

DOI: 10.61726/8410.2025.25.27.001

УДК 004.896

МЕТОД ФАЗЗИФИКАЦИИ ДАННЫХ СО СТОРОЖЕВЫМИ ТАЙМЕРАМИ ДЛЯ ДЕТЕКТИРОВАНИЯ АНОМАЛИЙ В СИСТЕМАХ ТЕЛЕМЕТРИИ

Т.А. РАДИШЕВСКАЯ

Учреждение образования «Белорусский государственный медицинский университет»
пр. Дзержинского, 83, Минск, 220083, Беларусь
ORCID: <https://orcid.org/0009-0001-9602-5696>

Поступила в редакцию 24 октября 2025

Представлен метод фаззификации данных со сторожевыми таймерами для детектирования аномалий в потоках телеметрии на основе метода оценки с учетом введения двух независимых сторожевых таймеров для каждого показателя. В методе используются пять лингвистических термов, причем аналитиком определяются три базовые точки, а еще две точки вычисляются. В методе фаззификации система функций степени принадлежности показателей лингвистическим термам является ортогональной системой, в которой скалярные произведения любых двух функций равны нулю.

Ключевые слова: фаззификация данных, лингвистический терм, детектирование аномалий, сторожевой таймер.

Введение. Обнаружение аномалий – важная проблема, изучаемая в различных областях исследований и областях применения. Обнаружение аномалий относится к проблеме поиска закономерностей в данных, которые не соответствуют ожидаемому поведению. Эти несоответствующие закономерности часто называют аномалиями, выбросами, несогласованными наблюдениями, исключениями, отклонениями, неожиданностями, особенностями или загрязнителями в различных областях применения. Из них аномалии и выбросы – два термина, наиболее часто используемые в контексте обнаружения аномалий; иногда взаимозаменяемые. Важность обнаружения аномалий обусловлена тем, что аномалии в данных преобразуются в важную (и часто критическую) практическую информацию в самых разных областях применения [1].

Важным аспектом любого метода обнаружения аномалий является способ сообщения об аномалиях. Как правило, результаты, получаемые методами обнаружения аномалий, относятся к одному из следующих двух типов [1]:

– оценки. Методы оценки присваивают оценку аномалии каждому экземпляру тестовых данных в зависимости от степени, в которой этот экземпляр считается аномалией. Таким образом, результатом таких методов является ранжированный список аномалий. Ранг аномалии чаще всего определяется при помощи механизма базовых точек [2], причем точки задаются в порядке возрастания значения переменной, а функция принадлежности между соседними точками является линейной. Степень принадлежности для каждого термина рассчитывается по четкому входному значению с помощью линейной интерполяции между двумя соответствующими соседними точками функции принадлежности. Базовые точки позволяют определить кусочно-линейную функцию степени принадлежности элемента x из четкого множества X нечеткому множеству A . Эта функция обозначается как $\mu_A(x):X \rightarrow [0,1]$. Аналитик может либо проанализировать несколько первых аномалий, либо использовать пороговое значение для отбора аномалий;

– метки. Методы этой категории присваивают метку (нормальный или аномальный) каждому тестовому экземпляру. Для этого в трендах телеметрии необходимо выявлять отклонения, позволяющие присвоить каждому показателю метку $IsAnomaly$, где значение 0 означает «норма», а 1 – «отклонение».

Методы обнаружения аномалий, основанные на оценке, позволяют аналитику использовать пороговое значение, специфичное для предметной области, для отбора наиболее значимых аномалий.

Методы, предоставляющие двоичные метки тестовым экземплярам, не позволяют аналитикам напрямую делать такой выбор, хотя это можно контролировать косвенно посредством выбора параметров в рамках каждого метода [1].

