескольких показателей качества в задаче разработки моделей прогнозирования при незначительном увеличении вычислительной сложности МАКО. При этом удастся расширить сферу применения моделей прогнозирования на основе СБД и МАКО.

Работа поддержана грантом РФФИ, номер заявки 16-08-00771.

Библиографический список

1. Демидова Л. А., Корячко А. В., Скворцова Т. С. Модифицированный алгоритм клонального отбора для анализа временных рядов с короткой длиной актуальной части // Системы управления и информационные технологии. 2010. Т. 42. № 4.1. С. 131-136.
2. Демидова Л. А. Модели прогнозирования временных рядов с короткой актуальной частью на основе модифицированного алгоритма клонального отбора // Вестник Рязанского государственного радиотехнического университета. 2012. № 39-2. С. 64-71.
3. Астахова Н. Н., Демидова Л. А. Использование почти полных строго бинарных деревьев и модифицированного алгоритма клонального отбора при разработке моделей прогнозирования временных рядов с короткой актуальной частью // Вестник

Рязанского государственного радиотехнического университета. 2013. № 4-2 (46). С. 89-96.

4. **Демидова Л. А.** Оценка качества моделей прогнозирования на основе строго бинарных деревьев и модифицированного алгоритма клонального отбора // *Cloud of Science*. 2014. Т. 1. № 2. С. 202-222.

5. **Demidova L. A.** Time Series Forecasting Models On The Base Of Modified Clonal Selection Algorithm // В сборнике: 2014 International conference on computer technologies in physical and engineering applications (ICCTPEA) Editor: E. I. Veremey. Санкт-Петербургский государственный университет; IEEE (IEEE Catalog number CFP14BDA-USB). 2014. p. 33-34.

6. **Демидова Л. А.** Подход к оценке качества моделей прогнозирования на основе строго бинарных деревьев и модифицированного алгоритма клонального отбора // *Бизнес-информатика*. 2015. № 1. С. 58-68.

7. **Astakhova N. N., Demidova L. A., Nikulchev E. V.** Forecasting Of Time Series' Groups With Application Of Fuzzy C-Mean Algorithm // *Contemporary Engineering Sciences*. 2015. Vol. 8. № 35. P. 1659-1677.

8. **Astakhova N., Demidova L., Nikulchev E., Pluzhnik E.** Forecasting Of Time Series' Groups With Application Of Fuzzy C-Mean Algorithm And Forecasting Models On The Base Of Strictly Binary Trees And Modified Clonal Selection Algorithm // В сборнике: 16th International Symposium on Advanced Intelligent Systems 2015. P. 861-873.

9. **Astakhova N. N., Demidova L. A., Nikulchev E. V.** Forecasting Method For Grouped Time Series With The Use Of K-Means Algorithm // *Applied Mathematical Sciences*. 2015. Vol. 9. No. 97. P. 4813-4830.

10. **Astakhova N., Demidova L., Konev V.** The Description Problem Of The Clusters' Centroids // В сборнике: 2015 International Conference «Stability and Control Processes» in Memory of V.I. Zubov (SCP) 2015. P. 448-451.

11. **Demidova L. A.** Genetic Algorithm For Optimal Parameters Search In The One-Factor Forecasting Model Based On Continuous Type-2 Fuzzy Sets // *Automation and Remote Control*. 2013. Vol. 74. No. 2. P. 313-320.

12. **Michalewicz Z.** Genetic algorithms, numerical optimization and constraints // *Proc. of the Sixth Int. Conf. on Genetic Algorithms and their Applications*, Pittsburgh, PA. 1995. P. 239-247.

13. **Schaffer J. D.** Multiple objective optimization with vector evaluated genetic algorithm // *J.J. Grefenstete (Ed.): Genetic Algorithms and Their Applications. Proc. of the First Int. Conf. on Genetic Algorithms*, Hillsdale, NJ: L. Erlbaum. 1985. P. 93-100.

