

УДК 004.896 + 004.428

Нейро-нечеткая модель детектирования аномалий в потоках телеметрии для распределенных телекоммуникационных систем

В статье представлена нейро-нечеткая модель детектирования аномалий в потоках телеметрии распределенных телекоммуникационных систем. Признаки состояния оборудования обрабатываются искусственной нейронной нечеткой сетью прямого распространения с помощью математического аппарата нечетких множеств. Основой модели является масштабатор/фаззификатор, который определяет входные лингвистические термы по пикам или провалам данных и/или по восходящим или нисходящим трендам с разными порогами скорости изменения данных. Настраиваемая длина скользящего окна для данных от отличающихся источников устраняет чувствительность предложенной модели к размеру окна при обнаружении аномалий разного типа. Нейро-нечеткая модель реализована в виде библиотеки подпрограмм в среде разработки EasyBuilder Pro для устройств сMT-SVR.

С. ПОЛОВЕНЯ,
зав. кафедрой
телекоммуникационных систем

Т. РАДИШЕВСКАЯ,
аспирант

Белорусская государственная
академия связи

Ключевые слова:

нейро-нечеткая модель,
детектирование аномалий,
масштабатор, фаззификатор,
лингвистический терм.

Введение. В середине 2021 года специалисты IoT Analytics опубликовали прогноз развития платформ «интернета вещей» (Internet of Thing, IoT) на 2021–2026 годы [1]. Они рассмотрели пять фактов, которые необходимо знать о платформах IoT. Сегодня платформу IoT определяют как разновидность программной цифровой платформы, предназначенную для создания и управления решениями IoT, и имеет широкие возможности настройки и расширения функционала, превосходящие возможности традиционных платформ разработки. Это определение отражает тот факт, что за последние 5 лет платформы IoT по своей природе стали более специализированными (например, с упором на управление

устройствами или телекоммуникациями), более модульными (например, с помощью микро-сервисов) и более ориентированными на данные (предлагают различные компоненты управления данными).

С другой стороны, технологиям IoT мешают развиваться высокая степень фрагментации рынка IoT и низкая оценка роли оборудования. Фрагментация рынка IoT проходит не по национальному или географическому принципу, не зависит от поставщиков оборудования и платформ IoT, и только четкое разграничение задач и потребностей B2B и B2C делит рынок IoT на промышленный «интернет вещей» (Industrial Internet of Thing – IIoT) и потребительский (IoT). В концепции IIoT также

недооценивается оборудование, хотя именно оно закладывает фундамент IoT-приложений, являясь технологической основой четвертой промышленной революции (Индустрия 4.0), предусматривающей цифровую трансформацию традиционных отраслей промышленности. Определяя нижний слой платформы IoT как слой управления устройствами специалисты IoT Analytics перенесли анализ данных в более высокие слои. С другой стороны, известные платформы IoT [2] относят компоненты «Связь и нормализация» («Connectivity & normalization») и «Обработка и управление действиями» («Processing & action management») к нижнему слою, который с этими дополнительными функциями часто называют слоем зондирования IoT.

Слой зондирования IoT, максимально приближенный к датчикам и устройствам управления IoT, кроме взаимодействия с этими устройствами должен обеспечивать приведение различных протоколов (использующих различные физические интерфейсы) и различных форматов данных к одному программному интерфейсу, мониторинг текущего состояния оконечных устройств и прогнозирование предаварийных состояний при гарантированной потоковой передаче данных. Индустрия 4.0 требует развертывания распределенных телекоммуникационных систем (РТКС 4.0), которые будут обеспечивать функционирование слоя зондирования IoT и передачу данных в более высокие слои платформ IoT, что позволит максимально приблизить вычисления к данным и к источникам, которые эти данные генерируют.

Развертывание РТКС 4.0 подразумевает, что на промышленном предприятии, помимо традиционных объектов основного производства, появятся новые объекты телекоммуникаций, в связи с чем общее число объектов диспетчеризации (контроля) возрастет на один, два, а то и три порядка по сравнению с уровнем Индустрии 3.0. С таким количеством объектов диспетчерский контроль пакетов телеметрии путем визуального анализа мнемосхем даст сбой и будет просто неоптимальным, так как эти показатели могут принимать различные значения и иметь различные траектории (тренды) в течение длительного времени.