Недостатком методов оценки и метки является то, что они позволяют выявлять однократные (точечные) аномалии, не учитывая их динамику, их наличия во времени. Выявление аномалий в режиме реального времени может потребовать дополнительной модификации этих методов. Наиболее простым в реализации является алгоритм скользящего окна [3]. Данная методика используется для временных рядов, которые разбиваются на некоторое число под-последовательностей – окон. Поиск аномальной под-последовательности осуществляется при помощи скользящего окна по всему ряду с шагом, меньшим длины окна. Нужно отметить, что для методов, требующих наличия всего объема данных (функционирующих в режиме распознавания без учителя) применение данного метода может привести к повышенной неточности результатов, так как вычисления будут проводиться только для экземпляров в пределах окна [3].

Детектирование аномалий основано на предположении о том, что нормальное поведение системы определяется одним или несколькими классами, в которых могут находиться значения поступающих данных при нормальном поведении. Система, основанная на классических нейронных сетях, может распознавать как один, так и несколько классов нормального поведения [4, 5].

Метод фаззификации со сторожевыми таймерами для детектирования аномалий. В работе представлен метод фаззификации со сторожевыми таймерами для детектирования аномалий в потоках телеметрии на основе метода оценки с учетом введения двух независимых сторожевых таймеров для каждого показателя.

В соответствии с [2] каждый показатель состояния оборудования описывается лингвистической переменной, которая может принимать значение одного из нескольких лингвистических термов. В модели используются следующие лингвистические термы:

- 1) LA («Low Alarm») – значение показателя ниже минимального порога аварии;
- 2) LW («Low Warning») – значение показателя ниже минимального порога предупреждения, но выше порога аварии;
- 3) G («Good») - значение показателя в норме;
- 4) HW («High Warning») – значение показателя выше максимального порога предупреждения, но ниже порога аварии;
- 5) HA («High Alarm») – значение показателя выше максимального порога аварии.

Такой алфавит позволяет привести все показатели разного типа (разного физического смысла) к набору однотипных лингвистических переменных, которые обрабатываются нейронной сетью. Для каждого из показателей определено минимальное и максимальное значение показателя состояния оборудования, а также базовые точки, определяющие функцию принадлежности значения показателя к одному из лингвистических термов. Функция принадлежности представляет собой кусочнолинейную функцию, определяемую таблицей точек. Точки задаются в порядке возрастания значения показателя, а функция принадлежности между соседними точками является линейной. Точки определяются при помощи базовых точек, причем число этих точек должно быть равно числу лингвистических термов [2]. В методе аналитиком задаются три базовых точки:

- Z-образная (минимальная) BP1;
- T-образная (треугольная) центральная BP2;
- S-образная (максимальная) BP3.

Две недостающие базовые точки TPL и TRH вычисляются из заданных базовых точек BP1, BP2, BP3. Вычисляемая базовая точка TPL определяется из условий пересечения линий Z-образной точки BP1 и T-образной точки BP2.

$$TPL = (BP2 - TPL) / (BP2 - BP1)$$

$$TPL = (TPL - BP1) / (BP2 - BP1)$$

Отбрасывая делители и получаем

$$BP2 - TPL = TPL - BP1 \text{ или}$$

$$BP2 + BP1 = 2 * TPL$$

В результате

$$TPL = (BP2 + BP1) / 2$$

Вычисляемая базовая точка TRH определяется аналогично из условий пересечения линий S-образной точки BP3 и T-образной точки BP2.

$$TRH = (BP2 + BP3) / 2$$

Эти базовые точки показаны на рис. 1.

Использование сторожевых таймеров в методе определяется установками начальных значений этих таймеров для верхнего (ТН) и нижнего (ТЛ) порога в секундах:

– если начальные значения равны максимальному значению 65535, то таймеры не используются (значения ТН и ТЛ загружаются в счетчики ТНс или ТЛс, которые не декрементируются) и для обработки показателей (обнаружения точечной аномалии) применяются базовые точки ВР1, ВР2, ВР3 и точки ТРЛ, ТРН;

– если начальные значения не равны максимальному значению 65535 (значения ТН и ТЛ загружаются в счетчики ТНс или ТЛс, которые декрементируются каждую секунду), то значения лингвистической переменной изменяются в зависимости от того, равен или нет счетчик нулю. Максимальное время счета ТНс и ТЛс составляет 65534 секунды, то есть 18 часов, 12 минут и 14 секунд.

