

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

ІВАЩЕНКО ГЕОРГІЙ СТАНІСЛАВОВИЧ

УДК 004.89

**ГІБРИДНІ МОДЕЛІ КОРОТКОСТРОКОВОГО ПРОГНОЗУВАННЯ
ЧАСОВИХ РЯДІВ НА ОСНОВІ ШТУЧНИХ ІМУННИХ СИСТЕМ**

05.13.23 – системи та засоби штучного інтелекту

АВТОРЕФЕРАТ

дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2016

Дисертацією є рукопис.

Робота виконана у Харківському національному університеті радіоелектроніки, Міністерство освіти і науки України.

Науковий керівник – доктор технічних наук, професор
Корабльов Микола Михайлович,
Харківський національний університет
радіоелектроніки, професор кафедри
електронних обчислювальних машин.

Офіційні опоненти: доктор технічних наук, професор
Бідюк Петро Іванович,
Інститут прикладного системного аналізу Націо-
нального технічного університету України «Київ-
ський політехнічний інститут», професор кафедри
математичних методів системного аналізу;

доктор технічних наук, професор
Субботін Сергій Олександрович,
Запорізький національний технічний університет,
завідувач кафедри програмних засобів.

Захист відбудеться « ____ » _____ 2016 р. о ____ годині на за-
сіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному
університеті радіоелектроніки за адресою: 61166, м. Харків, просп. Науки, 14.

З дисертацією можна ознайомитися в бібліотеці Харківського національного
університету радіоелектроніки за адресою: 61166, м. Харків, просп. Науки, 14.

Автореферат розіслано « ____ » _____ 2016 р.

Учений секретар
спеціалізованої вченої ради,
д.т.н, проф.

О.А. Винокурова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. В процесі розробки систем інтелектуальної обробки інформації все більше уваги приділяється використанню біоінспірованих підходів, таких як штучні нейронні мережі, еволюційні обчислення, штучні імунні системи (ШІС) та ін. ШІС відрізняються можливостями розпізнавання, паралельної децентралізованої обробки інформації та здатністю інтегруватися з іншими підходами. Застосування гібридних підходів, що реалізують комплексне застосування різних моделей і методів, дозволяє компенсувати недоліки одних моделей за допомогою врахування особливостей інших. Завдяки своїм особливостям, ШІС є об'єктом інтенсивних досліджень у напрямку створення гібридних систем інтелектуальної обробки інформації для вирішення різноманітних прикладних задач.

Однією з актуальних прикладних задач залишається вирішення завдання короткострокового прогнозування часових рядів, що дозволяє визначати майбутній стан різних систем (економічних, соціальних, технічних та інших) на основі аналізу вже наявних ретроспективних даних. Точність результату прогнозування залежить від таких факторів, як обсяг зібраної для отримання прогнозу інформації, наявність викривлень (аномальних викидів і пропусків), необхідна величина горизонту прогнозування, можливість використання зовнішніх факторів, що впливають на прогнозовану величину.

Одним з визначальних факторів отримання достовірного прогнозу є обґрунтований вибір серед існуючих методів прогнозування. Зараз найбільш поширені статистичні моделі прогнозування (експоненціального згладжування, регресійні та авторегресійні моделі), нейромережні моделі, моделі на основі виведення за прецедентами (Case Based Reasoning – CBR). На різних фрагментах часового ряду один і той самий метод прогнозування може мати різну ефективність, тому доцільне проведення сегментації вихідного ряду і використання набору різних підходів. Для вирішення цього завдання перспективне застосування гібридних моделей на основі ШІС. Використання комбінації методів у рамках гібридної моделі веде до збільшення трудомісткості процесу отримання прогнозу, і для забезпечення можливості обробки даних у режимі реального часу необхідно використовувати переваги паралельної обробки інформації гібридною моделлю на основі ШІС.

Різним аспектам використання ШІС присвячені роботи П.І. Бідюка, Д. Дасгупти, Л.Н. Де Кастро, В.І. Литвиненка, Д.І. Тімміса, Дж.-С.Р. Янга та інших. Серед зарубіжних вчених, роботи яких значно впливають на розвиток методів прогнозування часових рядів, можна відзначити Дж. Бокса, Г.М. Дженкінса, Р.Дж.А. Літтла, Ю.П. Лукашина, С. Макрідакіса, Е. Петерса, С. Сінга, І.А. Чучуєвої. Основні досягнення в галузі прискорення обчислень належать В.В. Воєводіну, Вл.В. Воєводіну, В.П. Гергелю та іншим. Слід зазначити невелику кількість робіт, присвячених проблемам використання ШІС для прогнозування часових рядів, особливо, з урахуванням присутності можливих викривлень вихідних даних.

У зв'язку з цим розробка гібридних моделей на основі ШС для вирішення завдання короткострокового прогнозування часових рядів є актуальною як з теоретичної, так і з практичної точки зору. Задачі, які при цьому виникають, обумовили напрямок досліджень дисертаційної роботи.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана відповідно до плану науково-дослідних робіт Харківського національного університету радіоелектроніки (ХНУРЕ) в рамках держбюджетних тем «Еволюційні гібридні системи обчислювального інтелекту зі змінною структурою для інтелектуального аналізу даних» (№ ДР 0110U000458), розділ «Еволюційні гібридні методи та моделі інтелектуальної обробки інформації зі змінною структурою за умов невизначеності», та «Нейро-фаззі системи для поточної кластеризації і класифікації послідовностей даних за умов їх викривленості відсутніми та аномальними спостереженнями» (№ ДР 0113U000361), розділ «Адаптивні методи та моделі класифікації даних і прогнозування часових рядів за умов їх викривленості відсутніми та аномальними спостереженнями на основі штучних імунних систем», затвердженими Міністерством освіти та науки України. Автор був одним з виконавців робіт за даними темами.

Мета і задачі дослідження. Метою дисертаційної роботи є розробка, дослідження та удосконалення гібридних моделей на основі штучних імунних систем, які застосовуються для вирішення задачі короткострокового прогнозування стохастичних часових рядів, що дозволяє підвищити як точність прогнозу, так і стійкість до викривлень вихідних даних.

Для досягнення цієї мети були поставлені та вирішені такі задачі:

1. Аналіз існуючих методів прогнозування часових рядів, виявлення основних недоліків і визначення перспективних підходів для їх усунення.
2. Розробка моделей прогнозування часових рядів на основі ШС і методу виведення за прецедентами.
3. Розробка гібридної моделі на основі ШС, методу виведення за прецедентами та найпростіших методів прогнозування, яка враховує можливі викривлення та вплив зовнішніх факторів.
4. Програмна реалізація запропонованих моделей з використанням засобів паралельних і розподілених обчислень, що забезпечує роботу ШС у режимі реального часу.
5. Експериментальні дослідження розроблених методів і моделей та вирішення за їх допомогою практичних задач.

Об'єктом дослідження є процеси короткострокового прогнозування часових рядів.

Предметом дослідження є гібридні моделі на основі штучних імунних систем, що застосовуються для вирішення задачі короткострокового прогнозування часових рядів.

