

ДОДАТОК А

Графічний матеріал кваліфікаційної роботи

Харківський національний університет радіоелектроніки

Модель деревовидної штучної імунної мережі для кластеризації об'єктів

Виконав:
магістрант групи СПМ-20-2 Носова О.Є.

Науковий керівник:
ст. викл. каф. ЕОМ Фомічов О.О.



Харківський національний університет радіоелектроніки
Кафедра ЕОМ

Актуальність проблеми

- Великий інтерес до методів інтелектуальної обробки даних
- Кластеризація даних є однією з найбільш актуальних проблем
- Традиційні методи кластеризації мають ряд недоліків
- Великий науковий інтерес до біологічних методів організації обчислень, в тому числі штучних імунних систем
- Існуючі імунні моделі мають низьку швидкість

1

Існуючі рішення

Найбільш поширеними класичними методами кластеризації є:

- C-means, що формує кластери-центроїди
- MST, що формує кластери-стрічки
- SVM, що формує гіперсфери для кластеризації

Найбільш поширеними імунними моделями є:

- Модель відбору клонів
- Модель штучної імунної мережі

2

Постановка задачі

- Розробка імунного методу кластеризації, що характеризується великою швидкістю роботи та точністю кластеризації
- Використати імунний метод aiNET як основу для нового методу кластеризації
- Використати теорію графів для побудови деревовидної імунної мережі для підвищення швидкості кластеризації
- Дослідити ефективність створеного методу кластеризації порівняно з еталонними методами

3

Використані технології

- Платформа .NET 4.7
- Архітектурний паттерн MVC
- Мова програмування C#
- Середовище розробки Visual Studio 2019 Community Edition
- Технологія WindowsForms для побудови інтерфеса користувача

4

Еталонні методи кластеризації

Еталонні класичні методи кластеризації:

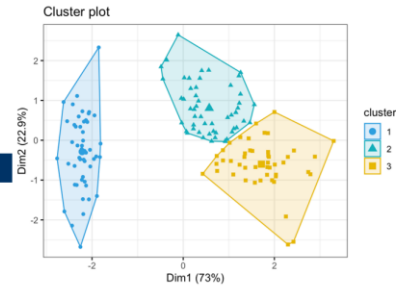
- C-means
- MST

Еталонні імунні методи:

- Clonalg
- aiNET

5

Метод C-means

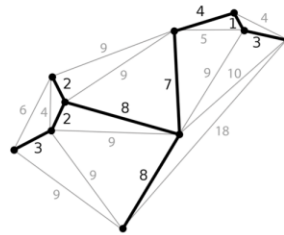


- Передбачає формування кластерів-центроїдів
- Кількість кластерів C задається користувачем
- На першій ітерації алгоритму центри кластерів обираються випадковим чином та перебираються на кожній ітерації
- На кожній ітерації всі об'єкти визначають належність до того чи іншого кластеру
- Алгоритм закінчує роботу коли дві ітерації одна за одною, кластери не змінюють своїх центрів та меж

6

Метод Minimal Spanning Tree, MST

- Функціонує на основі теорії графів
- Кількість кластерів визначається користувачем
- Спочатку формується зв'язаний граф з набору об'єктів, які стають вершинами цього графу
- Кластери формуються шляхом видалення ребер графу, які мають найбільшу довжину



7

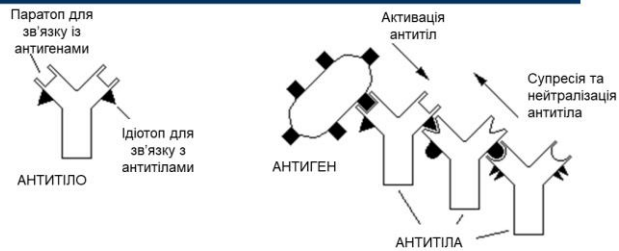
Модель клонального відбору, алгоритм Clonalg

