

МЕТОД КЛАСТЕР-АНАЛІЗУ НА ОСНОВІ ГІБРИДНОЇ МОДЕЛІ ШТУЧНОЇ ІМУННОЇ МЕРЕЖІ З ДИНАМІЧНИМ РЕГУЛЮВАННЯМ ПАРАМЕТРІВ ПРОЦЕСУ НАВЧАННЯ

ЗАЙЦЕВ С.О., СУББОТИН С.О.

Проводиться аналіз впливу параметрів керування на ефективність процесу навчання гібридної моделі штучних імунних мереж. Експериментально встановлюється характер їх впливу на процес навчання мережі, на основі чого запропоновано модифікацію методу навчання гібридної моделі з динамічним регулюванням цих параметрів.

1. Вступ

Кластерний аналіз являє собою інструмент, що дозволяє виділити з вибірки даних скупчення подібних спостережень і узагальнити їх, замінюючи кожен групу компактно розміщених екземплярів еталоном кластеру. Кластерний аналіз широко застосовується при побудові діагностичних та розпізнавальних моделей, оскільки, з одного боку, дозволяє замінити навчальну вибірку великої розмірності множиною еталонів меншої розмірності, а, з іншого боку, результати кластерного аналізу дозволяють представити в моделі топологію навчальних даних (наприклад, для нейро-нечітких моделей кластерний аналіз дозволяє виділити узагальнені правила та сформулювати нечіткі терми).

В [1] запропоновано гібридну модель штучних імунних мереж (ШІМ), яка основана на мережі aiNet [2]. Зокрема, в гібридній моделі було реалізовано механізм адаптації до даних на вході мережі, характерний для aiNet, та принцип визначення області мережі, яка підлягає зміні в процесі адаптації, запозичений з методу нейронного газу, що розширюється [3].

Експерименти з використанням гібридної моделі виявили, що вдалося підвищити швидкість збіжності методу навчання мережі. Проте запропонований метод навчання ШІМ [1] передбачає наявність великої кількості параметрів керування, що повинні заздалегідь задаватися експертом. При цьому вони не відображають природу задачі, що вирішується. Таким чином, якість процесу навчання мережі сильно залежить від того, наскільки успішно користувачем були встановлені параметри керування на початку навчання мережі.

Перспективним напрямком подальшого розвитку гібридної моделі ШІМ є дослідження впливу параметрів на процес навчання мережі. Автоматизація регулювання цих параметрів залежно від поточного стану мережі здатна скоротити час навчання мережі та раціонально використовувати обчислювальні ресурси.

Мета роботи полягає в модифікації методу навчання гібридної моделі ШІМ шляхом реалізації можливості динамічного регулювання параметрами керування процесу навчання, що дозволить самоорганізувати мережу.

2. Постановка задачі

У загальному вигляді задача кластер-аналізу може бути подана так. Нехай X – множина об'єктів, Y – множина міток кластерів. Задано метрику – функцію відстані між двома об'єктами $p(x, x')$, $x \in X$, $x' \in X$. Існує скінченна навчальна вибірка об'єктів $X^m = \{x_1, x_2, \dots, x_m\}$, $X^m \subset X$, де x_i – i -й екземпляр навчальної вибірки. Необхідно розбити вибірку X^m на підмножини (кластери), що не перетинаються, так, щоб кожний кластер складався з об'єктів, близьких за метрикою p , а об'єкти різних кластерів суттєво відрізнялися. При цьому кожному об'єкту $x \in X^m$ приписується номер кластера $y \in Y$. Слід вказати, що методи, які будуть розглядатися нижче, лише виділяють області кластерів, проте не ставлять їм мітки. Задача призначення міток кластерам є тривіальною і не розглядається у даній статті.

3. Штучна імунна мережа aiNet

Імунна пам'ять ШІМ aiNet [4] формується шляхом взаємодії імунних клітин. Вплив однієї імунної клітини на іншу може або стимулювати її, або пригнічувати.

ШІМ aiNet можна подати у вигляді зваженого графа, вузли якого називають клітинами, а ребра визначають зв'язок між клітинами. Кожному ребру графа відповідає деяке числове значення, що називається вагою або силою зв'язку.

В основі роботи ШІМ лежить постійна адаптація структури мережі до входних даних [5]. Процес навчання мережі супроводжується формуванням імунної відповіді шляхом клональної селекції щоразу при появі нових антигенів з навчальної вибірки на вході моделі. Для регуляції розміру популяції клонів використовують механізми клональної та мережної супресії. Мережа прагне скоротити кількість клітин у пам'яті та зробити їх більш специфічними, виконуючи таким чином компресію інформації у пам'яті.

Метод навчання моделі aiNet включає такі етапи.

