

УДК 004.021, 004.67, 338.984

С.А. Герасимов

МЕТОД ПРОГНОЗИРОВАНИЯ КОТИРОВОК ОФИСНЫХ АРЕНДНЫХ СТАВОК НА ОСНОВЕ АППАРАТА НЕЧЁТКИХ МНОЖЕСТВ И ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Рассмотрено применение аппарата нечётких множеств и генетических алгоритмов для прогнозирования котировок арендных ставок офисной недвижимости. Дано описание математической модели и ее практической реализации. Приведен пример использования предлагаемого метода для прогнозирования квартальных и годовых офисных арендных ставок.

Ключевые слова: прогнозирование временных рядов, арендные ставки, нечёткие системы, генетические алгоритмы.

Рынок недвижимости в нашей стране постоянно развивается: уровень инвестиций в строительство в России занимает второе место после энергоносителей. С точки зрения развития экономики, важным сектором рынка недвижимости является рынок офисных помещений, для которого актуальной задачей является прогнозирование и оценка изменения офисных арендных ставок.

Для прогнозирования котировок офисных арендных ставок может использоваться анализ временных рядов. Исходными данными для анализа служат: глобальные факторы (основные показатели развития макроэкономики); внешние факторы (показатели состояния рынка офисной недвижимости) и внутренние факторы (собственные параметры объекта недвижимости). Форма исходных данных для прогноза представлена в табл. 1, где значения факторов X_1, X_2, \dots, X_n , влияющих на прогнозируемый показатель Y , для нашей задачи определяются четыре раза в год (на начало каждого квартала).

Алгоритм построения математической модели (рис. 1) базируется на

использовании теории нечёткой логики [1].

В соответствии с методами нечёткой логики выделяются основные этапы моделирования:

- постановка задачи и построение дерева вывода, фазификация (выбор нечетких термов для лингвистической оценки факторов влияния). Рассматривается объект типа

$$y = f(x_1, x_2, \dots, x_n),$$

где n - число входов с показателями ($x_i, i=1:n$) и одним выходом (y).

С помощью экспертов формируется совокупность параметров, оказывающих влияние на прогнозируемую величину. Для каждого из параметров определяются лингвистические термы, дающие оценку данного параметра. Граф (дерево решений) отображает классификацию факторов (x_1, x_2, \dots, x_n), влияющих на прогнозируемый показатель (y). На этом же этапе осуществляется фазификация. Предлагаемый метод включает в себя:

- построение базы нечетких знаний, содержащей результаты планируемого эксперимента;

Таблица 1

	Время	Факторы				Показатель
		X ₁	X ₂	...	X _n	Y
История	t ₀	x ₁ (t ₀)	x ₂ (t ₀)	...	x _n (t ₀)	y(t ₀)
	t ₁	x ₁ (t ₁)	x ₂ (t ₁)	...	x _n (t ₁)	y(t ₁)
	t ₂	x ₁ (t ₂)	x ₂ (t ₂)	...	x _n (t ₂)	y(t ₂)

	t _{k-1}	x ₁ (t _{k-1})	x ₂ (t _{k-1})	...	x _n (t _{k-1})	y(t _{k-1})
Прогноз	t _k	x ₁ (t _k)	x ₂ (t _k)	...	x _n (t _k)	y(t _k)
	t _{k+1}	x ₁ (t _{k+1})	x ₂ (t _{k+1})	...	x _n (t _{k+1})	y(t _{k+1})

	t _{k+m}	x ₁ (t _{k+m})	x ₂ (t _{k+m})	...	x _n (t _{k+m})	y(t _{k+m})

