

В.И. Литвиненко

Кластерный анализ данных на основе модифицированной иммунной сети

Представлена новая версия искусственной иммунной системы для решения задач автоматической кластеризации данных. Алгоритм использует свойства самоорганизации иммунной системы и создает устойчивую иммунную сеть.

A new version of artificial immune system for solving the problems of data automatic clustering is presented. The algorithm uses the properties of self-organization of the immune system and creates a sustained immune network.

Представлено нову версію штучної імунної системи для рішення завдань автоматичної кластеризації даних. Алгоритм використовує властивості самоорганізації імунної системи й створює стійку імунну мережу.

Введение. Естественные иммунные системы обладают рядом привлекательных свойств с точки зрения обработки информации – способностью к самоорганизации [1]. Самоорганизация – это процесс, в результате которого на основе имманентных свойств функций и структур системы самостоятельно, без внешних управляющих воздействий, создается, воспроизводится или совершенствуется ее организация [2]. Принципы самоорганизации были предметом исследования многих ученых – Дж. Фон Неймана, Н. Винера, У.Р. Эшби и др. Значительный вклад в развитие этого направления внесли исследования украинских кибернетиков под руководством А.Г. Ивахненко [3–5], разработавших целый класс адаптивных самоорганизующихся моделей, по праву названный «интеллектуальным обобщением» эмпирико-статистических методов.

Итак, самоорганизация в иммунной системе – необратимый процесс, происходящий под действием внешней среды [3], в результате кооперативного взаимодействия множества элементов (подсистем), что приводит к образованию более эффективных (с позиции системы) структур. Известны две биологические самоорганизующиеся системы – нервная и иммунная, обладающие возможностями интеллектуальной обработки информации, включающими память, способность обучаться, распознавать и принимать решения относительно неизвестных ситуаций. Потенциал естественной нервной системы как биологического прототипа различных вычислительных схем уже достаточно хорошо развит в виде парадигмы

искусственных нейронных сетей и нейровычислений. Вычислительные возможности естественной иммунной системы в последнее время получили достойную оценку, что выразилось в создании нового направления – искусственные иммунные системы (ИИС). По результатам последних исследований, ИИС обладают высоким потенциалом в вычислительном плане и формируют основу иммунокомпьютинга – нового подхода, использующего принципы обработки информации иммунными системами [6].

Проблема кластеризации при помощи искусственных иммунных сетей

В основе алгоритмов кластеризации искусственными иммунными сетями лежит использование процесса самоорганизации. В них отсутствует явно выраженная целевая функция, а используются лишь критерии распознавания кластеризуемых данных множеством генерируемых детекторов. Существуют две наиболее известные системы анализа данных, основанные на иммунной метафоре: *aiNet* [7] и *AINE* [8]. Обе системы используют обучающее множество в качестве антигенов, а их цель – создание множества *B*-клеток или антител, представляющих эти антигены. В соответствии с гипотезой Н. Эрне, алгоритм *AINE* синтезирует сети (копии идиопатической сети), охватывающие основные свойства элементов данных из обучающей выборки. Система *AINE* использует почти все механизмы теории идиопатической сети:

- каждая *B*-клетка способна распознать несколько антигенов;

- схожие между собой *B*-клетки объединяются в иммунную сеть;

- все иммунные клетки подвергаются гипермутации и клональному отбору;

- *B*-клетки можно представить в виде искусственных распознающих сфер, или *ARB* (*Artificial Recognition Ball*).

Применение формализма *ARB* обусловлено стремлением улучшить устойчивость иммунной сети. В данном виде иммунной сети используется всего четыре управляющих параметра:

- порог аффинности сети;

- уровень мутации;

- число *ARB*;

- число создаваемых клонов, активированных *ARB*.

Влияние этих параметров на качество получаемых сетей исследовано в работе [9]. Иммунная сеть *aiNet* представляет собой систему, использующую более упрощенное представление. В данной ИИС вместо *B*-клеток или *ARB* используется популяция антител, представленных векторами вещественных, бинарных, категориальных или символьных значений.

В данной статье за основу взят алгоритм *aiNet* [7], что обусловлено двумя причинами:

- это один из самых известных в литературе методов кластеризации искусственными иммунными системами;

- его эффективность демонстрируют решения большого количества сложных задач.

В данном алгоритме клональный отбор управляет динамикой и метадиной сетью. Первоначально созданная случайным образом популяция антител видоизменяется с помощью операторов клонального отбора, гипермутации и апоптоза (удаление нестимулируемых клеток). Недостаток алгоритма – значительно большее количество определяемых пользователем параметров управления, а также необходимость использования стандартных инструментов кластерного анализа, таких как алгоритмы построения *минимального остовного дерева*. Преимуществом алгоритма является возможность сжа-

тия данных, которое в некоторых случаях достигает 90% [10].

