УДК 519-71

П.И. Бидюк, А.А. Фефелов, В.И. Литвиненко

ИММУННОСЕТЕВАЯ МОДИФИКАЦИЯ АЛГОРИТМА ОТРИЦАТЕЛЬНОГО ОТБОРА ДЛЯ РЕШЕНИЯ ЗАДАЧИ ОБНАРУЖЕНИЯ АНОМАЛИЙ

Введение. Поведение технической системы зачастую характеризуется дискретными временными рядами наблюдений. В этом случае проблему обнаружения изменения свойств или аномалий в поведении можно сформулировать как задачу поиска недопустимых отклонений технических характеристик системы. Одним из методов решения данной задачи является иммунный алгоритм, основанный на механизмах отрицательного отбора [1–4]. Отрицательный отбор в иммунной системе используется для распознавания чужеродных антигенов путем удаления тех клеток (антител), которые реагируют на собственные антигены. Этот процесс называется распознаванием "свой-чужой". Алгоритм отрицательного отбора, обобщающий данное свойство иммунной системы, состоит из следующих основных компонентов: 1. определяется множество "своих" строк S длины l, состоящих из символов конечного алфавита. Эти строки моделируют нормальное поведение системы; 2. генерируется множество R детекторов, которые не распознают (не совпадают) ни с одной строкой из множества S. При сравнении строк используется правило частичного совпадения, согласно которому две строки считаются совпадающими, если они находятся друг от друга на расстоянии, не превышающем некоторого значения r, рассчитанного с использованием определенной метрики (например, Евклидовой метрики);3.поступающие данные контролируются путем непрерывного сопоставления с каждым из детекторов. Обнаружение совпадения хотя бы с одним из детекторов рассматривается как аномалия в поведении контролируемой системы. Данный алгоритм является робастным методом, основным достоинством которого является способность обнаруживать новые аномалии, а не искать их среди заранее известного множества событий данного типа.

Постановка задачи. Классический вариант алгоритма отрицательного отбора обладает одним существенным недостатком. Генерация множества детекторов R в фазе обучения происходит случайно, вследствие чего заранее невозможно определить минимально необходимое количество детекторов, которое будет обеспечивать максимальное качество распознавания. Увеличение количества детекторов ведет к замедлению фазы распознавания, а его уменьшение – к ухудшению качества работы алгоритма, т.к. увеличивается вероятность образования "полостей", являющихся областями в пространстве "чужих", которые не распознаются ни одним из детекторов. Таким образом, задачей данного исследования является разработка усовершенствованного метода генерации детекторов, способного адаптивно подбирать их настройки, количество и распо-

ложение. В качестве такого метода в работе предложено использовать искусственную иммунную сеть.

Искусственные иммунные сети (ИИС). Теория иммунной сети утверждает, что даже в отсутствие внешнего раздражителя (болезнетворных микроорганизмов), иммунная система находится в постоянном движении. Ее клетки способны взаимодействовать не только с антигенами, но и сами с собой, т.е. распознавать себе подобные антитела. Как результат этого взаимного распознавания молекул антител, возникает связанная сеть в пределах иммунной системы. Ее называют иммунной сетью. Искусственная иммунная сеть [5,8] может быть представлена в виде графа, который состоит из множества узлов - клеток сети (антител) и множества взвешенных ребер, означающих связи между клетками. Значение веса ребра соответствует аффинности связи клеток друг с другом. В иммунных сетях различают два вида аффинности: а) аффинность связи “антиген-антитело” (*Ag-Ab*) – степень различия; б) аффинность связи “антитело-антитело” (*Ab-Ab*) – степень подобия. Формально алгоритм иммунной сети можно представить следующим образом:

$$immNET = (P^l, G^k, l, k, m_{Ab}, \delta, f, I, \tau, AG, AB, S, C, M, n, d, H, R) \qquad (1)$$

где P^l – пространство поиска (пространство форм); G^k– представление пространства; l – длина вектора атрибутов; k – длина рецептора клетки; m_{Ab} – размер популяции клеток; δ - функция экспрессии; f – функция аффинности; I – функция инициализации начальной популяции клеток сети; τ- условие завершения работы алгоритма; AG – подмножество антигенов; AB – популяция клеток сети (антител); S – оператор селекции; C – оператор клонирования; M – оператор мутации; n – количество лучших клеток, отбираемых для клонирования; d – количество худших клеток, подлежащих замене новыми; H – оператор клонального удаления; R– оператор сжатия сети.

Блок-схема алгоритма иммунной сети показана на рисунке 1.

Пошаговая реализация алгоритма представлена ниже.