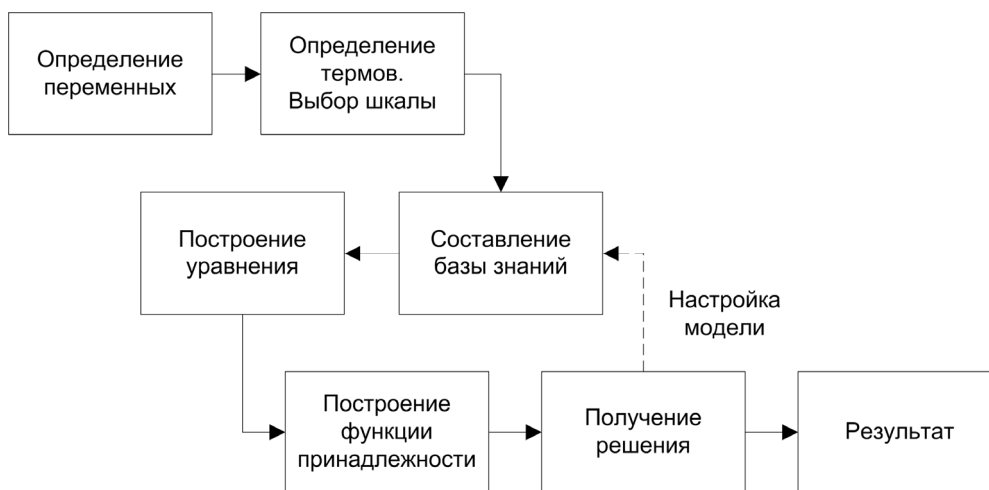


Рис. 1. Алгоритм построения математической модели

- построение функций принадлежности лингвистических переменных, которые задаются в параметрической форме;

- формирование нечеткого логического вывода;

- получение результатов моделирования, которое заключается в преобразовании нечеткого множества в четкое число (дефазификация);

- настройку нечеткой модели путем расширения объема базы знаний и настройки функций принадлежности.

Процедура дефазификации является процедурой получения решения. При этом величина Y рассчитывается:

$$Y = \frac{\sum_{i=1}^n [y_{\min} + (i-1)^{y_{\max} - y_{\min}}] \mu_i}{\sum_{i=1}^n \mu_i}$$

где n - число термов переменной y; y_{min}, y_{max} - шкала измерения; μ_i - значение функции принадлежности.

Предлагаемый метод использует генетическую нечеткую систему [2-3] для создания базы нечетких правил для прогнозирования временных рядов. База состоит из набора правил треугольного вида, связывающих значения факторов и прогнозируемого показателя. Хромосома состоит из двух частей.

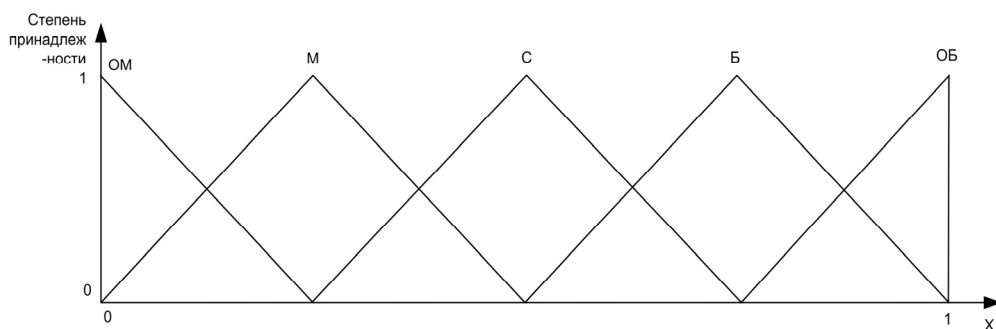


Рис. 2. Построение функций принадлежности, характеризующих термы лингвистической переменной "Размер"

Первая часть включает в себя номера термов лингвистических переменных, на основе которых строится набор правил. Пример такого набора термов (с именами: "очень маленький" (ОМ), "маленький" (М), "средний" (С), "большой" (Б), "очень большой" (ОБ)) для лингвистической переменной «Размер» приведен на рис. 2.

Вторая часть хромосомы уточняет вид правила и имеет вид $R=R_1R_2$, где R_1 – выбор конкретной лингвистической переменной для каждой входной, R_2 – точная настройка нечетких функций правила (координаты абсцисс левого края треугольника, его центра и его правого края). С помощью части хромосомы R_1 выполняется грубая настройка правила, с помощью части R_2 – точная. Если задействована только часть хромосомы R_1 , получаются лингвистические правила, которые можно объяснить словесно, например, вида «ЕСЛИ Рост ВВП Высокий ИЛИ Цены на нефть Высокие ТО Ставки аренды Высокие». Однако точность, полученная только лингвистическими средствами, зачастую оказывается недостаточной. Тогда необходимо задействовать и часть R_2 , которая более точно настраивает форму функций принадлежности под конкретные условия, увеличивая общую точность результата.

Для получения прогноза из нечетких правил используется процедура нечеткого вывода [4].

