

УДК 681.518

Л.А. Демидова

ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИЙ РЫНКА ТРУДА НА ОСНОВЕ ОДНОФАКТОРНЫХ НЕЧЕТКИХ ВРЕМЕННЫХ РЯДОВ И ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Рассматривается применение однофакторных нечетких временных рядов к задаче краткосрочного прогнозирования тенденций рынка труда. Для повышения точности прогноза предлагается использовать генетический алгоритм, обеспечивающий выбор оптимальных параметров модели прогнозирования – числа интервалов разбиения универсума и действительных чисел для корректировки границ универсума, а также учитывать особенности формирования групп нечетких логических зависимостей.

Введение. Важную роль при принятии управленческих решений играет прогнозирование, призванное выявить общие перспективы и тенденции организационно-структурного развития, обеспечить сбалансированность краткосрочных и долгосрочных программ. Изменчивость производственно-экономических отношений, неустойчивость, концептуальная непоследовательность и неполнота законодательной базы, регулирующей эти отношения, приводят к тому, что временные ряды, представляющие социально-экономические процессы, имеют короткую актуальную часть. В связи с этим существует необходимость в разработке методов прогнозирования, которые бы обеспечили получение адекватных прогнозов в данной ситуации.

1. Однофакторная модель прогнозирования на основе нечетких временных рядов. Анализ временных рядов играет важную роль в решении многих актуальных задач. Так как большинство реальных событий характеризуются некоторой неопределенностью, то каждому наблюдению временного ряда можно поставить в соответствие нечеткую переменную с некоторой функцией принадлежности. С помощью нечетких отношений и нечетких правил можно построить эффективную модель прогноза на основе большого числа входов и одного выхода (без значительных временных затрат и вычислительных сложностей).

Нечеткие временные ряды могут быть представлены с помощью нечетких множеств [1, 2]. Нечеткое множество (НМ) A , определенное на универсуме U , может быть определено в виде:

$$A = f_A(u_1)/u_1 + \dots + f_A(u_n)/u_n, \quad (1)$$

где $f_A(u)$ – функция принадлежности НМ A , $f_A(u): U \rightarrow [0, 1]$, $f_A(u_r)$ определяет степень принадлежности элемента u_r НМ A , $r = \overline{1, n}$.

Пусть $Y(t)$ ($t = \dots, 0, 1, 2, \dots$) – универсум, определенный на множестве действительных чисел. Предположим, что $F(t)$ представляет собой набор функций $f_i(t)$ ($i = 1, 2, \dots$), заданных на универсуме $Y(t)$ ($t = \dots, 0, 1, 2, \dots$). Тогда $F(t)$ называется нечетким временным рядом на универсуме $Y(t)$. Пусть $F(t) = F(t-1) \circ R(t, t-1)$, где $R(t, t-1)$ – нечеткое отношение и \circ – операция \max – \min -композиции. Обозначим зависимость $F(t)$ от $F(t-1)$ как: $F(t-1) \rightarrow F(t)$, где $F(t-1)$ и $F(t)$ – нечеткие множества. Если $F(t)$ зависит от $F(t-1)$, $F(t-2)$, ..., $F(t-k)$, то нечеткая логическая зависимость представляется как: $F(t-k), \dots, F(t-2), F(t-1) \rightarrow F(t)$ и называется однофакторной k -порядковой моделью прогноза на основе нечетких временных рядов. Рассмотрим однофакторную модель прогноза при $k=1$. При этом нечеткая логическая зависимость имеет вид:

$$F(t-1) \rightarrow F(t). \quad (2)$$

Представим нечеткие данные i -го и $(i+1)$ -го периодов (дней, месяцев, кварталов, годов и т.п.) как нечеткие множества A_j и A_k , определенные на универсуме U . Тогда нечеткая логическая зависимость может быть представлена в виде: $A_j \rightarrow A_k$, где A_j – текущее состояние, а A_k – следующее состояние нечеткой логической зависимости.

Пусть $f_i(t) = T_i$, ($i = 1, 2, \dots$) – реальные значения временного ряда для некоторого фактора. Определим универсум U для приращений значений фактора как $U = [D_{\min} - D_1, D_{\max} + D_2]$, где D_{\min} ($D_{\min} = \min_t (f(t) - f(t-1))$) и D_{\max}

$(D_{max} = \max_t(f(t) - f(t-1)))$ – минимальное и максимальное приращения значений фактора на основе известных данных временного ряда соответственно, а D_1 и D_2 – два действительных числа, использование которых позволяет разбить универсум U на n интервалов равной длины: u_1, u_2, \dots, u_n [1, 3].