1. Задано навчальну вибірку антигенів Ag . Ініціалізувати імунну пам'ять моделі $M = \emptyset$. Для кожного антигену з вибірки $ag_i \in Ag$ об'ємом n , де $i = 1, \dots, n$:

– обчислити афінність антигену кожному елементу з набору антитіл $ab_j \in Ab$, де $j = 1, \dots, m$; m – об'єм набору антитіл. Значення афінності обернено пропорційно евклідовій відстані між антигеном і антитілом, і обчислюється як:

$$\text{aff}(ag_i, ab_j) = \frac{1}{\|ab_j - ag_i\|};$$

– з набору антитіл Ab обрати підмножину Ab^* , що складається з q антитіл з найбільшим значенням афінності;

– сформуванню множини клонів C з q антитіл пропорційно їхньому рівню афінності: чим вище значення афінності антитіла антигену, тим більшу кількість клонів необхідно створити;

– клони з множини C проходять етап керованої мутації, результатом якого є множина мutowаних клонів C^* :

$$C_j^* = C_j + \alpha_j \cdot (ag_i - ab_j),$$

де α_j – коефіцієнт, що визначає ступінь мутації антитіла ab_j . Ступінь мутації кожного клону обернено пропорційна афінності його батьківської клітини антигену;

– виконати клональну та мережну супресію.

2. Визначити взаємну афінність кожної пари антитіл $s(ab_p, ab_r) = \|ab_p - ab_r\|$, де $p = 1, \dots, q$, $r = 1, \dots, q$.

3. Виконати мережну супресію – видалити антитіла з афінністю s , що перевищує поріг σ_s .

4. Якщо критерій зупину не досягнуто – перейти до етапу 1, в протилежному випадку – зупин.

На практиці використовують такі критерії зупину під час навчання мереж $aiNet$:

- досягнення максимальної кількості ітерацій;
- перевищення максимальної популяції мережі;
- зростання помилки впродовж k останніх послідовних ітерацій.

Процедура клональної та мережної супресії включає в себе такі етапи:

– обчислити взаємну афінність антигену ag_i клонам антитіл з C^* за формулою:

$$d_{ij} = \frac{1}{\|C_j^* - ag_i\|};$$

– відібрати δ клонів з найкращою афінністю d з C^* , занести їх до клональної пам'яті M_j ;

– видалити з клональної пам'яті M_j елементи з афінністю d , нижчою за деяке порогове значення σ_d ;

– виконати клональну супресію – визначити взаємну афінність клонів за формулою:

$$s(ab_j, ab_k) = \|ab_j - ab_k\|,$$

де ab_j, ab_k – антитіла з клональної пам'яті M_j ;

– видалити ті клони, для яких значення афінності s перевищує заданий поріг σ_d ;

– додати решту клонів до імунної пам'яті $M = M \cup M_j$.

Результатом навчання мережі є набір антитіл, які відображають внутрішню структуру антигенів з навчальної вибірки.

4. Гібридна модель ШІМ

В гібридній моделі ШІМ [1] антитіла характеризуються локальною помилкою, яка змінюється щоразу, коли антитіло має достатню афінність антигену з навчальної вибірки. Це дозволяє використовувати значення локальної помилки як параметр керування на етапі клональної селекції, тобто формувати популяцію клонів, спираючись не на афінність антитіла поточному антигену, а на його локальну помилку. Локальна помилка акумулює в собі результат попередніх взаємодій антитіла з антигенами, і її високе значення свідчить про те, що це антитіло перебуває в області з високою концентрацією антигенів, а тому необхідно провести клональну селекцію для більш точного покриття цієї області. З іншого боку, низька локальна помилка говорить про те, що в околі даного антитіла або знаходиться невелика кількість антигенів, або існує достатня кількість антитіл, щоб їх розпізнавати. Нижче наведено метод навчання гібридної моделі.

1. Ініціалізувати імунну пам'ять M двома випадковими антитілами з нульовою локальною помилкою.

2. Для кожного антигену з навчальної вибірки $ag \in Ag$ виконати:

- відібрати з імунної пам'яті M деякий процент δ антитіл Ab^* з найбільшою афінністю антигену ag ;
- накопичити для відібраних антитіл Ab^* локальну помилку

$$\varepsilon_{ab} = \varepsilon_{ab} + \|ab - ag\|,$$

де $ab \in Ab^*$, ab – антитіло, що розглядається; ε_{ab} – його локальна помилка;

– сформуванню множини клонів C для кожного з відібраних антитіл Ab^* пропорційно їх локальній помилці. Антитіла з більшою локальною помилкою будуть виробляти більше число клонів за формулою:

$$N_c = N_{\max} \left(1 - \frac{1}{\varepsilon_{ab}}\right),$$

де $N_{\max}$ – максимально можлива кількість клонів одного антитіла на одній ітерації;

– клони з множини C підлягають мутації. Ступінь мутації також прямо пропорційна локальній помилці батьківського антитіла;

– виконати клональну супресію і мережну супресію;

– видалити ті клони, локальна помилка яких перевищує максимально допустиму $\varepsilon_{\max}$;

– решту клонів C^* , що залишилися після етапу супресії, занести до імунної пам'яті $M = M \cup C^*$;

– встановити рівень локальної помилки отриманих клонів C^* рівним нулю.