Существует определенный класс задач, где способность иммунных сетей к сжатию информации может иметь отрицательное значение. Для них можно выделить следующие общие признаки:

- кластеры расположены слишком близко друг к другу;

- плотность расположения точек данных меняется от кластера к кластеру;

- границы кластеров имеют нечеткий или неравномерный характер и перекрываются, поэтому искажение структуры данных, вызванное их сжатием при создании «внутреннего образа» сети, может исказить ключевые свойства данных, необходимые для правильной идентификации кластеров.

В результате решения таких задач относительные расстояния между кластерами во внутреннем образе не будут соответствовать относительным расстояниям между исходными точками данных, что может значительно усложнить применение алгоритма кластеризации. Как показали эксперименты по исследованию данного класса задач с помощью сетей Кохонена (*SOM*) [11], там существует та же проблема. Кроме того, в работе показано, что *U*-матрица, в которой используется критерий расстояний для разделения образов на кластеры, не способна к решению подобных задач и что использование информационного представления плотности данных в комбинации с *U*-матрицей значительно улучшает качество кластеризации. В [11] подробно описаны практические задачи, где часто возникает подобная проблема.

В настоящее время известна одна работа [12], где описанная проблема решается с помощью иммунных сетей. Авторы предложили алгоритм иммунной сети с адаптивным подбором радиуса охвата, способного максимально сохранить информацию о плотности, подбирая значения радиуса обратно пропорционально локальной плотности данных. На основании этих значений проводится сжатие иммунной

сети, в результате чего области пространства с большей плотностью данных покрываются антителами меньшего радиуса, расположенными ближе друг к другу. Для областей с низкой плотностью антитела имеют больший радиус и не допускаются близко друг к другу процедурой сжатия.

Постановка задачи

Задача данного исследования – создание нового адаптивного алгоритма кластерного анализа данных, основанного на метафорах и принципах искусственной иммунной системы, использующего механизмы клонального расширения и супрессии иммунной сети и способного создавать внутренний образ с максимально полным сохранением информации о плотности данных, что позволит эффективно решать выделенный ранее класс труднорешаемых задач кластеризации.

Иммунные сети для решения задач кластеризации

Применительно к задачам распознавания и классификации в качестве популяции антигенов иммунной сети выступает набор данных (векторов), которые нужно распознать. Назовем этот набор $X = (x_1, x_2, \dots, x_N)$. Каждый вектор $x_i, i = 1, \dots, N_p$ состоит из p элементов. Элементами вектора могут быть наборы переменных, атрибутов или характеристик распознаваемого объекта. Для количественного расчета обоих видов взаимодействия векторов распознаваемого объекта в теории искусственных иммунных систем используется понятие «аффинности».

Определение 1. Под *аффинностью* будем понимать *меру взаимодействия* (или *силу связи*) соответствующих комплементарных участков антигена и антитела или двух антител, которая формально может быть представлена в виде одной из метрик (например, евклидова расстояния), указывающей на степень подобия или различия между соответствующими атрибутами строк, такой что $S^p \times S^p \rightarrow \mathfrak{R}^+$.

В текущей реализации аффинность определяется между 0 и 1 включительно и вычис-

ляется как евклидово расстояние *векторов свойств* двух объектов. Таким образом, низкие значения аффинности указывают на сильную близость двух объектов. Так как каждый элемент данных x_i представляет собой p -мерный вектор, то для моделирования связей «антиген-антитело» и «антитело-антитело» все антитела должны быть также представлены в форме p -мерных строк или векторов.

Определение 2. Множество стереохимических взаимодействий или множество свойств антител, определяющих их межмолекулярную аффинность, будем называть *многомерным пространством форм P* .

При использовании пространства форм предполагается возможность полного описания обобщенной формы антитела при помощи строки параметров длины p . Если при этом предположить, что и антиген, в свою очередь, имеет комплементарную антителу область, которую можно описать при помощи такого же количества параметров, то при объединении этих параметров в векторы антитело (Ab) и антиген (Ag) могут таким образом быть представлены как точки в p -мерном евклидовом пространстве.

Отображение множества форм во множество параметров решаемой задачи не имеет значения в вычислительном смысле и полностью определяется предметной областью использования ИИС. Каждое антитело способно распознавать не единственный антиген, а их множество, чья форма дополнения (комплемент) находится внутри определенного объема в пространстве форм. Этот объем характеризуется параметром ε , называемым кросс-реактивным порогом.