Методи дослідження. Для вирішення поставлених задач були використані: методи інтелектуального аналізу даних і апарат штучних імунних систем, що забезпечили побудову гібридних моделей для вирішення задачі прогнозу-

вання часових рядів; апарат математичної статистики, що дозволив виконувати систематизацію та аналіз отриманих результатів для формування наукових і практичних висновків; принципи паралельних та розподілених обчислень для підвищення ефективності роботи запропонованих імунних алгоритмів прогнозування; імітаційне моделювання, яке підтвердило ефективність запропонованих гібридних моделей.

Наукова новизна результатів дисертаційної роботи. Вирішення поставлених задач дозволило автору отримати такі результати:

1. Вперше запропонована гібридна модель на основі штучних імунних систем, яка використовує поєднання методу виведення за прецедентами, моделі клонального відбору та найпростіших методів прогнозування, що виконує фрагментацію вихідного часового ряду та підбір для кожного фрагмента свого методу прогнозування, що дозволяє скоротити обсяг навчальної вибірки та підвищити точність прогнозу.

2. Вперше запропонована гібридна модель на основі моделі клонального відбору та методу виведення за прецедентами, яка використовується для прогнозування викривлених часових рядів, враховує вплив зовнішніх факторів та використовує переваги розподіленої обробки даних, що дозволяє підвищити точність прогнозу часових рядів, які містять пропущені значення, та забезпечує отримання прогнозу у режимі реального часу.

3. Набула подальшого розвитку модель прогнозування часових рядів на основі штучної імунної мережі, яка відрізняється використанням мультиантигенів, які виконують роль прецедентів, що узагальнює її опис і дозволяє змінювати параметри та структуру імунної мережі в процесі навчання.

4. Набув подальшого розвитку метод прогнозування часових рядів на основі методу висновку за прецедентами, який відрізняється застосуванням імунних операторів для управління базою прецедентів, що дозволяє виконувати обробку аномальних викидів у часових рядах і підвищити точність отриманого прогнозу.

Практичне значення результатів дисертаційної роботи. Запропоновані моделі та методи прогнозування часових рядів дозволяють вирішувати широке коло завдань, які виникають у різноманітних практичних галузях. Розроблені методи і моделі дозволяють скоротити час на розробку інформаційних систем для вирішення задач прогнозування в науці та промисловості, забезпечують підвищення якості оцінок прогнозів. Експериментальні дослідження, проведені для оцінки запропонованих моделей і методів, підтверджують основні положення, які виносяться на захист.

Результати дисертаційної роботи були використані у ТОВ «Буденергомаш» (акт впровадження від 17.03.2015) для підтримки прийняття рішень у сфері планування продажів готової продукції та для вирішення завдань планування товарообігу у ТОВ РІК «СВ-Прилад» (акт впровадження від 19.10.2015). Результати дисертаційної роботи також були впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акт впровадження від 10.02.2016).

Особистий внесок здобувача. Усі положення, що виносяться на захист, основні результати теоретичних та експериментальних досліджень отримані здобувачем особисто. Їх основний зміст викладено у роботах [1–29]. Внесок автора в публікаціях, написаних у співавторстві, такий: в [1] – дослідження адаптаційних можливостей моделі клонального відбору для вирішення задачі комівояжера; в [2] – розроблена гібридна модель на основі моделі клонального відбору, що використовує метод висновку за прецедентами; в [3] – розроблена гібридна модель для прогнозування часових рядів із пропусками; в [4] – аналіз способів розпаралелювання імунного алгоритму прогнозування на основі гібридної моделі; в [5] – запропонована гібридна модель на основі моделей штучних імунних систем для прогнозування часових рядів, що містять аномальні викиди; в [6] – розроблена гібридна модель на основі моделі штучної імунної мережі, що використовує метод висновку за прецедентами, для вирішення задачі короткострокового прогнозування часових рядів; в [7] – розроблена гібридна модель прогнозування, яка використовує модель клонального відбору, метод висновку за прецедентами і найпростіші методи короткострокового прогнозування; в [8] – використання різнорідних антитіл у гібридній моделі, яка виконує фрагментацію часового ряду; в [9] – запропоновано гібридний підхід на основі штучних імунних систем і жадібного алгоритму, який використовується для вирішення задачі комівояжера; в [10] – агентно-орієнтований підхід на основі моделі клонального відбору для вирішення задачі комівояжера; в [11] – аналіз агентно-орієнтованого підходу для вирішення задачі комівояжера на основі моделі клонального відбору; в [12] – запропонована модель прогнозування часових рядів на основі моделі штучної імунної мережі; в [13] – запропонована модель прогнозування часових рядів на основі моделі клонального відбору, що використовує метод висновку за прецедентами; в [14] – порівняльний аналіз використання моделей на основі клонального відбору та імунної мережі в поєднанні з висновком за прецедентами для вирішення завдання короткострокового прогнозування; в [15] – запропоновано метод виявлення аномальних викидів на основі моделі клонального відбору; в [16] – запропоновано використання найпростіших методів прогнозування в гібридній моделі; в [17] – використання моделі клонального відбору для виявлення аномальних викидів у часових рядах; в [18] – розроблена гібридна модель на основі моделі клонального відбору для вирішення завдання короткострокового прогнозування, що дозволяє враховувати зовнішні фактори; в [19] – заповнення пропусків у часових рядах за допомогою моделі клонального відбору; в [20] – запропоновано метод апроксимації за допомогою моделі клонального відбору; в [21] – виявлення аномальних викидів за допомогою моделі позитивного відбору; в [22] – урахування зовнішніх факторів під час прогнозування викривлених часових рядів за допомогою гібридної моделі; в [23] – запропоновано паралельний імунний алгоритм короткострокового прогнозування на основі острівної моделі генетичного алгоритму; в [24] – виконані експериментальні дослідження послідовної та паралельної реалізації імунного алгоритму з використанням технології TPL; в [25] – проведено аналіз способів розпаралелювання імунного алгоритму; в [26] – запропоно-

вано метод настроювання параметрів імунного алгоритму прогнозування за допомогою спрощеної моделі штучної імунної мережі; в [27] – запропоновано використання показника застосовності імунного алгоритму короткострокового прогнозування; в [28] – аналіз показників для визначення прогнозованості часового ряду за допомогою гібридних моделей, що використовують висновок по прецедентах; в [29] – прогнозування за допомогою мультиагентних систем.

Апробація результатів дисертації. Основні положення та результати дисертаційної роботи доповідалися й обговорювалися на: 2-й Міжнародній науково-технічній конференції «Інформаційні технології в навігації і управлінні: стан та перспективи розвитку» (Київ, 2011); 9-11-й Міжнародних науково-практичних конференціях «Математичне та програмне забезпечення інтелектуальних систем» (Дніпропетровськ, 2011-2013 рр.); 16-19-му Міжнародних молодіжних форумах «Радіоелектроніка та молодь у XXI столітті» (Харків, 2012-2015 рр.); 8-й Міжнародній науково-технічній конференції «Інформатика та комп'ютерні технології» (Донецьк, 2012); 1-2-й Міжнародних науково-технічних конференціях «Проблеми інформатизації» (Київ, 2013-2014 рр.); Міжнародних наукових конференціях «Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту» – ISDMCI (Херсон, 2013-2015 рр.); 3-5-й Міжнародних науково-технічних конференціях «Сучасні напрямки розвитку інформаційно-комунікаційних технологій та засобів управління» (Київ, 2013-2015 рр.); Міжнародній науково-практичній конференції «Сучасні проблеми правового, економічного та соціального розвитку держави» (Харків, 2013); 16-й науково-технічній конференції з міжнародною участю «Нейроінформатика» (Москва, 2014); Міжнародних науково-практичних конференціях «Інформаційні управляючі системи та технології» (Одеса, 2014, 2015 рр.).

