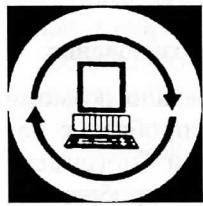


ванных процедур обработки знаний. Диалог пользователя с СБЗ осуществляется с помощью меню.

Список литературы

1. Цаленко М. Ш. Моделирование семантики в базах данных. М.: Наука, 1989. 288 с.
2. Миронов А. С. О подходе к организации и реализации одного семейства интеллектуальных информационных систем // Информационные технологии. 2003. № 7. С. 32–34.

3. Попов Э. В. Общение с ЭВМ на естественном языке. М.: Наука, 1982. 360 с.
4. Любарский Ю. Я. Интеллектуальные информационные системы. М.: Наука, 1990. 232 с.
5. Ефимов Е. И. Решатели интеллектуальных задач. М.: Наука, 1982. 320 с.
6. Искусственный интеллект. В 3-х кн. Кн. 2. Модели и методы: Справочник / Под. ред. Д. А. Поспелова. М.: Радио и связь. 1990. 304 с.
7. Петрушин В. А. Интеллектуальные обучающие системы: архитектура и методы реализации (обзор) // Изв. РАН. Техн. кибернетика. 1993. № 2. С. 164–189.



ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

УДК 004.3.06

М. Г. Мамедова, д-р техн. наук,
З. Г. Джабраилова, канд. техн. наук,
Институт информационных технологий
Нац.АН Азербайджана
depart 15@iit.ab.az; masuma@mail.az

Применение нечеткой логики в демографическом прогнозе

Исследованы возможности применения нечеткой логики в моделировании демографических процессов на примере задачи прогнозирования численности населения. Выделены особенности населения как динамической системы, функционирующей в условиях неопределенности, и дана нечеткая постановка задачи прогнозирования. Предложена методика прогнозирования численности населения с использованием модели нечетких временных рядов. Описаны эксперименты по ретроспективному и перспективному расчету населения и приведен сравнительный анализ результатов, полученных альтернативными методами прогнозирования.

Введение

Современный этап развития общества, характеризующийся внедрением рыночных механизмов во все сферы жизнедеятельности, в том числе и в сферу демографических процессов, предъявляет более высокие требования к перспективным расчетам ожидаемой численности и состава населения.

Демографический прогноз, будучи неотъемлемым элементом социально-экономического развития, позволяет определить на перспективу численность всего населения, численность отдельных половозрастных групп, населения в трудоспособном возрасте и т. д. Указанные показатели в той или иной мере учитываются при составлении научно обоснованной политики социально-экономического развития и формировании комплекса практических мер по ее реализации.

В настоящее время существуют различные подходы к постановке задач прогнозирования численности населения и предлагается широкий спектр методов моделирования для их решения. В качестве примера можно отметить статистические, адаптивные, имитационные модели, гладкие динамические ряды, авторегрессии и т. п.

Развитие информационных технологий и программных средств обусловили новые возможности для моделирования демографических процессов и решения прогнозных задач. Однако, как показывают прикладные исследования последних лет, внедрение традиционных методов анализа и моделирования процессов роста населения как объекта исследования, основанных на обработке численных данных, часто не дает желаемого эффекта и сопровождается значительными погрешностями и риском. Одной из причин сложившейся ситуации является недостаточная эффективность многих прогнозных моделей, что отчасти объясняется несовместимостью высокой точности количественных методов классического математического

аппарата с большой сложностью процессов роста численности населения. Другой причиной этого, на наш взгляд, является тот факт, что эти методы, будучи ориентированы на математический анализ точно определенных систем, не в состоянии учесть ряд существенных особенностей объекта исследования. Так, население, безотносительно к конкретной территории или определенной группе, представляет собой большую динамическую систему (типа социальной, экономической, экологической и др.). Отличительной особенностью данной системы является ее функционирование в условиях неопределенности, вызванной рядом обстоятельств. Это, прежде всего, невозможность выявления всех факторов, определяющих динамику развития человеческой популяции, колебание границ используемых в демографическом анализе показателей и возможность варьирования значений ряда из них в довольно широких пределах, отсутствие достаточно полной априорной информации о демографических процессах, связанное как с невозможностью регистрации всех демографических событий, так и с источниками информации. Здесь следует отметить неполноту фактических данных: неопределенности, возникающие при комбинировании данных, полученных из разных источников, таких как результаты переписи населения, выборочные демографические обследования, текущий учет движения населения и т. п. Определенные затруднения создают противоречия между официальными данными, неофициальными данными, а также экспертными оценками.

Процесс роста населения является многофакторным и многокомпонентным, причем этот процесс происходит во времени. Однако сознательно воздействовать на этот процесс путем варьирования одних параметров и слежения за изменением других практически невозможно. На процессы роста населения большое влияние оказывают разного рода внешние неконтролируемые воздействия (войны, межнациональные конфликты, стихийные бедствия, экологические факторы и т. д.). В условиях отсутствия систематизированных статистических материалов по затронутой проблеме единственным источником информации о демографических процессах представляются эмпирические данные.

В этой связи особую актуальность приобретает демографический анализ при неполной и неточной информации. Под ним понимаются как случаи, когда в силу различных причин демографическая информация неполная, так и те, когда часть или все имеющиеся данные такого рода вызывают сомнения. Неопределенность, неясность,

неточность, недостаточность, нечеткость данных о демографических событиях, процессах, оценках и превалирующая в господствующих подходах тенденция отвергать эти характеристики требуют новых подходов к анализу и оценке демографических ситуаций и, в частности, к прогнозу роста численности населения.