Определим лингвистические термы A_r ($r = \overline{1, n}$) в виде [2]:

$$A_1 = 1/u_1 + 0,5/u_2 + 0/u_3 + \dots + 0/u_{n-1} + 0/u_n,$$

$$A_2 = 0,5/u_1 + 1/u_2 + 0,5/u_3 + 0/u_4 + \dots + 0/u_n,$$

$$A_3 = 0/u_1 + 0,5/u_2 + 1/u_3 + 0,5/u_4 + \dots + 0/u_n,$$

$$\dots$$

$$A_n = 0/u_1 + 0/u_2 + \dots + 0/u_{n-2} + 0,5/u_{n-1} + 1/u_n.$$

Выполним фазификацию данных, определив интервалы u_r ($r = \overline{1, n}$), которым принадлежит приращения фактора. Если значение приращения фактора принадлежит интервалу u_1 , то соответствующее ему нечеткое значение имеет вид: $X_1 = 1/A_1 + 0,5/A_2$. Если значение приращения фактора принадлежит интервалу u_r , то соответствующее ему нечеткое значение имеет вид: $X_r = 0,5/A_{r-1} + 1/A_r + 0,5/A_{r+1}$, $r = \overline{2, n-1}$. Если значение приращения фактора принадлежит интервалу u_n , то соответствующее ему нечеткое значение имеет вид: $X_n = 0,5/A_{n-1} + 1/A_n$.

Пусть X_k и X_j – нечеткие значение фактора для i -го и $(i+1)$ -периодов соответственно. Тогда для i -го периода можно записать нечеткую логическую зависимость вида: $X_k \rightarrow X_j$. Определим нечеткие логические зависимости для всех известных значений временного ряда, а затем объединим нечеткие логические зависимости с одинаковой левой частью в одну группу. Так, если были сформированы нечеткие зависимости $X_k \rightarrow X_j$, $X_k \rightarrow X_l$, $X_k \rightarrow X_s$, то они объединяются в группу $X_k \rightarrow X_j, X_l, X_s$, функция принадлежности которой определяется как $f_X(u_r) = \max_{r=1, n}(f_{X_j}(u_r), f_{X_l}(u_r), f_{X_s}(u_r))$ [2].

Результирующее НМ для прогнозируемого значения временного ряда для $(i+1)$ -го периода находится как объединение нечетких множеств, входящих в правую часть группы нечетких зависимостей для i -го периода [2].

При этом не учитывается то обстоятельство, что при формировании групп нечетких логиче-

ских зависимостей могли быть выявлены повторяющиеся элементы (нечеткие множества) в правых частях групп.

Искомое значение прогнозируемой величины находится как сумма реального значения временного ряда фактора T_i для i -го периода и дефазифицированного (четкого) значения приращения фактора y_{i+1} :

$$F_{i+1} = T_i + y_{i+1}. \quad (3)$$

Четкое значение приращения фактора для $(i+1)$ -го периода находится по методу центра тяжести для одноточечных множеств

$$y_{i+1} = \sum_{r=1}^n c_r \cdot w_r / \sum_{r=1}^n c_r, \quad (4)$$

где n – количество интервалов u_r ($r = \overline{1, n}$), w_r – средняя точка r -го интервала, c_r – значение степени принадлежности для r -го интервала результирующего НМ, описывающего группу нечетких зависимостей.

Учет повторяющихся НМ в правых частях нечетких логических групп позволил бы повысить точность прогноза. Если при формировании групп нечетких зависимостей были выявлены повторяющиеся НМ в правых частях групп, то значение приращения для $(i+1)$ -го периода следует вычислять по формуле:

$$y_{i+1} = \sum_{r=1}^n v_r \cdot w_r / \sum_{r=1}^n v_r, \quad (5)$$

где w_r – дефазифицированное значение (по формуле (4)), соответствующее НМ X_r , v_r – количество повторений НМ X_r в правой части нечеткой логической зависимости, описывающей прогноз для $(i+1)$ -го периода.

Средняя относительная ошибка прогноза ($AFER$ – average forecasting error rate) может быть вычислена как:

$$AFER = \frac{\sum_{i=1}^m |(F_i - T_i)/T_i|}{m} \cdot 100\%, \quad (6)$$

где F_i и T_i – предсказанное и реальное значения для i -го периода, m – количество значений временного ряда.

