

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

**ФОМІЧОВ ОЛЕКСАНДР ОЛЕКСАНДРОВИЧ**

УДК 004.89

**МЕТОДИ ТА МОДЕЛІ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ  
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  
**Гороховатський Володимир Олексійович**,  
Харківський навчально-науковий інститут  
ДВНЗ «Університет банківської справи»,  
професор кафедри інформаційних технологій;

доктор технічних наук, професор  
**Литвиненко Володимир Іванович**,  
Херсонський національний технічний  
університет, завідувач кафедри інформатики  
та комп'ютерних наук.

Захист відбудеться « 08 » червня 2016 р. о 15-00 годині на засіданні 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 « 06 » травня 2016 р.

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

О.А. Винокурова

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** У сучасних умовах великим науково-практичним інтересом характеризуються задачі інтелектуальної обробки даних, до яких належать завдання класифікації і кластеризації даних, що можуть бути вирішені з використанням різних підходів і принципів. Існуючі методи класифікації залежно від постановленої задачі та обсягу апріорної інформації використовують різні моделі для групування даних і можуть бути реалізованими з контрольованим або неконтрольованим навчанням. Зараз для вирішення завдань інтелектуальної обробки інформації використовуються біологічні методи моделювання штучного інтелекту. До них належать штучні нейронні мережі, генетичні алгоритми, штучні імунні системи (ШИС), еволюційні обчислення тощо. Використання даних підходів в процесі вирішення задачі класифікації призводить до підвищення точності групування і швидкості класифікації та дозволяє автоматизувати процеси навчання моделей, що застосовуються для вирішення даного завдання. Існуючі методи класифікації, залежно від обраного способу навчання, поділяються на дві групи: методи класифікації з контрольованим навчанням, та методи кластеризації. Серед методів класифікації з контрольованим навчанням найбільш поширеними є метод найближчих сусідів, графові методи та машина опорних векторів. Основними їх недоліками є висока чутливість до характеристик навчальної вибірки, кількості найближчих сусідів, а для машини опорних векторів – складність модифікації. Існуючі імунні методи класифікації з контрольованим навчанням характеризуються низькою швидкістю або точністю групування об'єктів, залежно від обраної імунної моделі. Серед методів кластеризації найбільш поширеними є метод k-means та метод мінімального покриваючого дерева. Основними їх недоліками є велика чутливість до обрання початкових центрів кластерів і необхідність апріорного знання про кількість кластерів. Існуючі імунні методи кластеризації здебільшого характеризуються низькою швидкістю, що ускладнює їх застосування в ході вирішення практичних задач, але мають хорошу точність групування об'єктів.

Різними аспектами вирішення проблеми класифікації на основі систем моделювання штучного інтелекту присвячені роботи П.І. Бідюка, В.Н. Вапніка, В.В. В'югіна, Д.А. Вятченіна, В.О. Гороховатського, Д. Дасгупти, Р.О. Дуди, Л.Н. де Кастро, В.І. Литвиненка, Дж.В. де Олівейри, В. Педрича, Д.Г. Сторка, Дж. Тімміса, С. Форреста, А.О. Фефелова, П.Е. Харта та інших вчених.

Завдання навчання імунних моделей вирішують за допомогою різних підходів, серед яких найбільш поширеними є моделі клонального відбору, імунної мережі та негативного відбору. Основні труднощі застосування даних моделей пов'язані з їх низькою швидкістю та виникненням надлишкових обчислень на етапах мутації і редагування популяції імунних об'єктів. У зв'язку з цим, розробка методів і моделей класифікації, що функціонують на основі штучних імунних систем, або за допомогою формування імунних гібридів, є актуальною задачею 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 і науки України. Автор був одним з виконавців робіт за даними темами.

**Мета і задачі дослідження.** Метою дисертаційної роботи є розробка, дослідження та удосконалення методів і моделей класифікації об'єктів на основі штучних імунних систем з різними способами навчання, які дозволяють підвищити швидкість алгоритмів класифікації та кластеризації даних із забезпеченням високої точності угруповання.

Для досягнення цієї мети необхідно вирішити такі завдання:

- аналіз методів і моделей класифікації та кластеризації об'єктів;
- розробка методів і моделей класифікації об'єктів на основі імунного підходу з різними способами навчання;
- розробка моделі штучної імунної мережі з використанням стимулюючих антитіл для підвищення швидкості імунного навчання та визначення початкових центрів кластерів;
- розробка методу відбору клонів і методу їх мутації для підвищення швидкості імунного навчання моделей класифікації об'єктів;
- розробка гібридних моделей класифікації об'єктів на основі методів k-means та kNN з використанням імунного підходу;
- експериментальні дослідження розроблених методів і моделей та вирішення з їх допомогою практичних завдань.

*Об'єктом дослідження* є процеси класифікації та кластеризації об'єктів на основі штучних імунних систем.

*Предметом дослідження* є методи та моделі класифікації об'єктів на основі штучних імунних систем з різними способами навчання.

*Методи дослідження.* Для розв'язання поставлених задач були використані: теорії ШС, нечіткої логіки та теорія множин, що дозволило синтезувати нові методи та моделі класифікації об'єктів; теорії оптимізації, що дозволило розробити методи навчання та адаптації запропонованих моделей, а також апарат математичної статистики, що дозволив виконувати систематизацію та використання отриманих у роботі даних для наукових і практичних висновків.

**Наукова новизна результатів дисертаційної роботи.** В рамках вирішення задач дисертаційного дослідження отримано такі нові наукові результати:

1. Вперше запропоновано використання стимулюючих антитіл у моделі

штучної імунної мережі, які характеризуються високими значеннями афінностей до антигенів та презентують групи найближчих до них антитіл, що не підлягають клонуванню, для прискорення їх мутації та формування кластерів, що призводить до зростання швидкодії алгоритмів класифікації та кластеризації даних без зменшення точності групування.

2. Набув подальшого розвитку метод відбору клонів, який, на відміну від існуючих, використовує конкурентно-цільовий відбір, що призводить до зменшення кількості антигенів в процесі обчислення афінностей імунних об'єктів після мутації на етапі редагування популяції антитіл, наслідком чого є зростання швидкості класифікації та кластеризації даних.

3. Набув подальшого розвитку метод мутації клонів, який, на відміну від існуючих, використовує афінність далекого предка для клонів, що підлягають мутації, яка є мінімальним порогом в ході обчислення коефіцієнта мутації, що призводить до підвищення швидкості й точності класифікації об'єктів.

4. Набули подальшого розвитку гібридні моделі класифікації об'єктів на основі методів k-means та kNN, в яких, на відміну від існуючих, використовується імунний підхід для визначення кількості кластерів (метод k-means) та кількості найближчих сусідів (метод kNN), що без зменшення точності призводить до зростання швидкості групування даних.

**Практичне значення результатів дисертаційної роботи.** Отримані теоретичні результати були досліджені експериментально на тестових і реальних даних, де показали свою перевагу над відомими імунними методами класифікації та кластеризації даних. Результати дисертаційної роботи, які доведені до рівня програмних засобів, дозволяють у різних аспектах підвищити точність і швидкість класифікації об'єктів та вирішувати задачі автоматичної класифікації.

