

УДК 004.89

Н.М. Кораблев, Г.С. Иващенко

Харьковский национальный университет радиоэлектроники, Харьков

## ПРИМЕНЕНИЕ ИСКУССТВЕННЫХ ИММУННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ

В статье предложен комбинированный метод прогнозирования временных рядов при помощи искусственных иммунных систем. Рассмотрена модель на основе искусственных иммунных сетей и метода вывода по прецедентам (CBR). Оценка эффективности модели выполнена путем сравнительного анализа, представлены результаты экспериментальных исследований, демонстрирующие особенности предлагаемого подхода.

**Ключевые слова:** прогнозирование, временной ряд, вывод по прецедентам, искусственные иммунные системы, антитело, антиген, аффинность, клонирование, мутация.

### Введение

Прогнозирование временных рядов является важной научно-технической проблемой, которая позволяет решить актуальную задачу определения будущего состояния различных систем (экономических, социальных, экологических и др.) на основе анализа уже имеющихся ретроспективных данных. Точность прогноза зависит от многих параметров, таких как объем статистической информации (величина предистории), ее качество и достоверность, учет возможности изменения среды, в которой протекает процесс, число одновременно учтенных членов ряда, доступность необходимых вычислительных ресурсов. В настоящее время разработан широкий спектр методов прогнозирования, и выбор среди них наиболее подходящего для заданных условий является одним из определяющих факторов получения достоверного прогноза.

Несмотря на широкое распространение в экономической практике статистических методов прогнозирования [1, 2], в настоящее время активно развиваются подходы на основе методов искусственного интеллекта, таких как искусственные нейронные сети и искусственные иммунные системы (ИИС) [3 – 6], которые могут интегрироваться с другими подходами. Перспективным является применение метода вывода на основе прецедентов (case based reasoning, CBR), в

котором при рассмотрении новой задачи прогнозирования отыскивается подобный прецедент в предистории в качестве аналога [7, 8]. Проблема выбора подходящего прецедента среди имеющихся в базе прецедентов и его адаптация к текущим условиям является одной из самых важных в CBR-системах [9].

Цель проводимого исследования – разработка модели прогнозирования временных рядов на основе искусственной иммунной сети, использующей метод вывода по прецедентам, программная реализация разработанного подхода и проведение экспериментальных исследований.

### Постановка задачи

В данной работе рассматриваются дискретные временные ряды, значения которых получены в моменты времени  $t_1, t_2, t_3, \dots, t_N$ . Временной ряд  $Z = z(t_1), z(t_2), z(t_3), \dots, z(t_N)$  длины  $N$  обозначим  $Z_1^N = z_1, z_2, z_3, \dots, z_N$ . Набор последовательных значений  $Z_t^M = z_t, z_{t+1}, z_{t+2}, \dots, z_{t+M-1}$ , лежащих внутри временного ряда  $Z_1^N$ , назовем выборкой из этого ряда длины  $M$ , с моментом начала отсчета  $t$ ,  $M \in [1, N-1]$ ,  $t \in [1, N-M]$ . Задача прогнозирования временного ряда заключается в том, чтобы по его

известному участку оценить будущие значения. Целесообразность применения методов вывода по прецедентам определяется на основании следующих фундаментальных предположений о природе временного ряда [2]:

будущая динамика временного ряда зависит от его предыстории;

зависимость временного ряда может со временем меняться, но на некоторых участках она сохраняет определенное постоянство;

существуют участки временного ряда, на которых действуют одинаковые или близкие зависимости (история повторяется);

существуют участки временного ряда, для которых существует принципиальная возможность построения предикторов.

Согласно гипотезе, сформулированной в [9], если мера подобия между выборками  $Z_t^M$  и  $Z_{t-k}^M$  имеет значение, близкое к единице, то мера подобия между выборками длины  $P$ , следующими за ними,  $Z_{t+M}^P$  и  $Z_{t-k+M}^P$ , также близка к единице. Тогда путем определения выборки, максимально соответствующей последним известным значениям временного ряда, возможна оценка будущих значений ряда. В предложенной работе для решения поставленной задачи предлагается использование модели искусственной иммунной сети [3].

### Модель иммунной сети, использующей метод вывода по прецедентам

Метод вывода, основанный на прецедентах, представляет собой метод анализа данных, делающий заключения относительно данной ситуации по результатам поиска аналогий, хранящихся в базе прецедентов. Данный процесс включает в себя следующие этапы [7]:

- выбор из хранилища прецедентов наиболее уместного прецедента (или множества прецедентов), обычно на основе заданного отношения подобия;
- использование выбранных прецедентов для решения поставленной задачи;
- адаптация выбранного решения в связи с текущими условиями, если это необходимо;
- проверка корректности полученного решения;
- сохранение в хранилище для дальнейшего использования принятого решения и текущей ситуации в качестве нового прецедента, или соответствующее изменение выбранного прецедента.