При наступлении $(i+1)$ -го периода могут выполняться минимальные корректировки групп нечетких логических зависимостей, если новое значение временного ряда и соответствующая ему зависимость $X_k \rightarrow X_j$ ранее не были учтены в группах нечетких зависимостей.

2. Разработка генетического алгоритма для повышения точности прогнозирования. Самостоятельной задачей при прогнозировании

на основе нечетких временных рядов является определение оптимальных значений, обеспечивающих максимальную точность прогнозирования, для действительных чисел D_1 , D_2 , используемых при корректировке универсума U , и количества интервалов разбиения n универсума U . Применение генетического алгоритма (ГА) позволяет значительно сократить время поиска оптимальных значений параметров D_1 , D_2 и n [5]. При этом хромосома s будет иметь вид: $s = (D_1, D_2, n)$. Для каждого элемента хромосомы следует задать диапазоны их изменения: для D_1 – $[-d_1; 0]$, для D_2 – $[0; d_2]$, для n – $[2; n_{max}]$, где d_1 , d_2 – положительные действительные числа, $d_i = D_{max} - D_{min}$, $i = 1, 2$; n_{max} – натуральное число, $n_{max} \leq m - 1$ (m – количество значений временного ряда).

В качестве функции соответствия можно выбрать функцию (6), определяющую среднюю относительную ошибку прогноза. Однако особенностью данного ГА является то, что при вычислении функции соответствия по формуле (6) для хромосом как начальной популяции размера P , так и расширенной популяции размера $(P + P * P_c)$ (где P_c – коэффициент скрещивания) может быть получено значение вида «0/0» (если имеются группы логических нечетких зависимостей с неопределенными правыми частями, и поэтому невозможно вычислить прогнозное значение по формулам (3), (4) или (3)-(5), так как значение приращения фактора y_{i+1} определяется как «0/0»). Поэтому при удалении из популяции размером $(P + P * P_c)$ хромосом с худшими значениями функций соответствия следует предварительно оценить хромосомы со значением функции соответствия «0/0» как наихудшие (таким хромосомам можно поставить в соответствие максимально возможное значение ошибки, равное 100 %). Тогда при сортировке хромосом по возрастанию функций соответствия хромосомы со значением функции соответствия «0/0» будут занимать последние места в списке, и в результате отбраковки $P * P_c$ худших хромосом будут исключаться из популяции. Таким образом, могут возникать ситуации, когда функции соответствия хромосом принимают значения вида «0/0» (такие хромосомы исключаются из популяции), либо когда удается найти значения D_1 , D_2 и n , обеспечивающие минимальное значение средней относительной ошибки прогнозирования по формуле (6), однако нет гарантии, что удастся выполнить прогнозирование следующего значения временного ряда, так как

группа нечетких зависимостей для предпоследнего значения временного ряда может иметь неопределенную правую часть.

Чтобы обеспечить выполнение прогноза для следующего значения временного ряда, необходимо видоизменить функцию соответствия в ГА следующим образом. Если для некоторой хромосомы $s = (D_1, D_2, n)$ определены все правые части групп нечетких логических зависимостей, то функция соответствия определяется по формуле (6). Если для некоторой хромосомы $s = (D_1, D_2, n)$ не определена хотя бы одна правая часть в группах нечетких логических зависимостей, то значение функции соответствия находится как сумма средней относительной ошибки прогноза по формуле (6) и числа 100. В результате данная хромосома, несмотря на то, что для нее средняя относительная ошибка прогноза по формуле (6) может быть минимальной (если значение функции соответствия не определяется как «0/0»), будет признана одной из худших при реализации ГА и, возможно, будет исключена из популяции. Таким образом, видоизмененная функция соответствия имеет вид:

$$J(s) = \begin{cases} AFER, & \text{определены все} \\ & \text{правые части,} \\ AFER + 100, & \text{не определена хотя} \\ & \text{бы одна правая часть} \end{cases} \quad (7)$$

где $AFER$ определяется по формуле (6).

Хромосома, которая обеспечивает минимум целевой функции по формуле (7), имеет больше шансов быть признанной лучшей. Выбор родителя будет состоять в выборе лучшей хромосомы, которая минимизирует функцию (7), из двух случайно выбранных. Затем две выбранные таким образом хромосомы-родители будут использоваться для скрещивания.