Определение 3. Кросс-реактивным порогом называется область в многомерном пространстве форм, охватывающая антитело, внутри которой проявляется его активность по отношению к антигенам или другим антителам.

Определение 4. *Искусственная иммунная сеть* представляет собой *полносвязный граф, состоящий из множества узлов* – клеток, сети

(антител) и множества взвешенных ребер, устанавливающих связи между клетками.

Уровень изменения популяции	Стимуляция сети	Супрессия сети	Поступление новых элементов	Гибель нестимулируемых элементов
-----------------------------	-----------------	----------------	-----------------------------	----------------------------------

Рис. 1. Обобщенная схема алгоритма иммунной сети

Формально алгоритм иммунной сети можно представить следующим образом:

$$immNET = (X, C, M, N_c, D, S, n, \xi, \sigma_d, \sigma_s), \quad (1)$$

где X – набор данных, состоящий из N_p векторов размерности p ; C – матрица, содержащая все клетки (антитела) сети ($C \in \mathbb{R}^{N \times p}$) – популяция антител; M – матрица, состоящая из N клеток памяти ($M \subseteq C$); N_c – общее количество клонов, создаваемых стимулируемыми клетками в каждом поколении (при активации сети); D – матрица элементов d_{ij} Ag - Ab аффинностей; S – матрица элементов s_{ij} Ab - Ab аффинностей; n – количество лучших клеток, отобранных из C для клонирования и мутации; ξ – процент улучшенных клеток, отбираемых из популяции клонов для последующей обработки; σ_d – пороговый коэффициент (кросс-реактивный порог) гибели клетки в зависимости от ее Ag - Ab аффинности; σ_s – пороговый коэффициент сжатия сети.

Значение весового коэффициента ребра отвечает аффинности связи клеток друг с другом. В иммунных сетях различают два вида аффинности:

- аффинность связи «антиген-антитело» (Ag - Ab) – степень расхождения;
- аффинность связи «антитело-антитело» (Ab - Ab) – степень подобия.

В целом алгоритм искусственной иммунной сети представляет собой итерационную процедуру, которая может быть получена при реализации трех основных фаз:

1. *Созревание аффинности*: антигены (точки данных) представлены антителам, которые подвергаются гипермутации для лучшего соответствия антигенам (взаимодействие антитела и антигена).

2. *Клональное расширение*: те антитела, которые сильнее стимулируются, т.е. лучше соответствуют образу, отбираются для дальнейшего клонирования и роста иммунной сети.

3. *Супрессия сети*: взаимодействие между антителами определяется количественно, и если одно антитело распознает другое, то одно из них удаляется из пула клеток (взаимодействие антитела и антитела).

Описание нового иммунного алгоритма кластеризации

Разработанный иммунный алгоритм кластеризации использует вещественное кодирование антител (рис. 2), где Ab_{ij} – координаты центра антитела в евклидовом пространстве. При этом антитела формируют вокруг себя p -мерную радиальную область распознавания с радиусом r , который адаптивно подбирается в процессе работы алгоритма. Таким образом, иммунная сеть заполняет пространство «антигенов» распознающими гиперсферами разного радиуса, что дает возможность его более полного покрытия.

r	Ab_{i1}	Ab_{i2}	...	Ab_{ip}
-----	-----------	-----------	-----	-----------

Рис. 2. Представление антитела иммунного алгоритма кластеризации

Пошаговая реализация алгоритма представлена ниже.

1. Инициализация переменных (создание начальной популяции антител Ab)
2. *For* для каждой итерации $numGen$ *do*:
 - 2.1. Вычисление аффинности популяции Ab антител по формуле f_{Ab-Ag} ;
 - 2.2. Отбор n лучших антител в популяцию клонов Ab_C ;
 - 2.3. Клонирование антител в популяции клонов Ab_C ;
 - 2.4. Гипермутация антител Ab_C с использованием уровня мутации m_R ;
 - 2.5. Вычисление аффинности f_{Ab-Ag} антител в популяции клонов Ab_C^* ;
 - 2.6. Отбор n_c антител из популяции клонов Ab_C^* и формирование из них популяции клеток памяти Ab_M ;

2.7. *For* для каждого антитела из популяции Ab_M (супрессия памяти) *do*:

2.7.1. Вычисление аффинности межклеточного взаимодействия f_{Ab-Ag} ;

2.7.2. Отбор n_M антител из Ab_M в соответствии с порогом аффинности, меньшим σ_d , в популяцию Ab'_M ;

END

2.8. Объединение популяций Ab и Ab'_M ($Ab' = Ab \cup Ab'_M$);

2.9. *For* для каждого антитела из популяции Ab' (супрессия сети и апоптоз) *do*:

2.9.1. Вычисление аффинности межклеточного взаимодействия f_{Ab-Ab} ;

2.9.2. Отбор n' антител из Ab' в соответствии с порогом аффинности, меньшим σ_S , в популяцию Ab'' ;

2.9.3. Создание популяции Ab''' удалением из популяции Ab'' антител, чья $f_{Ab-Ag} \leq d_M$

END

2.10. Замещение d антител популяции Ab''' новыми, случайно сгенерированными антителами.

