УДК 621.372

А.А. Омельченко, А.Н. Полетайкин
Донецкая академия автомобильного транспорта

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ СОСТАВЛЕНИЯ ОПТИМАЛЬНОГО ПЛАНА РЕАЛИЗАЦИИ КОМПЬЮТЕРНОЙ ТЕХНИКИ

Розглядається задача складання оптимального плану реалізації комп'ютерної техніки та її пристроїв на регіональному ринку. Використовується імітаційне моделювання для генетичного кодування й декодування істот, що враховує особливості психології масового споживача.

Рассматривается задача составления оптимального плана реализации компьютерной техники и ее устройств на региональном рынке. Используется имитационное моделирование для генетического кодирования и кодирования существ, которое учитывает особенности психологии массового потребителя.

The problem of optimal planning of realization of computer equipment and its devices is considered at the regional market. A simulation technique is used for the genetic coding and coding of creatures, which takes into account the features of psychology of mass consumer.

Ключевые слова: генетический алгоритм, компьютерная техника, оптимальный план.

Постановка задач исследования. В условиях современной экономики объем реализации любого товара уже не определяется привычным соотношением количественных оценок спроса и предложения. В настоящее время реализация компьютерной техники (КТ) представляет собой сложный социально-экономический процесс, в основе которого заложено множество факторов неэкономического характера. Системный анализ влияния этих факторов реализации с использованием методов инженерии знаний и экспертных оценок [6] позволил выявить существенные и получить дополнительную информацию для их системного представления, а также для управления и оптимизации указанного процесса в структуре экспертной системы (ЭС). Учитывая слабую структурированность предметной области, совершенно очевидно, что представление факторов должно носить вероятностный характер. При этом, как известно из теории вероятностей, оценка случайной величины имеет более выразительный смысл в системе хотя бы двух координат *xy*, где отражен характер неопределенности функциональной зависимости $F=y(x)$. В [6] разработана модель экспертной оценки факторов реализации КТ в формате нечетких множеств, в связи с чем введено понятие нечеткого фактора (НФ). Здесь под зависимостью F понимается проекция НФ на характеристические признаки (ХП) дифференцирования товара. Множество рассмотренных зависимостей используются как входные данные для прогнозирования спроса КТ с использованием методов структурно-параметрической идентификации и композиционного правила вывода.

Следующим этапом является решение задачи определения варианта картины реализации, при фактическом выполнении которой потребность населения в КТ удовлетворялась бы в наибольшей степени. В [3] обосновано применение для этой цели технологии генетических алгоритмов, которая в последнее время все чаще используется для решения задач функциональной оптимизации. Преимущества ГА состоят в нахождении приблизительно оптимальных решений многопараметрических задач за относительно короткое время, и в возможности поиска всего множества экстремумов для многоэкстремальных функций [5].

Актуальной областью в технологии генетических алгоритмов является задача определения фитнес-функции, являющейся неотъемлемой компонентой отбора – важнейшего алгоритма ГА. Фактически, от успешной реализации фитнес-функции зависит эффективность работы ГА. По аналогии с современной эволюционной теорией в ГА могут быть реализованы принципы естественного и искусственного отбора. Критерий естественного отбора задается непрерывной, определенной на некотором промежутке, целевой функцией (ЦФ). Искусственный отбор задается дискретной функцией годности (ФГ) со значениями «годен» или «негоден», и отсеивает заведомо непригодные решения. При этом искусственный отбор может не иметь места, что замедлит сходимость поиска решения, но обеспечит планомерность развития популяции.

Проблемой применения ГА является слабая способность получения принципиально новых решений, лежащих вне области поиска алгоритма. В качестве средства для получения ОДР, в которой также формируется начальная популяция, используется имитационная модель удовлетворения потребности в компьютерной технике [3]. Модель учитывает особенности психики массового потребителя (человеческий фактор) и которая позволяет получать множественную картину реализации КТ. Инструментом фиксирования моделируемых параметров в области допустимых решений (ОДР), определяемой, главным образом, интеллектуальным механизмом параметризации системы, основанным на знаниях экспертов, является аппарат вероятностной математики: используются вероятностные распределения (моделирование непрерывно действующих факторов),

© *А.А. Омельченко, А.Н. Полетайкин, 2010*

марковские цепи (дискретная параметризация объектов и отношений) и матрицы вероятностей переходов (реализация отношений между объектами).