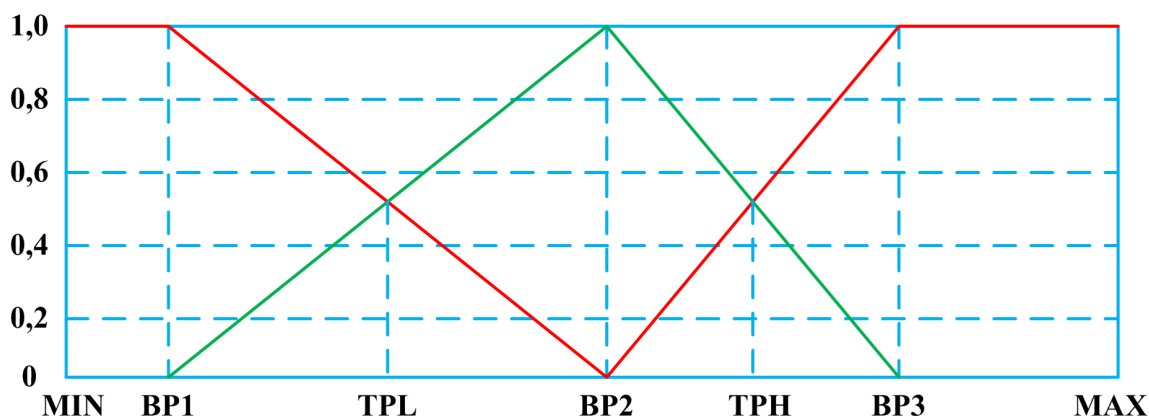


Рис. 1. Базовые точки в методе фаззификации

В таком случае можно определить кусочно-линейную функцию степени принадлежности элемента x из четкого множества X лингвистическому терму LT в зависимости от значения счетчиков ТНс или ТЛс. Такую функцию обозначим как $\mu_{LT}(x, THc, TLc): X \rightarrow [0, 1]$.

Степень принадлежности элемента x лингвистическому терму LA в зависимости от значения счетчика ТЛс определяется формулами, приведенными ниже.

$$\mu_{LA}(x, THc, TLc) = 1, x < MIN \tag{1.1}$$

$$\mu_{LA}(x, THc, TLc) = 1, MIN \leq x \leq ВР1, TLc = 65535 \tag{1.2}$$

$$\mu_{LA}(x, THc, TLc) = 0, x > ВР1, TLc = 65535 \tag{1.3}$$

$$\mu_{LA}(x, THc, TLc) = 0, MIN \leq x < ТРЛ, 0 < TLc < 65535 \tag{1.4}$$

$$\mu_{LA}(x, THc, TLc) = 1, MIN \leq x < ТРЛ, TLc = 0 \tag{1.5}$$

$$\mu_{LA}(x, THc, TLc) = 0, TLc = TL, x \geq ТРЛ \tag{1.6}$$

Степень принадлежности элемента x лингвистическому терму LW в зависимости от значения счетчика ТЛс определяется формулами, приведенными ниже.

$$\mu_{LW}(x, THc, TLc) = 1, ВР1 \leq x < ТРЛ, TLc = 65535 \tag{1.7}$$

$$\mu_{LW}(x, THc, TLc) = 0, x < ВР1, TLc = 65535 \tag{1.8}$$

$$\mu_{LW}(x, THc, TLc) = 1, MIN \leq x < ТРЛ, 0 < TLc < 65535 \tag{1.9}$$

$$\mu_{LW}(x, THc, TLc) = 0, MIN \leq x < ТРЛ, TLc = 0 \tag{1.10}$$

$$\mu_{LW}(x, THc, TLc) = 0, TLc = TL, x \geq ТРЛ \tag{1.11}$$

Степень принадлежности элемента x лингвистическому терму G определяется формулами, приведенными ниже.