В связи с тем, что объекты РТКС 4.0 сильнее поддаются типизации/унификации (номенклатура наблюдаемых признаков намного меньше, чем у объектов производственной инфраструктуры), дальнейшее изложение связано со спецификой объектов РТКС 4.0. При увеличении числа контролируемых объектов и сокращении числа сотрудников обнаружение аварий и классификация предаварийных состояний оборудования могут быть основаны только на технологиях «узкого» искусственного интеллекта

(Narrow AI), предназначенных для решения частной технической задачи. В РТКС 4.0 такой задачей является детектирование аномалий в потоках телеметрии от телекоммуникационного оборудования, которыми чаще всего считают пики или провалы в поступающих данных (точечная аномалия) и восходящие или нисходящие тренды с разными порогами скорости изменения данных.

Аномалии в данных могут быть отнесены к одному из трех основных типов [3]. Точечные аномалии (выбросы) возникают в ситуации, когда отдельный экземпляр данных может рассматриваться как аномальный по отношению к остальным данным. Контекстуальные аномалии наблюдаются, если экземпляр данных является аномальным лишь в определенном контексте. Для временных рядов (потоков телеметрии) контекстуальным атрибутом является время, в течение которого последовательность данных находится в аномальном состоянии по отношению к остальным данным. Такая временная аномалия (перебой) должна иметь различные таймеры для различных источников данных. Коллективные аномалии возникают, когда последовательность связанных между собой экземпляров данных (набор данных) от различных источников является аномальной по отношению к остальным наборам данных.

В статье описана нейро-нечеткая модель детектирования аномалий в потоках телеметрии распределенной телекоммуникационной сети на основе обработки признаков состояния оборудования искусственной нейронной нечеткой сетью прямого распространения с помощью математического аппарата нечетких множеств [4]. Основой модели является масштабатор/фазификатор, который определяет входные лингвистические термы по пикам или провалам данных и/или по восходящим или нисходящим трендам с разными порогами скорости изменения данных, что позволяет детектировать точечные и временные аномалии. Структура нейронной нечеткой сети определяется экспертом для каждого проекта IoT. Выходной слой дефазификатора в нейро-нечеткой модели отсутствует, так как она предназначена для решения задач классификации, а не аппроксимации, и на выходе присутствуют только лингвистические термы. Реализованный на основе нейро-нечеткой модели нейро-нечеткий классификатор (ННК) имеет в своем составе драйверы для связи с оконечными устройствами и журнал тревог и событий.

Архитектура РТКС 4.0. Архитектура РТКС 4.0 разработана на основе одного из трех эталонных (референсных) паттернов архитектуры IoT (Example Architecture Patterns), которые являются примерами и основой для концептуализации

архитектуры конкретных систем IoT [5]. Прото-типом выбран паттерн архитектуры взаимодействия и управления посредством шлюза, поскольку шлюз действует как оконечная точка для глобальной сети, изолируя граничные узлы от этой сети, что позволяет локализовать операции и контроль (обеспечить граничную аналитику и вычисления) [6]. Схема организации потоков данных в архитектуре РТКС 4.0 приведена на рисунке 1. Ключевым отличием архитектуры РТКС 4.0 от паттерна является то, что предлагаемая в паттерне единая точка сопряжения между LAN и WAN (устройство Edge Gateway/Hub) заменена на три типа устройств: Edge Hub, Edge/Sensing Gateway и Edgcell Gateway. Шлюзы беспроводной сети Edgcell Gateway работают под управлением Edge/Sensing Gateway по интерфейсу RS-485 и обеспечивают прием информации от датчиков с батарейным питанием. Реализованный на основе предлагаемой нейро-нечеткой модели ННК выполняется на устройствах Edge Hub.