Публікації. За темою дисертаційної роботи опубліковано 29 наукових праць, з них: 7 статей у фахових періодичних виданнях України з технічних наук (серед них 4 видання, що входять до міжнародних наукометричних баз); 22 публікації у збірниках праць і тез міжнародних наукових конференцій та семінарів.

Структура та обсяг дисертаційної роботи. Дисертація складається із вступу, чотирьох розділів, висновків, списку використаних літературних джерел з 138 найменувань, 1 додатка. Робота містить 29 рисунків, 12 таблиць. Загальний обсяг роботи складає 163 сторінки, з них 140 – основного тексту.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність обраної теми дисертаційної роботи, сформульовано мету та задачі дослідження, визначено об'єкт, предмет і методи досліджень, відзначено зв'язок роботи з науковими темами, охарактеризовано наукову новизну та практичне значення отриманих результатів, наведено відомості про апробацію результатів дисертації та кількість публікацій за темою дисертаційної роботи.

У першому розділі проведено аналіз проблеми прогнозування часових рядів, розглянуто існуючі методи короткострокового прогнозування, описані основні особливості апарату ШС та методу висновку за прецедентами. На основі проведеного аналізу визначено сукупність перспективних напрямків вирішення задачі короткострокового прогнозування на основі гібридних моделей та сформульована постановка задачі дослідження.

Внаслідок аналізу задачі короткострокового прогнозування були виявлені такі проблеми, як наявність викривлень (аномальних викидів і пропусків) в оброблюваних ретроспективних даних, складність урахування та відбору зовнішніх факторів, які впливають на прогнозовану величину, поданих у вигляді часових рядів, з їх апріорного набору, низьку точність зі збільшенням горизонту прогнозування, трудомісткість процесу отримання прогнозу, що унеможлиблює роботу в режимі реального часу.

Розглянуто найбільш поширені існуючі методи прогнозування. Недоліками статистичних методів прогнозування є складність ідентифікації параметрів, обмеження на мінімальну довжину ряду, складність адаптації до нових спостережень і роботи у режимі реального часу. Штучні нейронні мережі дозволяють створювати адаптивні моделі прогнозування, які підтримують паралельну обробку, але висувають високі вимоги до несуперечності навчальної вибірки, характеризуються складністю вибору топології, відповідної розв'язуваної задачі, та ресурсомісткістю процесу навчання. Прогнозування за допомогою класифікаційно-регресійних дерев дозволяє використовувати як безперервні, так і категоріальні ознаки, однак цей підхід схильний до «прокляття розмірності» та трудомісткості адаптації до нових даних. Перспективним є застосування методу виведення за прецедентами, який не накладає обмежень на стаціонарність ряду, легко піддається паралельній обробці та допускає роботу у режимі реального часу, але вимагає навчальної вибірки великого обсягу та може бути застосований для отримання лише короткострокового прогнозу.

Оскільки жоден з розглянутих підходів не відповідає повною мірою поставленим вимогам, пропонується використання гібридних моделей і методів, які використовують поєднання виведення за прецедентами та апарату штучних імунних систем, зокрема моделі клонального відбору та моделі штучної імунної мережі, для вирішення задачі короткострокового прогнозування стохастичних часових рядів.

У другому розділі для вирішення задачі короткострокового прогнозування пропонуються гібридні моделі на основі моделей штучної імунної мережі та клонального відбору, що використовують принципи методу висновку за прецедентами та найпростіші методи прогнозування часових рядів.

Розроблено модель прогнозування на основі моделі штучної імунної мережі, в якій антитіла формуються з відомих значень часового ряду, антигени – значення часового ряду, що безпосередньо передують прогнозованим, а афінність визначається не для окремої пари антиген-антитіло, а між кортежами антитіл і антигенів довжини L – мультиантитілом $mAb = ab_1, ab_2, ab_3, \dots, ab_{L+f}$ та

мультиантигеном $mAg = ag_1, ag_2, \dots, ag_L$. Мультиантитіло містить f прогнозованих значень, які не беруть участі у визначенні афінності. Мультиантитіло виконує роль прецеденту, описуючи задачу (послідовність відомих значень ряду) та прийняте раніше рішення (відповідний прогноз). Таким чином, для оцінки майбутніх значень ряду потрібно знайти мультиантитіла з найбільшою афінністю. Афінність визначається за формулою:

$$Aff = \frac{\sum_{i=1}^L (1 + d_i)^{-1}}{L} \in (0,1], \quad (1)$$

де $d_i = |ab_i - ag_i|$ – евклідова відстань між парами значень, що складають мультиантитіло mAb та мультиантиген mAg .

У загальному вигляді зміну популяції мультиантитіл у поточному поколінні gen внаслідок дії всіх імунних операторів можна представити в такий спосіб:

$$\begin{aligned} mAb^{gen+1} &= Suppres(Edit(Mutate(Clone(Sel(mAb^{gen})))))), \\ Clone &: mAb^{gen} \rightarrow mAb_C^{gen}, \\ Mutate &: mAb_C^{gen} \rightarrow mAb_{MC}^{gen}, \\ Edit &: (mAb_{MC}^{gen}, mAb^{gen}) \rightarrow mAb^{gen+1}. \end{aligned} \quad (2)$$

Мутації піддається лише та частина клону відібраного мультиантитіла, яка визначає його прогноз, і не бере участі у визначенні афінності. Потрібне визначення специфічного оператора мутації, величина кроку якої визначається за формулою:

$$ab_i = Aff(mAb) * (ab_i - D_i), \quad (3)$$

де ab_i – поточне значення, що складає антитіло, D_i – різниця між прогнозованим значенням ab_i' та відповідним реальним значенням, $i \in [L+1, L+f]$. Величина кроку мутації визначається афінністю цього мультиантитіла та величиною помилки відносно реального значення, тобто, у термінах виведення за прецедентами, найбільш відповідний прецедент піддається більш значній корекції. Особливістю моделі імунної мережі є взаємозв'язок між окремими антитілами. Корекція у процесі навчання ШС антитіл, що належали одному мультиантитілу, впливає на інших, близьких до нього.

Модель клонального відбору дозволяє поєднувати використання СВР та інших методів прогнозування, виконуючи у ході отримання прогнозу сегментацію ряду та підбір найбільш відповідного методу прогнозування для кожного фрагменту вихідного ряду. Тому крім антитіл, що реалізують виведення за прецедентами, на основі фрагментів часового ряду у рамках гібридної моделі на

основі моделі клонального відбору створюються антитіла інших типів, що формують свій варіант прогнозу самостійно, використовуючи найпростіші з методів прогнозування, такі як просте та експоненціальне середнє, наївні моделі та інші методи.