3. Проверка достижения критерия останова

END

На шаге 1 генерируется начальное множество популяции антител. На шаге 2.1. по формуле 2 вычисляется значение аффинности связи «антиген-антитело»:

$$f_{Ab-Ag} = \frac{k_r}{r} + \frac{r^2}{H \cdot nAg_{rec}} \cdot D_{E(Ab-Ag_{rec})}, \quad (2)$$

где k_r – параметр значимости радиуса антитела, т.е. значение его кросс-реактивного порога (увеличение значения этого параметра заставляет иммунную сеть поддерживать антитела большего радиуса, что дает более грубую кластеризацию); r – радиус антитела (антигены, попадающие внутрь данной области, считаются распознанными антителом); nAg_{rec} – количество распознанных данным антителом антигенов; $D_{E(Ab-Ag_{rec})}$ – евклидово расстояние между центром антитела и усредненным центром всех распознанных этим антителом антигенов; H – показатель равномерности распределения

антигенов в области пространства, ограниченной радиусом антитела.

Показатель H вычисляется следующим образом:

$$H = \sum_{i=1}^p H_i, \quad (3)$$

где p – размерность пространства данных (длина рецептора антитела); H_i – показатель равномерности распределения распознанных антигенов вдоль i -й координатной оси пространства данных.

$$H_i = - \sum_{j=1}^{nAg_{rec}} p_j^{Ag_{rec}} \cdot \log p_j^{Ag_{rec}}, \quad (4)$$

где $p_j^{Ag_{rec}}$ – вероятность попадания распознанного антигена в j -й интервал i -й координатной оси пространства данных (в рамках области, ограниченной радиусом антитела), $p_j^{Ag_{rec}} =$

$$= \frac{nAg_{rec}^j}{nAg_{rec}}, \quad nAg_{rec}^j - \text{количество распознанных}$$

антигенов, попавших в j -й интервал i -й координатной оси пространства данных.

Полученные антитела ранжируются и на шаге 2.2 отбирается n лучших антител, из которых создается популяция Ab_C . Далее на шаге 2.3. осуществляется клонирование антител в популяции клонов Ab_C в соответствии с их аффинностью по формуле

$$N_c = \sum_{i=1}^n \text{round} \left(\frac{\beta * N}{i} \right), \quad (5)$$

где N_c – общее количество клонов в популяции клонов; β – фактор умножения; N – общее количество антител в популяции; round – оператор, округляющий аргумент к ближайшему целому числу. Каждое слагаемое этой суммы соответствует размеру клона каждого отобранного антитела, например, для $N = 100$ и $\beta = 1$ антитело с самой высокой аффинностью ($i = 1$) производит 100 клонов, в то время как второе по аффинности антитело производит 50 клонов и т.д. На шаге 2.4. осуществляется мутация антител пропорционально их аффинности, при этом множество Ab_C подвергается направленному процессу созревания аффинности

(управляемая мутация), генерируя видоизмененное множество Ab_C^* , где каждое антитело k из Ab_C подвергается мутации с уровнем m_{Rk} , обратно пропорциональным значению аффинности $f_{Ab-Ag_{i,j}}$ его родительского антитела – чем выше аффинность, тем меньший уровень мутации:

$$Ab_C^* = Ab_{Ck} + m_{Rk} (Ag_j - Ab_{Ck}); \quad (6)$$

$$m_{Rk} \sim 1/f_{Ab-Ag_{i,j}}; k = 1, \dots, N_C; i = 1, \dots, N.$$

На шаге 2.5 вычисляется аффинность f_{Ab-Ag} антител из популяции Ab_C^* в соответствии с формулой (2) и создание новой популяции клонов $Ab_C^{*'}$. На шаге 2.6 осуществляется ранжирование антител по уровню их аффинности из полученной популяции клонов $Ab_C^{*'}$ и отбор n_C лучших антител для формирования из них популяции клеток памяти Ab_M . На шаге 2.7.1 по формуле (7) осуществляется вычисление аффинности межклеточного взаимодействия f_{Ab-Ag} :

$$f_{Ab-Ag} = - \frac{D_{E(Ab_1-Ab_2)} - (r_{Ab_1} + r_{Ab_2})}{2 \cdot \min(r_{Ab_1}, r_{Ab_2})}. \quad (7)$$