Разработка генетического алгоритма. Рассмотрим классический ГА, работу которого структурно демонстрирует рисунок 1. На первой итерации начальная популяция автоматически переходит в новое поколение, к которому применяется процедура естественного отбора — первая стадия оптимизации, — в результате которой из нового поколения отсеиваются менее приспособленные особи. Лучшие особи, прошедшие естественный отбор, получают право скрещивания и рекомбинации. Прежде всего, происходят турнирные розыгрыши пар для рекомбинации — вторая стадия оптимизации. Далее полученные пары подвергаются кроссинговеру, в результате которого образуются два потомка, с которыми далее происходит изменчивость. В результате получается множество потомственных особей, равное по величине множеству предков. Оба эти множества в совокупности подвергаются третьей стадии оптимизации — процедуре формирования нового поколения, которая на основании фитнес-функции производит отбор особей в новое поколение. Так происходит одна реализация ГА, в котором, как видно из его описания, каждый процесс выполняется также итеративно (на рисунке 1 итеративные процессы обведены двойной рамкой), где количество итераций определяется размерностью обрабатываемого множества особей. Алгоритм включает в себя три этапа оптимизации, каждый из которых опирается на вычисление значения ЦФ и ФГ. Функционирование ГА прекращается по истечении заданного количества реализаций, либо когда среди нового поколения будет получена особь, максимально точно соответствующая реальной картине, чему соответствует минимальное значение ЦФ при истинном значении функции годнос-ти.

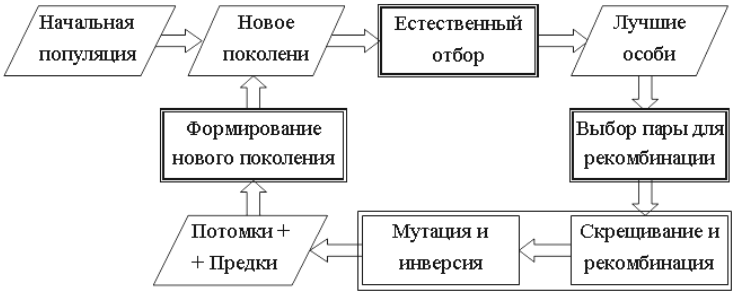


Рис. 1. Итерационный процесс функционирования ГА

Структура хромосомы. В данной задаче оптимального планирования хромосома кодирует развернутый план реализации, структура которого представлена в таблице 1. Кодирование производится в терминах объектно-структурного анализа [1], согласно которому предметная область решаемой задачи оптимизации дезагрегируется на несколько областей знаний – страт (вертикальная сегментация), которые, в свою очередь, иерархически детализируются по горизонтали от уровня проблемы до уровня подзадачи. Данный подход позволяет создавать структурные модели баз знаний любой сложности и степени детализации.

Развернутый план реализации представляет собой в общем виде множество кортежей $R_1\{S_1,...,S_N\}$ … $R_M\{S_1,...,S_N\}$, атрибуты S_i которого отражают параметры процесса реализации. Минимальный набор таких атрибутов: S_1: Код_товара и S_2: Количество_реализации. С целью определения качественных характеристик в множество $\{S\}$ включены дополнительные параметры, отражающие пространственный S_3, временной S_4 и функциональный S_5 аспекты процесса реализации (табл. 1). Параметр S_3: Канал_реализации представляет собой указатель локализации операции в моделируемом процессе. Параметр S_4: Дата_реализации определяет дату выполнения операции в данном временном периоде. Функциональная особенность реализации товара отражает способ, каким был отпущен товар: отдельным устройством или в составе компьютера. Этот аспект процесса задается параметром S_5: Способ_реализации.

Таблица 1

Структура хромосомной матрицы

Ген	S_1	S_2	S_3	S_4	S_5
Слой знаний	ЧТО	СКОЛЬКО	ГДЕ	КОГДА	КАК
Кортеж R_i	**{T}**	**Q(T)**	**{C}**	**{D}**	**{H}**

Алфавит всех параметров, за исключением S_2, есть конечное множество значений, определяемое в случае параметра S_1 из номенклатуры реализуемого товара {T}, для S_3 – из структуры подсистемы реализации товара {C}, а для S_4 – из планируемого периода, определяющего множество {D} возможных дат на шкале времени. Параметр S_5 характеризует способ ре-ализации товара из множества {H} мощностью 2. Параметр S_2, не выражаемый конечным множеством, описывается непрерывным вероятностным распределением Q(T) по экспертно определенному закону.

