

УДК 519.6

МЕТОД КЛАССИФИКАЦИИ НЕТИПИЗИРОВАННЫХ ОБЪЕКТОВ НА ОСНОВЕ КАСКАДНОГО НЕЙРОСЕТЕВОГО ФИЛЬТРА И КОНЕЧНОГО ДЕТЕРМИНИРОВАННОГО АВТОМАТА

Кремлева Э.Ш.

Казанский национальный исследовательский технический университет имени
А.Н. Туполева – КАИ, г. Казань

Поступила в редакцию 10.02.2019, после переработки 15.03.2019.

В статье описан метод поэтапного определения класса многомерного объекта в случае, когда множество классов заранее не известно. Разработанный метод поэтапно решает сначала задачу выделения классов из изначально нетипизированного разнородного множества объектов, а затем производит классификацию произвольного нового объекта по выделенным классам. Выделение классов осуществляется на основе авторского алгоритма каскадной нейросетевой фильтрации, а классификация объектов – при помощи авторской модели на базе конечного автомата.

Ключевые слова: классификация, кластеризация, типы объектов, нейронная сеть, конечный автомат.

Вестник ТвГУ. Серия: Прикладная математика. 2019. № 1. С. 83–100.
<https://doi.org/10.26456/vtppmk527>

Введение

Задачи распознавания, или классификации, в настоящее время приобрели особую актуальность в связи с широким распространением компьютерных систем идентификации и проверки подлинности данных. В первую очередь это коснулось систем распознавания изображений: распознавание наземных объектов с борта БПЛА [1], идентификация автомобильных номеров [2], проверка подлинности подписи [3], идентификация человека по фотографии [4-5] и т.д. В других областях задачи классификации не менее актуальны: распознавание неисправных состояний технических систем [6-8], классификация психотипов личности в психологии человека [9], определение класса опасности химических веществ [10] и т.д.

Всем задачам классификации присуща общая черта: множество классов, или типов, к которым теоретически может быть отнесен каждый новый объект, задается заранее. Так, всем известная классификация психотипов по Гиппократу включает 4 класса (типов темпераментов) - «холерик», «сангвиник», «флегматик», «меланхолик». Классификация опасности химических веществ: «вещества малоопасные», «вещества умеренно опасные», «высокоопасные» и «чрезвычайно опасные».

Однако на практике предложенные классификации часто являются лишь условным упрощенным обобщением, и не способны описать реальные свойства объекта в полной мере. Так, психологи отмечают, что чистые психотипы присущи лишь психически ненормальным личностям, в то время как рисунок личности здорового человека гораздо сложнее, и не вписывается в типологию Гиппократов. При классификации химических веществ по степени опасности важным неучтенным фактором является происхождение опасности и характер его влияния на организм человека, что также делает общепринятую классификацию недостаточной [11-12]. Относительно бинарной классификации состояния технических систем недостаток очевиден — мало знать, что система неисправна, важно иметь представление о характере неисправности, при этом диагностика системы проводится опосредованно через измерения набора ее выходных характеристик [13-19].

В такой постановке исследователи заранее могут не иметь заданного набора классов объектов, и должны выделить классы самостоятельно из имеющегося множества данных. Например, при стендовых испытаниях газотурбинного двигателя (ГТД) заранее может быть неизвестно, с какими неисправностями столкнутся испытатели. Диагностика состояния ГТД производится по наблюдениям за косвенными признаками его функционирования (частоте вращения одного из роторов, степени повышения давления воздуха за компрессором, температуре газов перед или за турбиной и др.) [20-23], и для классификации неисправностей вначале необходимо провести серию испытаний, и сопоставить наборы наблюдаемых параметров реальным техническим недостаткам двигателя в каждом конкретном случае. Только после этого накопленный опыт позволит классифицировать неисправности в будущем по определенным на предварительном этапе классам.

Подобная ситуация, когда изначально объекты для классификации не связаны с определенными типами, и типы (классы) необходимо выделить в ходе предварительных экспериментов, часто возникает и в других областях, таких как экология, где классификацию экологического состояния природных объектов можно провести только на основе анализа мониторинговой информации [24-28], и др.

Назовем объекты, для которых изначально не указаны их типы (классы), нетипизированными.

Задача классификации для нетипизированных объектов, таким образом, будет включать в себя две подзадачи:

1. Предварительное определение классов, к которым в будущем необходимо будет отнести некоторый неисследованный объект,
2. Непосредственно классификация, относящая заданный объект к одному из классов, определенных на предыдущем этапе.

1. Метод определения классов объектов на основе нейросетевого каскада

На начальном этапе имеется некоторое «облако» данных, не содержащих в себе меток классов. Для выделения классов из такого облака предлагается сгруппировать имеющиеся разнородные данные в однородные компактные множества по принципу евклидовой близости векторов-объектов, каждое из которых и будет

представлять собой некоторый класс. Такая группировка может реализовываться на основе различных алгоритмов кластеризации [29-30]. Качество группировки может оцениваться по критериям SWC (Silhouette Width Criterion - коэффициент силуэта) — чем выше значение, тем однороднее данные внутри кластеров, коэффициент FM (Folkes and Mallows) — чем ниже коэффициент, тем более отличны кластеры друг от друга, а также на основе экспертного оценивания [31].

Для повышения качества кластеризации в работе [32] описывался алгоритм построения модели каскадного фильтра, представляющего собой иерархическую структуру кластеризующих алгоритмов-нейронных SOM-карт Кохонена. Как показали численные эксперименты, предложенный в работе алгоритм позволяет кластеризовать данные с произвольной заданной точностью.

На основе разработанного в [32] алгоритма предлагается новый численный метод выделения классов из множества нетипизированных многомерных объектов.