При этом возможна следующая интерпретация значений f_{Ab-Ag} : ≤ 0 – распознающие области антител не перекрываются; $(0, 1)$ – области перекрываются оболочками, а само значение является степенью перекрытия; ≥ 1 – область меньшего радиуса (r) полностью находится внутри области большего радиуса. Далее на шаге 2.7.2 осуществляется отбор n_M антител из Ab_M в соответствии с порогом аффинности, меньшим σ_s , и создание новой популяции клеток памяти Ab'_M (супрессия антител клеток памяти). Пороговый коэффициент сжатия σ_s является максимально допустимой степенью перекрытия распознающих оболочек антител. На шаге 2.8. осуществляется объединение первоначальной популяции антител Ab , полученной на шаге 1, с популяцией клеток памяти, полученной на шаге 2.7.2: $Ab' = Ab \cup Ab'_M$. На шаге 2.9.1 для популяции Ab' вычисляется аф-

финность межклеточного взаимодействия f_{Ab-Ag} по формуле (7). На шаге 2.9.2 осуществляется супрессия сети Ab' , для этого отбирается n' антител из Ab' в соответствии с порогом аффинности, меньшим σ_s , и помещается в популяцию Ab'' . На шаге 2.9.3 создается популяция Ab''' посредством удаления из популяции Ab'' антител, чья аффинность $f_{Ab-Ag} \leq d_M$. На шаге 2.10 проводится замещение d антител популяции Ab''' новыми, случайно сгенерированными антителами.

В данной реализации алгоритм заканчивает работу после заданного пользователем числа итераций.

После окончания работы алгоритма образованная популяция клеток памяти используется для построения дерева минимального охвата (минимально остовного дерева), разделение которого (удаление максимальных связей) завершает процесс кластеризации.

Известно [15], что *охватывающее дерево* (*Spanning Tree*) графа G – это подграф, содержащий все вершины графа G . *Дерево минимального охвата* (*Minimal Spanning Tree*) – охватывающее дерево, имеющее минимальный суммарный вес всех дуг. *MST* применяется в иммунных сетях для идентификации кластерной структуры исследуемых данных.

Экспериментальные исследования

Для оценки эффективности разработанного алгоритма было использовано три набора задач для тестирования алгоритмов кластеризации Марбургского университета (*Philipps-Universität Marburg*), <http://www.uni-marburg.de/fb12/datenbionik/data>:

- *Lsun* – объем выборки 400 примеров, две переменных и три класса.
- *Target* – объем выборки 770 примеров, две переменных и шесть классов.
- *TwoDiamond* – объем выборки 800 примеров, две переменных и два класса.

На рис. 3,а,б и в показаны исходные данные наборов данных *Lsun*, *Target* и *TwoDiamond*. На рис. 4,а,б и в показаны построенные по данным иммунные сети совместно с *MST*-деревом.

На рис. 5 показана зависимость размера иммунной сети от количества итераций на исследуемых данных, на рис. 6 – зависимость качества кластеризации от размера сети. Из данного рисунка видно влияние количества синтезируемых антител на качество решения задачи кластеризации. В связи с этим мы попытались исследовать влияние основных параметров настройки на синтез антител.

Первым параметром, который был исследован, является порог супрессии σ_s на шаге 2.9.2 алгоритма (рис. 7). Данный параметр отображает допустимый уровень перекрытия оболочек. Его можно интерпретировать следующим образом – чем меньше его значение, тем сильнее супрессия, тем более разряженной будет сеть, т.е. антитела будут дальше расположены друг от друга. Из данного рисунка видно, что увеличение порога супрессии экспоненциально влияет на количество антител, а, следовательно, и на качество кластеризации. На рис. 8 показано влияние порогового кластерного удаления σ_d (шаг 2.7.2), который интерпретируется следующим образом: чем меньше его значение, тем больше будет удалений.

Несмотря на некоторое различие, наблюдается общая тенденция в поведении данного показателя – возрастание данного показателя значительно увеличивает синтез антител.