Таким образом, хромосома включает набор кортежей-подхромосом, состоящих из пяти генов, и имеет матричную структуру. Особи должны иметь одинаковую структуру, в которую включены кортежи для каждого

товара, упорядоченные по артикулу, то есть признак S_1 в хромосоме является обязательным. Остальные параметры заданы только для наборов ре-ализуемых товаров в соответствии с кодируемым планом. Популяция таких особей подвергается воздействию разрабатываемого ГА, который ее анализирует и преобразует на основе механизма натуральной эволюции. Ввиду значительной временной сложности оптимизационного алгоритма, $R_{ПОП}$ следует принять равным 100, а количество поколений $N_{ПОК}$ =20.

Естественный отбор. Проведенные исследования показали, что наиболее эффективным методом отбора особей в новую популяцию для данной задачи является метод турнирного отбора с высокой размерностью турнирной группы (k = 4), который предполагает в итоге формирование набора $M(n)$ лучших особей, где n можно положить чётным $0{,}7N$. Такая часть хорошо сочетается с методикой формирования нового поколения, которая, как будет показано далее, приводит к появлению двойной популяции.

Выбор пары для рекомбинации. Существует множество способов образования родительских пар [5], один из которых представляет для задачи оптимального планирования наибольший интерес – это аутбридинг. Сравнение генотипов осуществляется посредством функции аутбридинга $A(X,Y)$, построенной по принципу ЦФ (1). Выбор наиболее различающейся пары можно произвести методом турнирного отбора с k = 4.

$$A(X,Y)=\sum_{N} K_E\left(\frac{par_X - par_Y}{par_T}\right)^2. \qquad (1)$$

Здесь $par_{X.}$ и par_Y – параметры сравниваемых особей; par_T –заданные отклонения параметров; K_E – коэффициент, определяющийся точностью кодирования параметра. Уставки отклонений параметров par_T при вычислении функции аутбридинга бально оценены экспертами и представлены в таблице 2. Параметры 2 и 3 нормированы в соответствии с линейным преобразованием в единичный отрезок [0, 1] исходя из статистики, полученной из хромосомы, принимая минимальное значение величин равное нулю (как для идеальной системы реализации товара).

Таблица 2

Уставки отклонений параметров для функции аутбридинга

Код	Параметр	Задание
1	Вероятность отказа	0,15
2	Нормированная длительность обслуживания заявок	0,5
3	Нормированная длительность пребывания товара на складе	0,5
4	Точность предложения товара	0,3

Формирование нового поколения. В [2] для задач, использующих аутбридинг, предлагается идея элитного отбора, основанная на постро-ении новой популяции только из лучших особей репродукционной группы G с наилучшими показателями приспособленности. Надо отметить, что размерность группы G составляет $2n$ = $1{,}4N$, из чего следует, что $0{,}4N$ особей следует отсеять как заведомо лишние, так как новое поколение должно иметь размерность популяции N, принятую на подготовительном этапе ГА. Таким образом, массив G в первую очередь проверяется на функции годности. Особи, получившие отказ, исключаются из множества G. Если после этого массив G имеет размерность больше, чем N, запускается процедура элитного отбора.

Кроссинговер выполняется с вероятностью $P_{KP.}$ = 0,9. В [2] показано, что использование случайного чередования одно- и двухточечного кроссинговеров с равномерным распределением по хромосомной матрице понижает вероятность преждевременной сходимости ГА. Предлагаются такие варианты перестановок, опирающиеся на структуру поля знаний: меж-дутоварные (ЧТО-страта), междуканальные (ГДЕ-страта), междудневные (КОГДА-страта), а также произвольный разрыв. Наследование родительских частей для одноточечных операторов выполняется посредством обмена родительскими частями, как показано на рисунке 2, а.