В силу указанных причин значительный интерес представляет изучение возможностей применения в моделировании демографических процессов аппарата теории нечетких или размытых множеств, получивших также название нечеткой логики [1].

Нечеткая логика в задачах прогнозирования

Появление нечеткой логики сделало возможным нахождение решения многих проблем с нечеткой исходной информацией или "погруженных" в нечеткую среду. К числу таких проблем относятся и задачи прогнозирования. Так, большая часть составных элементов последних (исходные данные, зависимости между ними, оценки вероятных прогнозных исходов, критериальные оценки показателей и зависимостей между ними, оценки экспертов, суждения специалистов и т. п.), имея нечеткие значения или будучи в нечетких отношениях, способствуют нечеткому описанию задачи прогнозирования.

Применение нечеткой логики для решения прогнозных задач началось с исследований [2, 3], в которых дано описание математической модели временного ряда в нечеткой постановке для случая решения проблемы при нечеткой исходной информации. Впоследствии это направление было развито другими специалистами, занимающимися решением аналогичных проблем [4, 5]. При этом авторы, предлагая модель нечетких временных рядов, старались путем внесения в нее определенных изменений уменьшить среднюю погрешность прогнозирования. Так, в работе [6] проведен сравнительный анализ полученных результатов и сделан вывод о том, что усредненная погрешность прогнозирования по сравнению с другими методами является самой минимальной. В свете указанного в качестве одного из возможных путей использования математических средств для разработки демографических прогнозов, в частности для решения задачи прогнозирования численности населения, является модель нечетких временных рядов. Перечисленные выше особенности населения как динамической системы, функционирующей в условиях неопределенности, предопределяют нечеткость исходной информации,

т. е. "погружение" задачи прогнозирования численности населения в нечеткую среду. С этих позиций как с теоретической, так и с практической точек зрения представляется целесообразным решение данной проблемы в рамках нечетких временных рядов.

Таким образом, цель данной работы — методологическая: 1) на основе модели нечетких временных рядов дать метод оценки параметров модели; 2) проверить меру адекватности модели и исследуемого демографического процесса, т. е. вычислить погрешность метода; 3) провести сравнительный анализ результатов расчета; 4) выявить смысл модели как в теоретическом, так и в практическом плане.

Социально-экономический анализ полученных результатов в задачу авторов не входил.

Общие сведения о временных рядах

Временные ряды представляют собой последовательные ряды наблюдений (измерений определенных показателей), проведенных через одинаковые интервалы времени, и лежат в основе изучения реальных процессов в экономике, метеорологии, технике, естественных науках и др.

Анализ временного ряда наблюдений состоит: 1) из построения математической модели временного ряда наблюдений реального процесса; 2) идентификации модели, т. е. выбора метода количественной оценки параметров модели для проверки ее соответствия реальному процессу; 3) прогонки идентифицированной модели к временному ряду с помощью статистической оценки параметров. Формально временной ряд $\{F(t)\}, t = 1, 2, 3, \dots$ определяется как функция дискретного аргумента t .

Нечеткие временные ряды. Пусть $U = \{u_1, u_2, \dots, u_n\}$ — универсальное множество, например, интервал числовой оси, составленный из интервалов $u_i (i = 1, \dots, m)$. Нечеткое множество A универсального множества U определяется следующим образом:

$$A = \{(\mu_A(u_1)/u_1), (\mu_A(u_2)/u_2), \dots, (\mu_A(u_n)/u_n)\} \text{ или} \\ A = \{(\mu_A(u_i)/u_i)\}, u_i \in U, \mu_A(u_i) \in [0, 1].$$

Здесь $\mu_A(u_i)$ — функция принадлежности, $\mu_A(u_i): U \in [0, 1]$; $\mu_A(u_i)$ показывает степень принадлежности u_i множеству A , $\mu_A(u_i) \in [0, 1]$, "/" — знак разделения.

После "погружения" обычного временного ряда $\{F(t)\}, t = 1, 2, 3, \dots$ в нечеткую среду получим нечеткую функцию $\tilde{F}(t)$ аргумента t в универсальном множестве U со значениями в виде нечетких интер-

валов A^t с функцией принадлежности $\mu_{A^t}(u_i)$, т. е. $A^t = \{(\mu_{A^t}(u_i)/u_i)\}, u_i \in U, \mu_{A^t}(u_i) \in [0, 1]$. Это означает, что любая точка x (в том числе и $x = F(t)$) из интервала u_i будет принадлежать множеству A^t со степенью принадлежности $\mu_{A^t}(u_i) = \mu_i(t)$ с заданными числами $\mu_i(t)$, ($i = 1, \dots, m$) при каждом фиксированном $t = 1, 2, \dots$

Таким образом, получаем нечеткий временной ряд $\{\tilde{F}(t)\}$ в виде множества нечетких множеств A^t , $t = 1, 2, \dots$. Если составляющие этого ряда A^t принимают лингвистические значения, то $\tilde{F}(t)$ — нечеткая функция аргумента t , значениями которой являются нечеткие словесные понятия типа "много", "мало", "значительно" и т. п., которыми на практике оперирует человек. Лингвистические переменные позволяют ввести в рассмотрение качественные описания и учесть неопределенность решаемой задачи, достигнуть полного описания тех факторов, которые не поддаются точному количественному описанию.