Для построения модели прогнозирования на основе ИИС необходимо сопоставить биологические объекты и процессы с их аналогами из предметной области. Антитела представляют собой известные значения временного ряда. Антигены – значения временного ряда, непосредственно предшествующие прогнозируемым. Аффинность антител (основной

критерий отбора в алгоритме иммунной сети) – скалярная величина, оценивающая меру близости между антителом и антигеном. В рассматриваемом подходе аффинность определяется не для отдельной пары антиген-антитело, а между кортежами антител и антигенов длины  $M$  – мультиантителом и мультиантигеном:

$$mAb = ab_1, ab_2, ab_3, \dots, ab_{M+k},$$

$$mAg = ag_1, ag_2, ag_3, \dots, ag_M.$$

Мультиантитело включает в себя  $k$  прогнозируемых им значений – эта часть не участвует в определении аффинности. Число антител, которые в нее входят, соответствует величине горизонта прогнозирования. В терминах подхода СБР, мультиантитело исполняет роль прецедента, описывая текущую ситуацию (последовательность известных значений ряда) и принятое ранее решение (соответствующий прогноз). Таким образом, для оценки будущих значений ряда требуется найти мультиантитела с наибольшей аффинностью – наиболее соответствующие текущей задаче.

Аффинность определяется по формуле:

$$Aff = \sum_{i=1}^M (1 + d_i)^{-1} / M, \quad i \in [1, M], \quad (1)$$

где  $d_i$  – евклидово расстояние между парами значений  $mAb$  и  $mAg$ .

Формально модель ИИС можно представить следующим образом:

$$immNet = \langle X, C, C_m, D, S, n, N_c, \zeta, \sigma_d, \sigma_s \rangle, \quad (2)$$

где  $X$  – популяция антигенов;  $C$  – популяция антител;  $C_m$  – популяция клеток памяти,  $D$  – матрица  $mAb - mAg$  аффинностей;  $S$  – матрица  $mAb - mAb$  аффинностей;  $n$  – число клеток, отобранных из популяции антител  $C$  для клонирования и мутации. Мутация представляет собой изменение части мультиантитела, соответствующей прогнозируемому значению;  $N_c$  – количество клонов, создаваемых одним антителом;  $\zeta$  – процент клонов, отобранных для последующей обработки;  $\sigma_d$  – пороговый коэффициент стимуляции клетки;  $\sigma_s$  – пороговый коэффициент сжатия сети.

Алгоритм получения прогноза при помощи предлагаемой модели следующий:

Для создания иммунной сети используется часть известных значений временного ряда. Неиспользованные значения выполняют роль обучающей и контрольной выборок.

Если позволяет размер ряда, выполняется обучение сети.

При значительном размере сети целесообразно применение оператора супрессии – определение аффинности между мультиантителами и последующее сжатие сети, с целью устранения избыточности.

На основании поставленной задачи формируется мультиантиген – последовательность известных значений ряда, предшествующих прогнозируемым.

Определяется набор мультиантител, аффинность которых больше порогового значения.

Прогнозом иммунной сети является результат мультиантитела с наибольшей аффинностью.

После получения реального значения производится коррекция сети. Отобранные на этапе 5 мультиантитела подвергаются клонированию, в процессе которого выполняется оператор мутации. Клоны замещают мультиантитела, их породившие.

Мутации подвергается только та часть мультиантитела, которая определяет его прогноз, и не участвует в определении аффинности. Степень мутации опреде-

ляется аффинностью данного мультиантитела – наиболее подходящий прецедент подвергается более значительной коррекции. Если остались нерешенные задачи прогнозирования, возврат к пункту 4.

Шаги 4 – 9 представляют собой обучение иммунной сети. Шаг 7 выполняется при наличии реального значения прогнозируемой величины. На рис. 1 показана зависимость средней абсолютной ошибки (MAE) от аффинности для необученной (график 1) и обученной (график 2) ИИС. Таким образом, цель обучения ИИС заключается в настройке системы в соответствии с новыми значениями временного ряда.