Введем обозначения:

- $\Omega_x = \{X\}$ — множество векторов исходных данных $X = (x_1, x_2, \dots, x_N)$, из которых выделяются классы,
- N — размерность вектора X (каждый вектор содержит значения N признаков объекта),
- S_p — SOM-карта с номером p .

Под SOM-картой понимается нейронная сеть парадигмы «Карта Кохонена» [33]. Каждая SOM-карта формирует (порождает) кластеры из множества данных в процессе обучения по алгоритму WTA [34]. При этом каждый кластер ассоциируется с одним или несколькими топологически близко расположенными нейронами SOM-карты. Таким образом можно считать, что SOM-карта *содержит в себе кластеры* (их прототипы в виде набора нейронов). Одна SOM-карта порождает несколько кластеров.

- K_t — кластер с номером t .

В соответствии с алгоритмом [32], если кластер имеет недостаточную степень детализации (на основании одного или нескольких критериев оценки качества кластеризации), он будет разбит на несколько подкластеров. Разбиение производится при помощи SOM-карты для набора данных, содержащихся в кластере. Таким образом можно говорить, что кластер может *порождать дочернюю* SOM-карту. Каждый кластер порождает максимум одну SOM-карту.

- $V = \{t\}$ — множество номеров всех кластеров.
- $U_p = \{t\}$ — множество номеров кластеров, порождаемых SOM-картой S_p .
- $\xi(t)$ — функция перехода со значениями (аргументом функции является номер текущего кластера t): $\xi(t)$ равна номеру дочерней SOM-карты кластера K_t , если кластер недостаточно детализирован, и -1, если кластер K_t имеет достаточную степень детализации.
- $L = \{t\}$ — множество номеров кластеров с недостаточной степенью детализации, подлежащих разбиению на подкластеры в процессе вычислений.

Описываемый метод построит дерево, где каждая вершина представляет собой обученную SOM-карту с определенным набором порождаемых ею кластеров. Каждая SOM-карта в дереве имеет свой уникальный номер, и каждый кластер также имеет свой уникальный номер. Каждый кластер, не имеющий карт-потомков, представляет собой выделенный из изначального нетипизированного множества класс.

1.1 Идея метода

При инициализации определяется корневая вершина дерева, которая представляет собой SOM-карту S_0 с единственным кластером K_0 , в который записываются все кортежи исходного набора данных. SOM-карта S_0 становится текущей.

Далее определяется однородность данных в каждом кластере текущей SOM-карты. Если кластер K_t неоднороден, то он порождает новую SOM-карту S_p с заданным количеством кластеров. Назовем эту операцию *делением кластера*. Новая дочерняя SOM-карта обучается на наборе данных кластера-родителя.

Для запоминания неоднородных, но «неразделенных» кластеров, ведется список L . Процесс деления неоднородных кластеров продолжается до тех пор, пока не останется ни одного такого кластера, то есть все кластеры будут иметь достаточную степень детализации.

Каждому кластеру в построенном дереве приписывается метка класса. Свойства класса определяются аналитически исходя из интегральных свойств соответствующих терминальных кластеров.

Теперь опишем разработанный метод формально.

1.2 Численный метод выделения классов разнородных объектов на основе каскадной нейросетевой фильтрации

Шаг 1. Создать корневую вершину S_0 с единственным кластером K_0 , в который записываются все кортежи исходного набора данных Ω_x .

- Задать номер текущей SOM-карты $p = 0$;
- Задать номер текущего кластера $n = 0$;
- Определить множество $U_0 = \{0\}$;
- Определить список кластеров L как пустой: $L = \{\}$.

Шаг 2. Проверить однородность каждого кластера SOM-карты S_p :

Для всех $t \in U_p$:

- если K_t — однородный, то определить функцию перехода $\xi(t) = -1$,
- иначе включить номер этого кластера в список L : $L = L \cup \{t\}$.

Шаг 3. Если список L пуст, то переход к шагу 6, иначе — переход к шагу 4.

Шаг 4. Для всех кластеров $t \in L$:

4.1 удалить номер t из списка L , т.е. $L = L \setminus \{t\}$,

4.2 выполнить деление кластера K_t и организовать новую SOM-карту, дочернюю по отношению к кластеру K_t . Для этого:

- задать количество кластеров c , на которые следует поделить кластер K_t ,
- задать номер новой SOM-карты $p = p + 1$,
- определить множество кластеров $U_p = \{n + 1, n + 2, \dots, n + c\}$,
- создать и обучить SOM-карту S_p с количеством выделяемых кластеров, равным c ,
- обозначить выделенные кластеры как K_j , $j = n + 1, \dots, n + c$,
- задать функцию перехода $\xi(t) = p$. Связать кластер K_t с новой вершиной — SOM-картой S_p согласно значению функции перехода,
- задать номер последнего кластера $n = n + c$.

Шаг 5. Перейти к шагу 2.

Шаг 6. Принять:

- множество всех построенных SOM-карт $S = \{S_0, S_1, \dots, S_p\}$ в качестве вершин дерева каскадной кластеризации,
- множество ассоциированных с каждой вершиной (SOM-картой) кластеров как U_r , $r = 1, 2, \dots, p$,
- значения построенной дискретной функции $\xi(\cdot)$ в качестве закона перехода между родительским кластером и порождаемой SOM-картой.

Шаг 7. Все кластеры K_t , для которых $\xi(t) = -1$, считать представителями выделенных в результате вычислений классов. Завершить вычисления.

В результате будет построено дерево с множеством вершин-SOM-карт $S = \{S_0, S_1, \dots, S_p\}$.