Нечеткие временные ряды в демографическом прогнозе. Изменение интенсивности демографических процессов под воздействием социально-экономических факторов делает проблему определения перспектив изменения численности населения одной из важнейших при разработке демографических прогнозов. В настоящей работе для решения задачи прогнозирования численности населения в качестве модели принят нечеткий временный ряд и дана следующая **постановка задачи**. Считаются известными данные об общей численности населения Азербайджана за фиксированный период времени, т. е. динамика и соответствующие вариации численности населения за этот же отрезок времени. Задача состоит в определении перспективной численности населения с учетом вариаций за истекшие годы.

Для решения этой задачи руководствуемся приведенными ниже принципами.

1. Поскольку предлагаемый подход впервые применяется к демографическому прогнозу численности населения, то для идентификации модели, т. е. проверки ее соответствия реальному процессу, необходимо прежде всего дать "ретроспективный прогноз", суть которого заключается в следующем:

а) в качестве прогнозируемого года принять один из истекших годов ($t = s$) и с учетом вариаций за предшествующие ему годы ($s - 1, s - 2, \dots, s - k$) рассчитать численность населения на этот год;

б) провести сравнительный анализ результатов реализации предложенной модели с ретроспективными данными (фактом за год s) и оценить погрешность метода;

в) эксперимент распространить на определенный отрезок времени;

г) по значениям погрешностей сделать предварительные выводы о целесообразности применения метода.

2. В случае положительных результатов применить метод для перспективного расчета численности населения.

Методика прогнозирования численности населения

В соответствии с постановкой задачи предлагаются следующая методика прогнозирования.

1. Определение универсального множества U , которое представляет собой интервал между наименьшей и наибольшей вариациями численности населения.

2. Деление универсального множества U на несколько интервалов равной длины, включающих значения вариаций, соответствующих различным темпам роста населения.

3. Для словесного качественного описания в понятиях человека значений вариаций численности населения — лингвистической переменной — определение соответствующих лингвистических значений этой переменной, т. е. определение множества нечетких множеств $F(t)$.

4. Фазификация исходных данных, заключающаяся в преобразовании четких количественных значений в нечеткие. Эта операция позволяет в значениях функций принадлежности отразить соответствующие количественным значениям качественные представления о темпах роста населения.

5. Выбор параметра $w > 1$, соответствующего отрезку времени, предшествующему текущему году; вычисление матрицы нечетких отношений $R^W(t)$ и прогнозирование численности населения на последующий год.

6. Дефазификация полученного результата, т. е. переход от нечетких значений к четким (количественным).

Ниже на примере малой размерности приводится схема реализации предложенной методики для решения проблемы прогнозирования численности населения.

1-й шаг. В табл. 1 представлена динамика численности населения Азербайджана за 1980—2001 гг. (исходная база для "ретроспективного прогноза") и соответствующие вариации численности населе-

ния между каждым последующим и предыдущим годами. Под термином "вариация" для текущего года имеется в виду разница между фактической численностью населения текущего и предыдущего годов. Например, вариация для 1990 г. составляет $7\ 131\ 900 - 7\ 021\ 200 = 110\ 700$. Для определения универсального множества U прежде всего необходимо выделить наименьшее и наибольшее значения вариаций в отрезке времени 1980—2001 гг., а затем с целью получения более удобных границ интервала выбрать подходящие значения D_1 и D_2 (соответствующие положительные числа). Тогда универсальное множество U можно описать в виде $U = [V_{\min} - D_1, V_{\max} + D_2]$, где $V_{\min} = 62\ 800$ — наименьшая вариация (2000 г.), $V_{\max} = 115\ 900$ — наибольшая вариация (1993 г.), $D_1 = 1800$, $D_2 = 1100$. После этого универсальное множество U будет выглядеть следующим образом:

$$U = [61\ 000, 117\ 000].$$

2-й шаг. Универсальное множество U следует разделить на несколько интервалов равной длины. В данном случае универсальное множество U разделено на семь интервалов равной длины $u_1 = [61\ 000, 69\ 000]$; $u_2 = [69\ 000, 77\ 000]$; $u_3 = [77\ 000, 85\ 000]$; $u_4 = [85\ 000, 93\ 000]$; $u_5 = [93\ 000, 101\ 000]$; $u_6 = [101\ 000, 109\ 000]$; $u_7 = [109\ 000, 117\ 000]$.

Учитывая тот факт, что усредненная погрешность прогнозирования метода нечетких временных рядов является наименьшей, необходимо отметить средние точки этих интервалов:

$$u_{\text{ср}}^1 = 65\ 000, u_{\text{ср}}^2 = 73\ 000, u_{\text{ср}}^3 = 81\ 000,$$

$$u_{\text{ср}}^4 = 89\ 000;$$

$$u_{\text{ср}}^5 = 97\ 000, u_{\text{ср}}^6 = 105\ 000, u_{\text{ср}}^7 = 113\ 000. \quad (1)$$

3-й шаг. Этот шаг сводится к определению множества нечетких множеств в универсальном множестве U . В данном случае "вариация численности населения" — название лингвистической переменной, которая имеет следующие лингвистические значения: A_1 = (очень низкий уровень роста численности населения (ОНУРЧН)); A_2 = (низкий уровень роста численности населения (НУРЧН)); A_3 = (нет изменений роста численности населения (НИРЧН)); A_4 = (средний уровень роста численности населения (СУРЧН)); A_5 = (нормальный уровень роста численности населения (НорУРЧН)); A_6 = (высокий уровень роста численности населения (ВУРЧН)); A_7 = (очень

Таблица 1

Динамика и соответствующие вариации численности населения за 1980—2001 гг.