- Використовує афінність між антитілами та антигенами як міру подібності об'єктів
- Використовує популяцію антигенів у якості навчальної вибірки
- Використовує оператори відбору клонів та старіння антитіл задля редагування популяції антитіл
- Припиняє роботу у випадку коли всі антитіла відтворили ознаки антигенів



8

Модель штучної імунної мережі, алгоритм aiNET



- Використовує афінність як міру подібності імунних об'єктів
- Досліджує не тільки зовнішні зв'язки між антитілами та антигенами, а і внутрішні зв'язки між антитілами
- Зупиняється коли всі антитіла відтворили ознаки антигенів

9

Недоліки існуючих імунних методів кластеризації

- Велика кількість помилок кластеризації при високій швидкості процесу імунного навчання та саморегуляції мережі властива для мережевих методів та алгоритму aiNET
- Повільна швидкість роботи при високій точності групування об'єктів властива для методів відбору клонів та методу Clonalg
- Виконання надмірних обчислень під час мутації відбору клонів та супресії мережі антитіл
- Умовна сходиність алгоритмів

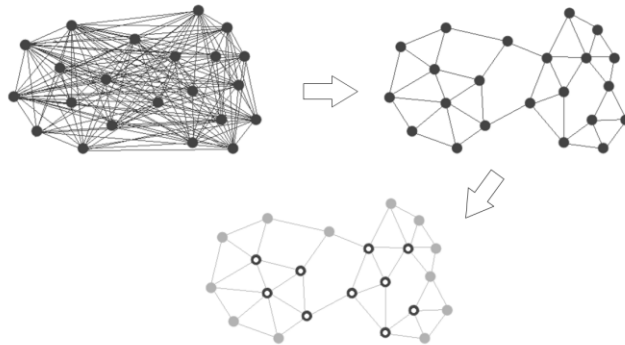
10

Особливості деревовидної штучної імунної мережі

- Використання афінності між антитілами у якості міцності зв'язку
- Можливість роботи за відсутності популяції антигенів, що використовуються у якості тренувальної вибірки об'єктів
- Формування графу повного зв'язку / дерева з набору антитіл, де антитіла є вершинами, а афінності між ними – ребрами
- Зменшення кількості зв'язків між антитілами за допомогою параметру k , кожне антитіло зберігає k зв'язків з максимальною афінністю, інші зв'язки видаляються
- Визначення рівней стимуляції для антитіл деревовидної мережі

11

Особливості деревовидної штучної імунної мережі



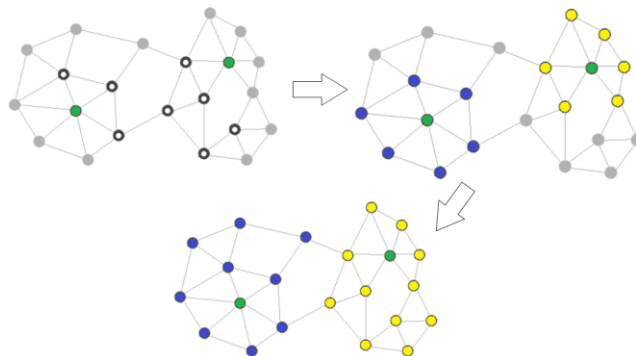
12

Використання деревовидної імунної мережі для кластеризації даних

- Використання кількості кластерів як вхідного параметру для алгоритма кластеризації
- Використання рівней стимуляції для визначення центрів кластерів
- Обчислення середнього порогу афінності NAT, який застосовується при виділенні центрів кластерів
- Групування антитіл, що мають прями зв'язки з центрами кластерів без запуску процесу імунної спомрегуляції мережі
- Формування кластерів сильного згущення у якості наборів цільових об'єктів для інших антитіл
- Використання аїдностей як міри належності до кластеру