С каждой SOM-картой S_r ассоциировано некое непустое множество кластеров, т.е. определено множество U_r , содержащее номера кластеров, входящих в S_r . При этом множества U_r не пересекаются, а объединение всех U_r дает в результате все множество номеров кластеров $V = \{0, 1, \dots, n\}$. Каждый кластер может быть связан с SOM-картой нижнего уровня. Связи между кластерами и SOM-картами хранятся в виде дискретной функции $\xi(\cdot)$.

Пример структуры, генерируемой разработанным методом, можно видеть на Рис. 1.

В данном примере дерево содержит 6 SOM-карт, т.е. $S = \{S_0, S_1, \dots, S_5\}$, 13 кластеров, т.е. $V = \{0, 1, \dots, 12\}$. SOM-карта S_3 содержит 3 кластера K_6, K_7, K_8 , т.е. $U_3 = \{6, 7, 8\}$. Кластер K_1 порождает SOM-карту S_2 , т.е. функция перехода от кластера 1 возвращает 2: $\xi(1)=2$. Кластер K_2 не имеет дочерней SOM-карты, т.е. функция перехода от кластера 2 и от всех таких кластеров возвращает -1: $\xi(2) = \xi(4) = \xi(6) = \xi(8) = \xi(9) = \xi(10) = \xi(11) = \xi(12) = -1$.

2. Модель классификации как конечный детерминированный автомат

По выделенным согласно описанному численному методу классам, для каждого вновь поступающего объекта данных $X = (x_1, x_2, \dots, x_N)$ необходимо принять решение, к какому из выделенных классов он относится.

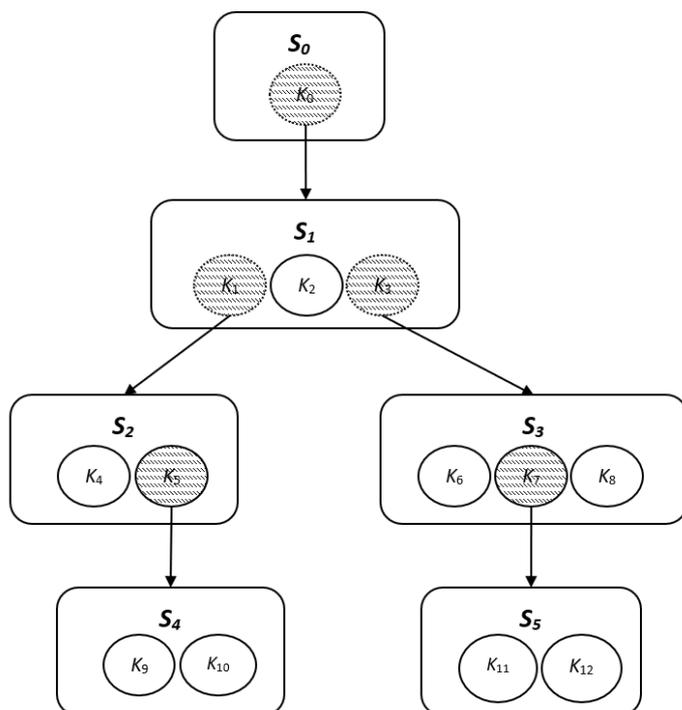


Рис. 1: Пример дерева SOM-карт, построенного согласно методу выделения классов разнородных объектов

Определим модель принятия решения о принадлежности некоторого входного вектора к одному из выделенных классов (*модель классификации на основе дерева SOM-карт*) как конечный детерминированный автомат вида:

$$A = (\Omega_x, V, S, S_0, \xi(\cdot));$$

здесь:

- Ω_x — входное множество,
- V — выходное множество,
- S — множество возможных состояний,
- S_0 — начальное состояние,
- $\xi(\cdot)$ — функция перехода .

Опишем эти составляющие более подробно.

На вход автомата подается вектор $X = (x_1, x_2, \dots, x_N)$, принадлежащий входному множеству Ω_x , т.е. декартовому произведению множеств $D_1 \times D_2 \times \dots \times D_N$, где каждое D_i является областью определения или доменом i -го признака. Иными словами, признак x_i должен принадлежать домену D_i .

Выходом автомата является метка класса — номер кластера v , к которому будет относиться входной вектор X . Таким образом, *выходным множеством* автомата является множество возможных номеров кластеров, т.е. целых чисел от 0 до n : $V = \{0, 1, \dots, n\}$, каждый из которых, в свою очередь, представляет один из выделенных ранее классов.

Множеством возможных состояний S является множество вершин (SOM-карт) дерева, построенного в результате метода выделения классов разнородных объектов: $S = \{S_0, S_1, \dots, S_p\}$.

Начальное состояние: корневая SOM-карта S_0 , содержащая кластер K_0 .

Работа автомата заключается в последовательном спуске от корневой вершины S_0 к вершинам нижнего уровня в соответствии с функцией перехода $\xi(\cdot) : V \rightarrow P \cup \{-1\}$, где P представляет собой множество номеров SOM-карт, т.е. $P = \{0, 1, \dots, p\}$, а -1 означает конец работы автомата.

Осталось определить *внутренние действия* начального состояния S_0 и всех остальных состояний S_r . Для начального состояния внутренними действиями будут:

$$S_0 \left| \begin{array}{l} v := 0; \\ r := \xi(v). \end{array} \right.$$

Для всех остальных состояний:

$$S_r \left| \begin{array}{l} v := \arg \min_{j \in U_r} \sum_{i=1}^N (x_i - w_{ij})^2; \\ r := \xi(v), \end{array} \right.$$

где w_{ij} — веса нейронов SOM-карты j .