Експериментальні дослідження, які проведені для оцінки працездатності та ефективності розроблених методів і моделей класифікації об'єктів, підтверджують основні положення, що виносяться на захист. Окремі положення, висновки та рекомендації дисертаційної роботи використано для вирішення задачі класифікації результатів тестування власників вогнепальної зброї НВП «Страж-Інвест» (акт від 15.12.2011), а також для визначення сум страхових платежів за характеристиками транспортних засобів ПАТ «Київський страховий дім» (акт від 15.08.2014). Результати дисертаційної роботи також були використані в навчальному процесі на кафедрі електронних обчислювальних машин в курсах «Інтелектуальні комп'ютерні системи», «Системи штучного інтелекту», «Методи та засоби обчислювального інтелекту» та «Імунні обчислювальні системи» Харківського національного університету радіоелектроніки (акт від 23.04.2015).

**Особистий внесок здобувача.** Всі основні результати, що виносяться на захист, отримано автором самостійно. Внесок автора в публікаціях, написаних у співавторстві, такий: в [1] – запропонована модель нечіткої кластеризації на основі ШС; в [2] – проведено дослідження імунних операторів у моделях класифікації; в [3] – запропоновано використання імунної моделі класифікації результатів тестування студентів; в [4] – запропоновані гібридні імунні моделі кластеризації на основі методу k-means; в [5] – запропоновано модифікацію імунного ме-

тоду RLAIS для автоматичної класифікації об'єктів; в [6] – запропоновано імунну модель автоматичної класифікації з використанням стимулюючих антитіл; в [7] – запропоновано використання імунного оператора конкурентно-цільового відбору клонів; в [8] – запропоновано модель класифікації на основі ШПС; в [9] – запропоновано модель нечіткої класифікації на основі ШПС; в [10] – проведено аналіз збіжності імунних алгоритмів; в [11] – запропонована гібридна модель класифікації на основі методу найближчих сусідів та ШПС; в [12] – запропоновані підходи підвищення ефективності імунних методів класифікації на основі моделі клонального відбору; в [13] – запропоновано модифіковану модель клонального відбору; в [14] – запропоновано імунну модель контрольованої класифікації; в [15] – розглянуто використання імунних методів для кластеризації; в [16] – розроблено метод кластеризації з використанням ШПС; в [17] – запропоновано імунну модель розпізнавання; в [18] – запропоновано імунну модель кластеризації на основі клонального відбору; в [19] – запропоновано імунний метод автоматичної класифікації об'єктів; в [20] – запропоновано гібридну імунну модель кластеризації на основі методу k-means; в [21] – запропонована модифікація функції визначення афінності між імунними об'єктами; в [22] – запропоновано модифіковану функцію визначення афінності на основі Манхеттенської відстані в моделях клонального відбору; в [23] – запропоновано використання стимулюючих антитіл в імунних методах класифікації; [24] – запропоновано модифікацію методу визначення коефіцієнта мутації в імунних методах класифікації об'єктів; в [25] – запропонована адаптивна імунна модель автоматичної класифікації об'єктів.

**Апробація результатів дисертації.** Основні положення та результати дисертаційної роботи доповідалися й обговорювалися на: 4-й міжнародній науково-практичній конференції «Математичне та програмне забезпечення інтелектуальних систем» (Дніпропетровськ, 2008 р.); 10-й і 11-й міжнародних науково-технічних конференціях «Системний аналіз та інформаційні технології» (Київ, 2008-2009 рр.); 9-й міжнародній конференції «Інтелектуальний аналіз інформації» (Київ, 2009 р.); 9-й міжнародній науково-технічній конференції «Проблеми інформатики і моделювання» (Харків, 2009 р.); 1-й науково-технічній конференції «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління» (Харків, 2010 р.); 2-й міжнародній науково-технічній конференції «Інформаційні технології в навігації і управлінні: стан та перспективи розвитку» (Київ, 2011 р.); 1-й і 2-й міжнародних науково-технічних конференціях «Проблеми інформатизації» (Харків, 2013-2014 рр.); 4-й і 5-й міжнародних науково-технічних конференціях «Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління» (Харків, 2014-2015 рр.); матеріалах всеукраїнської науково-практичної конференції «Інформаційні та моделюючі технології» (Черкаси, 2015 р.).

**Публікації.** За темою дисертаційної роботи опубліковано 25 наукових праць, з них: 9 статей у фахових періодичних виданнях України з технічних наук (серед них 4 видання, що входять до міжнародних наукометричних баз), 1 стаття у закордонному виданні (Польща); 15 публікацій у збірниках праць і тез міжнародних науково-технічних конференцій і семінарів.

**Структура та обсяг дисертаційної роботи.** Дисертація складається із вступу, чотирьох розділів, висновків, списку використаних літературних джерел з 154 найменувань і 1 додатка. Робота містить 10 рисунків, 18 таблиць. Загальний обсяг роботи складає 168 сторінок, з них 148 – основного тексту.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність обраної теми дисертації, сформульовано мету та задачі дослідження, визначено об'єкт, предмет і методи досліджень, охарактеризовано наукову новизну та практичне значення отриманих результатів, а також особистий внесок автора в роботах, виконаних у співавторстві, наведено відомості про апробацію результатів дисертації та кількість публікацій за темою дисертаційної роботи.

У **першому розділі** проведено аналіз проблеми класифікації об'єктів, який вказав на необхідність використання нових підходів до групування, які мають базуватися на основних принципах функціонування ШС. Проведений аналіз дозволив виділити основні способи оцінки ефективності методів і моделей класифікації об'єктів, які використовують різні підходи та принципи групування даних. Крім того, в процесі аналізу було виділено основні переваги та недоліки існуючих методів класифікації та кластеризації даних. Встановлено, що існуючим методам автоматичної класифікації, які не використовують біологічні принципи організації, потрібно більше інформації про об'єкти, що класифікуються, початкові класи та кількість кластерів, ніж методам, які функціонують на основі біологічних принципів, таких як штучні нейронні мережі, генетичні та еволюційні алгоритми, ШС. При цьому, серед виділених моделей штучного інтелекту ШС є найменш дослідженими, та мають найбільший потенціал.

Розглянуто методи штучного інтелекту, які функціонують на основі ШС та використовуються для класифікації об'єктів. Основними перевагами найбільш поширених імунних моделей (моделі клонального відбору, штучної імунної мережі та RLAIIS) є здатність до автоматичної класифікації, яка поєднує у собі основні властивості класифікації із контрольованим та неконтрольованим навчанням, залежно від особливостей початкових даних. Основними недоліками виділених імунних моделей є низька швидкодія та залежність точності класифікації від швидкодії. Аналіз методів і моделей класифікації показав, що для підвищення ефективності доцільною є розробка методів і моделей на основі поєднання різних технологій, в яких взаємно компенсуються їх недоліки та об'єднуються переваги. На основі проведеного аналізу сформульовано загальну задачу дослідження, визначено сукупність перспективних напрямків класифікації та кластеризації об'єктів на основі ШС і сформульовано задачі дисертаційної роботи.

У **другому розділі** запропоновано загальні імунні моделі класифікації та кластеризації об'єктів, які використовують переваги моделей клонального відбору та штучної імунної мережі, а також запропонована модель автоматичної класифікації, яка може не тільки розділяти множину початкових об'єктів, які відображаються популяцією антитіл, серед початкових наборів класів, а й формувати

нові кластери з об'єктів, які не можуть бути віднесені до жодного з класів.