Годы	Числ. населения, тыс. чел.	Вариация, тыс. чел.	Фазификация вариации
1980	6 114,3		$A^{81} = \{(0,12/u_1), (0,21/u_2), (0,43/u_3), (0,90/u_4), (0,83/u_5), (0,39/u_6), (0,19/u_7)\}$
1981	6 206,7	92,4	$A^{82} = \{(0,07/u_1), (0,10/u_2), (0,18/u_3), (0,36/u_4), (0,77/u_5), (0,94/u_6), (0,47/u_7)\}$
1982	6 308,8	102,1	$A^{83} = \{(0,09/u_1), (0,014/u_2), (0,27/u_3), (0,58/u_4), (1/u_5), (0,64/u_6), (0,29/u_7)\}$
1983	6 406,3	97,5	$A^{84} = \{(0,05/u_1), (0,08/u_2), (0,13/u_3), (0,24/u_4), (0,5/u_5), (0,96/u_6), (0,74/u_7)\}$
1984	6 513,3	107,0	$A^{85} = \{(0,05/u_1), (0,7/u_2), (0,11/u_3), (0,20/u_4), (0,41/u_5), (0,86/u_6), (0,87/u_7)\}$
1985	6 622,4	109,1	$A^{86} = \{(0,10/u_1), (0,16/u_2), (0,32/u_3), (0,70/u_4), (1/u_5), (0,53/u_6), (0,25/u_7)\}$
1986	6 717,9	95,5	$A^{87} = \{(0,06/u_1), (0,09/u_2), (0,15/u_3), (0,29/u_4), (0,62/u_5), (1/u_6), (0,60/u_7)\}$
1987	6 822,7	104,8	$A^{88} = \{(0,06/u_1), (0,09/u_2), (0,14/u_3), (0,27/u_4), (0,59/u_5), (1/u_6), (0,63/u_7)\}$
1988	6 928,0	105,3	$A^{89} = \{(0,11/u_1), (0,20/u_2), (0,40/u_3), (0,85/u_4), (0,87/u_5), (0,42/u_6), (0,26/u_7)\}$
1989	7 021,2	93,2	$A^{90} = \{(0,05/u_1), (0,07/u_2), (0,10/u_3), (0,18/u_4), (0,35/u_5), (0,75/u_6), (0,95/u_7)\}$
1990	7 131,9	110,7	$A^{91} = \{(0,18/u_1), (0,35/u_2), (0,76/u_3), (0,95/u_4), (0,48/u_5), (0,23/u_6), (0,13/u_7)\}$
1991	7 218,5	86,6	$A^{92} = \{(0,06/u_1), (0,09/u_2), (0,14/u_3), (0,27/u_4), (0,57/u_5), (1/u_6), (0,65/u_7)\}$
1992	7 324,1	105,6	$A^{93} = \{(0,0,4/u_1), (0,05/u_2), (0,08/u_3), (0,12/u_4), (0,22/u_5), (0,46/u_6), (0,93/u_7)\}$
1993	7 440,0	115,9	$A^{94} = \{(0,05/u_1), (0,07/u_2), (0,11/u_3), (0,19/u_4), (0,39/u_5), (0,83/u_6), (0,90/u_7)\}$
1994	7 549,6	109,6	$A^{95} = \{(0,11/u_1), (0,19/u_2), (0,38/u_3), (0,81/u_4), (0,91/u_5), (0,45/u_6), (0,22/u_7)\}$
1995	7 643,5	93,9	$A^{96} = \{(0,24/u_1), (0,52/u_2), (0,97/u_3), (0,72/u_4), (0,33/u_5), (0,17/u_6), (0,10/u_7)\}$
1996	7 726,2	82,7	$A^{97} = \{(0,57/u_1), (1/u_2), (0,65/u_3), (0,30/u_4), (0,15/u_5), (0,09/u_6), (0,06/u_7)\}$
1997	7 799,8	73,6	$A^{98} = \{(0,31/u_1), (0,68/u_2), (0,99/u_3), (0,55/u_4), (0,25/u_5), (0,14/u_6), (0,08/u_7)\}$
1998	7 879,7	79,9	$A^{99} = \{(0,57/u_1), (1/u_2), (0,65/u_3), (0,30/u_4), (0,16/u_5), (0,09/u_6), (0,06/u_7)\}$
1999	7 953,4	73,7	$A^{00} = \{(0,95/u_1), (0,49/u_2), (0,23/u_3), (0,13/u_4), (0,08/u_5), (0,05/u_6), (0,04/u_7)\}$
2000	8 016,2	62,8	$A^{01} = \{(1/u_1), (0,60/u_2), (0,28/u_3), (0,15/u_4), (0,09/u_5), (0,06/u_6), (0,04/u_7)\}$
2001	8 081,0	64,8	

высокий уровень роста численности населения (ОВУРЧН)). Каждому лингвистическому значению соответствует нечеткая переменная, которой по определенному правилу ставится в соответствие нечеткое множество, определяющее смысл этой переменной. Например, лингвистическое значение "очень низкий уровень роста численности населения" задается нечеткой переменной $\langle \text{ОНУРЧН}, [61\ 000, 69\ 000], A_1 \rangle$, где A_1 — нечеткое множество с областью определения $[61\ 000, 69\ 000]$ из универсального множества U может быть, например, описано в виде (3).