14. **Fonseca CM, Fleming PJ.** Multiobjective genetic algorithms. // *IEE colloquium on 'Genetic Algorithms for Control Systems Engineering*. 1993. P. 745-749.

15. **Knowles J., Corne D.** The Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimization // *Proceedings of the 1999*

Congress on Evolutionary Computation. Piscataway, New Jersey: IEEE Service Center. 1999. P. 98-105.

16. **Horn J., Nafpliotis N., Goldberg D. E.** A niched Pareto genetic algorithm for multiobjective optimization // *In Proceedings of the First IEEE Conference on Evolutionary Computation*. Piscataway, Vol. 1. 1994. P. 82-87.

17. **Deb K., Pratap A., Agarwal S., Meyarivan T.** A Fast and Elitist Multiobjective Genetic Algorithm: NSGA II // *KanGAL Report No. 200001*. Indian Institute of Technology. Kanpur. India. 2000. P. 182-197.

18. **Bentley P. J., Wakefield J. P.** Finding Acceptable Solutions in the Pareto-Optimal Range using Multiobjective Genetic Algorithms // *In Proceedings of the 2nd On-Line World Conference on Soft Computing in Engineering Design and Manufacturing*. 1997. P. 126-140.

19. **Deb K., Jain H.** An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints // *IEEE Transactions on Evolutionary Computation*. 2014. Vol. 18(4). P. 577-601.

20. **Zitzler E., Thiele L.** An evolutionary algorithm for multiobjective optimization: The strength pareto approach. // *Technical Report 43*. 1998. P. 494-508.

21. **Coello Coello C. A., Cruz Cortés N.** An approach to solve multiobjective optimization problems based on an artificial immune system // *Proceedings of the First International Conference on Artificial Immune Systems*, University of Kent at Canterbury. UK. September 9-11. 2002. P. 212-221.

22. **Luh G.-C., Chueh C.-H., Liu W.-W.** MOIA: Multi-Objective Immune Algorithm // *Computers and Structures* Vol. 82. 2004. P. 829-844.

23. **Campelo F., Guimarães F. G., Saldanha R. R., Igarashi H., Noguchi S., Lowther D. A., & Ramirez, J. A.** A novel multiobjective immune algorithm using nondominated sorting // *11th International IGTE Symposium on Numerical Field Calculation in Electrical Engineering*. 2004.

24. **Jiao L., Gong M., Shang R., Du H., Lu B.** Clonal selection with immune dominance and energy based multiobjective optimization // *3rd International Conference on Evolutionary Multi-Criterion Optimization*. 2005. P. 474-489.

25. **Wang X. L., Mahfouf M.** ACSAMO: An Adaptive Multiobjective Optimization Algorithm using the Clonal Selection Principle // *2nd European Symposium on Nature-Inspired Smart Information Systems*. 2006. P. 1-12.

26. **Zhang Z. H.** Constrained Multiobjective Optimization Immune Algorithm: Convergence and Application // *Computers and Mathematics with Applications*. 2006. Vol. 52(5). P. 791-808.

27. **Jiao L., Gong M., Du H., Bo L.** Multiobjective immune algorithm with nondominated neighbor-based selection // *Evolutionary Computation*. Vol. 16. Issue 2. Summer. 2008. P. 225-255.

UDC 004.855.5

MULTI-OBJECTIVE OPTIMIZATION FOR THE FORECASTING MODELS ON THE BASE OF STRICTLY BINARY TREES

L. A. Demidova, PhD (technical sciences), full professor, RSREU, Ryazan; liliya.demidova@rambler.ru
N. N. Astakhova, post-graduate student, RSREU, Ryazan; asnadya@yandex.ru

The problem of multi-objective optimization dealing with the development of forecasting models on the base of strictly binary trees is considered. The aim is to improve the search characteristics of a modified clonal selection algorithm which is applied for the development of forecasting models on the base of strictly binary trees by means of involvement in the process of selection of best forecasting models of two model quality indicators: the affinity indicator and the indicator of the tendencies' discrepancy. The accounting of two indicators of a forecasting model quality is carried out with the use of the notion «Pareto-set» which is applied in the formation process of new populations of forecasting models in a modified clonal selection algorithm. During the formation of new population of forecasting models for maintenance of its high variety it is offered to consider values of the crowding-distance of forecasting models. The results of experimental studies confirming the efficiency of the offered approach to improvement of the search characteristics of a modified clonal selection algorithm are given.