Автоматична класифікація виконується в декілька етапів. На першому етапі проводиться аналіз можливості класифікації об'єктів з контрольованим імунним навчанням. При цьому як об'єкти навчальні вибірки репрезентують початкові класи. Якщо навчальної вибірки немає, модель функціонує як модель класифікації. В іншому випадку на другому етапі починається уточнення можливості класифікації з контрольованим навчанням усієї множини початкових об'єктів. Якщо це можливо, модель функціонує як модель класифікації з контрольованим навчанням. В іншому випадку, модель розділяє початкові об'єкти на дві множини: об'єкти, що класифікуються, та об'єкти, для яких потрібно проведення класифікації. На третьому етапі для об'єктів, які можна класифікувати, встановлюється необхідність проведення імунного навчання, на якому відбувається розподілення об'єктів, що класифікуються, серед початкових класів, а для інших об'єктів відбувається формування кластерів. На останньому етапі відбувається уточнення меж початкових класів та сформованих кластерів для остаточного групування об'єктів, які не були класифіковані на етапі імунного навчання.

Для підвищення швидкості класифікації без втрат у точності групування об'єктів було запропоновано використання набору імунних операторів – структурних елементів, що виконують один з етапів роботи імунної системи (наприклад, оператор клонування, чи оператор мутації). Серед запропонованих імунних операторів і моделей можна виділити оператор конкурентно-цільового відбору клонів, модель стимулюючих антитіл і цільових об'єктів та оператор мутації клонів на основі афінності далекого предка.

Основною особливістю конкурентно-цільового відбору клонів на етапі редагування популяції антитіл та їх клонів після проведення мутації є зменшення кількості антигенів, які презентуються клонам для визначення афінностей між ними. Це досягається за допомогою використання цільових об'єктів, які обираються з початкової множини антигенів антитілами, що підлягають класифікації, та використовуються усіма їх клонами на етапі імунного навчання. Таким чином, при цільовому відборі відбувається визначення афінностей між клонами та їх цільовими антигенами, а не з усією множиною антигенів навчальної вибірки. Слід зазначити, що кількість цільових об'єктів відрізняється для кожного клонованого антитіла та його клонів. При цьому на етапі відбору клони конкурують між собою за можливість подальшої взаємодії з цільовими антигенами згідно зі значенням афінності між ними. Таким чином, процес відбору клонів полягає у поступовому зменшенні кількості клонів кожного клонованого антитіла до значення, яке залежить від афінності між клонами та їх цільовими антигенами.

Основною особливістю моделі стимулюючих антитіл та цільових об'єктів є реалізація можливості взаємної стимуляції антитіл у процесі функціонування імунної системи. Слід зазначити, що така поведінка антитіл може бути реалізована тільки для моделі штучної імунної мережі, тому що модель клонального відбору повністю виключає будь-яку взаємодію між антитілами в процесі функціонування імунної системи. Основна ідея стимулюючих антитіл полягає в тому, що антитіла, які вступили у взаємодію із антигенами, сигналізують про це іншим

антитілам, які характеризуються невеликим показником афінності до цього антигену, та стимулюють їх клонування та мутацію для детальнішого виявлення особливостей антигену. Слід зазначити, що цей принцип, притаманний імунним мережам, не було реалізовано у жодному з імунних методів. Таким чином, стимулюючим є антитіло, яке може взаємодіяти не тільки з антигенами навчальної вибірки, але ще стимулює антитіла, що класифікуються, на генерацію нових поколінь імунних об'єктів. Стимулюючі антитіла характеризуються максимальними показниками афінності з антигенами навчальної вибірки та формують області стимуляції з інших антитіл. При цьому вони для таких антитіл можуть бути обрані як цільові об'єкти. Таким чином, при виділенні цільових об'єктів з множини цільових антитіл використовується рівняння:

$$\begin{aligned} \forall ab_j \in AB : \exists ab_{si} \in AB_s, aff(ab_j, ab_{si}) \geq \frac{1}{\lambda} NAT_{AG} \rightarrow \\ \rightarrow ab_{si} \in Ts_j, \mu(Ts_j) \in [s; +\infty), \end{aligned} \quad (1)$$

де  $ab_j$  –  $j$ -те антитіло, що належить множині антитіл  $AB$ , для якого виділяються цільові об'єкти з множини стимулюючих антитіл  $AB_s$ ;  $ab_{si}$  –  $i$ -е стимулююче антитіло, для якого виконується перевірка включення до множини цільових об'єктів  $Ts_j$  для антитіла  $ab_j$ ;  $aff(ab_j, ab_{si})$  – афінність між антитілом, що класифікується  $ab_j$ , та стимулюючим антитілом  $ab_{si}$ ;  $NAT_{AG}$  – середня порогова афінність між антигенами навчальної вибірки;  $\lambda$  – коефіцієнт зростання класів, який є вхідним параметром;  $\mu$  – розмір множини класів цільових об'єктів для антитіла;  $s$  – мінімальна кількість цільових об'єктів з множини стимулюючих антитіл, яка є вхідним параметром. При цьому визначення афінності між двома імунними об'єктами виконується так:

$$aff(ab_j, ab_{si}) = (1 + d_{ji})^{-1}, \quad (2)$$

де  $d_{ji}$  – відстань між  $j$ -м та  $i$ -м імунними об'єктами.

Слід зазначити, що під час вирішення задачі класифікації використання принципу мережної взаємодії та методу стимулюючих антитіл дозволяє проводити кластеризацію антитіл  $AB$ , які не можуть бути віднесені до жодного початкового класу, що представляється антигенами навчальної вибірки, а також прискорити процес класифікації антитіл  $AB$ , які можуть бути віднесені до одного з початкових класів у процесі імунного навчання. Таким чином, задача кластеризації антитіл вирішується через відновлення характеристик стимулюючих антитіл, які є початковими центрами кластерів для об'єктів, що не можуть бути віднесені до жодного з початкових класів, або за їх відсутності завдяки клонуванню та мутації у процесі імунного навчання. При цьому етап виділення цільових об'єктів необхідно проводити для зменшення кількості обчислювальних операцій на етапі імунного навчання. Визначення кластера для кожного антитіла з множини  $AB$  в процесі імунного навчання завершується з виконанням умови:

$$\forall ab_l \in K_i : aff(ab_l, ab_{si}) = 1 - aff(ab_l, ab_{si}) > \arg \min(NAT_{AG_j}), \quad (3)$$

де  $ab_l$  – антитіло, що належить множині  $AB$ , для якого визначається належність до кластера  $K_i$ ;  $ab_{si}$  – стимулююче антитіло, яке є центром кластера  $K_i$ ;  $aff(ab_l, ab_{si}) = 1$  – умова досягнення специфічності між антитілом  $ab_l$  і центром кластера  $ab_{si}$ . Основною умовою кластеризації антитіла з множини  $AB$  є досягнення рівня специфічності з центром одного зі сформованих кластерів. Для прискорення процесу угруповання об'єктів використовується додаткова умова належності об'єкта кластеру, якщо його афінність з центром цього кластера перевершує мінімальне порогове значення  $NAT_{AG_j}$ , яке визначається для антигенів одного з вихідних класів на початковому етапі автоматичної класифікації.