Нечеткие множества A_1, A_2, \dots, A_7 в универсальном множестве U определяются с помощью формулы

$$\mu_{A_i}(u_i) = \frac{1}{1 + [C(U - u_{cp}^i)]^2}, \quad (2)$$

где U — вариации из табл. 1, u_{cp}^i — средние точки соответствующих интервалов, значения которых представлены в (1); C — постоянное число, которое подбирается таким образом, чтобы обеспечить преобразование четких количественных чисел в нечеткие, т. е. их вхождение в интервал $[0, 1]$.

(в данной работе $C = 0,0001$); $A_i = (\mu_{A_i}(u_i)/u_i)$, $u_i \in U$, $\mu_{A_i}(u_i) \in [0, 1]$ — нечеткие множества.

Если в формуле (2) в качестве значения переменной U принять средние точки соответствующих интервалов, тогда нечеткие множества A_i ($i = 1, 7$) будут определены следующим образом:

$$\begin{aligned} A_1 &= \{(1/u_1), (0,61/u_2), (0,27/u_3), (0,15/u_4), \\ &\quad (0,10/u_5), (0,06/u_6), (0,04/u_7)\}, \\ A_2 &= \{(0,61/u_1), (1/u_2), (0,61/u_3), (0,27/u_4), \\ &\quad (0,15/u_5), (0,10/u_6), (0,06/u_7)\}, \\ A_3 &= \{(0,27/u_1), (0,61/u_2), (1/u_3), (0,61/u_4), \\ &\quad (0,27/u_5), (0,15/u_6), (0,10/u_7)\}, \\ A_4 &= \{(0,15/u_1), (0,27/u_2), (0,61/u_3), (1/u_4), \\ &\quad (0,61/u_5), (0,27/u_6), (0,15/u_7)\}, \\ A_5 &= \{(0,10/u_1), (0,15/u_2), (0,27/u_3), (0,61/u_4), \\ &\quad (1/u_5), (0,61/u_6), (0,27/u_7)\}, \\ A_6 &= \{(0,06/u_1), (0,10/u_2), (0,15/u_3), (0,27/u_4), \\ &\quad (0,61/u_5), (1/u_6), (0,61/u_7)\}, \\ A_7 &= \{(0,04/u_1), (0,06/u_2), (0,10/u_3), (0,15/u_4), \\ &\quad (0,27/u_5), (0,61/u_6), (1/u_7)\} \end{aligned} \quad (3)$$

Примерный вид непрерывных функций принадлежности нечетких множеств A_i , описывающих значения лингвистической переменной "ва-

риация численности населения", представлен на рис. 1 (см. третью сторону обложки).

4-й шаг. Этот шаг сводится к фазификации вариаций, вычисленных в шаге 1. При этом, если для года i вариация будет V_i , $V_i \in u_j$, то для u_j ($u_j \in U$ — интервалы универсального множества U) функция принадлежности $\mu_{A_i}(u_j)$ вычисляется по формуле (2) с учетом $U = V_i$, т. е. из универсального множества выделяется тот интервал, в который попадает рассматриваемая вариация. Результаты фазификации для всех рассматриваемых годов представлены в табл. 1. Здесь A^{mn} — нечеткие множества соответствующих вариаций за год $t = mn$, где $1981 < t \leq 2001$ (во избежание громоздкости обозначения в табл. 1 указаны в индексе величины A последние две цифры года).

5-й шаг. Необходимо выбрать базис w ($1 < w < l$, где l — число лет, предшествующих текущему году и включенных в экспериментальную оценку). С учетом базиса, т. е. предыстории, вычислить матрицу нечетких отношений $\mathbf{R}^w(t)$, на основе которой выдается прогноз. С этой целью после выбора w строится операционная матрица $-i \times j$ $\mathbf{O}^w(t)$ (i — число строк, соответствующее числу годов в последовательности $t - 2, t - 3, \dots, t - w$; j — число столбцов, соответствующее количеству интервалов вариаций) и матрица-критерий $\mathbf{K}(t)$ для размера $1 \times j$ прогнозируемого года t (матрица-строка, соответствующая нечеткой вариации численности населения за год $t - 1$). Например, принимая $w = 7$, можно определить операционную матрицу $\mathbf{O}^7(t)$ размера 6×7 , т. е. матрицу нечеткой вариации численности населения за годы $t - 2, t - 3, t - 4, t - 5, t - 6, t - 7$, и матрицу-критерий $\mathbf{K}(t)$ размера 1×7 , т. е. матрицу нечеткой вариации численности населения за год $(t - 1)$. Значит при $w = 7$ фактически используются данные для восьми предыдущих годов.