При выполнении скрещивания выбирается коэффициент скрещивания R_c и генерируется случайное число $N_c = random([0, 1])$. Если $R_c > N_c$, то случайным образом выбирается точка скрещивания z и выполняется скрещивание.

При выполнении мутации выбирается коэффициент мутации R_m и генерируется случайное число $N_m = random([0, 1])$. Если $R_m > N_m$, то случайным образом выбирается точка мутации z .

Тогда генетический алгоритм имеет вид:

1. Создается начальная популяция размера P из случайным образом выбранных хромосом $s = (D_1, D_2, n)$.
2. При $g < G$ (g и G – текущее и максимальное количество поколений соответственно)

вычисляется функция соответствия по формуле (7) для каждой хромосомы, затем создается $P/2$ пар хромосом-родителей и осуществляется переход к шагу 3. При $g > G$ осуществляется переход к шагу 5.

3. Выполняются операции скрещивания и мутации для текущей популяции.

4. Создается новая генерация размера P , дополненная хромосомами-детьми, а хромосомы с худшими значениями функции соответствия по формуле (7) отбрасываются.

5. Выбирается хромосома, минимизирующая функцию соответствия по формуле (7).

3. Пример прогнозирования. Данные по фактору «численность безработных по методологии МОТ» в России для периодов 1-2004 – 6-2005 приведены в таблице 1 [4].

При разработке модели прогнозирования были определены приращения фактора для периодов 2-2004 – 5-2005, а также минимальное и максимальное значения приращений:

$$D_{min} = -0,45 \text{ и } D_{max} = 0,3.$$

Для периода 6-2005 были вычислены прогнозное значение фактора и относительная ошибка прогноза по сравнению с реальным значением в соответствии с формулами (3), (4) и (3)-(5) без применения и с применением генетического алгоритма.

Универсум U для приращений значений прогнозируемого фактора разбивался на 2-5 интервалов, на которых формировались нечеткие логические зависимости вида $X_k \rightarrow X_j$. В таблицах 3, 4 и 5, 6 приведены результаты прогнозирования без использования генетического алгоритма.

Если при формировании групп нечетких логических зависимостей не учитываются (или не выявлены) повторяющиеся НМ в правых частях зависимостей, то прогнозируемое значение может быть получено в результате дефазификации результирующего НМ [3], описывающего правую часть группы (таблица 2).

Таблица 1 – Численность безработных по методологии МОТ

| месяц-год | Фактор, млн. чел. | Приращение, млн. чел. | Функция принадлежности | Нечеткое множество |
|-----------|-------------------|-----------------------|-----------------------------------|--------------------|
| 1-2004 | 6,62 | - | - | - |
| 2-2004 | 6,92 | 0,3 | $0,5 / A_4 + 1 / A_5$ | X_5 |
| 3-2004 | 6,48 | -0,44 | $1 / A_1 + 0,5 / A_2$ | X_1 |
| 4-2004 | 6,03 | -0,45 | $1 / A_1 + 0,5 / A_2$ | X_1 |
| 5-2004 | 5,58 | -0,45 | $1 / A_1 + 0,5 / A_2$ | X_1 |
| 6-2004 | 5,53 | -0,05 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 7-2004 | 5,47 | -0,06 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 8-2004 | 5,42 | -0,05 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 9-2004 | 5,67 | 0,25 | $0,5 / A_3 + 1 / A_4 + 0,5 / A_5$ | X_4 |
| 10-2004 | 5,90 | 0,23 | $0,5 / A_3 + 1 / A_4 + 0,5 / A_5$ | X_4 |
| 11-2004 | 6,14 | 0,24 | $0,5 / A_3 + 1 / A_4 + 0,5 / A_5$ | X_4 |
| 12-2004 | 6,11 | -0,03 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 1-2005 | 6,08 | -0,03 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 2-2005 | 6,06 | -0,02 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 3-2005 | 6,00 | -0,06 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 4-2005 | 5,76 | -0,24 | $0,5 / A_1 + 1 / A_2 + 0,5 / A_3$ | X_2 |
| 5-2005 | 5,64 | -0,12 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |
| 6-2005 | 5,55 | -0,09 | $0,5 / A_2 + 1 / A_3 + 0,5 / A_4$ | X_3 |