Модель описанного конечного автомата можно представить в виде диаграммы состояний UML (State Machine Diagram) следующим образом:

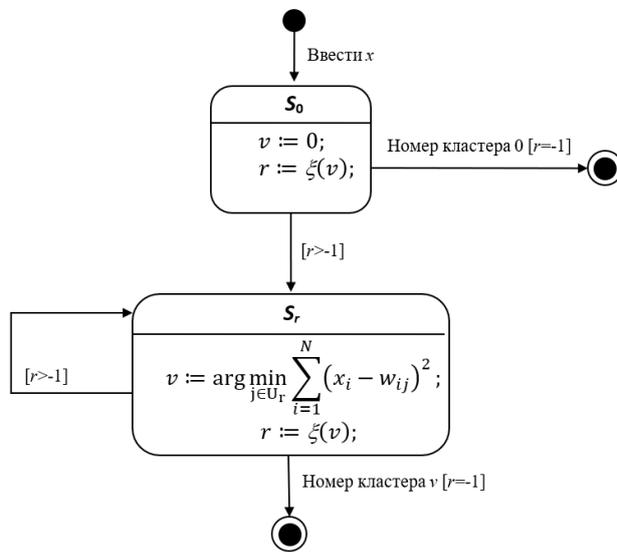


Рис. 2: Диаграмма состояний классифицирующего автомата по принципу нейросетевой каскадной фильтрации

3. Результаты вычислительных экспериментов

Разработанный метод классификации нетипизированных многомерных объектов был опробован на задаче определения типов поверхностных вод в 15 точках водоотбора на реках Кама, Волга, Свияга, Казанка, Ашит, Сулица, Меша, а также Куйбышевском водохранилище [35].

Каждый объект — проба поверхностных вод — характеризовался 33-мя физико-химическими показателями (содержание в воде металлов, показатели электропроводности, органических соединений и др.) и описывается вектором $X = (x_1, x_2, \dots, x_{33})$:

1. координата точки водоотбора,
2. координата точки водоотбора,
3. дата проведения анализа,
4. АПАВ,
5. алюминий,
6. аммоний ион,
7. БПК₅,
8. взвешенные вещества,
9. гидрокарбонаты,
10. железо,
11. жесткость,
12. кальций,
13. кислород раствор,
14. магний,
15. марганец,
16. медь,
17. натрий+калий,
18. нефтепродукты,
19. никель,
20. нитраты,
21. нитриты,
22. минерализация,

23. прозрачность,
24. свинец,
25. сульфаты,
26. температура,
27. фенол,
28. фосфат ион,
29. фосфор общий,
30. ХПК,
31. хлориды,
32. цинк,
33. электропроводность.

Всего множество данных содержало 151 объект.

На первом этапе при помощи *численного метода выделения классов разнородных объектов на основе каскадной нейросетевой фильтрации* было определено 11 классов, или типов, вод. Степень однородности кластеров оценивалась экспертно.

При этом был построен двухуровневый нейросетевой SOM-фильтр вида:

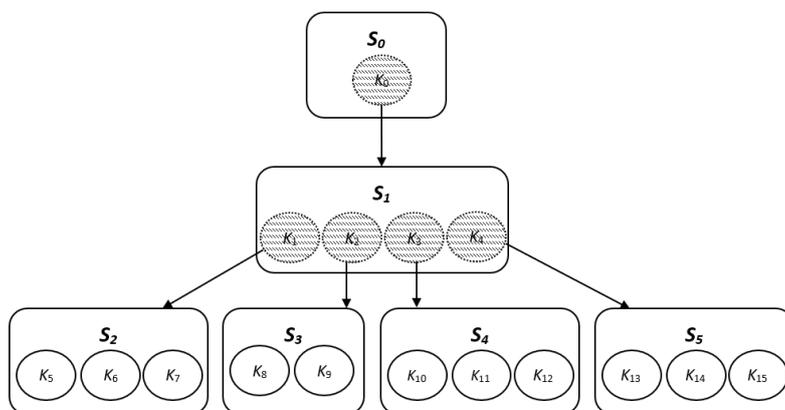


Рис. 3: SOM-дерево, соответствующее процессу выделения классов из множества данных о составе поверхностных вод

Характеристики каждого из 11 выделенных классов (типов) поверхностных вод были затем описаны экспертами предметной области на основании анализа векторов кластеров, представляющих соответствующие классы, например:

- Тип 1 (кластер K_5) — относительно чистые воды с высокой минерализацией;

- Тип 2 (кластер K_6) — воды с высоким загрязнением металлами и нефтепродуктами с низкой минерализацией;
- Тип 3 (кластер K_7) — воды низкой степени минерализации и незначительным общим загрязнением;

и так далее.

Для проверки адекватности *модели классификации на основе дерева SOM-карт* были определены классы 25-ти тестовых объектов, для которых модель правильно определила классы в 100% случаев.

Заключение

Разработанный метод классификации нетипизированных объектов на основе каскадной нейросетевой фильтрации позволяет выделять классы из множества данных, первоначально не имеющих меток классов. Дальнейшее применение модели классификации на основе разработанного детерминированного автомата позволяет классифицировать вновь поступившие на вход модели вектора данных в соответствии с разработанным методом. Практические вычислительные эксперименты подтверждают эффективность разработанных методов и алгоритмов.