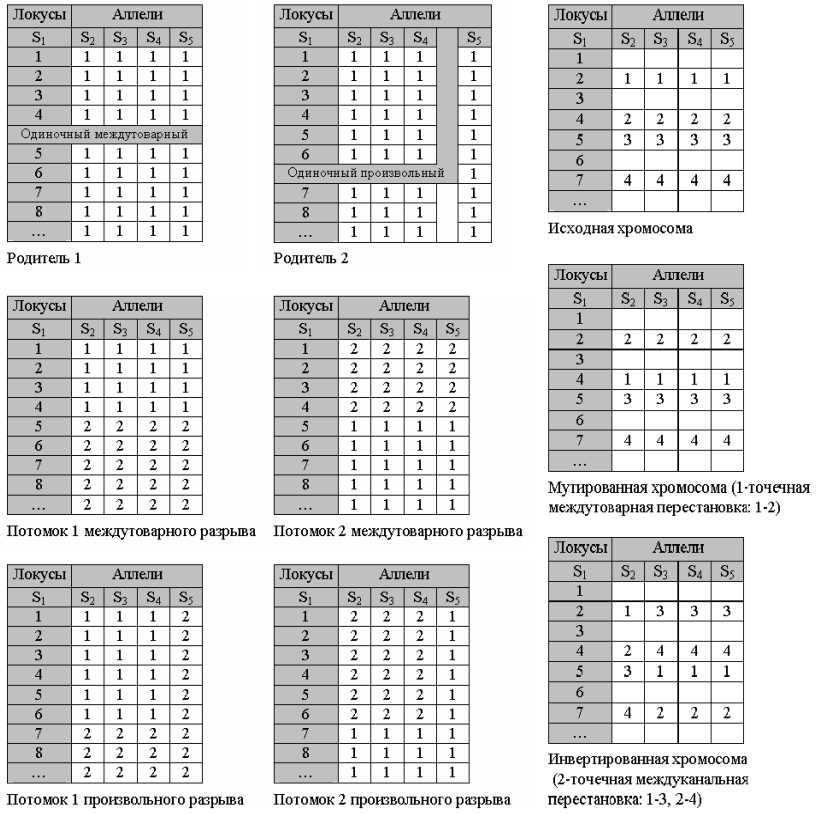
Специализированный оператор рекомбинации используется при скрещивании родителей и работает следующим образом. Рассматриваются гены S_1 каждого из родителей. Если у каждого из них этот ген является пустым, у потомков данный ген также не задействуется. Если же хотя бы один из особей-предков имеет обозначенный ген, то потомки, зависимо друг от друга, будут иметь задействованный ген с N-вероятностью (2), равной корню N-й степени вероятности его включения в хромосомах родителей в количестве N пар, скорректированной на интегральную оценку, полученную на основе приспособленности скрещиваемых особей.

$$P_N = \sqrt[N]{\frac{\sum_{i=1}^{N^S} C_i^S + \sum_{j=1}^{N^O}\left(1 - C_j^O\right)}{N_{CR}}}, \qquad (2)$$

где C_i^S – значение ЦФ *i*-й особи, включающей данный товар; C_j^O – значение ЦФ *j*-й особи, не включающей данный товар; N^S – количество особей, включающих данный товар в план; N^O – количество особей, не включающих данный товар в план; $N_{CR} = N^S + N^O = \text{even}(0.7N)$ – количество особей, подвергаемых рекомбинации.

Мутация. Полученные потомки подвергаются мутации с вероятнос-тью $P_{МУТ.}$=0,005…0,01. В классическом испонении данная операция также одно- или двухточечная по аналогии с операцией кроссинговера и такими же параметрами распределения точек в хромосоме (рис. 2, б). Кроме того, при рассмотрении кроссинговера были предложены два специализированных оператора мутации: 1) с дополнителями и 2) с заменителями.

Инверсия происходит с вероятностью $P_{ИНВ.}$=0,01…0,015. Механизм такой одно- или двухточечной изменчивости сводится к обмену частями хромосомы, включающей гены $S_2 – S_5$ при фиксированном положении гена S_1, в результате чего образуется качественно иная особь.