Устройства Edge Hub одновременно опрашивают устройства Edge/Sensing Gateway путем перебора их IP-адресов через локальную сеть по протоколу MODBUS RTU over Ethernet. Полученные данные сохраняются во внутренней памяти Edge Hub и обрабатываются масштабатором/фаззификатором, который программным модулем «ЦЗ/ЛП» преобразует цифровое значение (ЦЗ), представленное в виде

16-разрядного числа со знаком, в лингвистическую переменную (ЛП), которая может принимать значение одного из пяти лингвистических термов.

Дальнейшая обработка производится искусственной нейронной нечеткой сетью прямого распространения (ННС), структура которой определена экспертом. Прямое распространение сигнала в ННС обеспечивается через нейроны, имеющие два нечетких входа и один нечеткий выход. Каждый нейрон имеет собственную нечеткую передаточную функцию (НПФ), также определяемую экспертом, причем эти функции не обязательно должны совпадать. После обработки всех скрытых слоев ННС на выходе формируются одна или несколько лингвистических переменных, которые помещаются в ячейки «Сообщение», обрабатываемые журналом тревог и событий. Это сообщения могут отсылаться в случае аномалий на e-mail сотрудников предприятия через локальный SMTP-server, или быть доступны для облачных сервисов как топики на встроеном в Edge Hub MQTT-брокере.

Масштабатор/фаззификатор. Масштабатор/фаззификатор (МФ), разработанный в нейро-нечеткой модели, кроме точечных аномалий способен обрабатывать также временные аномалии (перебои), причем он определяет по два несвязанных между собой таймера для каждого регистра, считанного по протоколу MODBUS.