Процесс включает в себя следующие этапы, которые последовательно улучшают результат прогнозирования:

на первом этапе создается набор нечетких правил, на основании которого и получается прогноз. Правила состоят из двух частей: части для лингвистического (дескриптивного) описания (в правиле представлен индекс элемента из термножеств лингвистических переменных) и части для аппроксимативного описания (содержащей в себе формы функций принадлежности подусловий правила), которая уточняет базовую форму переменных из первой части правила. Используются лингвистические переменные треугольного вида, кодируемые тремя числами - координатами левого края, центра и правого края. Создание и подстройка правил осуществляется генетическим алгоритмом, который применяет как целочисленные генетические операторы для грубой настройки правил, используя первую часть правил, так и вещественные операторы для более точной их настройки с использованием второй части правил [3];

Таблица 2.

Квартал-год	Показатель, \$/кв. м/ г.	Прогноз \$/кв. м/ г.	Абсолютная ошибка	Относительная ошибка, %
1-2007	853	-	-	-
2-2007	890	-	-	-
3-2007	950	948	-2	-0,211
4-2007	1031	1009	-22	-2,134
1-2008	1072	1076	4	0,373
2-2008	1113	1123	10	0,898
3-2008	1100	1130	30	2,727
4-2008	952	940	-12	-1,261
1-2009	850	835	-15	-1,765
2-2009	780	771	-9	-1,154
3-2009	777	760	-17	-2,188
4-2009	780	776	-4	-0,513
1-2010	796	794	-2	-0,251
2-2010	799	809	10	1,252
3-2010	789	801	12	1,521
4-2010	801	799	-2	-0,250
1-2011	815	807	-8	-0,982
2-2011	879	860	-19	-2,162
3-2011	-	953	-	-

второй этап включает в себя процедуры, которые работают с уже созданным набором правил, последовательно улучшая его, настраивая для увеличения точности. Он включает в себя три процесса:

- процесс упрощения базы знаний (мультиимпликации). Производится путем удаления ненужных или малоинформативных правил из их множества, созданного на первом этапе. Это делается с помощью второго генетического алгоритма, который работает с двоичными хромосомами, каждый бит которых включает/не включает соответствующее по порядку правило, полученное в результате работы первого этапа – генерации, в итоговое множество;

- процесс взвешивания. Он улучшает точность прогноза путем присваивания каждому правилу своего «веса», с которым оно делает вклад в общий прогноз набора правил. Присвоение веса производится генетическим алгоритмом с вещественным кодированием хромосомы. При этом

гены хромосомы кодируют «веса» соответствующих правил;

- в процессе окончательной настройки правил базы знаний (тюнинга) форма правила корректируется для получения оптимального по точности результата. Это производится генетическим алгоритмом, который работает с мультихромосомами, первая из которых является «склеенными» вместе правилами, полученными в результате работы второй части, остальные созданы по подобию первой, но с небольшими отклонениями в пределах разрешенных интервалов. Для работы здесь используются те же операторы, что и при работе со второй частью хромосомы R_2 на этапе генерации популяции.

Результатом работы является набор правил, по которым можно получить прогноз следующего элемента временного ряда на основании подаваемых на вход влияющих факторов.

Из базы данных были получены временные ряды с временным шагом данных (размером лага) в 1 квартал.