3. Якщо критерій зупину не досягнуто – перейти до етапу 2, в протилежному випадку – зупин.

Сумарна помилка моделі розраховується за формулою: $\varepsilon^* = \sum_{ab \in Ab} \varepsilon_{ab}$.

Коли всі антитіла будуть достатньо близько знаходитися до антигенів з навчальної вибірки, сумарна помилка мережі буде залишатися практично незмінною. Дана властивість використовується в гібридній моделі як критерій зупину.

Розглянемо вплив основних параметрів керування методу на його роботу. Під параметрами керування тут слід розуміти δ (кількість антитіл, що допускаються до клонування), $\varepsilon_{\max}$ (максимально допустима локальна помилка антитіла), $N_{\max}$ (максимальна кількість клонів), α (максимальний ступінь мутації). Коефіцієнти, що відповідають за проведення клональної і мережної супресії, детально розглянуті в [6] і не мають специфічного впливу на модель.

5. Модифікована гібридна модель ШІМ

З метою підвищення швидкості збіжності процесу навчання моделі авторами пропонується додати до методу навчання критерій, на основі якого можна було б змінювати параметри керування в процесі навчання моделі. Таким критерієм може виступати зміна середньої помилки моделі, причому середня помилка повинна враховувати не лише локальні помилки антитіл у складі моделі, але і їх вік, оскільки старші антитіла зазвичай мають більшу локальну помилку. Вік антитіла $\text{age}(ab)$ визначається як кількість ітерацій методу з моменту створення антитіла.

Середня локальна помилка моделі визначається як:

$$\bar{\varepsilon} = \frac{\sum_{ab \in Ab} \varepsilon_{ab}}{|\text{age}(ab)|},$$

де $|Ab|$ – потужність множини антитіл Ab .

Модифікований метод навчання гібридної моделі ШІМ включає додатковий етап (виконується після етапу 2). Нехай $\bar{\varepsilon}$ – локальна помилка на попередній ітерації. Визначити середню локальну помилку на поточній ітерації $\bar{\varepsilon}'$. Якщо $\bar{\varepsilon}' > \bar{\varepsilon}$, тоді покласти:

$$\delta = \max(\delta - 1, 2),$$

$$\text{age}(ab) = \min(\text{age}(ab) + 1, |Ag|).$$

У протилежному випадку покласти:

$$\delta = \delta + 1, \text{age}(ab) = \max(\text{age}(ab) - 1, \frac{|Ag|}{N_c}),$$

де N_c – очікувана кількість кластерів. Покласти $\bar{\varepsilon}' = \bar{\varepsilon}$.

Такий етап дозволить варіювати параметри керування залежно від стану моделі. Введення на цьому етапі граничних значень пояснюється тим, що розмаїття популяції можна забезпечити, коли існує хоча б два різних антитіла, які можуть створювати клони. Максимальний та мінімальний вік антитіл обрано залежно від імовірності відбору антитіла для створення клонів.

Завдяки модифікованому оператору мутації в гібридному методі можна уникнути коефіцієнту мутації α . Оскільки клони створюються вздовж відрізка, що з'єднує батьківське антитіло з антигеном, який викликає процес мутації, то максимально можливе значення α призведе до створення клону, координати якого збігатимуться з антигеном в просторі ознак. Тоді положення мутованого клону C можна визначити як:

$$C = ab + (ag - ab)\varepsilon_{ab} \cdot \text{rand}(),$$

де ag – антиген; ab – антитіло; $\text{rand}()$ – функція, яка повертає псевдовипадкове число з діапазону $[0,1)$; ε_{ab} – локальна помилка батьківського антитіла ab .

Таким чином, в разі використання модифікованого методу навчання потрібно задати лише один параметр керування – очікувану кількість кластерів N_c .

6. Експерименти та результати

Модифікований метод навчання гібридної ШІМ було реалізовано мовою Python. Розроблене програмне забезпечення дозволяє проводити кластер-аналіз даних, а також подавати в графічній та текстовій формах інформацію про динаміку середньої локальної помилки моделі та параметрів керування процесом її навчання.