Шаг 1. *Инициализация*.

Шаг 1.1. Создание начальной популяции клеток памяти (M_R).

Шаг 1.2. Создание популяции антител (AB).

Шаг 2. *Антигенное присутствие*. Начиная с этого блока, алгоритм осуществляет по одному проходу для каждого антигена.

Шаг 2.1. *Вычисление аффинности*. Вычисляется аффинность всех клеток памяти $m_j, \quad m_j \in M_R$ для очередного антигена $Ag_i, \quad Ag_i \in AG$ и выбирается одна лучшая клетка m_b.

Шаг 2.2. *Клонирование*. Выбранная клетка памяти клонируется пропорционально своей аффинности с образованием популяции клонов M_C.

Шаг 2.3. *Созревание аффинности*. Производится мутация клонов из M_C. Измененные клоны добавляются к популяции антител, т. е. $AB \leftarrow AB \cup M_C$. Вычисляется аффинность популяции антител AB с антигеном Ag_i.

Шаг 2.4. *Метадинамика*. Производится клональное удаление нестимулируемых клеток в соответствии с порогом σ_d.

Шаг 2.5. Повторное клонирование части антител из популяции AB с образованием популяции клонов M_C и переход к шагу 2.3., если средняя аффинность популяции AB ниже заданного порогового значения.

Шаг 2.6. Из популяции AB выбирается клетка-кандидат (лучшее антитело) в популяцию клеток памяти Ab_b.

Шаг 2.7. Переход к шагу 3, если $f(Ab_b, Ag_i) < f(m_b, Ag_i)$.

Шаг 2.8. Добавление антитела Ab_b в популяцию M_R

Шаг 2.9. *Межклеточное взаимодействие*. Определяется аффинность взаимодействия всех клеток популяции M_R друг с другом, т. е. $f(m_i, m_j), \quad m_i, m_j \in M_R$.

Шаг 2.10. *Сжатие сети*. Удаляются распознающие друг друга клетки популяции M_R в соответствии с заданным порогом σ_s.

Шаг 3. Проверка выполнения условия останова алгоритма и переход к шагу 2, если условие останова не выполняется.

В данном типе алгоритма оператор H использует пороговый коэффициент гибели (σ_d) как управляющий параметр, уменьшая размер сети за счет удаления нестимулируемых клеток. Оператор сжатия сети R использует пороговый коэффициент сжатия σ_s как управляющий параметр, также уменьшая размер сети за счет удаления самораспознанных (подобных) клеток. Таким образом, работа иммунной сети основана на смене этапов активации и сжатия. Во время активации сеть создает большое количество клеток, которые обрабатываются клональным алгоритмом (т.е. вычисление аффинности, мутация). Во время сжатия (супрессии), сеть уничтожает те клетки, которые находятся слишком близко друг к другу, т.е. распознают друг друга и те клетки, которые находятся слишком далеко от антигенов (данных), т.е. плохо распознают антигены. Такое чередование этапов активации и сжатия приводит к тому, что сеть за счет изменения своей топологии формирует образ распознаваемых данных (так называемый внутренний образ антигена).

Модификация фазы обучения алгоритма отрицательного отбора. Решение конкретной задачи при помощи обобщенного алгоритма требует описания специфической реализации некоторых операторов и функций. В данном случае иммунная сеть использует вещественное кодирование антител (рис. 2), при котором для вычисления расстояния применяется Евклидовая метрика. При этом антитела формируют вокруг себя l-мерную радиальную область распознавания с радиусом r, который называется кросс-реактивным порогом. Как показано на рисунке 2, кросс-реактивный порог включен в состав структуры антитела, что дает возможность адаптивной настройки его значения. Таким образом, иммунная сеть заполняет пространство "чужих" распознающими гиперсферами разного радиуса, что дает возможность его более полного покрытия.

Для расчета значений аффинности связи "антиген-антитело" используется следующее соотношение:

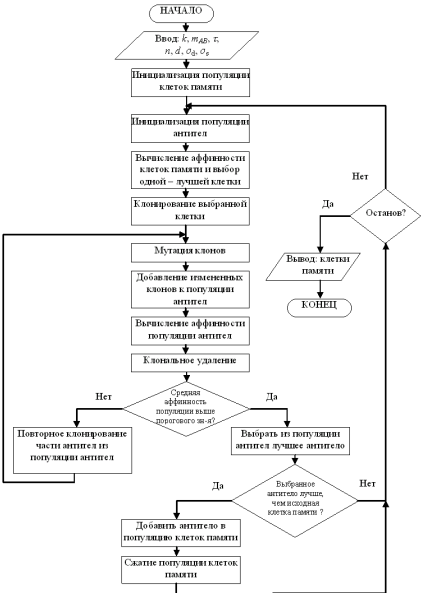


Рис. 1 – Блок-схема алгоритма иммунной сети

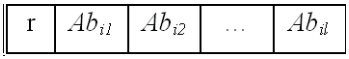


Рис. 2 – Представление антитела иммунной сети

 ISSN 1562-9945

$$f_{Ab-Ag} = \frac{k_r}{r} + D_{E(Ab-Ag)} \tag{2}$$

где r – кросс-реактивный порог антитела (детектора); k_r – коэффициент значимости кросс-реактивного порога (параметр настройки алгоритма). Параметр k_r является очень важным параметром обучения. Он управляет робастностью получаемого решения. Увеличение этого параметра заставляет иммунную сеть поддерживать детекторы большего радиуса, что дает более грубое, но при этом более устойчивое решение. Однако чрезмерное увеличение k_r отрицательно сказывается на точности решения. На рисунке 3 демонстрируется влияние параметра k_r на способ генерации детекторов.

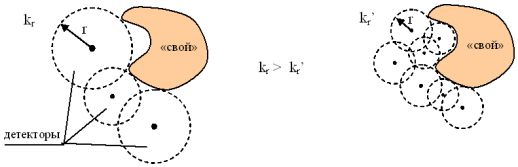


Рис. 3 – Влияние значения параметра k_r на способ генерации детекторов

Сжатие иммунной сети осуществляется на основании самораспознавания клеток, которое численно выражается в виде афинности связей антител друг с другом. Для расчета значений аффинности связи “антитело-антитело” предложена следующая формула:

$$f_{Ab-Ab} = -\frac{D_{E(Ab_1-Ab_2)} - (r_{Ab_1} + r_{Ab_2})}{2 \cdot \min(r_{Ab_1}, r_{Ab_2})} \tag{3}$$

При этом возможна следующая интерпретация значений f_{Ab-Ab}: $\leqslant 0$ – распознающие гиперсферы детекторов не перекрываются. Этот вариант не требует сжатия, т.к. антитела не распознают друг друга (рис. 4 а); $(0, 1)$ – гиперсферы перекрываются оболочками, а само значение является степенью перекрытия (рис. 4 б). При этом сжатие осуществляется в зависимости от величины параметра порога сжатия σ_s, который является параметром алгоритма обучения; $\geqslant 1$ – гиперсфера меньшего радиуса (r) полностью находится внутри гиперсферы большего радиуса (рис. 4 в). В данном случае сжатие, безусловно, необходимо, т. к. наблюдается избыточность распознающих элементов.

В данной реализации оператор клонального удаления H действует только на те антитела, которые распознают хотя бы один антиген. Таким образом, результирующая иммунная сеть в конце каждого поколения гарантировано не содержит детекторов, распознающих “свои” антигены.

Решение задачи обнаружения аномалий. В общем виде задачу обнаружения аномалий можно представить следующим образом [6,7]. Пусть дан дискретный временной ряд значений переменной процесса: $y_1, y_2, \ldots, y_n$. Предполагается, что дискретный контроль значений пере-

Рис. 4 – Различное взаимное расположение распознающих гиперсфер детекторов

менной $y(t)$ производится на интервале $[t_1, t_n]$. На данном ряде выбирается временное окно шириной kизмерений ($k < n$). Значения временного ряда, попавшие внутрь временного окна образуют вектор признаков $Y_t\,(y_t, y_{t-1}, \ldots, y_{t-k+1})$. Вектору признаков ставится в соответствие принадлежность к одному из двух классов: $normal(1)$, если фрагмент временного ряда, соответствующий данному вектору не содержит аномалий (т.е. относится к классу “своих”), $abnormal(0)$ в противном случае. Временное окно смещается (скользит) вдоль временного ряда на величину Δl шагов, образовывая множество векторов, которые делят признаковое пространство на две части: с аномалиями и без них. Задача заключается в отнесении любого вектора Y_j, образованного скользящим временным окном, к одному из двух выделенных классов. С другой стороны, если рассматривать изучаемый процесс как динамическую систему, то получаемое при помощи скользящего окна множество векторов представляет собой восстановленный фазовый портрет динамической системы, а сами вектора – точки, принадлежащие фазовой траектории этой системы. При нормальном поведении, данная траектория может восприниматься как эталонный образ, всякое отклонение от которого является признаком аномалии (рис. 5). Следует подчеркнуть, что при использоваии отрицательного отбора нет необходимости включать в обучающую выборку вектора, соответствующие аномальному поведению, что дает возможность фиксировать любые, даже не известные аномалии.