Антитіло – основний елемент гібридної моделі, являє собою набір параметрів, що описують поставлене завдання (набір відомих значень ряду) та прийняте рішення (пропонований варіант прогнозу): $Ab_i = ab_1, ab_2, \dots, ab_L, \dots, ab_{L+f}$, де i – індекс в популяції, L – довжина вибірки відомих значень ряду, f – величина горизонту прогнозування. Антиген включає до себе вибірку відомих значень часового ряду, які безпосередньо передують прогнозованим, і множину відповідних вибірок значень всіх зовнішніх чинників. Афінність визначається з урахуванням значень вагових коефіцієнтів для вибірок супутніх часових рядів, що описують зовнішні фактори:

$$Aff(Ab) = \eta * (\eta_{Ab} * Aff_{Ab} + \eta_{Ab'} * Aff_{Ab'}), \quad (4)$$

де η – значення коефіцієнта відбору; η_{Ab} і $\eta_{Ab'}$ – коефіцієнти, що визначають вплив вибірок вихідного і супутнього рядів на афінність антитіла, при цьому $\eta_{Ab} + \eta_{Ab'} = 1$. В антигені можуть бути не представлені деякі зовнішні чинники, а в окремому антитілі представлена лише одна супутня вибірка.

Коефіцієнт відбору η призначений для визначення пріоритету антитіл в залежності від їх типів, оскільки антитіла, створені на основі однієї й тієї самої вибірки, матимуть однакове значення афінності. Міри подібності вибірок прогнозованого часового ряду Aff_{Ab} і вибірок рядів значень зовнішніх чинників $Aff_{Ab'}$ визначаються аналогічно (1).

Після отримання реальних значень прогнозу відбувається корекція коефіцієнтів, які впливають на значення афінності антитіл. На рис. 1 показано зміну кількості антитіл, створених за допомогою методу виведення за прецедентами (CBR), наївної моделі (Naive1), простого середнього (Avg) і експоненціального середнього (ExpAvg) у процесі навчання гібридної моделі на основі клонального відбору.

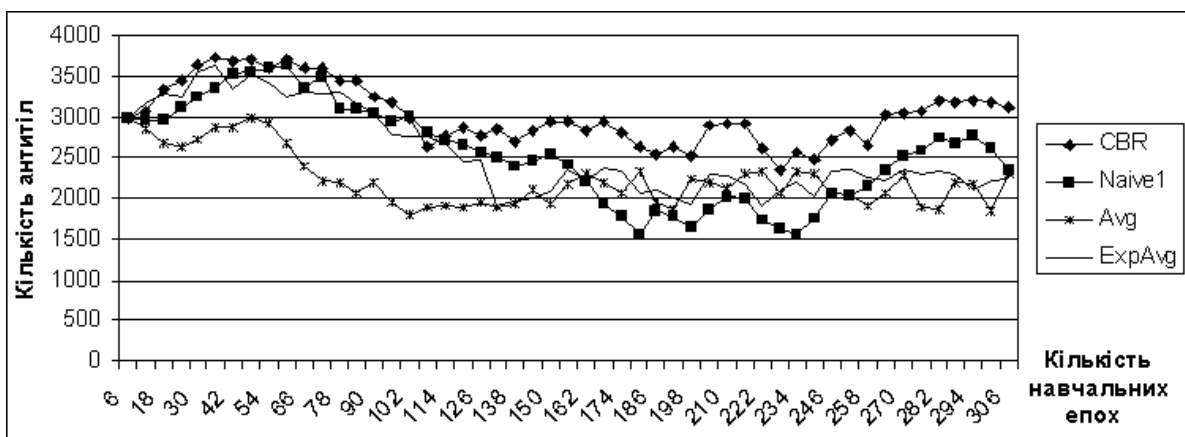


Рисунок 1 – Зміна кількості антитіл різного типу в ході навчання ШС

У гібридній моделі на основі моделі клонального відбору мутації піддаються всі клони відібраних антитіл, значення кроку мутації антитіла σ_t для поточного покоління визначається так:

$$\sigma_t = \sigma_{t-1} r_m \frac{Aff(Ab_i) - Aff_{best}}{Aff_{worst} - Aff_{best}}, \quad (5)$$

де σ_{t-1} – величина кроку мутації цього антитіла у попередньому поколінні популяції, r_m – випадкове значення в діапазоні від 0 до 1, $Aff(Ab_i)$ – значення афінності антитіла Ab_i , Aff_{best} – краще значення афінності, отримане в поточному поколінні, Aff_{worst} – найгірше значення афінності, отримане в поточному поколінні. Таким чином, крок мутації параметрів антитіл регулюється в процесі роботи імунного алгоритму.

Проведено порівняння отриманих результатів (AINet і Clonal Selection) з результатами прогнозування за допомогою експоненціального згладжування (Exp.Smoothing), моделі Хольта (HoltWinters), моделі ARIMA, експертних систем (ForecastPro і SmartFcs), штучних нейронних мереж. В процесі навчання гібридної моделі виконується вибір методу прогнозування для кожного фрагменту вихідного ряду, що дозволяє скоротити обсяг навчальної вибірки і підвищити точність прогнозу на коротких часових рядах. Дана модель дозволяє враховувати та проводити відбір зовнішніх чинників, які мають найбільший вплив на прогнозовану величину, що дозволяє підвищити точність прогнозу, але веде до значного збільшення необхідної популяції антитіл.

У третьому розділі розглянуто питання обробки викривлень у прогнозованих часових рядах (аномальних викидів та пропущених значень), проблеми організації паралельної роботи імунного алгоритму та пошук критерію застосовності розробленого підходу для прогнозування часових рядів.

Запропоновано метод, який дозволяє виявити аномалії та враховувати пропущені значення у часовому ряді в процесі вирішення задачі короткострокового прогнозування. У разі застосування поєднання моделі клонального відбору та СВР для вирішення задачі виявлення аномалій, прецедент містить вибірку значень часового ряду та її характеристику (чи містить дана вибірка аномальні викиди). Кожен елемент множини Ag – приклад з навчальної вибірки, що перевірявся фрагмент часового ряду, представлений у вигляді вектора фіксованої довжини $Ag_j = \langle ag_1, ag_2, ag_3, \dots, ag_L \rangle$, де j – індекс у популяції антигенів, L – число значень у складі антигену – довжина вибірки значень ряду, що перевіряються на наявність аномальних викидів.