За допомогою розробленого математичного забезпечення вирішувалася задача діагностування лопаток газотурбінних авіадвигунів [7]. Вибірка експериментально отриманих спостережень містила екземпляри, що характеризувалися спектрами потужності вільних затухаючих коливань лопаток після ударного збудження. Виходячи з цих даних, необхідно було виконати класифікацію лопаток на групи кондиційних та дефектних (некондиційних). Кожний екземпляр вибірки характеризувався 10240 ознаками. Вибірка складалася з 32 екземплярів. Для навчання моделі використовувалися згортки, що дозволило скоротити розмірність задачі до 100 ознак. В результаті роботи моделі було виділено 6 кластерів, і після призначення кластерам міток досягнуто 91% точності класифікації.

Отримані результати кластер-аналізу можуть бути використані для формування навчальної множини при побудові діагностичних моделей на основі нейронних мереж шляхом виділення з вхідної вибірки екземплярів, що відповідають еталонам кластерів, та екземплярів, розташованих на границях кластерів. Також центри і границі кластерів, виділені за допомогою запропонованого методу, можуть бути спроектовані

на осі ознак та використані для формування нечітких термів вхідних ознак, що дозволить визначити для кожного кластеру нечітке продукційне правило. Сукупність таких правил складатиме нечітку базу знань. У свою чергу, нечітка база знань може бути використана для побудови діагностичної моделі на основі нейро-нечіткої мережі шляхом відображення нечітких правил в структуру мережі.

На рисунку наведено графік динаміки середньої локальної помилки моделі та розміру популяції антитіл під час навчання.



Динаміка середньої локальної помилки гібридної моделі та кількості антитіл в моделі

Як видно з рисунка, модифікований метод спочатку створює надмірне число антитіл, що призводить до зростання популяції без суттєвого зменшення середньої локальної помилки. Оскільки метод автоматично адаптує керуючі параметри, це призводить до зменшення популяції та зниження середньої локальної помилки.

7. Висновки

З метою вирішення актуального завдання автоматизації кластер-аналізу даних розроблено математичне забезпечення на основі методів імунокомп'ютингу.

Наукова новизна роботи полягає в тому, що дістав подальшого розвитку метод кластеризації, який використовує гібридну модель ШІМ шляхом адаптації його параметрів керування, що дозволило самоорганізувати процес навчання моделі та за рахунок цього підвищити швидкість збіжності методу навчання.

Практична цінність роботи полягає в тому, що розроблено програмне забезпечення для проведення кластер-аналізу на основі запропонованої гібридної моделі ШІМ.

Література: 1. Зайцев С.А., Субботин С.А. Кластерный анализ с использованием гибридной модели на основе искусственной иммунной сети // Бионика интеллекта. 2010. №3. С. 70-75. 2. Tang N., Rao Vemuri V. An artificial immune system approach to document clustering // The 2005th ACM symposium on Applied computing: proceedings (13-17 March 2005). New Mexico. New York: ACM, 2005. P. 918-922. 3. Jirayusakul A., Auwatanamongkol S. A supervised growing neural gas algorithm for cluster analysis // International Journal of Hybrid Intelligent Systems. 2007. №4. P. 217-229. 4. Jerne N.K. Towards Network Theory of the Immune System // Ann. Immunol. 1974. №125C. P. 373-389. 5. De Castro L.N., Von Zuben F.J. An evolutionary immune network for data clustering // The Sixth Brazilian Symposium on Artificial Neural Networks: proceeding of the IEEE SBRN (22-25 November 2000). Rio de Janeiro. Zurich: Springer-Verlag, 2000. P. 84-89. 6. De Castro L.N. The immune response of an artificial immune network (aiNet) // The 2003th Congress on Evolutionary Computation: proceedings (8-12 December 2003). Canberra. Chicago: IEEE Press, 2003. P. 146-153. 7. Дубровин В.И., Субботин С.А., Богуслаев А.В., Яценко В.К. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей: монография. Запорожье: ОАО "Мотор-Сич", 2003. 279 с.

Надійшла до редколегії 14.06.2011

Рецензент: д-р техн. наук, проф. Бодянский С.В.

Зайцев Сергій Олексійович, аспірант кафедри програмних засобів Запорізького національного технічного університету. Наукові інтереси: інтелектуальні системи технічного діагностування, штучні імунні системи. Адреса: Україна, 69063, Запоріжжя, вул. Жуковського, 64, тел. (061) 769-82-67.

Субботін Сергій Олександрович, канд. техн. наук, лауреат премії Президента України, доцент, докторант Запорізького національного технічного університету. Наукові інтереси: інтелектуальні системи технічного діагностування, нейро-нечіткі мережі, оптимізація. Адреса: Україна, 69063, Запоріжжя, вул. Жуковського, 64, тел. (061) 769-82-67.