а) одноточечный кроссинговер **б) изменчивость**

Рис. 2. Схематическая демонстрация операторов рекомбинации ГА

Параметрическая фитнес-функция. Рассматриваемая задача оптимизации является многопараметрической, где параметры, определяющие план реализации, формируют состав ЦФ:

$$F^* = \sum_N K_E \left(\frac{par_{зад.} - par_{рез.}}{par_{зад.}} \right)^2 \rightarrow \min_{par_{рез.}}, \qquad (3)$$

где $par_{рез.}$ – параметры, выделенные из решения ГА; $par_{зад.}$ – заданная величина, определяющая оптимум. Как и для функции аутбридинга, параметры 2 и 3 нормированы в единичный отрезок.

Таблица 3

Уставки параметров оптимизации целевой функции

Код	Параметр	Задание
1	Вероятность отказа	0,001
2	Нормированная длительность обслуживания заявок	0,01
3	Нормированная длительность пребывания товара на складе	0,01
4	Точность предложения товара	0,01

Дискретная функция годности отбраковывает неподходящие особи. Выполняется проверка выборки по параметру S_2 в хромосоме на соответствие заданному закону распределения по критерию согласия Пирсона. Если гипотеза не подтверждается, план оказывается вне области адекватности. Также выборка анализируется по частотному распределению каналов реализации (S_3), по равномерности реализации (S_4) и соответствию

прошлым временным рядам, пропорционально относительно способа ре-ализации (S_5). Интегральная оценка адекватности модели образуется посредством алгоритма «И» четырех критериев $C_{S_i}^{\alpha_i}$ типа $\chi^2 \le \chi^2_{n,\alpha}$ в соответствии с экспертными оценками уровня значимости α_i:

$$AD = \sum C_{S_i}^{\alpha_i} = C_{S_2}^{\alpha_2} \cap C_{S_3}^{\alpha_3} \cap C_{S_4}^{\alpha_4} \cap C_{S_5}^{\alpha_5} \tag{4}$$

Рассмотренные в данном разделе процедуры ГА выполняются последовательно в рамках схемы, показанной на рисунке 1 с циклическим проигрыванием ряда поколений, количество которых определяется фиксировано, либо достижением определенного уровня оптимальности, определяемого исходя из значения параметрической целевой функции. Укрупненный алгоритм работы ГА схематически представлен на рисунке 3.

Исследования и результаты работы генетического алгоритма. Поскольку разработанный ГА встроен в экспертную систему, эффективность его работы целесообразно оценивать в структуре ЭС. С этой целью в [4] предложена концепция подсистемы самооценки системы по 4 независимым оценкам: обобщенной полезности, вероятности достижения цели по интегральному критерию оптимизации, интегральному показателю эффективности и коэффициенту упорядоченности. Результаты оценивания представлены на рис. 4, где под циклом управления подразумевается полный цикл бизнес-процессов планирования реализации КТ.