Так, например, для прогноза численности населения на 1990 г. операционная матрица $\mathbf{O}^7(t)$ и матрица-критерий $\mathbf{K}(t)$ будут определены следующим образом:

$$\mathbf{O}^7(1990) = \begin{array}{|c|c|} \hline & \text{нечеткая вариация численности населения на 1983 г.} & A^{83} \\ & \text{нечеткая вариация численности населения на 1984 г.} & A^{84} \\ & \text{нечеткая вариация численности населения на 1985 г.} & A^{85} \\ & \text{нечеткая вариация численности населения на 1986 г.} & A^{86} \\ & \text{нечеткая вариация численности населения на 1987 г.} & A^{87} \\ & \text{нечеткая вариация численности населения на 1988 г.} & A^{88} \\ \hline \end{array}$$

$$\mathbf{O}^7(1990) = \begin{array}{|c|c|c|c|c|c|c|c|} \hline & \text{онурчн} & \text{нурчн} & \text{нирчн} & \text{сурчн} & \text{норурчн} & \text{вурчн} & \text{овурчн} \\ \hline 0.09 & 0.14 & 0.27 & 0.58 & 1 & 0.64 & 0.29 \\ 0.05 & 0.08 & 0.13 & 0.24 & 0.50 & 0.96 & 0.74 \\ 0.05 & 0.07 & 0.11 & 0.20 & 0.41 & 0.86 & 0.87 \\ 0.10 & 0.16 & 0.32 & 0.70 & 1 & 0.53 & 0.25 \\ 0.06 & 0.09 & 0.15 & 0.29 & 0.62 & 1 & 0.60 \\ 0.06 & 0.09 & 0.14 & 0.27 & 0.59 & 1 & 0.63 \\ \hline \end{array}$$

$$\mathbf{K}(1990) = \begin{array}{|c|c|c|c|c|c|c|} \hline & \text{нечеткая вариация численности населения на 1989 г.} & & & & & & \\ \hline & & & & & & & - [A^{89}] \\ \hline \end{array}$$

т. е.

$$\mathbf{K}(1990) = \begin{array}{|c|c|c|c|c|c|c|} \hline & \text{онурчн} & \text{нурчн} & \text{нирчн} & \text{сурчн} & \text{норурчн} & \text{вурчн} & \text{овурчн} \\ \hline 0.11 & 0.20 & 0.40 & 0.85 & 0.87 & 0.42 & 0.26 \\ \hline \end{array}$$

Согласно методу следующий шаг сводится к вычислению матрицы отношений

$\mathbf{R}(t)$:

$$\mathbf{R}(t)[i, j] = \mathbf{O}^w(t)[i, j] \cap \mathbf{K}(t)[j],$$

или

$$\mathbf{R}(t) = \mathbf{O}^w(t) \otimes \mathbf{K}(t) = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1j} \\ R_{21} & R_{22} & \dots & R_{2j} \\ \dots & \dots & \dots & \dots \\ R_{i1} & R_{i2} & \dots & R_{ij} \end{bmatrix},$$

где $\mathbf{O}^w(t)$ — операционная матрица; $\mathbf{R}(t)$ — матрица нечетких отношений; \otimes — операция (\cap) Min.

Далее определяется прогнозируемое значение $F(t)$ для года t , представленное в виде нечеткого множества

$$F(t) = [\text{Max}(R_{11}, R_{21}, \dots, R_{i1}) \text{ Max}(R_{12}, R_{22}, \dots, R_{i2}) \dots \text{ Max}(R_{1j}, R_{2j}, \dots, R_{ij})].$$

Для нашего примера $1 \leq i \leq 6$ и $1 \leq j \leq 7$.

$$\mathbf{R}(1990) = \begin{array}{|c|c|c|c|c|c|c|} \hline & \text{онурчн} & \text{нурчн} & \text{нирчн} & \text{сурчн} & \text{норурчн} & \text{вурчн} & \text{овурчн} \\ \hline 0.09 & 0.14 & 0.27 & 0.58 & 0.87 & 0.42 & 0.26 \\ 0.05 & 0.08 & 0.13 & 0.24 & 0.5 & 0.42 & 0.26 \\ 0.05 & 0.07 & 0.11 & 0.20 & 0.41 & 0.42 & 0.26 \\ 0.10 & 0.16 & 0.32 & 0.70 & 0.87 & 0.42 & 0.25 \\ 0.06 & 0.09 & 0.15 & 0.29 & 0.62 & 0.42 & 0.26 \\ 0.06 & 0.09 & 0.14 & 0.27 & 0.59 & 0.42 & 0.26 \\ \hline \end{array}$$

Окончательный результат для прогноза роста численности населения для 1990 г. будет выглядеть следующим образом:

$$\mathbf{F}(1990) = \begin{array}{|c|c|c|c|c|c|c|} \hline & \text{онурчн} & \text{нурчн} & \text{нирчн} & \text{сурчн} & \text{норурчн} & \text{вурчн} & \text{овурчн} \\ \hline 0.10 & 0.16 & 0.32 & 0.70 & 0.87 & 0.42 & 0.26 \\ \hline \end{array}$$

Для остальных годов прогнозные результаты вычисляются аналогичным образом.

6-й шаг. Для дефазификации результатов, полученных на 5-м шаге, предлагается следующая формула:

$$V(t) = \frac{\sum_{i=1}^7 \mu_t(u_i) u_{\text{ср}}^i}{\sum_{i=1}^7 \mu_t(u_i)},$$

где $\mu_t(u_i)$ — вычисленные значения функций при надежностей для прогнозируемого года; $u_{\text{ср}}^i$ — средние точки интервалов.