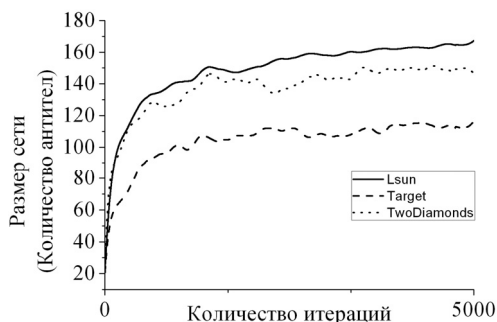


Рис. 5. Зависимость размера сети от количества итераций

Существенно влияет на синтез антител минимально допустимое значение радиуса детектора (антитела). На рис. 9 показано, как данное минимально допустимое значение параметра r в формуле (2) влияет на количество синтезируемых детекторов. Значимость радиуса детектора (параметр k_r в формуле (2)) по данным проведенных нами экспериментов также существенно влияет на качество кластеризации (рис. 10).

Заключение. В ходе проведенных исследований получены следующие результаты.

Разработан алгоритм кластеризации, основанный на принципах функционирования искусственной иммунной сети с модифицированной структурой.

В рамках разработанного алгоритма предложены оценки взаимосвязей «антиген-антитело» и «антитело-антитело», позволяющие сократить количество вложенных циклов алгоритма, что существенно ускорило его работу по сравнению с подобными алгоритмами.

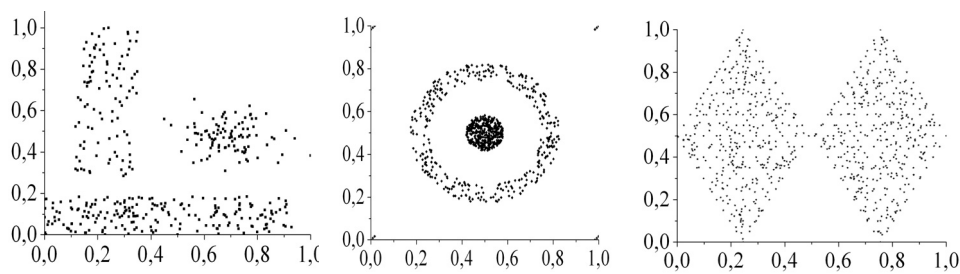


Рис. 3. Данные: а – Lsun; б – Target; в – TwoDiamond

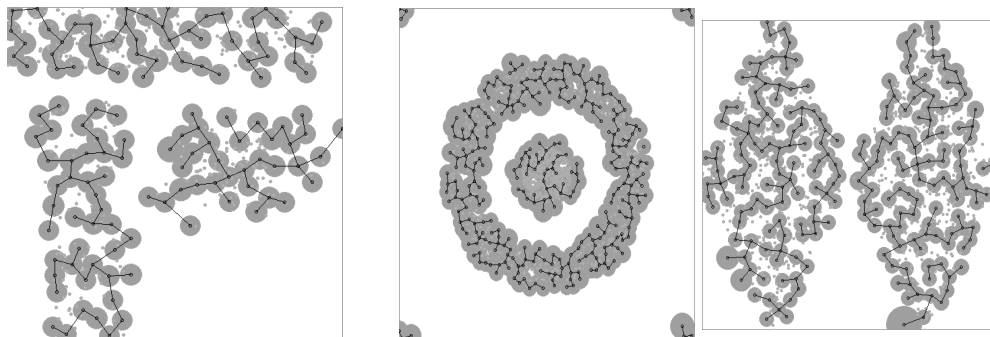


Рис. 4. Окончательная иммунная сеть для данных: а – Lsun; б – Target; в – TwoDiamond

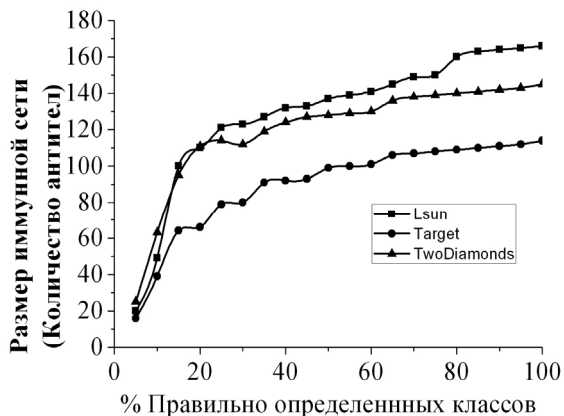


Рис. 6. Зависимость правильности классификации от размера иммунной сети

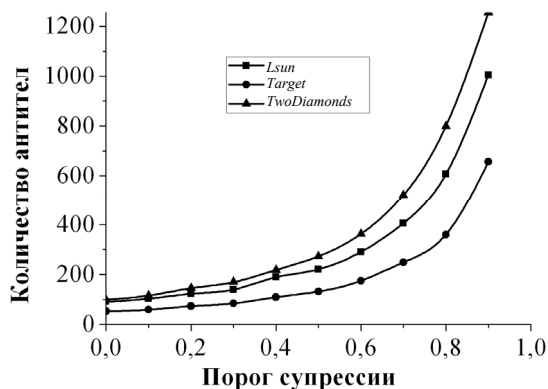


Рис. 7. Влияние порогового значения супрессии сети на количество синтезируемых антител