Список литературы

- [1] Shlyemovich M.P., Medvedev M.V., Lyasheva S.A. Image analysis in unmanned aerial vehicle on-board system for objects detection and recognition with the help of energy characteristics based on wavelet transform // Proc. SPIE 10342, Optical Technologies for Telecommunications 2016. ID 1034210. <https://doi.org/10.1117/12.2270141>
- [2] Кирпичников А.П., Ляшева С.А., Обухов А.В., Шлеймович М.П. Автоматическое распознавание автомобильных номеров // Вестник технологического университета. 2015. Т. 18, № 4. С. 218–222.
- [3] Катасёв А.С., Катасёва Д.В., Кирпичников А.П., Гумерова Р.И. Нейросетевая модель распознавания рукописных символов в системах биометрической идентификации и аутентификации // Вестник технологического университета. 2016. Т. 19, № 4. С. 122–126.
- [4] Ахметвалеев А.М., Катасёв А.С. Схема бесконтактной идентификации лиц находящихся в состоянии опьянения // Информация и безопасность. 2015. Т. 18, № 3. С. 360–365.
- [5] Ахметвалеев А.М., Катасёв А.С., Шлеймович М.П. Повышение эффективности обнаружения лица и глаз человека на видеоизображении в задачах бесконтактного выявления потенциально опасных лиц // Информация и безопасность. 2016. Т. 19, № 4. С. 519–522.
- [6] Костин В.А. Теория и практика прочностной отработки конструкции ЛА / под ред. В.А. Костин, Н.Л. Валитова. Казань: Издательство Казанского государственного технического университета, 2014. 140 с.

- [7] Веденькин Д.А., Седельников Ю.Е., Потапова О.В. Свойства сфокусированных антенн для задач акустической эхолокационной диагностики // *Нелинейный мир*. 2014. Т. 12, № 10. С. 64–68.
- [8] Костин В.А., Снегуренко А.П. Теория и практика решения обратных задач прочности летательных аппаратов. Казань: Издательство Казанского государственного технического университета, 2004.
- [9] Щеголева М.Е. Определение психотипов личности с использованием экспресс-диагностической методики // *Психопедагогика в правоохранительных органах*. 2006. № 1. С. 102–105.
- [10] Тунакова Ю.А., Новикова С.В., Габдрахманова Г.Н. Способ оценки качества доходящих до потребителей питьевых вод с позиций химической безопасности и физиологической полноценности // *Сборник материалов Международной научно-практической конференции «Хартия Земли – практический инструмент решения фундаментальных проблем устойчивого развития»*. Казань: Татарское книжное издательство, 2016. С. 378–381.
- [11] Тунакова Ю.А. Прикладная экология: учебное пособие / под ред. Ю.А. Тунаковой, С.В. Новиковой. Казань: Изд-во КГТУ им. А.Н. Туполева, 2008. 155 с.
- [12] Тунакова Ю.А., Новикова С.В., Шагидуллина Р.А., Кремлева Э.Ш. Роль качественных оценок в задачах экологического управления в зонах действия полимерных производств // *Вестник Казанского технологического университета*. 2013. Т. 16, № 20. С. 276–279.
- [13] Костин В.А., Валитова Н.Л. Идентификация изгибной жесткости балки с использованием функции чувствительности // *Сборник докладов Всероссийской научно-практической конференции с международным участием*. Т. 1. Казань: АН РТ, 2016. С. 86–91.
- [14] Снегуренко А.П. Идентификация тонкостенных конструкций по данным прочностного эксперимента: дисс. ... канд. техн. наук. Казань, 2001.
- [15] Костин В.А., Снегуренко А.П. Идентификация поля цилиндрических жесткостей изотропных и ортотропных пластин // *Вестник Казанского государственного технического университета им. А.Н. Туполева*. 2001. № 2. С. 3–9.
- [16] Костин В.А., Хуан Ш., Валитова Н.Л. Применение дискретно-континуальной модели расчета на прочность для решения задачи идентификации теплонагруженной конструкции // *Известия высших учебных заведений. Авиационная техника*. 2017. № 3. С. 3–7.
- [17] Rodnishev N.E., Novikova S.V., Denisov K.G. Developing Methods and Computer Technologies for Learning, Identification and Optimization of Nonlinear Stochastic Systems // *IFAC Proceedings Volumes*. 2012. Vol. 45, № 11. Pp. 51–56.
- [18] Костин В.А., Хуан Ш., Валитова Н.Л. Численные методы анализа чувствительности в задачах идентификации конструкций // *Вестник Казанского государственного технического университета им. А.Н. Туполева*. 2017. № 1. С. 78–83.