Популяція антитіл під час вирішення задачі виявлення аномальних викидів представлена у вигляді множини $Ab = \{Ab_1, Ab_2, \dots, Ab_N\}$, де N – розмір популяції антитіл. Перша частина ab_1, ab_2, \dots, ab_L аналогічна за структурою антигену, визначає міру близькості антитіла антигену за допомогою визначення значення функції афінності. Друга частина антитіла містить його ідентифікатор,

значення міри відхилення D_{Ab} , що показує близькість антитіла до аномального, $D_{Ab} \in (0,1]$ та ідентифікатор того антитіла, на підставі порівняння з яким визначається значення D_{Ab} . Ця частина не бере участь у визначенні афінності. Визначення антитіла як аномального проводиться шляхом співставлення з іншими антитілами у популяції, зазначеними як такі, що містять аномальні значення. Аномальним відзначається усе антитіло, незалежно від того, скільки аномальних значень воно містить. Можливе навчання ШІС без вчителя шляхом розрахунку матриці афінності між антитілами та подальшого урахування антитіл, що мають найменшу середню афінність.

Під час отримання прогнозу та навчання ШІС афінність визначається з урахуванням кількості пропущених значень у складі антитіла та антигену:

$$Aff(Ab) = (1 - n_m^{-1}) * \eta * (\eta_{Ab} * Aff_{Ab} + \eta_{Ab'} * Aff_{Ab'}), \quad (6)$$

де n_m – кількість пропусків у вибірці значень, на основі якої створюється антитіло; η – значення коефіцієнта відбору; η_{Ab} і $\eta_{Ab'}$ – коефіцієнти, що визначають вплив вибірок вихідного та супутнього рядів на афінність антитіла, при цьому $\eta_{Ab} + \eta_{Ab'} = 1$.

Таблиця 1 – Симетрична середня абсолютна помилка (SMAPE,%) для різних методів прогнозування за різної кількості пропущених значень

Метод	N704 (44)	N736 (44)	N1366 (63)	N2830 (104)	N2841 (104)	N2867 (79)	Meteo (21337)
Exp.Smooth	4,08	12,11	0,42	2,47	0,5	20,52	4,56
RobustTrend	4,76	8,88	0,41	2,18	0,52	18,83	–
HoltWinters	4,92	10,68	1,04	3,27	0,39	20,08	2,9
CombSHD	4,36	9,65	0,5	2,73	0,46	20,09	–
Box–Jenkins	3,66	7,35	0,57	2,45	0,5	26,13	2,99
ForecastPro	3,13	6,5	0,41	2,47	0,5	20,52	–
SmartFcs	5,18	8,62	0,29	2,47	0,5	23,85	–
AutoANN	3,42	8,73	0,27	1,56	0,53	26,11	–
ClonAlg	5,20	7,65	0,41	1,83	0,14	16,26	2,44
ClonAlg (10%)	5,92	9,74	0,91	2,34	0,21	18,01	-
ClonAlg (10%), control	5,20	8,45	0,41	1,83	0,23	18,35	2,44
ClonAlg (15%)	6,42	7,31	0,88	2,98	0,21	19,01	2,56
ClonAlg (20%)	6,30	8,15	0,56	3,31	0,53	20,09	3,77

У ході обчислювального експерименту аналізувалося два варіанти розташування пропусків – у всьому ряду, та лише у контрольній частині (для випадку, коли була можливість перед використанням створити та навчити ШІС на непошкоджених даних). Симетричні середні абсолютні помилки прогнозування наведені у табл. 1, а для ряду Meteo приведена середня абсолютна помилка

(MAE, °C). Наявність пропусків в аналізованих рядах веде до збільшення помилки на 2-8 відсотків, але якщо пропущені значення розташовані лише у контрольній частині ряду (результати для ClonAlg (10%), control), помилка прогнозу збільшується лише на 1-3 відсотки.

Використання у рамках гібридної моделі комбінації різних методів веде до збільшення трудомісткості отримання прогнозу, що у поєднанні з вимогою обробки даних у режимі реального часу призводить до необхідності розпаралелювання гібридних підходів прогнозування. Проведено аналіз двох способів розпаралелювання, що відрізняються порядком обміну повідомленнями між обчислювальними вузлами, на основі врахування комунікаційної трудомісткості. До переваг першого способу розпаралелювання належать унікальність антитіл у популяціях на різних обчислювальних вузлах, кожен з яких має незалежну популяцію клітин пам'яті, та можливість контролю параметрів алгоритму для кожного вузла на кожній ітерації.

Для оцінки часу передачі даних між обчислювальними вузлами використана модель, запропонована Хокні. Об'ємом контрольних даних можна знехотити. Тоді тривалість виконання комунікаційних операцій у процесі навчання ШС для першого варіанта паралельного алгоритму може бути виражена як:

$$T_1 = (N_{cn} - 1)N_g \left(\alpha + \frac{V_{ag} + V_{ab}n_c n' + V_{ab}n_m}{\beta} \right), \quad (7)$$

де N_{cn} – кількість обчислювальних вузлів, N_g – кількість ітерацій алгоритму у ході навчання ШС, V_{ag} – обсяг даних для передачі одного антигену, V_{ab} – обсяг даних для передачі одного антитіла, n_c – кількість клонів одного антитіла, n' – кількість антитіл, відібраних для клонування, n_m – розмір популяції клітин пам'яті, α і β – параметри моделі Хокні (латентність та пропускна здатність мережі передачі даних).

Для реалізації обміну даними між обчислювальними вузлами використовується технологія MPI.NET, для прискорення розрахунків у межах популяцій окремих вузлів застосовується оптимізація керованого коду для багатоядерних процесорів з використанням технології TPL. Обчислювальний експеримент полягав у запуску паралельної версії алгоритму з однаковим критерієм зупинки та максимальним (для послідовної версії імунного алгоритму) розміром популяції на різній кількості доступних обчислювальних вузлів кластера. Використовувалося від одного до двадцяти обчислювальних вузлів (Intel Core 2 Quad Q6600 2,400 МГц). Як показано на рис. 2, використання розподілених обчислень дозволяє досягти прискорення, близького до лінійного. Оптимізація коду для багатоядерної архітектури дозволяє отримати прискорення близько 7%. Низький результат використання технології TPL обумовлений малою кількістю паралельних регіонів у імунному алгоритмі.



Рисунок 2 – Залежність часу навчання від кількості процесорів

Запропоновано новий показник χ для визначення застосовності імунного підходу, заснованого на моделі клонального відбору та CBR, для короткострокового прогнозування часових рядів. Запропонований показник визначається шляхом урахування кількості схожих антитіл та міри подібності їх варіантів прогнозів для вже відомих значень ряду та дозволяє на етапі обробки вихідних даних перед прогнозуванням визначити оптимальне значення довжини антитіл. Прогнозований часовий ряд можна характеризувати показником χ на основі комбінації частки відібраних антитіл та їх середньої точності:

$$\chi = \frac{n'}{\sum_{i=1}^{n-1} i} \cdot \frac{\sum_{j=0}^{n'} Aff_j}{n'} = \frac{\sum_{j=0}^{n'} Aff_j}{\sum_{i=1}^{n-1} i} \in (0,1], \quad (8)$$

де Aff_j – точність прогнозу (міра подібності кортежів прогнозованих значень).

Завдяки корекції коефіцієнтів відбору у ході обчислення показника χ можливе скорочення часу навчання ШС та зменшення розміру популяції антитіл. В процесі виконання оператора супресії аналогічно відбувається формування матриці афінності між антитілами, та ці дані можна використовувати не тільки для стиснення популяції, але й для розрахунку показника χ , що дозволяє у режимі реального часу коригувати параметри алгоритму.