Key words: time series, forecasting model, strictly binary tree, modified clonal selection algorithm, genetic algorithm, multi-objective optimization, Pareto-dominance, crowding-distance.

References

1. **Demidova L.A., Koryachko A.V., Skvorcova T.S.** Modificirovannyj algoritm klonal'nogo otbora dlya analiza vremennyh ryadov s korotkoj dlinoj aktual'noj chasti. *Sistemy upravleniya i informacionnye tekhnologii*. 2010, vol. 42, no. 4.1, pp. 131-136 (in Russian).
2. **Demidova L.A.** Modeli prognozirovaniya vremennyh ryadov s korotkoj aktual'noj chast'yu na osnove modificirovannogo algoritma klonal'nogo otbora. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2012, no. 39-2, pp. 64-71 (in Russian).
3. **Astakhova N.N., Demidova L.A.** Ispol'zovanie pochti polnyh strogo binarnyh derev'ev i modificirovannogo algoritma klonal'nogo otbora pri razrabotke modelej prognozirovaniya vremennyh ryadov s korotkoj aktual'noj chast'yu. *Vestnik Ryazanskogo gosudarstvennogo radiotekhnicheskogo universiteta*. 2013, no. 4-2 (46), pp. 89-96 (in Russian).
4. **Demidova L.A.** Ocenka kachestva modelej prognozirovaniya na osnove strogo binarnyh derev'ev i modificirovannogo algoritma klonal'nogo otbora. *Cloud of Science*. 2014, vol. 1, no. 2, pp. 202-222 (in Russian).
5. **Demidova L.A.** Time Series Forecasting Models On The Base Of Modified Clonal Selection Algorithm. *V sbornike: 2014 International conference on computer technologies in physical and engineering applications (ICCTPEA)* Editor: E. I. Veremey. Sankt-Peterburgskij gosudarstvennyj universitet; IEEE (IEEE Catalog number CFP14BDA-USB). 2014, pp. 33-34.
6. **Demidova L.A.** Podhod k ocenke kachestva modelej prognozirovaniya na osnove strogo binarnyh derev'ev i modificirovannogo algoritma klonal'nogo otbora. *Biznes-informatika*. 2015, no. 1, pp. 58-68 (in Russian).
7. **Astakhova N.N., Demidova L.A., Nikulchev E.V.** Forecasting Of Time Series' Groups With Application Of Fuzzy C-Mean Algorithm. *Contemporary Engineering Sciences*. 2015, vol. 8, no. 35, pp. 1659-1677.
8. **Astakhova N., Demidova L., Nikulchev E., Plu-zhnik E.** Forecasting Of Time Series' Groups With Application Of Fuzzy C-Mean Algorithm And Forecasting Models On The Base Of Strictly Binary Trees And Modified Clonal Selection Algorithm. *V sbornike: 16th International Symposium on Advanced Intelligent Systems* 2015, pp. 861-873.
9. **Astakhova N.N., Demidova L.A., Nikulchev E.V.** Forecasting Method For Grouped Time Series With The Use Of K-Means Algorithm. *Applied Mathematical Sciences*. 2015, vol. 9, no. 97, pp. 4813-4830.
10. **Astakhova N., Demidova L., Konev V.** The Description Problem Of The Clusters' Centroids. *V sbornike: 2015 International Conference «Stability and Control Processes» in Memory of V.I. Zubov (SCP)* 2015, pp. 448-451.
11. **Demidova L.A.** Genetic Algorithm For Optimal Parameters Search In The One-Factor Forecasting Model Based On Continuous Type-2 Fuzzy Sets. *Automation and Remote Control*. 2013, vol. 74, no. 2, pp. 313-320.