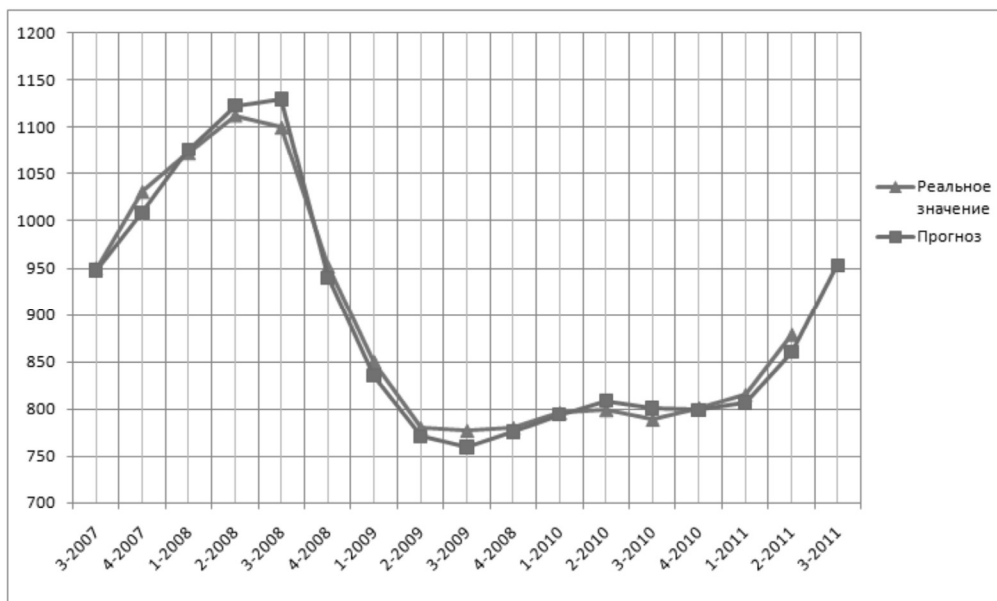


Рис. 3. Графические зависимости для реальных и долгосрочных прогнозируемых значений арендных ставок офисной недвижимости класса А в пределах Бульварного Кольца в Москве на основе генетического алгоритма

В результате анализа автокорреляционной функции рядов «количество подключений» была определена значащая глубина автокорреляционной функции – 3 лага (предыдущих элементов ряда). Поэтому максимально возможной глубиной прогноза является 4 квартала (1 год). Таким образом, производилось создание двух баз правил для нахождения двух видов прогноза: краткосрочного (на 1 квартал) и долгосрочного (на 1 год).

Отбор значащих факторов производился с помощью компактного генетического алгоритма. Этот алгоритм порождает двоичные векторы, с помощью которых кодируется набор факторов (каждый аллель отвечает за присутствие определенного атрибута в этом наборе). Фитнес-функцией при этом является оценка качества такого набора, определяемая погрешностью прогнозирования. Эта погрешность

при отборе значащих факторов находится с помощью нейронных сетей, для чего может быть использован пакет matlab [4].

Первоначально экспертами было отобрано 15 факторов, важных для решения данной задачи. Для долгосрочного прогнозирования компактными генетическими алгоритмами были отобраны факторы: Темп роста ВВП, Цена на нефть, Значение индекса ММВБ, Курс доллара США. Результаты краткосрочного прогнозирования арендных ставок офисной недвижимости класса А в пределах Бульварного Кольца в Москве приведены в табл. 2.

Графические зависимости для реальных и долгосрочных прогнозируемых значений арендных ставок офисной недвижимости класса А в пределах Бульварного Кольца в Москве на основе генетического алгоритма приведены на рис. 3.

Предложенный метод был положен в основу реализации автоматизированной системы оценки и прогнозирования изменения офисных арендных ставок объектов недвижимости, которая успешно использовалась в практической работе при управлении комплексом зданий в

центральных кварталах г. Москвы [5]. Использование предложенного метода позволило путем учета тенденций и прогнозирования состояния рынка недвижимости при заключении договоров аренды офисных помещений повысить рентабельность бизнеса.

СПИСОК ЛИТЕРАТУРЫ

1. *Заде Л.А.* Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976. – 167 с.
2. *Ярушкина Н.Г.* Основы теории нечетких и гибридных систем. М. Финансы и статистика, 2009. – 320 с.: ил.
3. *Cordon O.* Hybridizing Genetic Algorithms with Sharing Scheme and Evolution Strategies for Designing Approximate Fuzzy Rule-Based Systems //Fuzzy Sets and Systems. – 2001. – №118:2. – pp. 235-255.
4. *Леоненков А.* Нечеткое моделирование в среде Matlab и fuzzyTECH. – СПб., 2005. – 736 с.: ил.
5. *Герасимов С.А., Огиренко А.Г., Смирнов М.И.* Автоматизированная система прогнозирования изменения офисных арендных ставок. **ИИАС**

КОРОТКО ОБ АВТОРЕ

Герасимов Сергей Александрович – аспирант, prrto@gmail.com, Московский государственный горный университет, Moscow State Mining University, Russia, ud@msmu.ru



ПАМЯТНИКИ ГОРНЯКАМ

МЕМОРИАЛ ПОДВИГУ ШАХТЕРОВ МАКЕЕВКИ

Памяти тех, кто не вернулся из шахты и подвига тех, кто сейчас ежедневно спускается в забой.
Год установки — 2003.
Автор памятника — В.Ф. Пискун.