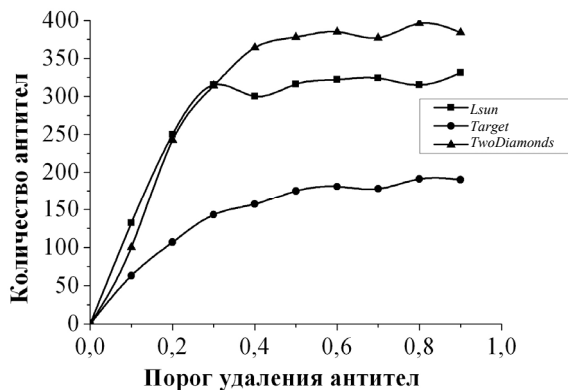


Рис. 8. Влияние значения порогового клонального удаления на количество синтезируемых антител

Проведены экспериментальные исследования влияния различных параметров алгоритма на качество решения задач кластеризации.

Для рассмотренных общепринятых тестовых задач обнаружена сходимость размера се-

ти, обеспечивающая практически 100%-ное распознавание.

Выявленные особенности поведения параметров количества детекторов от минимального радиуса и значимости радиуса детекторов имеют локальные минимумы, что говорит о возможности дальнейшей оптимизации структуры иммунной сети.

Сходимость процедуры увеличения размера сети имеет асимптотический характер, т.е. сходится к конечному размеру сети.

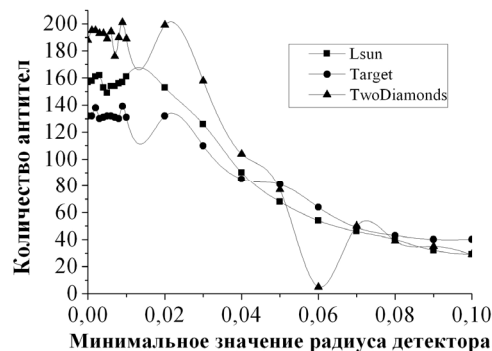


Рис. 9. Влияние минимального значения радиуса детектора на количество синтезируемых антител

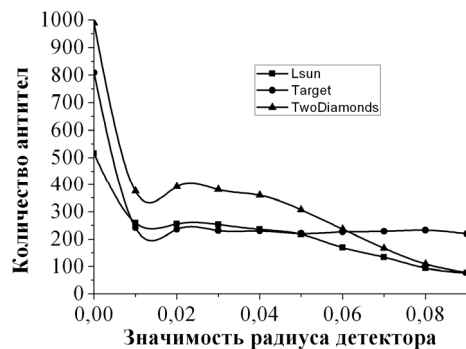


Рис. 10. Влияние значимости радиуса детектора на количество синтезируемых антител

Дальнейшие исследования планируется проводить по изучению свойств самоорганизации и практических приложений разработанного иммунного алгоритма.

1. Литвиненко В.И. Искусственные иммунные системы как средство индуктивного построения оптимальных моделей сложных объектов // Проблемы управления и информатики. – 2008. – №3. – С. 43–61.

Окончание на стр. 85.