У четвертому розділі розглянуто існуючі програмні засоби отримання прогнозу та виявлено їх основні недоліки, обґрунтовано необхідність створення інструментального середовища для аналізу імунного алгоритму та вирішення завдань короткострокового прогнозування часових рядів.

На базі розробленої гібридної моделі на основі ШС, яка застосовується для розв'язання задачі короткострокового прогнозування часових рядів, запропоновано структурну схему та виконано програмну реалізацію інструментального середовища. До його складу входять модуль обробки вихідних даних, який здійснює перетворення вихідних даних до формату, придатного для використання іншими модулями; модуль прогнозування часових рядів, що здійснює

побудову ШІС, виконання навчання та взаємодію з користувачем; модуль оцінки результатів. Для реалізації розподілених обчислень у даній роботі була використана технологія MPI.NET.

Проведена перевірка ефективності запропонованої гібридної моделі на основі ШІС шляхом розробки інтелектуальної системи підтримки прийняття рішень для прогнозування обсягу товарообороту підприємств «Буденергомаш» та РІК «СВ-Прилад». Дана система призначена для побудови короткострокового прогнозу часових рядів динаміки продажів та дозволяє підвищити ефективність господарської діяльності підприємства шляхом зниження обсягу статистичного врахування та раціонального використання грошових ресурсів. Величина помилки для прогнозу обсягу товарообігу «СВ-Прилад» для ряду з 22 значень та горизонту прогнозування, що дорівнює 2, склала 3,15%, що є хорошим показником в економічній практиці. Оцінка прогнозу за даними підприємства «Буденергомаш» виконувалася з урахуванням зовнішніх чинників та величина помилки склала 4,28% в процесі прогнозування на 4 значення вперед для часового ряду, що складається з 30 значень.

У додатку наведено акти про використання результатів дисертаційної роботи для вирішення практичних задач та у навчальному процесі.

ВИСНОВКИ

У дисертаційній роботі наведені результати дослідження, які у сукупності є рішенням актуальної наукової задачі – розробки гібридних моделей на основі штучних імунних систем для розв'язання задачі короткострокового прогнозування часових рядів, що має велике значення для підвищення точності прогнозу та забезпечення стійкості до викривлень вихідних даних.

Вирішення поставлених у роботі завдань дозволило отримати такі основні результати:

1. Проведено аналіз існуючих підходів прогнозування часових рядів на основі статистичних методів і методів на основі обчислювального інтелекту, який виявив недоліки, які пов'язані з вимогами до обсягу навчальної вибірки, ресурсомісткістю процесу навчання, складністю налаштування прогнозуючих моделей, чутливістю до викривлень, а також відсутністю можливості роботи у режимі реального часу. Актуальним є використання гібридних моделей, які компенсують недоліки одних підходів за допомогою врахування особливостей інших.

2. Запропонована гібридна модель, яка використовує поєднання методу виведення по прецедентах, моделі клонального відбору та найпростіших методів прогнозування, яка виконує фрагментацію вихідного часового ряду та підбір для кожного фрагмента свого методу прогнозування, що дозволяє скоротити обсяг навчальної вибірки та підвищити точність прогнозу. Дана модель дозволяє враховувати і виконувати відбір зовнішніх факторів, які мають найбільший вплив на прогнозовану величину, що дозволяє підвищити точність прогнозу, але веде до значного збільшення необхідної популяції антитіл.

3. Досліджено вплив викривлень часових рядів (аномальних викидів і пропущених значень) на точність прогнозування за допомогою розробленої гібридної моделі, та запропоновано метод, що дозволяє під час навчання гібридної моделі короткострокового прогнозування виявляти та враховувати викривлення в процесі побудови прогнозу.

4. Запропоновано показник для визначення застосовності імунного підходу короткострокового прогнозування часових рядів, заснованого на моделі клонального відбору та СBR, розрахунок якого дозволяє перед виконанням прогнозу виконувати налаштування параметрів гібридної моделі прогнозування, що дозволяє скоротити час навчання гібридної моделі та зменшити розмір популяції антитіл.

5. Розроблено інструментальне середовище для аналізу імунного алгоритму та вирішення задачі короткострокового прогнозування часових рядів, що дозволяє виконувати порівняльний аналіз для оцінки ефективності пропонованих підходів.

6. Проведено аналіз способів розпаралелювання імунного алгоритму прогнозування на основі гібридної моделі. Використання засобів розподілених обчислень дозволило добитися прискорення, близького до лінійного, тоді як оптимізація коду для багатоядерної архітектури дозволяє отримати прискорення близько 7%.

7. Запропоновані моделі та методи прогнозування стохастичних часових рядів були використані для вирішення завдань короткострокового прогнозування значень економічних показників підприємств «Буденергомаш» та «СВ-Прилад». Результати роботи також використані у навчальному процесі Харківського національного університету радіоелектроніки.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Кораблев, Н.М. Агентно-ориентированный подход на основе искусственных иммунных систем для решения задачи коммивояжера / Н.М. Кораблев, Г.С. Иващенко, М.В. Кушнарев // Бионика интеллекта: науч.-техн. журнал. – 2012. – № 2 (79). – С. 33-37.

2. Кораблев, Н.М. Применение модели клонального отбора, использующей вывод по прецедентам, для прогнозирования временных рядов / Н.М. Кораблев, Г.С. Иващенко // Бионика интеллекта: науч.-техн. журнал. – 2013. – № 1(80). – С. 108-111.

3. Кораблев, Н.М. Применение модели клонального отбора для прогнозирования временных рядов, имеющих пропущенные значения / Н.М. Кораблев, Г.С. Иващенко // Электротехнические и компьютерные системы: науч.-техн. журнал. – 2014. – № 13 (89). – С. 170-177. (Входит до міжнародних наукометричних баз Index Copernicus International, Ulrich's Periodicals Directory, Electronic Journals Library.)

4. Korablev, N.M. Parallel immune algorithm of short-term forecasting based on model of clonal selection / N.M. Korablev, G.S. Ivaschenko // Radio Electronics,

Computer Science, Control: scientific journal. – 2014. – № 2(31). – Pp. 73-78. ISSN 1607-3274. (Входит до міжнародних наукометричних баз Index Copernicus International, Ulrich's Periodicals Directory, INSPEC, DOI, DOAJ, EBSCO та інші.)

5. Кораблев, Н.М. Краткосрочное прогнозирование временных рядов, содержащих аномальные значения, при помощи моделей искусственных иммунных систем / Н.М. Кораблев, Г.С. Иващенко // Бионика интеллекта: науч.-техн. журнал. – 2015. – № 2(85). – С. 95-99.

6. Кораблев, Н.М. Применение искусственных иммунных сетей для прогнозирования временных рядов / Н.М. Кораблев, Г.С. Иващенко // Системы обработки информации: сборник научных трудов. – 2012. – № 9(107). – С. 42-45. (Входит до міжнародних наукометричних баз Index Copernicus International.)