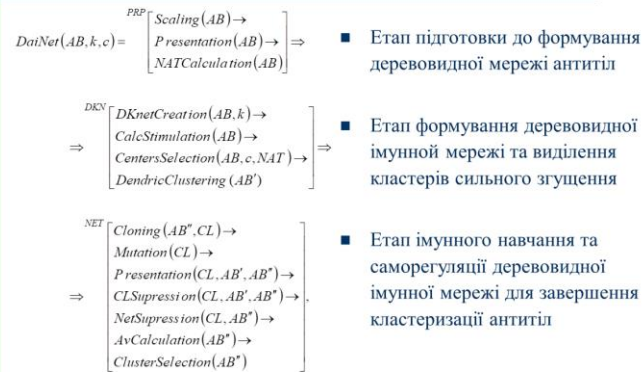
13

Формування кластерів за допомогою деревовидної імунної мережі



14

Загальна схема роботи деревовидної імунної мережі при кластеризації даних



15

Особливості роботи імунних операторів в моделі деревовидної імунної мережі

- Пропорційне клонування антитіл з посиленням афінності антитіла задля підвищення кількості клонів
- Зворотно-пропорційна мутація клонів із обмеженням нижнього порогу при виділенні коефіцієнту мутації

$$\mu = rand \left[\frac{1}{2}(1 - aff(ab, AG)); 1 - aff(ab, AG) \right]$$

- Використання аівдності до цільових об'єктів під час визначення належності антитіл сформованим кластерам

$$av_j = \sum_{i=1}^m aff(ab_i, ab_j)$$

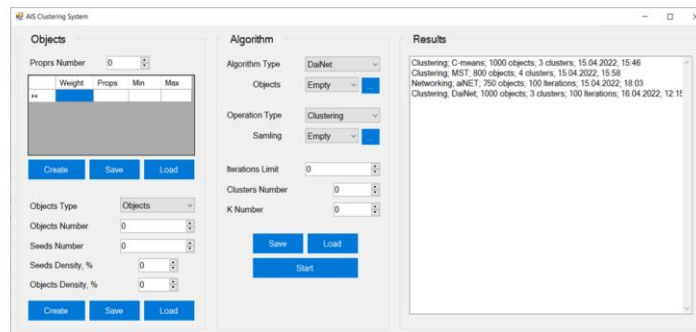
16

Результати досліджень

Алгоритм		Набір 1	Набір 2	Набір 3
MST	T	38%	36%	39%
	A	88%	85%	82%
C-means	T	72%	74%	72%
	A	100 %	100 %	100 %
Clonalg	T	100 %	100 %	100 %
	A	80%	83%	81%
aiNet	T	98%	95%	93%
	A	52%	50%	50%
Dendric-aiNet	T	48%	46%	47%
	A	95%	93%	96%

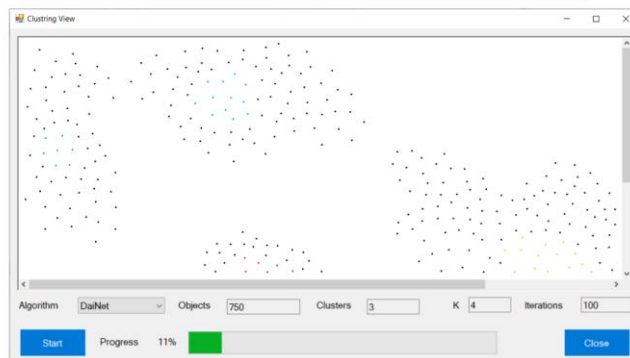
17

Застосунок для моделювання роботи деревоподібної штучної імунної мережі



18

Застосунок для моделювання роботи деревоподібної штучної імунної мережі



19

Висновки

- Проведено аналіз найбільш поширених класичних методів та моделей штучних імунних систем для кластеризації даних
- Виділено набір універсальних імунних операторів та проаналізовано можливості їх модифікації задля підвищення якості кластеризації
- Сформовано модель деревоподібної штучної імунної мережі та алгоритм Dendric-aiNet для кластеризації даних
- Розроблено програмний застосунок для модулювання алгоритмів кластеризації

20