Таблица 2 – Представление значений фактора нечеткими временными рядами

| месяц-год | Приращение, млн. чел. | НМ | Выходное НМ, описывающее правую часть нечеткой логической группы |
|-----------|-----------------------|-------|--|
| 1-2004 | - | - | - |
| 2-2004 | 0,3 | X_5 | - |
| 3-2004 | -0,44 | X_1 | $1/A_1 + 0,5/A_2 + 0/A_3 + 0/A_4 + 0/A_5 + 0/A_6$ |
| 4-2004 | -0,45 | X_1 | $1/A_1 + 0,5/A_2 + 1/A_3 + 0,5/A_4 + 0/A_5 + 0/A_6$ |
| 5-2004 | -0,45 | X_1 | $1/A_1 + 0,5/A_2 + 1/A_3 + 0,5/A_4 + 0/A_5 + 0/A_6$ |
| 6-2004 | -0,05 | X_3 | $1/A_1 + 0,5/A_2 + 1/A_3 + 0,5/A_4 + 0/A_5 + 0/A_6$ |
| 7-2004 | -0,06 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 8-2004 | -0,05 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 9-2004 | 0,25 | X_4 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 10-2004 | 0,23 | X_4 | $0/A_1 + 0,5/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 11-2004 | 0,24 | X_4 | $0/A_1 + 0,5/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 12-2004 | -0,03 | X_3 | $0/A_1 + 0,5/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 1-2005 | -0,03 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 2-2005 | -0,02 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 3-2005 | -0,06 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 4-2005 | -0,24 | X_2 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |
| 5-2005 | -0,12 | X_3 | $0/A_1 + 0,5/A_2 + 1/A_3 + 0,5/A_4 + 0/A_5 + 0/A_6$ |
| 6-2005 | -0,09 | X_3 | $0,5/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 0,5/A_5 + 0/A_6$ |

Наилучшие результаты - с минимальным значением средней относительной ошибки прогноза - были получены при $n=5$. При этом $D_1 = 0,2$ и $D_2 = 0,2$. Группы нечетких логических зависимостей (без учета повторов НМ в правой части группы) имеют следующий вид:

Группа 1: $X_1 \rightarrow X_1, X_3$.

Группа 2: $X_2 \rightarrow X_3$.

Группа 3: $X_3 \rightarrow X_2, X_3, X_4$.

Группа 4: $X_4 \rightarrow X_3, X_4$.

Группа 5: $X_5 \rightarrow X_1$.

Приращения фактора при разбиении на 5 интервалов могут быть описаны лингвистическими терминами: X_1 – «большое отрицательное приращение», X_2 – «малое отрицательное приращение», X_3 – «приращение, близкое к нулю», X_4 – «малое положительное приращение», X_5 – «большое положительное приращение».

Прогнозируемое значение («Прогноз 1») F_i для i -го периода вычислялось в соответствии с формулами (3), (4). При этом средняя относительная ошибка прогноза по формуле (6) равна 1,942345 %, а относительная ошибка для периода 6-2005 составила 0,270270 %.

Результаты прогнозирования при $n=5$ и значения ошибок прогнозирования при различных n отображены в таблицах 3 и 4 соответственно.

Как только становится известно новое значение временного ряда (например, для периода 6-2005), для него вычисляется значение приращения и определяется соответствующее НМ (таблица 1).

Далее осуществляется корректировка одной из групп нечетких логических зависимостей (если для вновь определенного НМ возникла новая зависимость вида $X_k \rightarrow X_j$, которая не была учтена ранее ни в одной из групп). Для периода 6-2005 было определено НМ X_3 и зависимость $X_3 \rightarrow X_3$, которая уже была учтена ранее в группе 3. В данном случае добавление нового значения к временному ряду и вычисление прогнозируемого значения по формулам (3), (4) не отразились ни на средней ошибке прогноза, ни на ошибке для периода 6-2005.

Однако возможно изменение прогнозируемых значений для тех периодов временного ряда, для которых реализуется нечеткая логическая зависимость, описываемая измененной (дополненной) группой нечетких логических зависимостей. При необходимости может быть выполнена

корректировка диапазона задания универсума U .

С учетом повторяющихся НМ группы нечетких логических зависимостей при $n=5$ имеют следующий вид:

Группа 1: $X_1 \rightarrow X_1(2), X_3(1)$.

Группа 2: $X_2 \rightarrow X_3(1)$.

Группа 3: $X_3 \rightarrow X_2(1), X_3(5), X_4(1)$.

Группа 4: $X_4 \rightarrow X_3(1), X_4(2)$.

Группа 5: $X_5 \rightarrow X_1(1)$.