7. Ivaschenko, G. S. Time series forecasting on the basis of the case-based reasoning using the models of artificial immune systems / G. S. Ivaschenko, N. M. Korablev // Системные технологии: сборник научных трудов. – 2014. – № 6(95). – С. 43-51. (Входит до міжнародних наукометричних баз Index Copernicus International.)

8. Кораблев, Н.М. Гибридный метод краткосрочного прогнозирования временных рядов на основе модели клонального отбора / Н.М. Кораблев, Г.С. Иващенко // Нейроинформатика: научно-техническая конференция с международным участием, 27-31 января 2014 г.: сборник научных трудов. – Москва, 2014. – Т. 1. – С. 79-89.

9. Кораблев, Н.М. Применение гибридного иммунного алгоритма для решения задачи коммивояжера / Н.М. Кораблев, Г.С. Иващенко // Информационные технологии в навигации и управлении: состояние и перспективы развития: международная научно-техническая конференция, 16-17 июля 2011 г.: тезисы докладов. – Киев, 2011. – С. 38.

10. Кораблев, Н.М. Использование агентно-ориентированного подхода для решения задачи коммивояжера / Н.М. Кораблев, Г.С. Иващенко, М.В. Кушнарев // Математическое и программное обеспечение информационных систем: международная научно-практическая конференция, 23-25 ноября 2011 г.: тезисы докладов. – Днепропетровск, 2011. – С. 137-138.

11. Кораблев, Н.М. Периодичность характеристик агентов искусственной иммунной системы, используемой для решения задачи коммивояжера / Н.М. Кораблев, Г.С. Иващенко // Радиоэлектроника и молодежь в XXI ст.: международный молодежный форум, 17-19 апреля 2012 г.: тезисы докладов. – Харьков, 2012. – Т. 10. – С. 37-38.

12. Иващенко, Г.С. Модель прогнозирования временных рядов на основе искусственных иммунных сетей / Г.С. Иващенко, Н.М. Кораблев // Информатика и компьютерные технологии: международная научно-техническая конференция, 27-29 сентября 2012 г.: тезисы докладов. – Донецк, 2012. – Т. 1. – С. 265-270.

13. Кораблев, Н.М. Модель клонального отбора для решения задачи краткосрочного прогнозирования / Н.М. Кораблев, Г.С. Иващенко // Математическое и программное обеспечение информационных систем: международная научно-практическая конференция, 21-23 ноября 2012 г.: тезисы докладов. – Днепропетровск, 2012. – С. 161-162.

14. Кораблев, Н.М. Применение искусственных иммунных систем с использованием вывода по прецедентам для решения задачи краткосрочного прогнозирования / Н.М. Кораблев, Г.С. Иващенко, О.Е. Власенко // Радиоэлектроника и молодежь в XXI ст.: международный молодежный форум, 22-24 апреля 2013 г.: тезисы докладов. – Харьков, 2013. – Т. 5. – С. 198-199.

15. Кораблев, Н.М. Применение модели клонального отбора для выявления аномальных выбросов во временных рядах / Н.М. Кораблев, Г.С. Иващенко // Современные направления развития информационно-коммуникационных технологий и средств управления: международная научно-техническая конференция, 11-12 апреля 2013 г.: тезисы докладов. – Киев, 2013. – С. 49.

16. Кораблев, Н.М. Применение разнородных антител в модели клонального отбора для решения задачи краткосрочного прогнозирования / Н.М. Кораблев, Г.С. Иващенко // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: международная научная конференция, 20-24 мая 2013 г.: тезисы докладов. – Херсон: ХНТУ, 2013. – С. 454-456.

17. Кораблев, Н.М. Выявление аномальных значений во временных рядах при помощи модели клонального отбора, использующей вывод по прецедентам / Н.М. Кораблев, Г.С. Иващенко // Сучасні проблеми правового, економічного та соціального розвитку держави: міжнародна науково-практична конференція, 22 листопада 2013 г.: тезисы докладов. – Харьков, 2013. – С. 421-423.

18. Кораблев, Н.М. Прогнозирование временных рядов с учетом внешних факторов при помощи модели клонального отбора / Н.М. Кораблев, Г.С. Иващенко // Математическое и программное обеспечение информационных систем: международная научно-практическая конференция, 20-22 ноября 2013 г.: тезисы докладов. – Днепропетровск, 2013. – С. 121-122.

19. Кораблев, Н.М. Восстановление пропущенных значений во временных рядах при помощи модели клонального отбора / Н.М. Кораблев, Г.С. Иващенко, Т.В. Гайдамака // Проблемы информатизации: международная научно-техническая конференция, 19-20 декабря 2013 г.: тезисы докладов. – Киев, 2013. – С. 51.

20. Иващенко, Г.С. Прогнозирование временных рядов путем восстановления функциональной зависимости при помощи модели клонального отбора / Г.С. Иващенко, Т.В. Гайдамака // Радиоэлектроника и молодежь в XXI ст.: международный молодежный форум, 22-24 апреля 2014 г.: тезисы докладов. – 2014. – Т. 5. – С. 170-171.

21. Кораблев, Н.М. Обнаружение аномальных выбросов во временных рядах при помощи модели положительного отбора / Н.М. Кораблев, Г.С. Иващенко, Т.В. Гайдамака // Проблемы информатизации: международная научно-техническая конференция, 12-13 ноября 2014 г.: тезисы докладов. – Киев, 2014. – С. 72.

22. Кораблев, Н.М. Учет внешних факторов в гибридной модели прогнозирования искаженных временных рядов / Н.М. Кораблев, Г.С. Иващенко // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: международная научная конференция, 28-31 мая 2014 г.: тезисы докладов. – Херсон: ХНТУ, 2014. – С. 288-290.

23. Кораблев, Н.М. Распараллеливание иммунного алгоритма краткосрочного прогнозирования на основе островной модели генетического алгоритма / Н.М. Кораблев, Г.С. Иващенко // Информационные управляющие системы и технологии: международная научно-практическая конференция, 23-24 сентября 2014.: тезисы докладов. – Одесса, 2014. – С. 210-212.

24. Кораблев, Н.М. Применение технологии TPL для распараллеливания иммунного алгоритма прогнозирования временных рядов / Н.М. Кораблев, Г.С. Иващенко, В.К. Федосов // Современные направления развития информационно-коммуникационных технологий и средств управления: международная научно-техническая конференция, 4-5 декабря 2014 г.: тезисы докладов. – Киев, 2014. – С. 38-39.

25. Иващенко, Г.С. Применение технологий распределенного программирования при реализации иммунных алгоритмов краткосрочного прогнозирования временных рядов / Г.С. Иващенко, В.К. Федосов // Радиоэлектроника и молодежь в XXI ст.: международный молодежный форум, 20-22 апреля 2015 р.: тезисы докладов. – Харьков, 2015. – Т. 5. – С. 150-151.

26. Кораблев, Н.М. Настройка параметров иммунного алгоритма прогнозирования временных рядов с помощью упрощенной модели искусственной иммунной сети / Н.М. Кораблев, Г.С. Иващенко, Р.Г. Еськов // Современные направления развития информационно-коммуникационных технологий и средств управления: международная научно-техническая конференция, 23-24 апреля 2015 г.: тезисы докладов. – Киев, 2015. – С. 28-29.