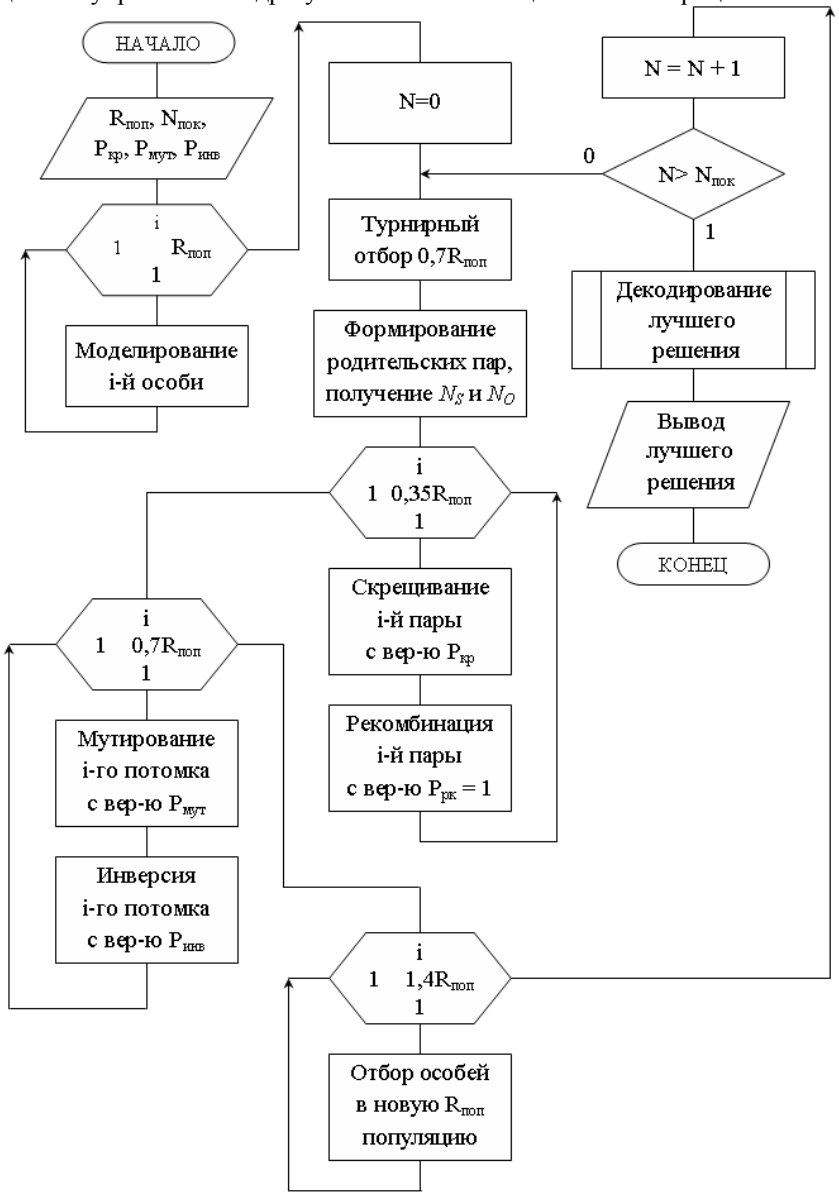


Рис. 3. Схема укрупненного алгоритма работы ГА

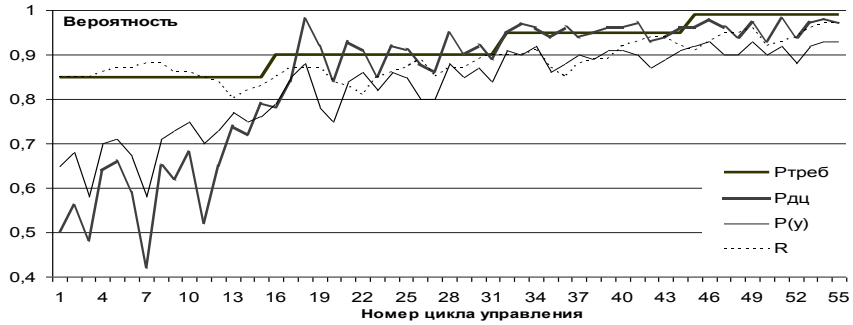


Рис. 4. График изменения эффективности ЭС

Выводы. 1. Предложен новый подход к оптимизации многопараметрических систем, который заключается в совместном использовании аппарата имитационного моделирования и генетических алгоритмов. 2. Разработан генетический алгоритм для получения оптимально составленного плана реализации, который встраивается в структуру экспертной системы поддержки принятия решений при реализации компьютерной техники.

Библиографические ссылки

1. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. – СПб., 2001. – 384 с.
2. **Гладков, Л.А.** Генетические алгоритмы / Л.А. Гладков, В.В. Курейчик, В.М. Курейчик; под ред. В.М. Курейчика. – [2-е изд., испр. и доп.] – М., 2006. – 320 с.
3. **Полетайкин, А.Н.** Имитационная модель процесса удовлетворения потребности в компьютерной технике. / А.Н.Полетайкин // Наукові праці Донецького національного технічного університету. Серія: Обчислювальна техніка та автоматизація, випуск 12 (118). – Донецьк, 2007. – с. 92 – 101.
4. **Полетайкин, А.Н.** Модель оценивания эффективности автоматизированных информационных систем. / А.Н.Полетайкин // Сборник научных трудов НГТУ. – 2009 №3 (53) С.58 – 65.
5. **Скобцов, Ю.А.** Основы эволюционных вычислений. / Ю.А.Скобцов – Учебное пособие. – Донецк, 2008. – 326 с.
6. **Спорыхин, В.Я.** Концепция экспертной системы оценки факторов реализации компьютерной техники. / В.Я.Спорыхин, А.Н.Полетайкин // Вестник Херсонского национального технического университета.− Херсон, 2007. − №4(27). − С.286−292.

Надійшла до редколегії 12.01.10