В круглых скобках указано, сколько раз НМ участвовало в зависимости вида $X_k \rightarrow X_j$ для имеющихся значений временного ряда.

Прогнозируемое значение («Прогноз 2») F_i для i -го периода вычислялось в соответствии с формулами (3)-(5).

Средняя относительная ошибка прогноза равна 1,797013 %, а относительная ошибка для периода 6-2005 – 0,270270%.

Результаты прогнозирования при $n=5$ и

значения ошибок прогнозирования при различных n отображены в таблицах 5 и 6 соответственно.

На рисунке 1 приведены графические зависимости для реальных и прогнозируемых значений временного ряда фактора «численность безработных по методологии МОТ» при $n=5$ без учета и с учетом количества повторяющихся НМ в правых частях групп логических зависимостей («Прогноз 1» и «Прогноз 2» соответственно).

Анализ ошибок прогнозирования из таблиц 3 и 4, а также из таблиц 5 и 6 показал увеличение точности прогноза при учете количества повторяющихся НМ в правых частях групп логических зависимостей и подтвердил преимущество предлагаемого метода по сравнению с методом, изложенным в [2].

Исходя из таблиц 5 и 6, можно сделать вывод, что количество интервалов разбиения существенным образом влияет на результаты прогнозирования.

Таблица 3 – Прогноз численности безработных по формулам (3), (4)

| месяц-год | Фактор, млн. чел. | «Прогноз 1», млн. чел. | Абсолютная ошибка | Относительная ошибка, % |
|-----------|-------------------|------------------------|-------------------|-------------------------|
| 1-2004 | 6,62 | - | - | - |
| 2-2004 | 6,92 | - | - | - |
| 3-2004 | 6,48 | 6,461667 | -0,018333 | 0,282922 |
| 4-2004 | 6,03 | 6,251667 | 0,221667 | 3,676064 |
| 5-2004 | 5,58 | 5,801667 | 0,221667 | 3,972521 |
| 6-2004 | 5,53 | 5,351667 | -0,178333 | 3,224834 |
| 7-2004 | 5,47 | 5,455000 | -0,015000 | 0,274223 |
| 8-2004 | 5,42 | 5,395000 | -0,025000 | 0,461255 |
| 9-2004 | 5,67 | 5,345000 | -0,325000 | 5,731922 |
| 10-2004 | 5,90 | 5,710000 | -0,190000 | 3,220339 |
| 11-2004 | 6,14 | 5,940000 | -0,200000 | 3,257329 |
| 12-2004 | 6,11 | 6,180000 | 0,070000 | 1,145663 |
| 1-2005 | 6,08 | 6,035000 | -0,045000 | 0,740132 |
| 2-2005 | 6,06 | 6,005000 | -0,055000 | 0,907591 |
| 3-2005 | 6,00 | 5,985000 | -0,015000 | 0,250000 |
| 4-2005 | 5,76 | 5,925000 | 0,165000 | 2,864583 |
| 5-2005 | 5,64 | 5,685000 | 0,045000 | 0,797872 |
| 6-2005 | 5,55 | 5,565000 | 0,015000 | 0,270270 |

Таблица 4 – Значения ошибок прогнозирования по формулам (3), (4)

| n ошибка | $n=2$ | $n=3$ | $n=4$ | $n=5$ |
|-----------------------------|----------|----------|-----------|----------|
| средняя относительная, % | 2,647391 | 2,823276 | 2,744280 | 1,942345 |
| относительная для 6-2005, % | 0,270270 | 1,651652 | 2,319820 | 0,270270 |
| абсолютная для 6-2005 | 0,015000 | 0,091667 | -0,128750 | 0,015000 |

Таблица 5 – Прогноз численности безработных по формулам (3)-(5)