27. Кораблев, Н.М. Показатели применимости иммунного подхода для краткосрочного прогнозирования временных рядов / Н.М. Кораблев, Г.С. Иващенко // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: международная научная конференция, 25-28 мая 2015 г.: тезисы докладов. – Херсон: ХНТУ, 2015. – С. 288-290.

28. Кораблев, Н.М. Использование показателя Херста для определения прогнозируемости временного ряда с помощью модели клонального отбора / Н.М. Кораблев, Г.С. Иващенко // Информационные управляющие системы и технологии: международная научно-практическая конференция, 22-24 сентября 2015.: тезисы докладов. – Одесса, 2015. – С. 275-277.

29. Кораблев, Н.М. Применение мультиагентных систем для краткосрочного прогнозирования временных рядов / Н.М. Кораблев,

Г.С. Иващенко, Р.Г. Еськов // Проблемы информатизации: международная научно-техническая конференция, 12-13 ноября 2015 г.: тезисы докладов. – Киев, 2015. – С. 21-22.

АНОТАЦІЯ

Иващенко Г.С. Гібридні моделі короткострокового прогнозування часових рядів на основі штучних імунних систем. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Харків, 2016.

Дисертаційна робота присвячена розробці гібридних моделей на основі штучних імунних систем, які об'єднують переваги моделі клонального відбору, моделі штучної імунної мережі, методу виведення за прецедентами і існуючих методів прогнозування, що дозволяє підвищити як точність прогнозу, так і стійкість до викривлень у вихідних даних.

Розглянуто використання мультиантитіл у якості прецедентів в моделі штучної імунної мережі. Отримана модель відрізняється можливістю змінювати параметри роботи і структуру імунної мережі в процесі навчання ШІС.

Запропоновано гібридну модель, яка виконує фрагментацію вихідного часового ряду та підбір для кожного фрагмента свого методу прогнозування, що дозволяє скоротити обсяг навчальної вибірки і підвищити точність прогнозу. Для проведення сегментації дана модель використовує поєднання методу висновку за прецедентами та моделі клонального відбору.

Вперше запропоновано гібридну модель на основі моделі клонального відбору, яка використовується для прогнозування викривлених часових рядів, враховує вплив зовнішніх факторів, дозволяє підвищити точність прогнозування та забезпечує отримання прогнозу в режимі реального часу. Для забезпечення обробки даних в режимі реального часу в роботі розглянуто підходи до розпаралелювання гібридних моделей прогнозування на основі систем з загальною та індивідуальною пам'яттю.

Запропоновано показник для визначення застосовності імунного підходу, заснованого на моделі клонального відбору і СВР, який дозволяє перед вирішенням завдання короткострокового прогнозування виконувати налаштування параметрів гібридної моделі прогнозування, що дозволяє скоротити час навчання ШІС і зменшити розмір популяції антитіл.

Розроблено інструментальне середовище для аналізу імунного алгоритму і вирішення завдання короткострокового прогнозування часових рядів, що дозволяє виконувати порівняльний аналіз для оцінки ефективності запропонованих підходів.

Ключові слова: часовий ряд, моделі прогнозування, висновок за прецедентами, штучні імунні системи, клональний відбір, імунна мережа, паралельні обчислення.

АННОТАЦИЯ

Иващенко Г.С. Гибридные модели краткосрочного прогнозирования временных рядов на основе искусственных иммунных систем. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Харьков, 2016.

Диссертационная работа посвящена разработке гибридных моделей на основе искусственных иммунных систем, объединяющих преимущества модели клонального отбора, модели иммунной сети, метода вывода по прецедентам и простейших методов прогнозирования, что позволяет повысить как точность прогноза, так и устойчивость к искажениям исходных данных.

Рассмотрено использование мультиантител в качестве прецедентов в модели искусственной иммунной сети. Полученная модель отличается возможностью изменять параметры работы и структуру иммунной сети в процессе обучения.

Предложена гибридная модель, выполняющая фрагментацию исходного временного ряда и подбор для каждого фрагмента своего метода прогнозирования, что позволяет сократить объем обучающей выборки и повысить точность прогноза. Для проведения сегментации данная модель использует сочетание метода вывода по прецедентам и модели клонального отбора.

Впервые предложена гибридная модель на основе модели клонального отбора, используемая для прогнозирования искаженных временных рядов, которая учитывает влияние внешних факторов, позволяет повысить точность прогноза и обеспечивает получение прогноза в режиме реального времени. Для обеспечения обработки данных в режиме реального времени в работе рассмотрены подходы к распараллеливанию гибридных моделей прогнозирования на основе вычислительных систем с общей и индивидуальной памятью.

Предложен показатель для определения применимости иммунного подхода краткосрочного прогнозирования временных рядов, основанного на модели клонального отбора и CBR, который позволяет на этапе предпрогнозного анализа выполнять настройку параметров гибридной модели прогнозирования, что позволяет сократить время обучения ИИС и уменьшить размер популяции антител.

Разработана инструментальная среда для анализа иммунного алгоритма и решения задачи краткосрочного прогнозирования временных рядов, позволяющая выполнять сравнительный анализ для оценки эффективности предлагаемых подходов.

Ключевые слова: временной ряд, модели прогнозирования, вывод по прецедентам, искусственные иммунные системы, клональный отбор, иммунная сеть, параллельные вычисления.

ABSTRACT**Ivaschenko G. S. Hybrid models of short-term time series forecasting based on artificial immune systems. – Manuscript.**

This is for candidates degree in engineering science by specialty 05.13.23 – systems and means of artificial intelligence. – Kharkiv National University of Radio Electronics, Kharkiv, Ministry of Education and Science of Ukraine, 2016.

Dissertation work is devoted to development of hybrid models of short-term time series prediction based on a combination of the principles of artificial immune systems and existing forecasting methods that can improve the forecast accuracy and resistance to distortion of the original data.

The hybrid model of time series forecasting is proposed. This model performs fragmentation of the original time series and selection for each fragment of his prediction method, which reduces of learning sample and increases the accuracy of the forecast. For the time series segmentation the model uses a combination of the case based reasoning and the model of clonal selection. Considered using of multiantibodies as precedents in the artificial immune network model. The resulting model can change the parameters of the work and structure of the immune network in the learning process.

The hybrid model based on the model of clonal selection is proposed. This model is used to predict of the distorted time series, takes into account the impact of external factors and allows increasing the accuracy of the forecast, and provides a possibility of forecast in real time. To ensure data processing in real time, these approaches use parallel systems based on common and individual memory.

Criterion is proposed for determining the applicability of the immune approach of short-term time series forecasting based on the model of clonal selection and a CBR, which allows configuring parameters of hybrid model, that reduces the time for training of AIS and reduces the size of the antibody population.

It was developed instrumental environment for the analysis of the immune algorithm and solving the problem of short-term time series forecasting, which allows to perform a comparative analysis for evaluating effectiveness of the proposed approaches.

Keywords: time series, forecasting model, case based reasoning, artificial immune system, clonal selection, immune network, parallel computing.