$$\mu_G(x, THc, TLc) = 1, TLc = TL, THc = TH, ТРЛ \leq x \leq ТРН \tag{1.12}$$

$$\mu_G(x, THc, TLc) = 0, x < TPL \tag{1.13}$$

$$\mu_G(x, THc, TLc) = 0, x > TRH \tag{1.14}$$

Степень принадлежности элемента x лингвистическому терму HW в зависимости от значения счетчика THc определяется формулами, приведенными ниже.

$$\mu_{HW}(x, THc, TLc) = 1, TRH < x \leq BP3, THc = 65535 \tag{1.15}$$

$$\mu_{HW}(x, THc, TLc) = 0, x > BP3, THc = 65535 \tag{1.16}$$

$$\mu_{HW}(x, THc, TLc) = 1, TRH < x \leq MAX, 0 < THc < 65535 \tag{1.17}$$

$$\mu_{HW}(x, THc, TLc) = 0, TRH < x \leq MAX, THc = 0 \tag{1.18}$$

$$\mu_{HW}(x, THc, TLc) = 0, THc = TH, x < TRH \tag{1.19}$$

Степень принадлежности элемента x лингвистическому терму HA в зависимости от значения счетчика THc определяется формулами, приведенными ниже.

$$\mu_{HA}(x, THc, TLc) = 1, x > MAX \tag{1.20}$$

$$\mu_{HA}(x, THc, TLc) = 1, BP3 \leq x \leq MAX, THc = 65535 \tag{1.21}$$

$$\mu_{HA}(x, THc, TLc) = 0, x < BP3, THc = 65535 \tag{1.22}$$

$$\mu_{HA}(x, THc, TLc) = 0, TRH < x \leq MAX, 0 < THc < 65535 \tag{1.23}$$

$$\mu_{HA}(x, THc, TLc) = 1, TRH < x \leq MAX, THc = 0 \tag{1.24}$$

$$\mu_{HA}(x, THc, TLc) = 0, THc = TH, x \leq TRH \tag{1.25}$$

В методе фаззификации система функций степени принадлежности показателей лингвистическим термам является ортогональной системой, в которой скалярные произведения любых двух функций равны нулю.