| месяц-год | Фактор, млн. чел. | «Прогноз 2», млн. чел. | Абсолютная ошибка | Относительная ошибка, % |
|-----------|-------------------|------------------------|-------------------|-------------------------|
| 1-2004 | 6,62 | - | - | - |
| 2-2004 | 6,92 | - | - | - |
| 3-2004 | 6,48 | 6,461667 | -0,018333 | 0,282922 |
| 4-2004 | 6,03 | 6,149444 | 0,119444 | 1,980837 |
| 5-2004 | 5,58 | 5,699444 | 0,119444 | 2,140581 |
| 6-2004 | 5,53 | 5,249444 | -0,280556 | 5,073337 |
| 7-2004 | 5,47 | 5,455000 | -0,015000 | 0,274223 |
| 8-2004 | 5,42 | 5,395000 | -0,025000 | 0,461255 |
| 9-2004 | 5,67 | 5,345000 | -0,325000 | 5,731922 |
| 10-2004 | 5,90 | 5,748333 | -0,151667 | 2,570621 |
| 11-2004 | 6,14 | 5,978333 | -0,161667 | 2,633008 |
| 12-2004 | 6,11 | 6,218333 | 0,108333 | 1,773050 |
| 1-2005 | 6,08 | 6,035000 | -0,045000 | 0,740132 |
| 2-2005 | 6,06 | 6,005000 | -0,055000 | 0,907591 |
| 3-2005 | 6,00 | 5,985000 | -0,015000 | 0,250000 |
| 4-2005 | 5,76 | 5,925000 | 0,165000 | 2,864583 |
| 5-2005 | 5,64 | 5,685000 | 0,045000 | 0,797872 |
| 6-2005 | 5,55 | 5,565000 | 0,015000 | 0,270270 |

Таблица 6 – Значения ошибок прогнозирования по формулам (3)-(5)

| ошибка n | $n = 2$ | $n = 3$ | $n = 4$ | $n = 5$ |
|-----------------------------|-----------|----------|-----------|----------|
| средняя относительная, % | 2,473131 | 2,363196 | 2,922409 | 1,797013 |
| относительная для 6-2005, % | 0,593093 | 0,845846 | 2,319820 | 0,270270 |
| абсолютная для 6-2005 | -0,032917 | 0,046944 | -0,128750 | 0,015000 |

Таблица 7 – Прогноз численности безработных по формулам (3)-(5) на основе генетического алгоритма

| месяц-год | Фактор, млн. чел. | «Прогноз с ГА», млн. чел. | Абсолютная ошибка | Относительная ошибка, % |
|-----------|-------------------|---------------------------|-------------------|-------------------------|
| 1-2004 | 6,62 | - | - | - |
| 2-2004 | 6,92 | - | - | - |
| 3-2004 | 6,48 | 6,460215 | -0,019785 | 0,305322 |
| 4-2004 | 6,03 | 6,153451 | 0,123451 | 2,047287 |
| 5-2004 | 5,58 | 5,703451 | 0,123451 | 2,212391 |
| 6-2004 | 5,53 | 5,253451 | -0,276549 | 5,000878 |
| 7-2004 | 5,47 | 5,469924 | -0,000076 | 0,001389 |
| 8-2004 | 5,42 | 5,409924 | -0,010076 | 0,185904 |
| 9-2004 | 5,67 | 5,359924 | -0,310076 | 5,468712 |
| 10-2004 | 5,90 | 5,769808 | -0,130192 | 2,206651 |
| 11-2004 | 6,14 | 5,999808 | -0,140192 | 2,283264 |
| 12-2004 | 6,11 | 6,239808 | 0,129808 | 2,124510 |
| 1-2005 | 6,08 | 6,049924 | -0,030076 | 0,494671 |
| 2-2005 | 6,06 | 6,019924 | -0,040076 | 0,661320 |
| 3-2005 | 6,00 | 5,999924 | -0,000076 | 0,001266 |
| 4-2005 | 5,76 | 5,939924 | 0,179924 | 3,123681 |
| 5-2005 | 5,64 | 5,699924 | 0,059924 | 1,062483 |
| 6-2005 | 5,55 | 5,579924 | 0,029924 | 0,539172 |

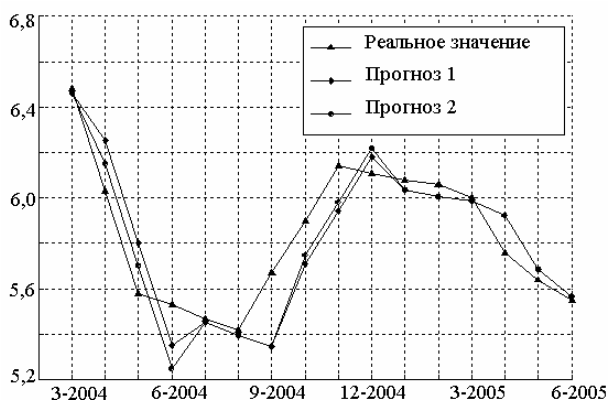


Рисунок 1 – Графические зависимости для реальных и прогнозируемых значений фактора «численность безработных по методологии МОТ»