- [19] Костин В.А., Снегуренко А.П. К вопросу уточнения внешней нагрузки по заданным деформациям // Вестник Казанского государственного технического университета им. А.Н. Туполева. 1999. № 4. С. 3–8.
- [20] Новикова С.В., Снегуренко А.П. Решение обратной задачи для динамической стохастической системы на примере управления газотурбинным авиационным двигателем // Дифференциальные уравнения и их приложения в математическом моделировании: материалы XIII Международной научной конференции (Саранск, 12-16 июля 2017 г.). Саранск: СВМО, 2017. С. 119–130.
- [21] Новикова С.В. Структурная оптимизация нейросетевой модели мониторинга ГТД // Известия высших учебных заведений. Авиационная техника. 2016. № 2. С. 101–107.
- [22] Новикова С.В. Применение алгоритма адаптивной коррекции для регулирования газотурбинного двигателя в условиях помех // Известия высших учебных заведений. Авиационная техника. 2006. № 4. С. 37–40.
- [23] Кожевников Ю.В., Новикова С.В. Алгоритм автоматизированной доводки многомерных параметрических систем регулирования ГТД // Известия высших учебных заведений. Авиационная техника. 2002. № 2. С. 61–65.
- [24] Тунакова Ю.А., Шагидуллина Р.А., Новикова С.В. Методология определения нормативов качества для приоритетных загрязняющих веществ // Вестник НЦБЖД. 2014. № 1 (19). С. 122–127.
- [25] Тунакова Ю.А., Шагидуллина Р.А., Новикова С.В., Шмакова Ю.А. Оценка воздействия полимерных производств на общий уровень загрязнения атмосферного воздуха (на примере г. Нижнекамска) // Вестник Казанского технологического университета. 2012. Т. 15, № 12. С. 76–79.
- [26] Тунакова Ю.А., Новикова С.В., Галимова А.Р. Способы оценки воздействия потребляемых питьевых вод на здоровье детского населения и обоснование способов повышения их качества // Известия Самарского научного центра Российской академии наук. 2016. Т. 18, № 5 (3). С. 500–504.
- [27] Тунакова Ю.А., Шагидуллина Р.А., Новикова С.В., Шмакова Ю.А. Определение времени формирования высокого уровня загрязнения атмосферного воздуха в зонах действия полимерных производств (на примере Нижнекамска). Сообщение 2 // Вестник Казанского технологического университета. 2012. Т. 15, № 16. С. 115–119.
- [28] Тунакова Ю.А., Шагидуллина Р.А., Новикова С.В., Шмакова Ю.А. Распознавание групп неблагоприятных метеорологических условий формирования высокого уровня загрязнения атмосферного воздуха в зонах действия полимерных производств (на примере г. Нижнекамска). Сообщение 3 // Вестник Казанского технологического университета. 2012. Т. 15, № 16. С. 119–121.
- [29] Marsh R.J. Lecture notes on cluster algebras // Bulletin of the American Mathematical Society (New Series). 2016. Vol. 53. Pp. 325–330. <https://doi.org/10.1090/bull/1514>

- [30] Guojun G., Chaoqun M., Jianhong W. Data Clustering: Theory, Algorithms, and Applications. Series: ASA-SIAM Series on Statistics and Applied Probability. Philadelphia, Alexandria, VA: SIAM, ASA, 2007. 488 p. <http://doi.org/10.1137/1.9780898718348>
- [31] Сивоголовко Е.В. Методы оценки качества чёткой кластеризации // Компьютерные инструменты в образовании. 2011. № 4. С. 14–31.
- [32] Кремлёва Э.Ш., Кирпичников А.П., Новикова С.В., Валитова Н.Л. Алгоритм построения модели каскадной нейросетевой фильтрации данных с различной степенью детализации // Вестник технологического университета. 2018. Т. 21, № 8. С. 109–115.
- [33] Кремлёва Э.Ш., Кирпичников А.П., Новикова С.В., Валитова Н.Л. Алгоритм построения модели каскадной нейросетевой фильтрации данных с различной степенью детализации // Вестник Технологического университета. 2018. Т. 21, № 8. С. 109–115.
- [34] Kohonen T. Essentials of the self-organizing map // Neural Networks. 2013. Vol. 37. Pp. 52–65. <https://doi.org/10.1016/j.neunet.2012.09.018>
- [35] Oster M., Douglas R., Liu S.-C. Computation with spikes in a winner-take-all network // Neural Computation. 2009. Vol. 21, № 9. Pp. 2437–2465. <https://doi.org/10.1162/neco.2009.07-08-829>
- [36] Новикова С.В., Тунакова Ю.А., Шагидуллин А.Р., Валиев В.С., Габдрахманова Г.Н. Синергетическая нейро-экспертная модель на основе SOM-КАРТ для интегрального анализа экологических объектов // Сборник трудов международной научной конференции (школа молодых ученых) «Химия и инженерная экология». 2018. С. 318–322.

Образец цитирования

Кремлева Э.Ш. Метод классификации нетипизированных объектов на основе каскадного нейросетевого фильтра и конечного детерминированного автомата // Вестник ТвГУ. Серия: Прикладная математика. 2019. № 1. С. 83–100. <https://doi.org/10.26456/vtpmk527>

Сведения об авторах

1. Кремлева Эльмира Шамильевна

старший преподаватель кафедры прикладной математики и информатики Казанского национального исследовательского технического университета им А.Н. Туполева – КАИ.

Россия, 420111, г. Казань, ул. К. Маркса, д. 10, КНИТУ-КАИ.

E-mail: e-smile29.04@mail.ru

METHOD FOR CLASSIFYING UNTYPED OBJECTS BASED ON CASCADE NEURAL NETWORK FILTER AND FINITE DETERMINISTIC AUTOMATON

Kremleva Elmira Shamilyevna

Senior Lecturer at the Department of Applied Mathematics and Computer Science,
Kazan National Research Technical University named after A.N. Tupolev – KAI
Russia, 420111, Kazan, 10 Karla Marx str., KNRTU-KAI.
E-mail: e-smile29.04@mail.ru

Received 10.02.2019, revised 15.03.2019.

The article describes a method for the phased finding of the multidimensional object class in the case when the set of classes is not known in advance. The developed method first solves the problem of selecting classes from an untyped heterogeneous set of objects, and then classifies an arbitrary new object into the determined classes. Classes are found on the basis of the author's algorithm of cascade neural filtering, and the objects classification is performed using the author's model based on a finite automaton.

Keywords: classification, clustering, object types, neural network, finite-state automaton.