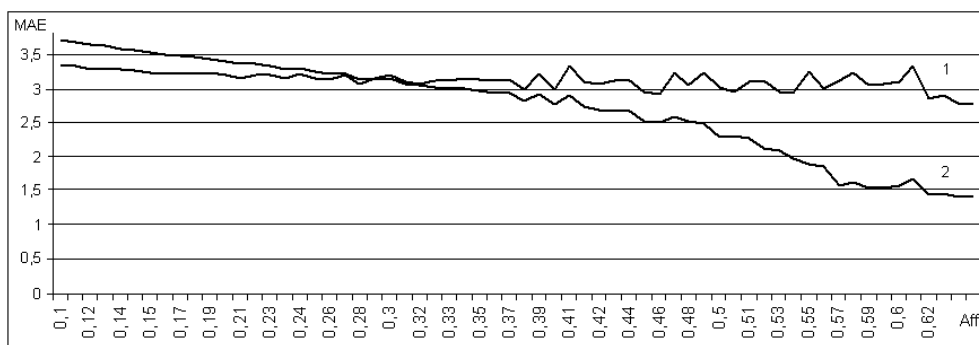


Рис. 1. Зависимость величины ошибки от аффинности мультиантител

Настройка параметров иммунной сети, таких как пороговая аффинность, определяющая отбор мультиантител, величина сети и коэффициент сжатия, на конкретное приложение в подавляющем количестве применений осуществляется методами вычислительного эксперимента. Идентификация

модели (определение величины мультиантитела) выполняется на основе рекомендаций, изложенных в [9], в качестве критерия используется средняя абсолютная ошибка. График на рис. 2 демонстрирует влияние величины мультиантитела  $M$  на точность получаемого прогноза.

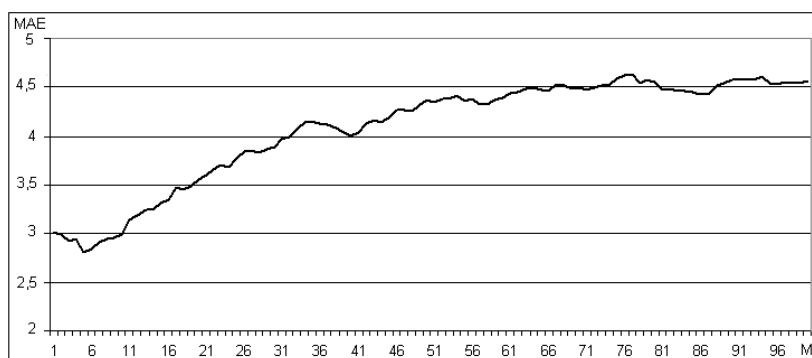


Рис. 2. Зависимость средней ошибки прогноза от величины мультиантитела

## Результаты сравнительного анализа

В ходе экспериментальных исследований было проведено краткосрочное прогнозирование на примере рядов, используемых в M3-Competition [10], а также ряда среднесуточных показаний температуры. Для оценки прогноза использовалось значение средней абсолютной ошибки (MAE), результаты применения методов прогнозирования, отличных от AINet, рассматриваемого в работе, взяты из [10]. Результаты прогнозирования разными методами приведены в табл. 1. Результат прогнозирования ряда среднесуточных показаний температуры (Meteo) демонстрирует преимущество использования обученной иммунной

сети. Однако, в случае малого количества значений временного ряда, не позволяющего создать иммунную сеть достаточного размера, рассмотренный подход может значительно уступать существующим методам. Малый размер сети ведет к невозможности найти мультиантитела, обладающие высокой аффинностью (т.е. в базе отсутствуют подходящие прецеденты), и как следствие, к недостаточной обученности системы.

## Выводы

В работе предложена модель прогнозирования временных рядов на основе искусственной иммунной сети, использующей метод вывода по прецедентам.

Основные особенности предложенной модели заключаются в следующем:

- целесообразность применения для решения задачи краткосрочного прогнозирования;

- необходимость обучающей выборки значительного объема;

- возможен низкий результат при наличии тренда, желательна постоянство математического ожидания;

Таблица 1

Средняя ошибка прогнозирования при использовании различных методов

Метод	N2832 (104)	N2841 (104)	N2856 (92)	N2859 (79)	N2863 (79)	N2864 (79)	N2869 (79)	Meteo (21337)
Exp.Smoothing	3148,09	68,70	253,08	132,71	672,5	1028,13	497,5	2,92
Holt-Winters	3676,51	52,72	296,47	149,2	504,79	1146,32	580,31	2,9
Box-Jenkins	2814,27	68,7	260,21	125,09	639,94	896,62	210,65	2,99
RBF	3759,34	37,47	381,63	145,57	832,50	1091,78	628,15	–
ForecastPro	3098,99	68,58	260,26	132,71	672,5	1028,13	185,79	–
SmartFcs	3183,35	67,71	255,99	132,71	746,12	1077,06	525,2	–
Automat ANN	1937,68	72,95	300,77	111,76	604,83	976,70	349,1	–
AINet	2848,68	151,46	216,50	130,51	707,64	1463,12	225,76	2,45

- чувствительность к выбросам на ранних этапах обучения сети.