Таким чином, використання стимулюючих антитіл в поєднанні з міжпопулярційною супресією, дозволяє проведення кластеризації досліджуваних об'єктів виключно імунними методами без використання інших підходів угруповання даних. Крім того, застосування стимулюючих антитіл не суперечить основним принципам організації функціонування моделі штучної імунної мережі, що використовується в різних імунних моделях. Завдяки цьому в процесі імунного навчання відбувається не тільки розподіл об'єктів між початковими класами, але й виділення нових кластерів без проведення додаткових обчислень.

Основною особливістю принципу використання афінності далекого предка в ході обчислення коефіцієнта мутації є спосіб встановлення його мінімальної межі, що призводить до підвищення швидкості імунного навчання. Суть даного підходу полягає в підвищенні нижнього порога в діапазоні визначення коефіцієнта мутації для збільшення кількості змін параметрів мутуючих клонів. Як нижня межа пропонується використання афінності антитіла-предка з попереднього покоління, який породив у процесі імунного навчання антитіло, для клонів якого визначається коефіцієнт мутації. Даний підхід може використовуватися для пропорційної та обернено пропорційної мутації клонів. Завдяки такій зміні нижньої межі діапазону відбувається зниження ймовірності вибору зневажливо малого значення коефіцієнта мутації. Наслідком цього є підвищення швидкості імунного навчання. В процесі виконання обернено пропорційної мутації використання афінності далекого предка з визначенням коефіцієнта мутації відбувається інакше, для чого пропонується такий вираз:

$$\mu = 1 - \text{rand}[aff(ab_p, AG); aff(cl, AG)], \quad (4)$$

де  $aff(cl, AG)$  – афінність клона, що підлягає мутації, до його цільових антигенів;  $aff(ab_p, AG)$  – афінність далекого предка клона до цільових антигенів.

Крім використання афінності далекого предка для підвищення швидкості імунного навчання може використовуватися інший підхід, який полягає у збільшенні значення афінності під час визначення коефіцієнта мутації. Це обумовлюється нерівномірним зростанням значення афінності на проміжку  $[0,5; 1]$  з вико-

ристанням цілих чисел для подання параметрів об'єктів, що класифікуються. Наслідком цього є збільшення кількості популяцій імунних об'єктів, необхідних для проведення імунного навчання, що негативно позначається на швидкості класифікації. Для усунення вказаного недоліку пропонується збільшення афінності, яка використовується як поріг в ході визначення коефіцієнта мутації. Відповідно до цього афінність клону визначається в такий спосіб:

$$aff(cl, AG) = \begin{cases} 2 \cdot aff(cl, AG), & \text{при } aff(cl, AG) < 0.5, \\ aff(cl, AG), & \text{при } aff(cl, AG) \geq 0.5. \end{cases} \quad (5)$$

Використання даного підходу дозволяє підвищити кількість змін параметрів клонів, оскільки при цьому підвищується верхня межа коефіцієнта мутації. Наслідком цього є зростання афінності між перетвореними в процесі мутації клонами і антигенами навчальної вибірки, або досягнення стану специфічності між ними. Таким чином, зростання афінності в ході визначення коефіцієнта мутації призводить до скорочення кількості популяцій імунних об'єктів, необхідних для проведення імунного навчання. Даний підхід може застосовуватися в поєднанні з використанням афінності далекого предка як нижньої межі діапазону під час визначення коефіцієнта мутації. В такому випадку зростанню підлягає не тільки афінність клону, але й афінність його далекого предка, незалежно від виду оператора мутації. Найбільше прискорення процесу імунного навчання досягається в ході використання даного оператора мутації для зміни параметрів клонів.

**У третьому розділі** запропоновано імунні методи класифікації, кластеризації та автоматичної класифікації, які використовують різні імунні моделі (модель клонального відбору, модель штучної імунної мережі та модель RLAIIS), а також гібридні методи класифікації, що функціонують на основі імунного підходу та моделей kNN і k-means.

Основною особливістю запропонованого методу автоматичної класифікації aiNETma є зміна поведінки стимулюючих антитіл, які взаємодіють з різними множинами об'єктів, що класифікуються в процесі імунного навчання. Відповідно до цього, під час роботи з об'єктами, які не можуть бути віднесені до жодного з початкових класів у процесі навчання, використовуються стимулюючі антитіла як початкові центри нових кластерів. У такому разі з цих стимулюючих антитіл формується множина цільових об'єктів, що підлягають кластеризації в процесі імунного навчання. Для антитіл, які можуть бути віднесені до одного з початкових класів, використовуються стимулюючі антитіла для формування множини цільових об'єктів. Таким чином, метод aiNETma використовує два набори стимулюючих антитіл для забезпечення класифікації та кластеризації вихідної множини об'єктів. При цьому в процесі імунного навчання бере участь вся множина початкових об'єктів, що класифікуються, і навчання не проводиться окремо для множини антитіл, які класифікуються, і антитіл, які не можуть бути віднесені до жодного з початкових класів і підлягають кластеризації.

Відповідно до цього, робота запропонованого методу автоматичної класифікації aiNETma на рівні імунних операторів подається виразом:

$$\begin{aligned}
 aiNETma(AB, AG, T, \lambda, n) = & \left[ \begin{array}{l}
 NatCalculation(AG, AB) \rightarrow \\
 \rightarrow Presentation(AB, AG) \rightarrow \\
 \rightarrow ClassDetection(AB, \lambda) \rightarrow \\
 \rightarrow StimSelection(AB'') \rightarrow \\
 \rightarrow ClusterDetection(AG, AB'', AB''', \lambda) \rightarrow \\
 \rightarrow TrgSelection(AG, AB'', AB''') \rightarrow \\
 \rightarrow SpacesSelection(AB''', AB^s)
 \end{array} \right]^{PRP} \rightarrow \\
 \rightarrow & \left[ \begin{array}{l}
 \left. \begin{array}{l}
 \begin{array}{l}
 Cloning(ab_i, CL_i, n) \rightarrow \\
 \rightarrow Mutation(CL_i) \rightarrow \\
 \rightarrow Presentation(CL_i, Trg_i) \rightarrow \\
 \rightarrow Supression(ab_i, CL_i) \rightarrow \\
 \rightarrow ClassDetection(ab_i, Trg_i, \lambda)
 \end{array} \\
 \left. \begin{array}{l}
 \begin{array}{l}
 ClassDetection(AB'', \lambda) \rightarrow \\
 \rightarrow ClusterDetection(AB''', AG) \rightarrow \\
 \rightarrow Reconstruction(AB, AB')
 \end{array}
 \end{array} \right]^{CLs}
 \end{array} \right]^{LRN} \rightarrow
 \end{aligned} \tag{6}
 \end{math>$$