Экспериментальные исследования. Для первого эксперимента [9] выбран периодический сигнал, график которого представлен на рисунке 6а. Обучающая выборка насчитывает 200 значений ряда. Обучающий сигнал не содержит аномалий. Для создания обучающей выборки использовалось скользящее окно шириной в два значения. В этом случае признаковое пространство будет двухмерным: (y_t, y_{t+1}). Внешний вид восстановленного фазового портрета показан на рисунке 6б.

Как видно из рисунка, фазовая траектория такой динамической системы не является всюду плотной, т. к. обучающая выборка ограничена 200 значениями. Показанная на рисунке 6б геометрическая фигура является отображением данных класса “свой” и может использоваться как обучающий образ для искусственной иммунной сети. Единственной трудностью при этом является наличие просветов на восстановленной траектории, что для корректного распознавания нормального сигнала требует получения устойчивого решения. Ниже показаны результаты обу-

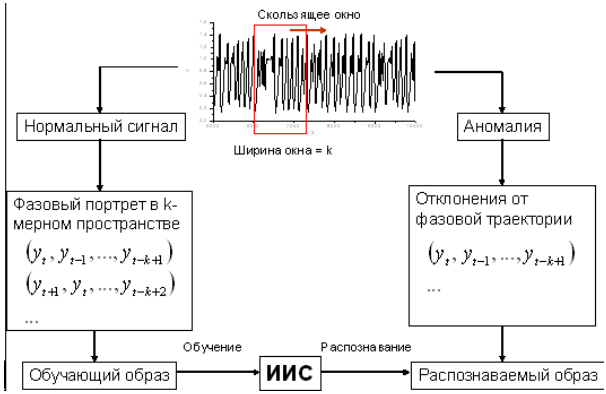
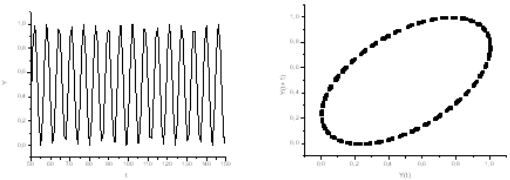


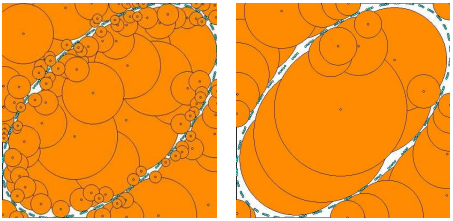
Рис. 5 – Задача обнаружения аномалий



а б

Рис. 6 – Экспериментальный сигнал; а – график сигнала, не содержащего аномалий; б – восстановленный фазовый портрет сигнала, не содержащего аномалий

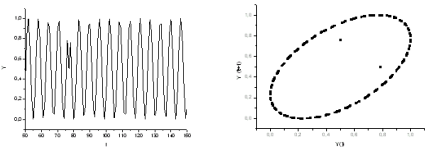
чения искусственной иммунной сети для двух значений параметра значимости кросс-реактивного порога k_r, управляющего устойчивостью решения. Рисунок 7 а демонстрирует менее устойчивое решение, т.к. часть плоскости, соответствующая пропущенным элементам фазовой траектории перекрыта детекторами. Это приведет к их ложному срабатыванию. Решение на рисунке 7 б является более устойчивым, и взято для дальнейшего тестирования.



а б

Рис. 7 – Конфигурация обученной иммунной сети для двух значений параметра k_r: а – при $k_r = 0.01$, б – при $k_r = 0.1$

Далее в сигнал была введена локальная аномалия (рис. 8 а), что нашло свое отражение на фазовом портрете, как показано на рисунке 8 б. В данном случае 2 точки отклонились от фазовой траектории и были распознаны соответствующими детекторами.

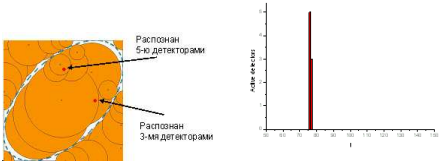


а б

Рис. 8 – Тестовый сигнал с локальной аномалией и его фазовый портрет

При обнаружении вполне возможно многократное распознавание аномального явления сразу несколькими детекторами (рис. 9 а).