Представленные в работе результаты подтверждают эффективность использования рассмотренного подхода для краткосрочного прогнозирования временных рядов. Использование модели для прогнозирования ряда в условиях недостатка исходной информации требует дополнительных исследований и доработки модели.

## Список литературы

1. Дуброва Т.А. Статистические методы прогнозирования в экономике: Учебное пособие, практикум, тесты, программа курса / Т.А. Дуброва, М.Ю. Архипова. – М.: Московский государственный университет экономики, статистики и информатики, 2004. – 136 с.
2. Батуру А.П. Финансовые временные ряды: курсовое прогнозирование и проблема обнаружения предвестников существующего изменения закономерности / А.П. Батуру, Н.М. Еременко // Банковские технологии. – 2001. – № 12. – С. 70-77.
3. Дасгупта Д. Искусственные иммунные системы и их применение: пер. с англ. под ред. А.А. Романюхи / Д. Дасгупта. – М.: ФИЗМАТЛИТ, 2006. – 344 с. – ISBN 5-9221-0706-2.
4. Прасолов С.В. Постановка задачи прогнозирования, основанная на применении упрощенной модели искусственной иммунной сети / С.В. Прасолов, Д.В. Шойтов //

Ученые записки. Электронный научный журнал Курского государственного университета. – 2009. – Т. 2. – С. 1-3.

5. Бидюк П.И. Алгоритм клонального отбора для прогнозирования нестационарных динамических систем / П.И. Бидюк, В.И. Литвиненко, И.В. Баклан, А.А. Фефелов // Искусственный интеллект. – 2004. – № 4. – С. 89-99.

6. Самигулина Г.А. Разработка интеллектуальных экспертных систем прогнозирования и управления на основе искусственных иммунных систем / Г.А. Самигулина // Проблемы информатики. – Н-ск, 2010. – № 1. – С. 15-22.

7. Черный С.Г. Применение case based reasoning для поддержки принятия решений / С.Г. Черный // Вестник ХНТУ. – 2010. – № 2(38). – С. 336–342.

8. Поминчук Е.В. Метод интеллектуального анализа данных прецедентов для прогнозирования временных рядов / Е.В. Поминчук, А.Б. Иващенко // Мат. 4-й междунаучно-технической конференции "Моделирование и компьютерная графика – 2011". – Донецк, 2011. – С. 193-197.

9. Чучуева И.А. Модель экстраполяции временных рядов по выборке максимального подобия / И.А. Чучуева // Информационные технологии. – 2010. – № 12. – С. 43-47.

10. Makridakis S. The M-3 Competition: Results, Conclusions and Implications / S. Makridakis, M. Hibon // International of Forecasting. – 2000. – № 16. – С. 451-476.

Поступила в редколлегию 10.12.2012

**Рецензент:** д-р техн. наук, проф. О.Г. Руденко, Харьковский национальный университет радиоэлектроники, Харьков.

## ВИКОРИСТОВУВАННЯ ШТУЧНИХ ІМУННИХ МЕРЕЖ ДЛЯ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ

М.М. Корабльов, Г.С. Івашенко

У статті запропоновано комбінований метод прогнозування часових рядів за допомогою штучних імунних систем. Розглянуто модель на основі моделі штучних імунних мереж і методу виведення з прецедентів (СВР). Оцінка ефективності моделі виконана шляхом порівняльного аналізу, представлені результати експериментальних досліджень, що демонструють особливості запропонованого підходу.

**Ключові слова:** прогнозування, часовий ряд, вивід по прецедентах, штучні імунні системи, антитіло, антиген, афінність, клонування, мутація.

## ARTIFICIAL IMMUNE NETWORK FOR TIME SERIES PREDICTION

N.M. Korablev, G.S. Ivaschenko

This paper proposes a combined method of time series prediction using artificial immune systems. Considered a model based on artificial immune network and the case based reasoning method. Evaluation of the effectiveness the model is made by a comparative analysis, the results of experimental research demonstrate the features of the proposed approach.

**Keywords:** prognostication, temporal row, conclusion on precedents, artificial immune systems, antibody, antigen, cloning, mutation.