де  $AB$  – множина об’єктів, що класифікуються, представлених популяцією анти-тіл;  $AG$  – набір антигенів навчальної вибірки, що представляє початкові класи;  $T$  – гранична кількість популяцій анти-тіл, які формуються в процесі імунного навчання;  $\lambda$  – коефіцієнт зростання кластерів;  $n$  – мінімальна кількість об’єктів, які створюються в процесі клонування анти-тіл, що класифікуються;  $PRP$  – підготовчий етап;  $LRN$  – етап імунного навчання;  $CLS$  – етап завершення класифікації;  $NatCalculation(AG, AB)$  – оператор визначення порогової афінності  $NAT$  для популяції антигенів в разі використання навчальної вибірки, або для популяції анти-тіл, в разі її відсутності;  $Presentation(AB, AG)$  – оператор визначення анти-тіл, що класифікуються за допомогою популяції антигенів;  $ClassDetection(AB, \lambda)$  – оператор відбору та класифікації анти-тіл без імунного навчання;  $StimSelection(AB'')$  – оператор визначення стимулюючих анти-тіл;  $ClusterDetection(AG, AB'', AB''', \lambda)$  – оператор перевірки можливості кластеризації анти-тіл без проведення імунного навчання;  $TrgSelection(AG, AB'', AB''')$  – оператор визначення цільових антигенів для анти-тіл, які класифікуються, і анти-тіл, які формують нові кластери в процесі імунного навчання;  $SpacesSelection(AB''', AB^s)$  – оператор визначення областей стимуляції для стимулюючих анти-тіл  $AB^s$ ;  $Cloning(ab_i, CL_i, n)$  – оператор пропорційного клонування анти-тіла з обмеженням мінімальної кількості клонів;  $Mutation(CL_i)$  – оператор обернено пропорційної мутації клонів на основі афінності далекого предка;  $Presentation(CL_i, Trg_i)$  – оператор визначення афінностей між клонами та їх цільовими об’єктами;  $Supression(ab_i, CL_i)$  – оператор супресії анти-

тіл та їх клонів;  $ClassDetection(ab_i, Trg_i, \lambda)$  – оператор визначення класу для антитіла;  $ClassDetection(AB'', \lambda)$  – оператор визначення класів для неспецифічних антитіл множини  $AB''$ ;  $ClusterDetection(AB''', AG)$  – оператор визначення кластерів для неспецифічних антитіл множини  $AB'''$ ;  $Reconstruction(AB, AB')$  – оператор відновлення вихідних об'єктів і визначення їх приналежності класам і кластерам.

Для підвищення характеристик імунних методів класифікації об'єктів пропонується їх гібридизація. Для цього в роботу імунної моделі додаються особливості роботи не імунних методів класифікації та кластеризації даних. Це може бути виражено у додаванні імунних операторів до не імунних моделей класифікації, або навпаки. Оскільки метод CLONALG реалізує модель клонального відбору, яка виключає можливості мережної взаємодії антитіл у процесі імунного навчання, цей алгоритм майже не використовується в ході створення гібридних імунних методів класифікації. Метод aiNET, реалізований на основі моделі штучної імунної мережі, найбільш придатний для гібридизації порівняно з іншими імунними методами, тому що він може поєднувати різні особливості функціонування імунних моделей на етапах редагування популяції антитіл, клонування та мутації. Тому для підвищення ефективності класифікації об'єктів шляхом створення імунних гібридів за основу був обраний модифікований метод aiNETm.

На основі даного методу було запропоновано кілька гібридних методів для класифікації об'єктів: 1) метод з контрольованим навчанням aiNETmkn, який поєднує ідею найближчих сусідів методу kNN з принципами мережної взаємодії; 2) метод кластеризації даних aiNETmkm, який формує кластери за допомогою ідеї центроїдів, яка реалізована у методі k-means; 3) метод автоматичної класифікації, що поєднує в собі основні принципи функціонування імунної мережі з елементами нечіткої логіки.

На першому етапі функціонування моделі aiNETmkn завдяки використанню порогової афінності між антигенами відбувається визначення стимулюючих антитіл, які є найближчими сусідами для об'єктів, що характеризуються низькими показниками афінності до антигенів навчальної вибірки. Слід зазначити, що більшість стимулюючих антитіл класифікуються за допомогою імунного навчання, проте, для деяких антитіл передбачається можливість класифікації без навчання в разі, якщо вони характеризуються максимальними значеннями афінності до антигенів. Такі антитіла не підлягають дії імунних операторів клонування, мутації та відбору. Класифікація антитіл, що входять до області стимуляції, проходить після завершення імунного навчання, яке відбувається тільки для не класифікованих стимулюючих антитіл. При цьому класифіковані стимулюючі антитіла утворюють множину найближчих сусідів для антитіл, які знаходилися в областях стимуляції та не брали участі у процесі імунного навчання. Відповідно до цього, для кожного антитіла, що знаходиться в області стимуляції, на початковому етапі роботи моделі aiNETmkn визначається  $k$  найближчих стимулюючих антитіл, на основі яких обирається належність цього антитіла тому чи іншому початковому класу. При цьому значення афінності між антитілами з областей стимуляції та їх найближчими стимулюючими антитілами використовуються як вагові коефіцієнти. Викорис-

тання такого підходу призводить до збільшення швидкості класифікації порівняно з іншими імунними методами, та збільшенню точності групування, порівняно з методом kNN.

Функціонування гібридної моделі aiNETmkm починається з виділення стимулюючих антитіл як центрів кластерів та визначення областей стимуляції за допомогою значення порогової афінності, яке обчислюється для множини антитіл. В ході кластеризації популяція навчальних антигенів відсутня і формування кластерів відбувається виключно завдяки взаємодії між антитілами в процесі імунного навчання. Таким чином, в ході використання стимулюючих антитіл як центрів на початку імунного навчання формуються області стимуляції, що перетинаються та містять початкові антитіла. У процесі імунного навчання антитіла, які знаходяться в областях стимуляції, вступають у взаємодію зі стимулюючими об'єктами як з антигенами. Слід зазначити, що стимулюючі антитіла не підлягають дії імунних операторів клонування та мутації і змінюються тільки після перевизначення центрів для кожного кластера. Процес імунного навчання в моделі aiNETmkm завершується в разі досягнення стану специфічності між усіма антитілами, які перебувають в областях стимуляції, і виділеними центрами кластерів, або з досягненням максимальної кількості популяцій імунних об'єктів, які формуються в процесі навчання шляхом клонування антитіл. Використання такого підходу призводить до підвищення швидкості кластеризації, порівняно з методом k-means без втрати точності групування об'єктів.

Використання елементів нечіткої логіки в імунних гібридних методах дозволяє обчислити належність антитіл до декількох класів чи кластерів у процесі імунного навчання. При цьому основні етапи роботи імунного методу нечіткої кластеризації майже не відрізняються від етапів методу aiNETm. Слід зазначити, що використання цього підходу призводить до зниження швидкості класифікації зі збільшенням точності групування об'єктів.

**У четвертому розділі** розглянуті питання створення програмних засобів для моделювання роботи розроблених методів і моделей класифікації, що функціонують на основі ШС, та проведення експериментальних досліджень. Запропоновано інструментальне середовище моделювання, яке забезпечує можливість впровадження розроблених методів і моделей у різних задачах класифікації об'єктів із різним способом навчання. Окрім того, розроблене середовище дозволяє оцінити характеристики методів класифікації шляхом визначення їх швидкодії й точності групування об'єктів, та має модуль візуалізації процесу класифікації на різних етапах навчання.