Гистограмма на рисунке 9б демонстрирует количество детекторов, которые активируются в месте обнаружения аномалии при движении сканирующего окна по временному ряду сигнала. Второй эксперимент отражает возможность обнаружения аномальных изменений в параметрах математических зависимостей, описывающих поведение системы. Этот способ весьма полезен для раннего обнаружения дрейфа контролируемых параметров технической системы. В данном случае в качестве исследуемой модели было взято логистическое уравнение $y_{t+1} = \lambda \cdot y_t(1-y_t)$ с нормальным параметром $\lambda = 4.0$. График нормального сигнала показан на рисунке 10а. В какой-то момент времени, значение параметра



а б

Рис. 9 – а – распознавание аномальных векторов иммунной сетью; б – гистограмма активации детекторов

λскачкообразно изменилось до 3.6, а затем опять восстановилось до нормальной величины, что демонстрирует рисунок 10 б.

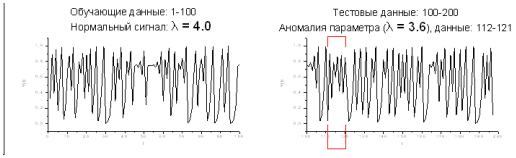


Рис. 10 – а – нормальный сигнал с параметром $\lambda = 4.0$; б – сигнал, содержащий аномалию параметра

Структура обученной искусственной иммунной сети показана на рисунке 11а. Иммунная сеть зафиксировала аномальные изменения параметра, активировав соответствующие детекторы (рис. 11б).

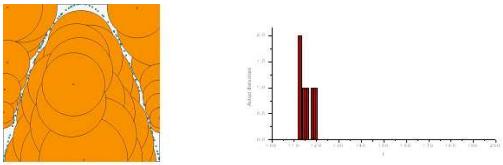


Рис. 11 – Рисунок 11. а – структура обученной искусственной иммунной сети; б – активация детекторов в месте возникновения аномалии

Как видно на рисунке 11 б, активация детекторов происходит без задержки во времени, что является существенным достоинством данной методики.

Выводы. Разработан модифицированный алгоритм отрицательного отбора для решения задачи обнаружения аномалий в работе сложных технических систем. Этот алгоритм для обучения использует механизмы работы искусственных иммунных сетей. Отличительной особенностью алгоритма является модификация процесса обучения, благодаря которой реализована возможность адаптивного подбора настроек, количества и расположения детекторов. Экспериментальные исследования показали высокую эффективность предложенного алгоритма, которая

проявляется в его устойчивости, благодаря адаптивному подбору значения кросс-реактивного порога; оптимальности, вследствие адаптивной настройки размера иммунной сети, т.е. количества необходимых детекторов; точности, вследствие уменьшения количества и размеров образуемых "полостей".

Литература

1. Искусственные иммунные системы и их применение / Под ред. Д. Дасгупты. Пер. с англ. под ред. А. А. Романюхи. – М.: ФИЗМАТЛИТ, 2006.
2. S. Forrest, A. S. Perelson, L. Allen, and R. Cherukuri. Self-nonself discrimination in a computer. In Proceedings of the IEEE Symposium on Research in Security and Privacy, IEEE Computer Society Press, Los Alamitos, CA, pp. 202-212, 1994.
3. Литвиненко В. И. Иммунный классификатор для решения задач бинарной классификации (теоретические основы) //Системні технології. регіональний міжвузівський збірник наукових праць. Випуск 1(42). –Дніпропетровськ, 2006. с.114-130
4. Литвиненко В. И. Иммунный классификатор для решения задач классификации (практические аспекты)// Системні технології. Регіональний міжвузівський збірник наукових праць. Випуск 5(46). –Дніпропетровськ, 2006. с.113-126.
5. D. Dasgupta, S. Forrest. An anomaly detection algorithm inspired by the immune system. In: D. Dasgupta (eds) Artificial Immune Systems and Their Applications, Springer-Verlag, pp. 262-277, 1999.
6. P. D'haeseleer, S. Forrest, P. Helman. An immunological approach to change detection: algorithms, analysis and implications. In Proceedings of the IEEE Symposium on Computer Security and Privacy, IEEE Computer Society Press, Los Alamitos, CA, 1996.
7. Литвиненко В.І. Вирішення задач классифікації з використанням механізмів ідіопатичної мережі// Наукові праці Миколаївського державного гуманітарного університету імені Петра Могили Випуск 44, том 57, 2006 р. с.136-146.
8. Фефелов А.А., Литвиненко В.И., Бидюк П.И. Модификация алгоритма отрицательного отбора на основе механизмов искусственной иммунной сети для решения задач обнаружения аномалий // Збірник наукових праць у п'яти томах другої міжнародної наукової конференції Інтелектуальні системи прийняття рішень та прикладні аспекти інформаційних технологій / Євпаторія 2007, Том 3, с.73-78.

 ISSN 1562-9945