2. Давидов М.В., Никольський Ю.В., Пасічник В.В. Вибір ефективного методу опрацювання зображень на основі еталону для ідентифікації елементів жестової мови // Вісн. Харківського нац. ун-ту радіоелектроніки. Сер. «АСУ і прилади автоматики». – 2008. – № 139. – С. 59–68.
3. Годич О.В., Никольський Ю.В., Щербина Ю.М. Застосування штучної нейронної мережі типу SOM для розв'язування задачі діагностування // Вісн. Нац. ун-ту «Львівська політехніка». – 2002. – № 464. – С. 31–43.
4. Дослідження ефективності алгоритмів навчання мереж Кохонена / В.В.Пасічник, О.В.Годич, Ю.В.Никольський та ін. // УСiМ. – 2006. – № 2. – С. 63–80.
5. Jander Moreira, Luciano Da Fontoura Costa Neural-based color image segmentation and classification using self-organizing maps, 1996. – <http://mirror.impa.br/sibgrapi96/trabs/pdf/a19.pdf>
6. Campbell N.W., Thomas B.T., Troscianko T. Neural Networks for the Segmentation of Outdoor Images, International Conference on Engineering Applications of Neural Networks. – 1996. – P. 343–346.
7. Campbell N.W., Thomas B.T., Troscianko T. Segmentation of Natural Images Using Self-Organising Feature Maps, University of Bristol, 1996. – P. 223–232.
8. Brown D., Craw I., Lewthwaite J. A SOM Based Approach to Skin Detection with Application in Real Time Systems, University of Aberdeen, 2001. – http://www.bmva.ac.uk/bmvc/2001/papers/33/accepted_33.pdf
9. Akgul C.B. Cascaded self-organizing networks for color image segmentation, 2004. – http://www.tsi.enst.fr/~akgul/oldprojects/CascadedSOM_cba.pdf
10. Jiang Y., Chen K.-J., Zhou Z.-H. SOM Based Image Segmentation, Lecture Notes in Artificial Intelligence 2639, Springer, 2003. – P. 640–643.
11. Reyes-Aldasoro C.C. Image Segmentation with Kohonen Neural Network Self-Organising Maps, 2004. – <http://www.cs.jhu.edu/~cis/cista/446/papers/SegmentationWithSOM.pdf>
12. Wu Y., Liu Q., Huang T.S. An Adaptive Self-Organizing Color Segmentation Algorithm with Application to Robust Real-time Human Hand Localization // Proc. Asian Conf. on Computer Vision, Taiwan, 2000.
13. Kohonen T. Self-Organizing Maps. – Springer, 2001. – 501 p.
14. Ford A., Roberts A. Colour Space Conversions, 1998. – <http://www.poynton.com/PDFs/coloureq.pdf>
15. Hunt R.W.G. Measuring Colour Fountain Pr Ltd, 1998. – 336 p.
16. Hoffmann G. CIELab Color Space, 2003. – <http://www.fho-emden.de/~hoffmann/cielab03022003.pdf>
17. Hoffmann G. CIE Color Space, 2000. – <http://www.fho-emden.de/~hoffmann/ciexyz29082000.pdf>
18. Highdimensional data structure analysis using Self-Organising Maps / O. Hodych, Y. Shcherbyna, Y. Nikolskyi. et. al. // CADSM 9th Intern. Conf., 19–24 Feb. 2007. – P. 218–221.

© О.В. Годич, К.Н. Гуцин, Ю.В. Никольський, В.В. Пасічник, Ю.Н. Щербина, 2009

2. Философский энциклопедический словарь / Гл. редакция: Л.Ф. Ильичева, П.Н. Федосеев, С.М. Ковалев и др. – М.: Сов. Энциклопедия, 1983. – 840 с.
3. Ивахненко А.Г. Самообучающиеся системы распознавания и автоматического управления. – Киев: Техника, 1969. – 392 с.
4. Ивахненко А.Г., Степанко В.С. Помехоустойчивость моделирования. – К.: Наук. думка, 1985. – 216 с.
5. Ивахненко А.Г., Зайченко Ю.П., Димитров В.Д. Принятие решений на основе самоорганизации. – М.: Сов. радио, 1976. – 280 с.
6. Tarakanov A.O., Skormin V.A., Sokolova S.P. Immunocomputing: principles and applications New York: Springer-Verlag, 2003. – 193 p.
7. de Castro, L.N., Von Zuben, F.J. aiNet: An artificial Immune Network for Data Analysis // Data Mining: A Heuristic Approach / Eds. H.A. Abbass, R.A. Saker, C.S. Newton, Idea Group Publ., USA, Chapter XII. – 2001. – P. 231–259.
8. Timmis J., Neal M. A resource Limited Artificial Immune System for Data Analysis. Knowledge Based Syst. – 2001. – 14(3–4)– P. 121–130.
9. Thomas Knight, Jon Timmis. AINE: An Immunological Approach to Data Mining. ICDM 2001. – P. 297–304.
10. Wierzchoc S.T. Artificial Immune Systems. Theory and Appl. Akademicka Oficyna Wydawnicza EXIT. – (Polish) Warszawa, 2001. – 282 p.
11. Ultsch A. U*-Matrix: a Tool to Visualize Clusters in High Dimensional Data // Technical Report N 36, Department of Mathematics and Computer Science Philipps-University Marburg, 2003. – P. 290–303.
12. Adaptive radius immune algorithm for data clustering / G.B. Bezerra, T.V. Barra, L.N. de Castro et al. // Proc. of 4th Intern. Conf. on Artificial Immune Systems – ICARIS. 3627 of Lecture Notes in Comp. Sci. – New York: Springer-Verlag, 2005. – P. 290–303.

© В.И. Литвиненко, 2009