Для порівняння розроблених методів і моделей класифікації у середовищі моделювання було сформовано набори класифікованих даних, які відрізняються між собою кількістю об'єктів, класів та характеристик кожного з них. Тестові набори даних наведені в табл. 1. Для аналізу методів класифікації з контрольованим навчанням набори даних, наведені в табл. 1, використовувалися без змін, а для аналізу методів кластеризації групуванню підлягали тільки об'єкти навчальної вибірки. Результати моделювання методів класифікації та кластеризації наведені в табл. 2. Слід зазначити, що в табл. 2 використовуються декілька скорочень, які ха-

рактикують ефективність методу класифікації: Т – характеризує швидкість класифікації (в % до самого повільного методу), А – характеризує точність групування (в % до вихідних даних, що класифікуються).

Таблиця 1 – Характеристики тестових наборів даних

Набір	Кількість об'єктів, що класифікуються	Навчальна вибірка	Кількість класів	Матриця характеристик
Н 1	5000	400	5	5 × 5
Н 2	5000	400	5	10 × 10
Н 3	10 000	500	10	5 × 5

Таблиця 2 – Результати моделювання методів класифікації

Методи класифікації	Набори даних			Методи кластеризації	Набори даних				
	Н 1	Н 2	Н 3		Н 1	Н 2	Н 3		
kNN (k = 5)	Т	22.3	22.9	23.9	k-means	Т	74.7	75.1	75.3
	А	95.5	95.5	92.2		А	99.5	99.4	99.5
CART	Т	71.3	74.1	68.5	MST	Т	44.4	44.0	44.1
	А	96.1	96.2	96.3		А	98.0	98.1	98.0
SVM	Т	97.9	96.8	96.6	aiNET	Т	100	100	100
	А	99.8	99.9	99.9		А	65.3	65.0	64.9
aiNET	Т	100	100	100	aiNETmc	Т	69.1	68.0	68.2
	А	60.9	60.3	59.8		А	98.7	98.2	98.0
aiNETm	Т	69.8	63.5	63.6	aiNETmkm	Т	60.2	60.7	60.9
	А	98.5	98.4	98.3		А	99.3	99.2	99.3
aiNETmkn (k = 5)	Т	68.0	61.6	61.2	aiNETmfc	Т	61.3	61.5	61.4
	А	98.5	98.4	98.3		А	99.5	99.3	99.5

Згідно з результатами моделювання, серед методів класифікації з контрольованим навчанням найбільшою точністю характеризується метод SVM, а найбільшою швидкістю класифікації – метод kNN. Слід зазначити, що отримані шляхом модифікації та гібридизації імунні методи aiNETm і aiNETmkn майже не відрізняються від методу SVM у точності, але мають над ним перевагу у швидкодії майже на 30 %. Серед методів кластеризації найбільшою точністю характеризується метод k-means, а найбільшою швидкістю – метод MST. Імунний гібридний метод aiNETmkm майже не відрізняється від методу k-means у точності групування об'єктів, але має над ним перевагу у швидкості майже на 15 %.

Розроблені в роботі методи та моделі класифікації використані для розв'язання практичних задач. Модель класифікації з контрольованим навчанням на основі модифікованого методу клонального відбору застосовано в задачі визначення результатів тестування власників вогнепальної зброї в компанії «СТРАЖ». Основною особливістю задачі класифікації результатів тестування є велика кількість контрольних питань, що поділяються на декілька основних тем,

які відрізняються між собою вагою помилки. При цьому результати тестування розподіляються на три групи: «Склав», «Необхідна співбесіда», «Не склав». Відповідно до цього, сукупність результатів тестування формує початкову множину антитіл, а сукупність правил віднесення результатів до тієї чи іншої групи – множину навчальних антигенів. Оскільки в ході вирішення цієї задачі відсутня необхідність мережевої взаємодії між антитілами в процесі імунного навчання, для класифікації результатів тестування було використано модель клонального відбору та модифікований метод CLONALGcls, у якому застосовується оператор конкурентно-цільового відбору клонів на етапі редагування популяції антитіл.

Результати роботи застосовано для визначення сум страхових платежів в ході страхування автотранспорту клієнтів компанії «Київський страховий дім», що володіють транспортними засобами. В більшості страхових компаній України для розв'язання цієї задачі використовуються таблиці з визначеними сумами залежно від характеристик автотранспорту, що підлягають страхуванню, та кваліфікації власника-водія. Під час обчислення сум страхового платежу використовувався модифікований імунний метод клонального відбору з контрольованим навчанням. При цьому ідентифікатором кожного класу є сума страхового платежу, а вся сукупність характеристик транспортного засобу та стаж водія формує об'єкт, що класифікується. Якщо всі необхідні характеристики вказані, то процес імунного навчання не проводиться, але якщо характеристики об'єкта подані в неповному обсязі – для визначення суми страхового платежу необхідне використання імунної моделі класифікації. В даному випадку кожному об'єкту відповідатиме декілька класів, які формують діапазон суми страхового платежу, що обумовлює використання методів нечіткої логіки. Слід зазначити, що в процесі імунного навчання в ході класифікації об'єкти не взаємодіють між собою, тому для вирішення цієї задачі було використано модель клонального відбору та модифікований метод нечіткої класифікації CLONALGfcls, у якому для визначення діапазону сум страхових платежів використовується функція належності об'єктів, що класифікуються, множині початкових класів.

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

## ВИСНОВКИ

У дисертаційній роботі наведено результати, які, згідно з метою дослідження, в сукупності є вирішенням актуальної наукової задачі – розробки методів і моделей класифікації об'єктів на основі штучних імунних систем, що має велике значення для підвищення ефективності систем інтелектуальної обробки інформації. Вирішення поставлених у роботі завдань дозволило отримати такі основні результати:

1. Проведено аналіз методів і моделей класифікації об'єктів, який вказав на необхідність використання нових підходів до групування, які мають базуватися на основних принципах функціонування штучних імунних систем.

2. Розроблено та формалізовано імунні моделі класифікації об'єктів з різ-

ними способами навчання, а також імунні моделі автоматичної класифікації об'єктів, що формують кластери з множини об'єктів навчальної вибірки.

3. Запропоновано використання стимулюючих антитіл для підвищення швидкості навчання штучної імунної мережі та визначення початкових центрів кластерів, які формуються в ході вирішення задач класифікації з неконтрольованим навчанням, або автоматичної класифікації об'єктів.

4. Запропоновано метод конкурентно-цільового відбору клонів, який може використовуватися в різних імунних моделях на етапі редагування множини антитіл, що дозволяє підвищити швидкість імунного навчання та класифікації об'єктів без втрати точності угруповання.

5. Запропоновано метод мутації клонів, що полягає у використанні афінності далекого предка як нижньої межі діапазону допустимих значень для визначення коефіцієнта мутації, що призводить до підвищення швидкості імунного навчання без втрати точності класифікації об'єктів.

6. Розроблено гібридні моделі класифікації об'єктів на основі методів k-means та kNN, у яких використовується імунний підхід для визначення кількості кластерів (метод k-means) та кількості найближчих сусідів (метод kNN), що без зменшення точності призводить до зростання швидкості групування даних.

7. Проведено експериментальні дослідження розроблених методів і моделей класифікації об'єктів, які показали високу швидкість і точність класифікації об'єктів, зіставні з результатами еталонних методів класифікації, що використовують не імунні принципи угруповання об'єктів.