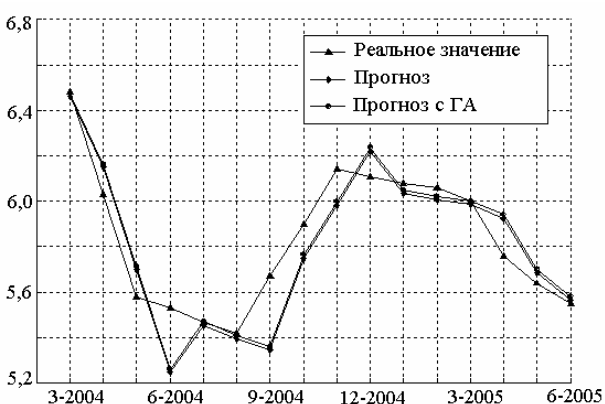


Рисунок 2 – Графические зависимости для реальных и прогнозируемых значений фактора «численность безработных по методологии МОТ» на основе генетического алгоритма

Однако стремление использовать чрезмерно большое количество интервалов может привести к невозможности применения данного метода ввиду того, что для i -го периода может быть не сформирована логическая зависимость вида $X_k \rightarrow X_j$ (так как не задана ее правая часть: $X_k \rightarrow ?$), то есть на основе имеющихся данных временного ряда невозможно будет сказать, какое НМ вероятно для $(i+1)$ -го периода.

Применение ГА для прогнозирования фактора «численность безработных по методологии МОТ» по формулам (3)-(5) при $D_1 \in [-d_1; 0]$, $D_2 \in [0; d_2]$, $d_i = D_{max} - D_{min} = 0,75$, $i = 1, 2$; $n \in [2; n_{max}]$, $n_{max} = m - 1 = 17$, $m = 18$ позволило определить оптимальное число разбиений универсума U на 5 интервалов, а также числа $D_1 = 0,2096393$ и $D_2 = 0,2394873$ для корректировки границ универсума.

Для найденных значений D_1 , D_2 и n дости-

гается минимальное значение средней относительной ошибки прогнозирования по формуле (7), равное 1,732431 %. Относительная ошибка для периода 6-2005 составила 0,539172 %. В таблице 7 приведены результаты прогнозирования по формулам (3)-(5) с применением ГА при $n = 5$. На рисунке 2 приведены графические зависимости для реальных и прогнозируемых значений временного ряда фактора «численность безработных по методологии МОТ» при $n = 5$ с учетом количества повторяющихся НМ в правых частях нечетких логических зависимостей без использования и с использованием ГА («Прогноз» и Прогноз с ГА» соответственно).

Таким образом, применение ГА обеспечило более высокую точность прогноза за счет оптимального выбора параметров D_1 , D_2 и n (при одновременном сокращении времени подбора оптимальных значений параметров D_1 , D_2 и n), а использование функции соответствия в виде (7) позволило исключить из рассмотрения группы нечетких логических зависимостей с пустыми правыми частями.

Заключение. Предлагаемый метод прогнозирования тенденций рынка труда обеспечивает получение более высоких результатов прогноза, чем предложенный в [3], и может быть рекомендован для проведения краткосрочных прогнозов. Учет повторяющихся элементов в правых частях групп нечетких логических зависимостей и применение ГА для поиска оптимальных значений числа разбиений n универсума U и действительных чисел D_1 , D_2 корректировки универсума обеспечили более высокую точность прогноза. Для повышения точности прогноза можно, кроме того, использовать многофакторные высокого порядка ($k > 1$) нечеткие временные ряды (где один фактор является главным – прогнозируемым, а другие – вспомогательными) и представлять значения временных рядов с помощью нечетких множеств второго типа.

Библиографический список

1. Демидова Л.А. Прогнозирование тенденций рынка труда на основе однофакторных нечетких временных рядов // Системы управления и информационные технологии. 2007. № 3.2 (29). - С. 241-246.
2. S.M. Chen. Forecasting enrollments based on fuzzy time series // Fuzzy Sets Systems, 1996, vol. 81, no. 3, pp. 311-319.
3. <http://users.ecs.soton.ac.uk/~ms305r/icci.ppt>
4. <http://www.mzsrrf.ru/stat/40.html>
5. Ярушклина Н.Г. Основы теории нечетких и гибридных систем: учеб. пособие. – М.: Финансы и статистика. 2004. 320 с.: ил.