12. **Michalewicz Z.** Genetic algorithms, numerical optimization and constraints. *Proc. of the Sixth InVol. Conf. on Genetic Algorithms and their Applications, Pitts-burgh, PA.* 1995, pp. 239-247.
13. **Goldberg D.E.** Genetic Algorithms in Search, Optimization, and Machine Learning. *Reading, Massachusetts: Addison-Wesley.* 1989, p. 372.
14. **Fonseca C.M., Fleming P.J.** Multiobjective optimization and multiple constraint handling with evolutionary algorithms - Part I: A unified formulation. *Technical report 564. University of Sheffield, Sheffield.* UK. January. 1995, pp. 1-16.
15. **Horn J., Nafpliotis N., Goldberg D.E.** A niched Pareto genetic algorithm for multiobjective optimization. *In Proceedings of the First IEEE Conference on Evolutionary Computation. Piscataway.* 1994, vol. 1, pp. 82-87.
16. **Knowles J., Corne D.** The Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimization. *Proceedings of the 1999 Congress on Evolutionary Computation. Piscataway. New Jersey: IEEE Service Center.* 1999, pp. 98-105.
17. **Deb K., Pratap A., Agarwal S., Meyarivan Vol. A** Fast and Elitist Multiobjective Genetic Algorithm: NSGA II. *KanGAL Report No. 200001. Indian Institute of Technology.* Kanpur, India. 2000, pp. 182-197.
18. **Bentley P.J., Wakefield J.P.** Finding Acceptable Solutions in the Pareto-Optimal Range using Multiobjective Genetic Algorithms. *In Proceedings of the 2nd On-Line World Conference on Soft Computing in Engineering Design and Manufacturing.* 1997, pp. 126-140.
19. **Deb K., Jain H.** An evolutionary many-objective optimization algorithm using reference-point based non-dominated sorting approach, Part I: Solving problems with box constraints. *IEEE Transactions on Evolutionary Computation.* 2014, vol. 18(4), pp. 577-601.
20. **Deb K.** Multi-objective Optimization using Evolutionary Algorithms. Chichester. UK: Wiley. 2001, pp. 221-232.
21. **Coello Coello C.A., Cruz Cortés N.** An approach to solve multiobjective optimization problems based on an artificial immune system. *Proceedings of the First International Conference on Artificial Immune Systems,* University of Kent at Canterbury, UK, September 9–11. 2002, pp. 212–221.
22. **Luh G.-C., Chueh C.-H., Liu W.-W.** MOIA: Multi-Objective Immune Algorithm. *Computers and Structures.* 2004, vol. 82, pp. 829-844.
23. **Campelo, F., Guimarães, F. G., Saldanha, R. R., Igarashi, H., Noguchi, S., Lowther, D. A., & Ramirez, J. A.** A novel multiobjective immune algorithm using non-dominated sorting. In 11th International. *IGTE Symposium on Numerical Field Calculation in Electrical Engineering.* 2004.
24. **Jiao L., Gong M., Shang R., Du H., Lu B.** Clonal selection with immune dominance and energy based multiobjective optimization. *3rd International Conference on Evolutionary Multi-Criterion Optimization.* 2005, pp. 474-489.
25. **Wang, X. L., & Mahfouf, M.** ACSAMO: An Adaptive Multiobjective Optimization Algorithm using the Clonal Selection Principle. *2nd European Symposium on Nature-Inspired Smart Information Systems.* 2006, pp. 1-12.
26. **Zhang, Z.** Constrained Multiobjective Optimization Immune Algorithm: Convergence and Application. *Computers and Mathematics with Applications.* 2006, vol. 52(5), pp. 791-808.
27. **Jiao L., Gong M., Du H., Bo L.** Multiobjective immune algorithm with nondominated neighbor-based selection. *Evolutionary Computation.* Vol. 16. Issue 2. Summer. 2008, pp. 225-255.