8. Розроблені методи та моделі використані для вирішення задачі класифікації результатів тестування учнів власників зброї в компанії «Страж» (м. Харків), в процесі визначення розміру страхового платежу під час страхування автотранспорту шляхом класифікації характеристик транспортних засобів у філії страхової компанії «Київський страховий будинок» (м. Харків), а також у навчальному процесі Харківського національного університету радіоелектроніки.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Кораблев Н.М. Нечеткая кластеризация данных на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичёв // Вестник ХНТУ. – 2010. – № 2 (38). – С. 323–328.

2. Кораблев Н.М. Исследование иммунных операторов в задаче кластеризации объектов / Н.М. Кораблёв, А.А. Фомичёв // Бионика интеллекта. – 2010. – № 1 (72). – С. 70–74.

3. Кораблев Н.М. Обработка результатов тестирования обучающихся с использованием искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Бионика интеллекта. – 2011. – № 1 (75). – С. 84–87.

4. Кораблев Н.М. Кластеризация данных методом k-means с использованием иммунных операторов / Н.М. Кораблев, А.А. Фомичев // Бионика интеллекта. – 2011. – № 3 (77). – С. 102–106.

5. Кораблев Н.М. Модификация иммунного метода RLAIIS для автоматиче-

ской классификации объектов / Н.М. Кораблев, А.А. Фомичев, М.В. Кушнарев // Проблемы інформаційних технологій: наук.-техн. журнал. – 2014. – № 02(016). – С. 29-38. (Входит до міжнародних наукометричних баз Research Bible, Open Academic Journals Index (ОАЖ)).

6. Кораблев Н.М. Автоматическая классификация данных на основе иммунного похода / Н.М. Кораблев, А.А. Фомичев // Бионика интеллекта: науч.-техн. журнал. – 2014. – № 2 (83). – С. 83-90.

7. Кораблев Н.М. Кластеризация данных на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Системы обробки інформації. – 2009. – Вип. 4 (78). – С. 77–82. (Входит до міжнародних наукометричних баз Index Copernicus International).

8. Кораблев Н.М. Классификация объектов на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Системы обробки інформації. – 2010. – Вип. 6 (87). – С. 13–17. (Входит до міжнародних наукометричних баз Index Copernicus International).

9. Кораблев Н.М. Нечеткая классификация объектов на основе искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев, М.В. Кушнарев // Комп'ютерні системи та компоненти. – 2010. – Том 1. – Вип.2. – С. 88–94.

10. Кораблев Н.М. Анализ сходимости иммунных алгоритмов / Н.М. Кораблев, А.Э. Макогон, А.А. Фомичев // Системы обробки інформації. – 2011. – Вып. 2 (92). – С. 29–33. (Входит до міжнародних наукометричних баз Index Copernicus International).

11. Кораблев Н.М. Классификация объектов с помощью иммунного метода ближайших соседей / Н.М. Кораблев, А.А. Фомичев // Системы управління, навігації та зв'язку. – 2011. – Вип. 2 (18). – С. 114–118.

12. Кораблёв Н.М. Повышение эффективности иммунных методов классификации объектов с помощью целевого клонального отбора / Н.М. Кораблев, А.А. Фомичев // Автоматизированные системы управления и приборы автоматики. – 2012. – Вып. 161. – С. 45-50.

13. Korablyov M. The immune method for classifying objects on the basis of the target clonal selection (Immunologiczne metody klasyfikacji obiektów bazujące na selekcji klonalnej) / M. Korablyov, O. Fomichov, M. Kushnaryov, W. Wójcik // Elektronika (LIV). – 2013. – № 8. – P. 36-39.

14. Кораблев Н.М. Восстановление классов с помощью искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Тези доповідей VI міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2008)», 12-14 листопада 2008 р., м. Дніпропетровськ. – ДНУ, 2008. – С. 180.

15. Кораблев Н.М. Использование искусственных иммунных систем для кластеризации данных / Н.М. Кораблев, А.А. Фомичев // Матеріали X Між нар. Наково-техн. конфер. «Системний аналіз та інформаційні технології САІТ-2008». – К.: НТУУ «КП», 2008. – С. 210.

16. Кораблев Н.М. Решение задачи кластеризации данных при помощи искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичев // Сборник трудов

IX міжнародної конференції «Інтелектуальний аналіз інформації – ІАІ-2009». – К.: Просвіта, 2009. – С. 171-177.

17. Кораблев Н.М. Распознавание образов на основе искусственных иммунных систем / Н.М. Кораблев, Т.Н. Кротенко, А.А. Фомичев // Системный анализ и информационные технологии: XI Междунар. науч.-техн. конф.: материалы конф. – К.: ННК «ІПСА» НТУУ «КПІ», 2009. – С. 326.

18. Кораблев Н.М. Кластеризация данных с помощью нечетких искусственных иммунных систем / Н.М. Кораблев, А.А. Фомичёв, Т.Н. Кротенко // Проблемы информатики и моделирования: 9-я междунар. науч.-техн. конф.: материалы конф. – Харьков: НТУ «ХПИ». – 2009. – С. 28.

19. Кораблев Н.М. Иммунный подход к классификации объектов / Н.М. Кораблев, А.А. Фомичев, А.А. Москаленко // Современные направления развития информационно-коммуникационных технологий и средств управления: 1-я науч.-техн. конф.: тез. докладов. – Х: ДП «ХНДИ ТМ»; К.: ДП «ЦНДИ НиУ», 2010. – С. 71.

20. Кораблев Н.М. Кластеризация данных методом k-means при использовании иммунного подхода / Н.М. Кораблев, А.А. Фомичев // Информационные технологии в навигации и управлении: состояние и перспективы развития: 2-я междунар. науч.-техн. конф.: материалы конф. – К.: 2011. – С. 38.

21. Кораблев Н.М. Модификация функции афинности в иммунных методах обработки информации / Н.М. Кораблев, А.А. Фомичёв, С.Ю. Ломаненко // Перша міжнародна науково-технічна конференція «Проблеми інформатизації». – Черкаси–Київ–Тольятті–Харків, 2013. – С. 26.

22. Кораблев Н.М. Использование аффинности на основе Манхэттенского расстояния в методах клонального отбора / Н.М. Кораблев, А.А. Фомичёв, С.Ю. Ломаненко // Друга міжнародна науково-технічна конференція «Проблеми інформатизації». Тези доповідей. – Київ–Полтава–Катовице–Париж–Білгород–Черкаси–Харків, 2014. – С. 72.

23. Кораблев Н.М. Повышение быстродействия алгоритмов, функционирующих на основе модели иммунной сети / Н.М. Кораблев, А.А. Фомичёв, А.Ю. Арутюнов // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: Матеріали четвертої міжнар. наук.-техн. конференції, 4-5 грудня 2014 р. – Полтава: ПНТУ; Баку: ВА ЗС АР; Белгород: НДУ «БелДУ»; Кіровоград: КЛА НАУ; Харків: ДП «ХНДИ ТМ», 2014. – С. 38.