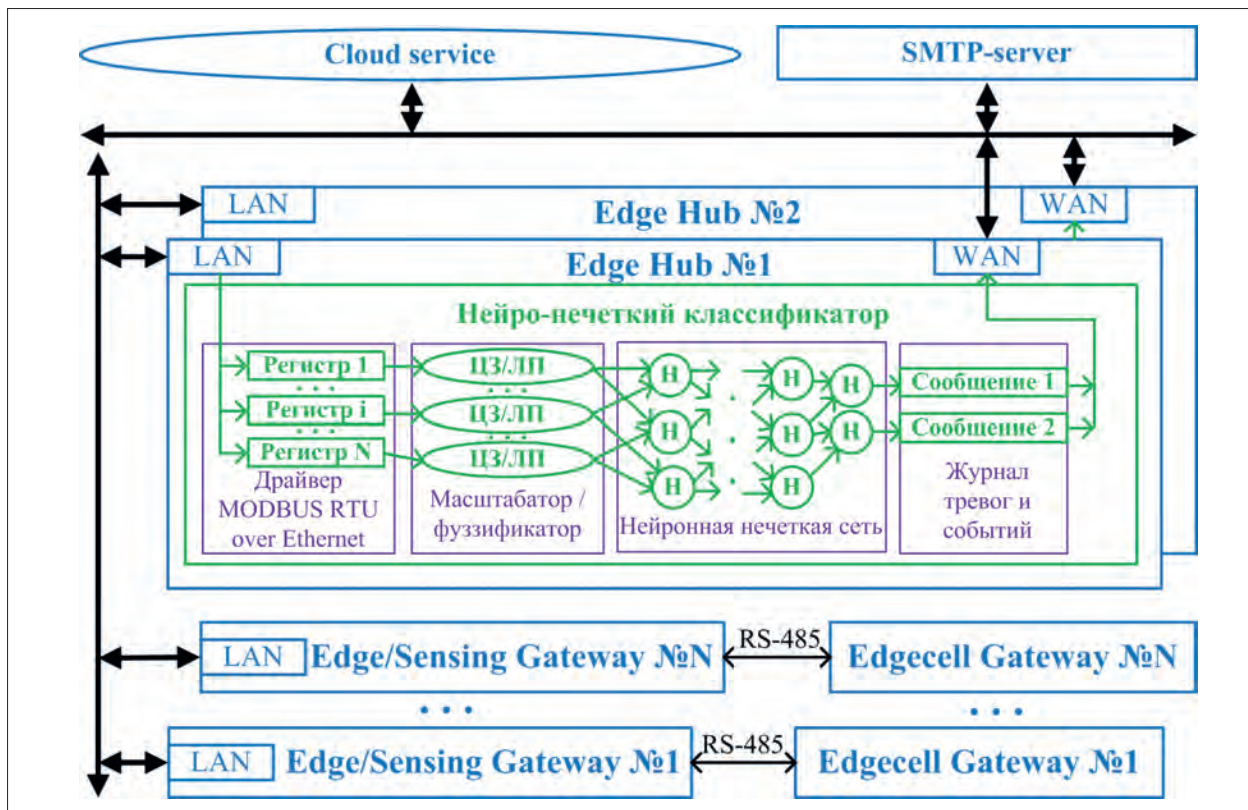


Рисунок 1 – Схема организации потоков данных в архитектуре РТКС 4.0

На рис. 2 показана экранная форма настройки параметров регистра № 7 для Edge/Sensing Gateway № 1. Значения Min и Max применяются к данным регистра первыми и могут быть в диапазоне от -32768 до 32767. На основании этих параметров проверяется, что полученные данные регистра № 7 (значение Value) приводятся к масштабу от Min до Max и над этим значением может проводиться фаззификация. В ННК используется алфавит лингвистических термов LA (Low Alarm), LW (Low Warning), G (Good), HW (High Warning), HA (High Alarm), что позволяет привести все параметры разного физического типа к набору лингвистических переменных в едином алфавите.

Метод фаззификации определяется таймерами верхнего и нижнего порога в секундах: если таймеры равны максимальному значению 65535, то для обработки данных регистра (обнаружения точечной аномалии) и получения значения лингвистической переменной LV применяются базовые точки BP1, BP3 и вычисляемые нижний и верхний пороги TPL, TRH. Если один или оба таймера не равны максимальному значению 65535, то для обработки данных регистра (обнаружения временной аномалии) значения лингвистической переменной LV изменяют свои значения в зависимости от того, равен или нет таймер 0, причем каждый таймер декрементируется ежесекундно.

Детектирование аномалий по выбросам или перебоям с учетом различных установок таймеров для различных регистров позволяет выявлять временные аномалии при помощи метода скользящего окна, причем размер окна для различных параметров различен. В таймеры загружается начальное значение TL или TH, если параметр перешел в состояние Good1, что позволяет фильтровать кратковременные перебои.

На рис. 2 показан пример настройки параметра «Сетевое напряжение на входе ИБП», который считывается устройством Edge/Sensing Gateway из ИБП по интерфейсу RS-232, затем значение параметра записывается в регистр MODBUS и передается в устройство Edge Hub по команде чтения. Для сетевого напряжения в 230 Вольт с допустимым диапазоном напряжений от 198 до 253 Вольт рассчитаны значения TPL=198 и TRH=241. Поскольку диспетчера интересуют факт превышения входного напряжения над нормой (выброс напряжения) и факт превышения времени отсутствия входного напряжения (перебой питания), то таймеры равны TL=120 и TH=65535.