Например, после расчетов для $F(1990)$: $V(1990) = 93\ 300$, т. е. ожидаемый на 1990 г. прирост населения составляет 93 300 чел. А теперь для того, чтобы получить прогнозируемое значение численности населения на 1990 г., необходимо к численности населения за 1989 г. прибавить полученное расчетное значение прироста населения, т. е.

$$N(1990) = 7\ 021\ 200 + 93\ 300 = 7\ 114\ 500.$$

Прогнозируемые значения численности населения на основе базиса $w = 7$ за период 1988—2001 гг., вычисленные в соответствии с изложенной методикой, представлены в табл. 2.

Результаты экспериментальных расчетов. Для оценки эффективности применения предложенного подхода к решению задачи демографического прогноза проведены эксперименты по расчету общей численности населения на определенные отрезки времени. Эксперименты осуществлены в два этапа, которые условно названы ретроспективный и перспективный.

1. На первом этапе в качестве экспериментальной базы принят отрезок времени 1988—2001 гг., т. е. ретроспектива, предшествующая текущему 2002 г. Естественно, что в данном случае статистические данные об общей численности населения на указанный период известны.

Суть проводимого эксперимента заключалась в том, что:

а) динамика численности населения за рассматриваемый период времени считалась неизвестной;

б) для каждого отдельно взятого года выбранного временного отрезка [1988—2001 гг.] с учетом соответствующей предыстории изменения темпов роста населения был рассчитан прогноз численности населения, базирующийся на предложенной методике;

в) для проверки меры адекватности модели фактическая динамика численности населения и соответствующие вариации за 1988—2001 г. были сопоставлены с реализацией математической модели и была рассчитана погрешность аппроксимирующего модель метода прогнозирования.

В работе погрешность метода рассчитана по следующей формуле:

$$\delta(t) = \frac{N_{\text{факт}}^t - N_{\text{прогн}}^t}{N_{\text{факт}}^t} 100 \%,$$

где $N_{\text{факт}}^t$ — фактическая численность населения за год t ; $N_{\text{прогн}}^t$ — прогнозная численность населения за год t ; $1988 \leq t \leq 2001$.

Таблица 2

Результаты ретроспективного анализа прогноза численности населения

Годы	Фактическая		Прогнозируемая		Погрешность, %	Средняя погрешность, %
	численность населения, тыс. чел.	вариация, тыс. чел.	численность населения, тыс. чел.	вариация, тыс. чел.		
1998	6 928,0	105,3	6 926,7	104,0	0,02	
1989	7 021,2	93,2	7 028,0	100,0	0,10	
1990	7 131,9	110,7	7 114,5	93,3	0,25	
1991	7 218,5	86,6	7 234,9	103,0	0,23	
1992	7 324,1	105,6	7 308,5	90,0	0,22	
1993	7 440,0	115,9	7 425,1	101,0	0,20	
1994	7 549,6	109,6	7 544,3	104,3	0,07	
1995	7 643,5	93,9	7 647,9	98,3	0,06	
1996	7 726,2	82,7	7 736,5	93,0	0,13	
1997	7 799,8	73,6	7 812,0	85,8	0,03	
1998	7 879,7	79,9	7 884,0	84,2	0,05	
1999	7 953,4	73,7	7 962,6	82,3	0,11	
2000	8 016,2	62,8	8 034,4	81,0	0,23	
2001	8 081,0	64,8	8 093,4	77,2	0,15	

В табл. 2 приведены фактические значения и вариации общей численности населения за 1988—2001 гг., результаты прогнозных расчетов и соответствующие значения погрешностей за этот же период, а также значение средней погрешности. Прогнозные расчеты проведены с учетом базиса $w = 7$, т. е. расчет прогноза для каждого года ($t = s_i$) из выбранного временного отрезка [1988—2001] проводился с учетом вариации за восемь предшествующих лет ($s_i - s_j = 8$; $1988 \leq s_i \leq 2001$; $1980 \leq s_j \leq 1993$).

Сравнительный анализ фактических и расчетных данных и полученные значения погрешности метода аппроксимации свидетельствуют о достаточно хорошем качестве модели и позволяют сделать предварительный вывод о целесообразности использования последней в демографическом прогнозе.

Графическое изображение фактической и прогнозируемой динамики численности населения, представленное на рис. 2 (см. третью сторону обложки), также наглядно демонстрирует достаточную близость указанных данных, что, в свою очередь, предопределяет необходимость продолжения исследований в данном направлении.

Таблица 3

Прогнозируемая динамика численности населения Азербайджана, полученная с применением различных моделей

Годы	PROST	SPSS Статистические модели	Модель нечеткого временного ряда (МНЧВР)
2000			
2001		8081	
2002	8200,189	8175,016	8155,3
2003	8277,965	8247,278	8233,8
2004	8357,393	8277,211	8315,2
2005	8438,496	8356,911	8398
2006	8521,262	8450,611	8481,9
2007	8606,523	8540,611	8566,7
2008	8694,348	8646,278	8652,3
2009	8784,848	8758,811	8738,5
2010	8878,074	8876,011	8825,3
2011	8974,047	8989,394	8912,5
2012	9070,694	9099,561	9000
2013	9168,131	9209,906	9087,8
2014	9266,445	9322,628	9175,9
2015	9365,678	9437,628	9264,2
2016		9544,478	9352,6
2017		9645,178	9441,1
2018		9742,611	9529,7
2019		9833,778	9618,4
2020		9925,978	9707,1
2021			9795,8
2022			9884,5
2023			9973,2
2024			10061,9
2025			10150,6

2. В свете изложенного на втором этапе проведены эксперименты по перспективному расчету численности населения до 2025 г. включительно. Прогнозные данные сопоставлены с результатами, полученными с применением других прогнозных моделей. Так, табл. 3 демонстрирует прогнозные значения численности населения Азербайджана, полученные с применением предложенной модели, модели расчета социального бюджета Всемирного банка PROST и пакета программ SPSS, базирующегося на статистических моделях прогнозирования.