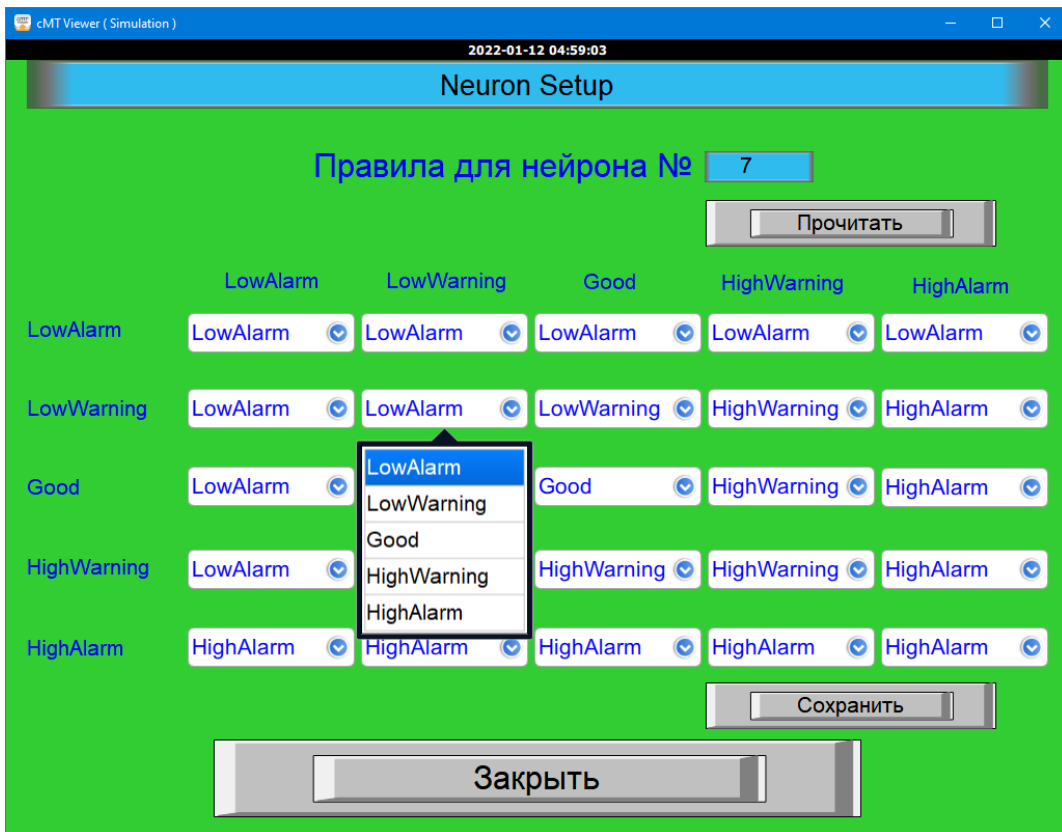


Рис. 2. Экранная форма определения функций активации для нейрона

На основе метода фаззификации построен масштабатор/фаззификатор (МФ), который определяет входные лингвистические переменные для искусственной нейронной сети прямого распространения. В этой сети каждый нейрон имеет два входа (две лингвистические переменные), один выход (лингвистическую переменную) и функцию активации (передаточную функцию) в виде таблицы правил 5x5 элементов. МФ реализован в составе программного средства «Библиотека управляющих программ для устройства Edge Hub» [6]. На рисунке 2 показан интерфейс экранной формы, обеспечивающей определение функций активации для нейрона.

Заключение. На основе метода фаззификации разработана нейро-нечеткая модель (ННМ) детектирования аномалий в потоках телеметрии [7]. На основе ННМ реализован нейро-нечеткий классификатор, который предназначен для организации систем мониторинга инженерной инфраструктуры сети передачи данных на промышленных предприятиях для автоматизации процесса оповещения служб эксплуатации.

A METHOD OF DATA FUZZY-FIXING WITH WATCHDOG TIMERS FOR DETECTING ANOMALIES IN TELEMETRY SYSTEMS

T.A. RADZISHEUSKAYA

Abstract

A data fuzzification method with watchdog timers for detecting anomalies in telemetry streams is presented. This method is based on an estimation method that takes into account the introduction of two independent watchdog timers for each metric. The method uses five linguistic terms, with the analyst defining three base points and calculating two more base points. In the fuzzification method, the system of functions for the membership degrees of metrics to linguistic terms is an orthogonal system in which the scalar products of any two functions are equal to zero.

Список литературы

1. Chandola, V. Anomaly detection : A survey / V. Chandola, A. Banerjee, V. Kumar // ACM Computing Surveys. – 2009. – Vol. 41 (3). – P. 1–58.
2. Programmable controllers – Part 7 : Fuzzy control programming : IEC 61131-7 ; введ. 01.09.2018. – P. 122.
3. Антипов, С. Проблема обнаружения аномалий в наборах временных рядов / С. Антипов, М. Фомина // Программные продукты и системы. – 2012. – № 2. – С. 78–82.
4. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский. – М. : Горячая линия – Телеком, 2006. – 452 с.
5. Нечеткие системы и мягкие вычисления. Промышленные применения. Fuzzy Technologies in the Industry (FTI-2017) : Первая Всероссийская научно-практическая конференция (Россия, г. Ульяновск, 14–15 ноября, 2017 г.) : сборник научных трудов. – Ульяновск : УлГТУ, 2017. – 406 с.
6. Радишевская, Т. А. Библиотека управляющих программ для устройства Edge Hub / Т. А. Радишевская, // Свидетельство о регистрации компьютерной программы №1556-КП, дата внесения записи о регистрации и депонировании в Реестр 21.10.2022.
7. Половня, С. И. Нейро-нечеткая модель детектирования аномалий в потоках телеметрии распределенных телекоммуникационных систем / С. И. Половня, Т. А. Радишевская // Веснік сувязі. – 2022. – № 2. С. – 50–55.