Citation

Kremleva E.Sh., "Method for classifying untyped objects based on cascade neural network filter and finite deterministic automaton", *Vestnik TvGU. Seriya: Prikladnaya Matematika [Herald of Tver State University. Series: Applied Mathematics]*, 2019, № 1, 83–100 (in Russian). <https://doi.org/10.26456/vtpmk527>

References

- [1] Shleymovich M.P., Medvedev M.V., Lyasheva S.A., "Image analysis in unmanned aerial vehicle on-board system for objects detection and recognition with the help of energy characteristics based on wavelet transform", *Proc. SPIE 10342, Optical Technologies for Telecommunications 2016*, <https://doi.org/10.1117/12.2270141>.
- [2] Kirpichnikov A.P., Lyasheva S.A., Obukhov A.V., Shleymovich M.P., "Automatic license plate recognition", *Vestnik Tekhnologicheskogo Universiteta [Bulletin of Technological University]*, **18:4** (2015), 218–222 (in Russian).
- [3] Katasyov A.S., Katasyova D.V., Kirpichnikov A.P., Gumerova R.I., "Neural network model of handwriting recognition in biometric identification and authentication systems", *Vestnik Tekhnologicheskogo Universiteta [Bulletin of Technological University]*, **19:4** (2016), 122–126 (in Russian).

- [4] Akhmetvaleev A.M., Katasyov A.S., “Contactless identification scheme for persons intoxicated”, *Informatsiya i Bezopasnost [Information and Security]*, **18**:3 (2015), 360–365 (in Russian).
- [5] Akhmetvaleev A.M., Katasyov A.S., Shlejmovich M.P., “Improving the efficiency of detecting a person’s face and eyes in a video image in the tasks of non-contact detection of potentially dangerous people”, *Informatsiya i Bezopasnost [Information and Security]*, **19**:4 (2016), 519–522 (in Russian).
- [6] Kostin V.A., *Teoriya i Praktika Prochnostnoj Otrabotki Konstruktsii LA [Theory and Practice of Strength Testing of Aircraft Design]*, eds. V.A. Kostin, N.L. Valitova, Kazan State Technical University Publ., Kazan, 2014 (in Russian), 140 pp.
- [7] Vedenkin D.A., Sedelnikov Yu.E., Potapova O.V., “Properties of focused antennas for acoustic echolocation diagnostics tasks”, *Nelinejnyj Mir [Nonlinear World]*, **12**:10 (2014), 64–68 (in Russian).
- [8] Kostin V.A., Snegurenko A.P., *Teoriya i Praktika Resheniya Obratnykh Zadach Prochnosti Letatelnykh Apparatov [Theory and Practice of Solving Inverse Problems of Aircraft Durability]*, Kazan State Technical University Publ., Kazan, 2004 (in Russian).
- [9] Shchegoleva M.E., “Definition of psycho-types of personality using an express diagnostic method”, *Psikhopedagogika v Pravoohranitelnykh Organakh [Psychopedagogy in Law Enforcement]*, 2006, № 1, 102–105 (in Russian).
- [10] Tunakova Yu.A., Novikova S.V., Gabdrakhmanova G.N., “The method of assessing the quality of drinking water reaching consumers from the standpoint of chemical safety and physiological usefulness”, *Sb. materialov Mezhdunarodnoj nauchno-prakticheskoy konferentsii “Khartiya Zemli – prakticheskij instrument resheniya fundamentalnykh problem ustojchivogo razvitiya” [Proceedings of the International scientific-practical conference “Earth Charter as a practical tool for solving the fundamental problems of sustainable development”]*, Tatar Book Publishing House, Kazan, 2016, 378–381 (in Russian).
- [11] Tunakova Yu.A., *Prikladnaya Ekologiya: Uchebnoe Posobie [Applied Ecology: Tutorial]*, eds. Yu.A. Tunakova, S.V. Novikova, Publishing House of KNRTU-KAI, Kazan, 2008 (in Russian), 155 pp.
- [12] Tunakova Yu.A., Novikova S.V., Shagidullina R.A., Kremleva E.Sh., “The role of qualitative assessments in the tasks of environmental management in the areas of polymer production”, *Vestnik Kazanskogo tekhnologicheskogo universiteta [Bulletin of Kazan Technological University]*, **16**:20 (2013), 276–279 (in Russian).
- [13] Kostin V.A., Valitova N.L., “Identification of flexural stiffness of a beam using the sensitivity function”, *Sbornik докладov Vserossijskoj nauchno-prakticheskoy konferentsii s mezhdunarodnym uchastiem [Proceedings of the All-Russian scientific-practical conference with international participation]*. V. 1, AS RT, Kazan, 2016, 86–91 (in Russian).