Следует отметить, что по данным Государственного комитета по статистике Азербайджана к началу 2002 г. фактическая численность населения республики в 2001 г. составила 8141,41 млн чел., а к началу 2003 г. фактическая численность населения республики в 2002 г. составила 8228,3 млн человек [8, 9]. Сравнительный анализ данных с фактом (см. табл. 3) показывает, что прогнозные значения общей численности населения за 2001 и 2002 годы, полученные применением нечетких временных рядов, наиболее близки к реальным значениям.

На рис. 3 (см. третью сторону обложки) представлена графическая иллюстрация прогнозных данных, полученных с применением указанных моделей.

Заключение

Методика, предложенная в статье, позволяет получить демографический прогноз на основе нечетких временных рядов. Особенностью методики является возможность выдачи прогнозов при неполной, нечеткой исходной информации. Описанный подход допускает при расчете прогноза на какую-либо дату включение в экспериментальную базу истекшей динамики численности населения на любую отдаленную перспективу. Это, в свою очередь, способствует учету при расчете прогноза предыстории тенденций изменения темпов роста населения и, соответственно, получению более точных прогнозов.

На базе предложенной методики разработаны соответствующие алгоритмы и программные средства, позволившие провести обработку представительного статистического материала, сравнить результаты, полученные на основе нечеткой прогнозной модели, с результатами других методов прогнозных расчетов и сделать окончательный вывод о перспективности метода при решении задач демографического прогнозирования.

Как известно, исследование динамики общей численности населения дает первую, обобщенную характеристику населения и не раскрывает полностью процесс его воспроизводства. Поэтому в настоящее время ведутся работы в направлении расширения прогнозируемых характеристик населения за счет определения перспектив половозрастного состава, рождаемости, смертности, миграции и других демографических показателей.

Список литературы

1. Заде Л. А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М.: Мир, 1976.
2. Song Q., Chissom B. S. Fuzzy time series and its models // Fuzzy Sets and Systems. 1993. 54.

3. Song Q., Chissom B. S. Forecasting enrollments with fuzzy time series — part II // Fuzzy Sets and Systems. 1994. 62.
4. Chen S. M. Forecasting enrollments based on fuzzy time series // Fuzzy Sets and Systems. 1996. 81.
5. Hwang J. R., Chen S. M., Lee C. H. A new method for handling forecasting problems based on fuzzy time series // Proc. 7th Internat. Conf. On Information Management. Chungli, Taiwan, ROC, 1996.
6. Мамедова М. Г., Джабраилова З. Г. Применение нечетких временных рядов для прогнозирования численности населения // Сб. тр. НИУЦ по труду и социальным проблемам. Вып. 1. Баку. 2002. С. 41—63.
7. Abbasov A. M., Mamedova M. H., Gasimov V. A. Fuzzy Relational Model for Knowledge Processing and Decision Making // Advances in Mathematics. New York. 2002. Vol. 1. P. 191—223.
8. Statistical Yearbook of Azerbaijan — 2001. Statistical Committee of Azerbaijan Republic. Baku, 2001.
9. Statistical Yearbook of Azerbaijan — 2002. Statistical Committee of Azerbaijan Republic. Baku, 2002.

УДК 004.3.06

С. Д. Коровкин, канд. техн. наук,
И. Д. Ратманова, канд. техн. наук,
Л. В. Щавелев, канд. техн. наук,
И. А. Левенец, канд. техн. наук, М. А. Козырев,
Ивановский государственный энергетический
университет

Система мониторинга и анализа социально-экономического положения региона

Описан опыт реализации информационно-аналитической системы по социально-экономическому положению региона. В основу положена концепция хранилищ данных и систем поддержки принятия решений на их основе. Система построена в среде комплекса инструментальных средств ИнфоВизор.

Введение

Концепция хранилищ данных и систем поддержки принятия решений открывает перспективы новых подходов к организации управления. Отработанные принципы интеграции информации и построенные на их основе аналитические модели позволяют по-новому взглянуть на объект управления, выявить закономерности в ретроспективе показателей, характеризующих объекты и события, локализовать аномальные явления и

состояния объектов, а также построить прогнозистические модели для проигрывания сценариев "а что будет, если".

Предлагается к рассмотрению решение организации мониторинга и анализа социально-экономического положения региона на основе концепции хранилищ данных. Консолидация информации по уровням: Российской Федерации, федерального округа, субъекта федерации и муниципального образования позволяет обеспечить информационную поддержку управления регионом. Рассмотрим последовательно основные составляющие полученного решения:

- модель хранилища данных;
- систему сбора и консолидации информации;
- организацию оперативной аналитической обработки накопленной в хранилище информации;
- обеспечение комплексного мониторинга и анализа сложившейся социально-экономической ситуации в регионе на основе пакета аналитических отчетов;
- технологическую основу разработки.

Модель хранилища данных по социально-экономическому положению региона

Хранилище, как правило, представляет собой совокупность информационных объектов, где каждый объект на метауровне СУБД представлен рядом взаимосвязанных сущностей, а на уровне межданных хранилища — соответствующими моделями