24. Кораблев Н.М. Повышение скорости иммунного обучения при использовании минимального ограничительного порога при клонировании / Н.М. Кораблев, А.А. Фомичёв, О.Ю. Крук // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління: Матеріали п'ятої міжнар. наук.-техн.конференції, 23-24 квітня 2015 р. – Полтава: ПНТУ; Баку: ВА ЗС АР; Белгород: НДУ «БелДУ»; Кіровоград: КЛА НАУ; Харків: ДП «ХНДИ ТМ», 2015. – С. 28.

25. Кораблев Н.М. Адаптивная модель автоматической классификации объектов на основе иммунного подхода / Н.М. Кораблев, А.А. Фомичёв // Інформаційні та моделюючі технології: Матеріали всеукраїнської науково-практичної

конференції – ІМТ-2015. – Черкаси, 2015. – С. 73.

## АНОТАЦІЯ

**Фомічов О.О. Методи та моделі класифікації об'єктів 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, Міністерство освіти і науки України, Харків, 2016.

Дисертаційна робота присвячена розробці методів і моделей класифікації об'єктів на основі штучних імунних систем з різними способами навчання, які дозволяють підвищити швидкість класифікації та кластеризації даних із забезпеченням високої точності угруповання.

Проведено аналіз методів класифікації об'єктів, який виявив основні недоліки, пов'язані з проблемами низької швидкості класифікації за високої точності поділу об'єктів, які стають найбільш очевидними зі збільшенням розмірності параметрів об'єктів, що класифікуються. Запропоновано метод конкурентно-цільового відбору клонів, який може використовуватися в алгоритмах, що реалізують різні імунні моделі на етапі редагування множини антитіл. Використання даного методу дозволяє підвищити швидкість імунного навчання, що призводить до підвищення швидкості класифікації об'єктів без втрати точності угруповання.

Запропоновано узагальнену модель автоматичної класифікації на основі імунних методів і моделей організації обчислень, яка виконує не тільки розподіл об'єктів між вихідним набором класів, а й виділення нових кластерів для об'єктів, які за своїми характеристиками не можуть бути віднесені до жодного класу. Розроблено гібридні методи класифікації, які використовують не тільки імунні моделі обробки даних, а й класичні методи класифікації (kNN і k-means), а також принципи організації м'яких обчислень на основі нечіткої логіки.

Запропоновано метод мутації клонів, що полягає у використанні афінності далекого предка як нижньої межі діапазону допустимих значень, що використовується для визначення коефіцієнта мутації. Використання даного оператора мутації призводить до підвищення швидкості імунного навчання без втрати точності класифікації. Запропоновано використання стимулюючих антитіл для підвищення швидкості імунного навчання і визначення початкових центрів кластерів, які формуються на етапі імунного навчання в ході вирішення задачі класифікації з неконтрольованим навчанням, або автоматичної класифікації об'єктів.

Проведено експериментальні дослідження розроблених методів класифікації об'єктів, які показали їх високу ефективність. Розроблені методи і моделі використані для вирішення задачі класифікації результатів тестування учнів-власників зброї в компанії «Страж» (м. Харків), а також в ході визначення розміру страхового платежу за обов'язкового страхування автотранспорту шляхом класифікації характеристик транспортних засобів у філії страхового компанії «Київський страховий дім» (м. Харків).

Ключові слова: класифікація, антитіла, антигени, коефіцієнт мутації, штуч-

на імунна система, імунна оператор, імунна модель.

### **АННОТАЦИЯ**

**Фомичёв А.А. Методы и модели классификации объектов 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, Министерство образования и науки Украины, Харьков, 2016.

Диссертация посвящена разработке методов и моделей классификации объектов на основе искусственных иммунных систем с различными способами обучения, которые позволяют повысить быстродействие алгоритмов классификации и кластеризации данных при обеспечении высокой точности группировки.

Проведен анализ методов классификации объектов, который выявил основные недостатки, связанные с проблемами низкой скорости классификации при высокой точности разделения объектов, которые становятся наиболее очевидными при увеличении размерности параметров классифицируемых объектов. Предложен метод конкурентно-целевого отбора антител, который может использоваться в алгоритмах, реализующих различные иммунные модели на этапе редактирования множества клонов. Использование данного метода позволяет повысить скорость иммунного обучения, что приводит к повышению скорости классификации объектов без потери точности группировки.

Предложена обобщенная модель автоматической классификации на основе иммунных методов и моделей организации вычислений, которая выполняет не только распределение объектов между исходным набором классов, но и выделение новых кластеров для объектов, которые по своим характеристикам не могут быть отнесены ни к одному из классов. Разработаны гибридные методы классификации, которые используют не только иммунные модели обработки данных, но и классические методы классификации (kNN и k-means), а также принципы организации мягких вычислений на основе нечеткой логики.

Предложен метод мутации клонов, заключающийся в использовании аффинности дальнего предка в качестве нижней границы диапазона допустимых значений, который используется для определения коэффициента мутации. Использование данного метода мутации приводит к повышению скорости иммунного обучения без потери точности классификации. Предложено использование стимулирующих антител для повышения скорости иммунного обучения и определения начальных центров кластеров, которые формируются на этапе иммунного обучения при решении задачи классификации с неконтролируемым обучением, или автоматической классификации объектов.

Проведены экспериментальные исследования разработанных методов и моделей классификации объектов, которые показали их высокую эффективность. Разработанные методы и модели использованы для решения задачи классификации результатов тестирования учащихся владельцев оружия в компании «Страж» (г. Харьков), а также при определении размера страхового платежа при обяза-

тельном страховании автотранспорта путем классификации характеристик транспортных средств в филиале страховой компании «Киевский страховой дом» (г. Харьков).

Ключевые слова: классификация, антитела, антигены, коэффициент мутации, искусственная иммунная система, иммунный оператор, иммунная модель.

## ABSTRACT

### **Fomichev AA Methods and models of object classification based on artificial immune systems. - Manuscript.**

Thesis for candidate's 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.

The dissertation is devoted to development of methods and models of objects classification using artificial immune systems with different ways of learning that will improve the performance and high accuracy of classification and clustering.

The analysis methods for classifying objects, which revealed major shortcomings associated with the problems of classification of low speed high precision separation of objects. A method of a competitive target selection can be applied in algorithms with different immune models in the stage of editing a set of clones or antibodies. Usage of this method allows to increase the speed of immune learning stage, and object classification without loss of accuracy. A method for mutation clones comprising using affinities of elder antibodies as a lower allowable value range can increase the immune learning speed without losing the classification accuracy. Proposed use of stimulating antibodies to enhance immune speed training and the definition of the initial cluster centers, which are formed at the stage of the unsupervised learning, or automatic classification.

The generalized model of automatic classification using immune models, can perform not only the distribution of objects between the original set of classes, but also the creation the new clusters for objects, that cannot be attributed to any class. Developed hybrid classification methods which use not only immune data model, but also classical methods of classification (kNN and k-means), as well as soft principles algorithms based on fuzzy logic.

Experimental studies of the developed methods and models of objects classification, which have shown their high efficiency. The developed methods and models used to solve the classification problem in the test results of gun owners in «The Guardian» (Kharkiv), as well as determining the amount of insurance payment under compulsory insurance of vehicles classification by their characteristics in the company «Kievan insurance house» (Kharkiv).

Keywords: classification, antibodies, antigens, mutation rate, artificial immune system, immune operator, immune model.