- [14] Snegurenko A.P., *Identifikatsiya tonkostennykh konstruksij po dannym prochnostnogo eksperimenta*, thesis for the degree of candidate of technical sciences, Kazan, 2001 (in Russian).
- [15] Kostin V.A., Snegurenko A.P., “Identification of the field of cylindrical stiffness of isotropic and orthotropic plates”, *Vestnik Kazanskogo Gosudarstvennogo Tekhnicheskogo Universiteta im. A.N. Tupoleva [Bulletin of Kazan State Technical University named after A.N. Tupolev]*, 2001, № 2, 3–9 (in Russian).
- [16] Kostin V.A., Khuan Sh., Valitova N.L., “Application of a discrete-continuous model of strength analysis to solve the identification problem of a thermally loaded structure”, *Russian Aeronautics*, **60**:3 (2017), 321–326.
- [17] Rodnishev N.E., Novikova S.V., Denisov K.G., “Developing Methods and Computer Technologies for Learning, Identification and Optimization of Nonlinear Stochastic Systems”, *IFAC Proceedings Volumes*, **45**:11 (2012), 51–56.
- [18] Kostin V.A., Khuan Sh., Valitova N.L., “Numerical methods of sensitivity analysis in the problems of identification of structures”, *Vestnik Kazanskogo Gosudarstvennogo Tekhnicheskogo Universiteta im. A.N. Tupoleva [Bulletin of Kazan State Technical University named after A.N. Tupolev]*, 2017, № 1, 78–83 (in Russian).
- [19] Kostin V.A., Snegurenko A.P., “To the question of clarification of the external load for a given deformation”, *Vestnik Kazanskogo Gosudarstvennogo Tekhnicheskogo Universiteta im. A.N. Tupoleva [Bulletin of Kazan State Technical University named after A.N. Tupolev]*, 1999, № 4, 3–8 (in Russian).
- [20] Novikova S.V., Snegurenko A.P., “Solution of the inverse problem for a dynamic stochastic system on the example of control of a gas turbine aircraft engine”, *Differentsialnye uravneniya i ikh prilozheniya v matematicheskom modelirovanii: materialy XIII Mezhdunarodnoj nauchnoj konferentsii [Differential equations and their applications in mathematical modeling: materials of the XIII International Scientific Conference]* (Saransk, July 12-16, 2017), SVMO Publ., Saransk, 2017, 119–130 (in Russian).
- [21] Novikova S.V., “Structural optimization of the neural network model for the gas turbine engine monitoring”, *Russian Aeronautics*, **59**:2 (2016), 263–270.
- [22] Novikova S.V., “Application of an adaptive correction algorithm to control a gas turbine engine under conditions of noises”, *Russian Aeronautics*, **49**:4 (2006), 55–62.
- [23] Kozhevnikov Yu.V., Novikova S.V., “An Algorithm for Computer-Aided Development of Multidimensional Parametric Systems of GTE Control”, *Russian Aeronautics*, 2002, № 2, 61–65 (in Russian).
- [24] Tunakova Yu.A., Shagidullina R.A., Novikova S.V., “Methodology for determining quality standards for priority pollutants”, *Vestnik NTsBZhD [Bulletin of the National Center for Life Safety]*, 2014, № 1 (19), 122–127 (in Russian).

- [25] Tunakova Yu.A., Shagidullina R.A., Novikova S.V., Shmakova Yu.A., “Evaluation of the impact of polymer production on the overall level of air pollution (on example of the city of Nizhnekamsk)”, *Vestnik Kazanskogo Tekhnologicheskogo Universiteta [Bulletin of Kazan Technological University]*, **15**:12 (2012), 76–79 (in Russian).
- [26] Tunakova Yu.A., Novikova S.V., Galimova A.R., “Ways to assess the impact of consumed drinking water on the health of the child population and the rationale for ways to improve their quality”, *Izvestiya Samarskogo nauchnogo tsentra Rossijskoj akademii nauk [News of the Samara Scientific Center of the Russian Academy of Sciences]*, **18**:5 (3) (2016), 500–504 (in Russian).
- [27] Tunakova Yu.A., Shagidullina R.A., Novikova S.V., Shmakova Yu.A., “Determination of the time of formation of a high level of air pollution in the areas of polymer production (on example of Nizhnekamsk). Post 2”, *Vestnik Kazanskogo Tekhnologicheskogo universiteta [Bulletin of Kazan Technological University]*, **15**:16 (2012), 115–119 (in Russian).
- [28] Tunakova Yu.A., Shagidullina R.A., Novikova S.V., Shmakova Yu.A., “Recognition of groups of adverse meteorological conditions for the formation of a high level of air pollution in the areas of polymer production (on example of the city of Nizhnekamsk). Post 3”, *Vestnik Kazanskogo tekhnologicheskogo universiteta [Bulletin of Kazan Technological University]*, **15**:16 (2012), 119–121 (in Russian).
- [29] Marsh R.J., “Lecture notes on cluster algebras”, *Bulletin of the American Mathematical Society (New Series)*, **53** (2016), 325–330, <https://doi.org/10.1090/bull/1514>.
- [30] Guojun G., Chaoqun M., Jianhong W., *Data Clustering: Theory, Algorithms, and Applications*, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia; ASA, Alexandria, VA, 2007, 488 pp., <http://doi.org/10.1137/1.9780898718348>.
- [31] Sivogolovko E.V., “Methods for assessing the quality of clear clustering”, *Kompyuternye instrumenty v obrazovanii [Computer tools in education]*, 2011, № 4, 14–31 (in Russian).
- [32] Kremlyova E.Sh., Kirpichnikov A.P., Novikova S.V., Valitova N.L., “Algorithm of constructing a neural cascade data filtration model with different detail degrees”, *Vestnik Tekhnologicheskogo Universiteta [Bulletin of Technological University]*, **21**:8 (2018), 109–115 (in Russian).
- [33] Kremlyova E.Sh., Kirpichnikov A.P., Novikova S.V., Valitova N.L., “Algoritm postroeniya modeli kaskadnoj nejrosetevoj filtratsii dannykh s razlichnoj stepenyu detalizatsii”, *Vestnik Tekhnologicheskogo Universiteta [Bulletin of Technological University]*, **21**:8 (2018), 109–115 (in Russian).
- [34] Kohonen T., “Essentials of the self-organizing map”, *Neural Networks*, **37** (2013), 52–65, <https://doi.org/10.1016/j.neunet.2012.09.018>.

- [35] Oster M., Douglas R., Liu S.-C., “Computation with spikes in a winner-take-all network”, *Neural Computation*, **21**:9 (2009), 2437–2465, <https://doi.org/10.1162/neco.2009.07-08-829>.
- [36] Novikova S.V., Tunakova Yu.A., Shagidullin A.R., Valiev V.S., Gabdrakhmanova G.N., “Synergetic neuro-expert model based on SOM-KART for the integral analysis of ecological objects”, *Sbornik trudov mezhdunarodnoj nauchnoj konferentsii (shkola molodykh uchenykh) “Khimiya i inzhenernaya ekologiya” [Collection of works of the international scientific conference (school of young scientists) “Chemistry and engineering ecology”]*, 2018, 318–322 (in Russian).