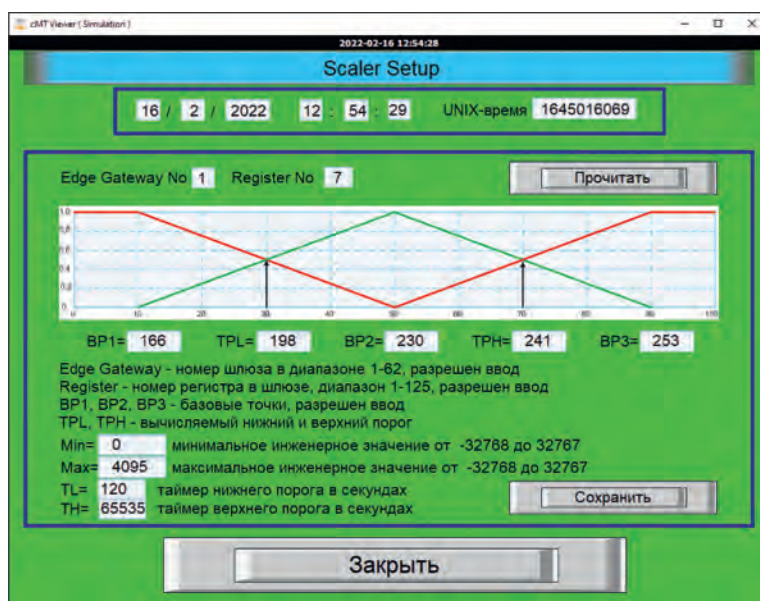


Рисунок 2 – Экранная форма настройки параметров регистра

Метод масштабирования/фаззификации описывается алгоритмом, представленным на рис. 3. Алгоритм описывает только логику метода масштабирования/фаззификации и не содержит дополнительных проверок, связанных с организацией вычислительных процессов в устройстве Edge Hub.

Нейронная нечеткая сеть. Детектирование аномалий основано на предположении, что нормальное поведение системы может определяться одним или несколькими классами, в которых могут находиться значения поступающих данных при нормальном поведении. Экземпляр, не принадлежащий ни к одному из классов, является аномалией. Система, основанная на классических нейронных сетях [7], может распознавать как один, так и несколько классов нормального поведения.

В работе предложен подход к организации ННС, использующий механизм нечеткого логического вывода, основанный на базе знаний, формируемой специалистами предметной области (экспертами) в виде нечеткой передаточной функции (НПФ) для каждого нейрона в ННС. Поскольку ННС может иметь несколько выходов, обрабатываемых входящими в ее состав подсетями, то она может распознавать несколько классов нормального поведения, относящихся к разным подсистемам контролируемой системы. Функционирование ННС обеспечивается нейронами, имеющими два нечетких входа и один нечеткий выход и собственную НПФ, состоящую из 25 правил, организованных в матрицу 5x5. Настройка НПФ производится путем определения значения выходной лингвистической переменной для 25 правил (всех комбинаций двух входных лингвистических переменных).

Эксперимент. Тестовый набор данных содержит шесть классов потоков данных по 6000 отсчетов в каждом, которые были созданы синтетическим путем, и выбран для тестирования программного обеспечения как адекватный пример потока телеметрии от шести различных источников данных следующих типов [8]: нормальный поток данных, циклический поток данных, поток данных с тенденцией к росту, поток данных с тенденцией к снижению, поток данных со сдвигом вверх, поток данных со сдвигом вниз.

Тестирование нейро-нечеткой модели, созданного на ее основе классификатора и разработанного программного обеспечения проводилось

на инструментальном стенде, в состав которого входило устройство Edge/Sensing Gateway для генерации шести каналов в потоке телеметрии на основе тестового набора данных и инструментальная ПЭВМ для симуляции Edge Hub в режиме анализа потока телеметрии и визуализации процессов, происходящих в стенде. Устройство Edge/Sensing Gateway выполняет преобразование данных из исходного формата в плавающей запятой (float) в 16-битный целочисленный формат со знаком с учетом коэффициента усиления 100.

На рис. 4 показан экран инструментальной ПЭВМ, в левой части которого видна WEB-страница устройства Edge/Sensing Gateway со списком выполняемых задач и состоянием каналов

Reg1-Reg6 в потоке телеметрии, а в правой части представлен экран устройства Edge Hub с получаемыми данными в цифровом виде и на индикаторе, а также результаты работы масштабатора/фаззификатора в виде лингвистических переменных LV. Визуальный контроль функционирования инструментального стенда осуществляется при помощи основных параметров, причем синие окружности с цифрой внутри и исходящей стрелкой обозначают тип контролируемого параметра, а в черном эллипсе находится значение этого параметра.

Параметр 1 показывает число обновленных WEB-страниц на текущий момент времени. Параметр 2 показывает текущее значение индекса, по которому считываются данные для одного из шести регистров. Параметры 3 и 4 показывают значения Min и Max в целочисленном формате

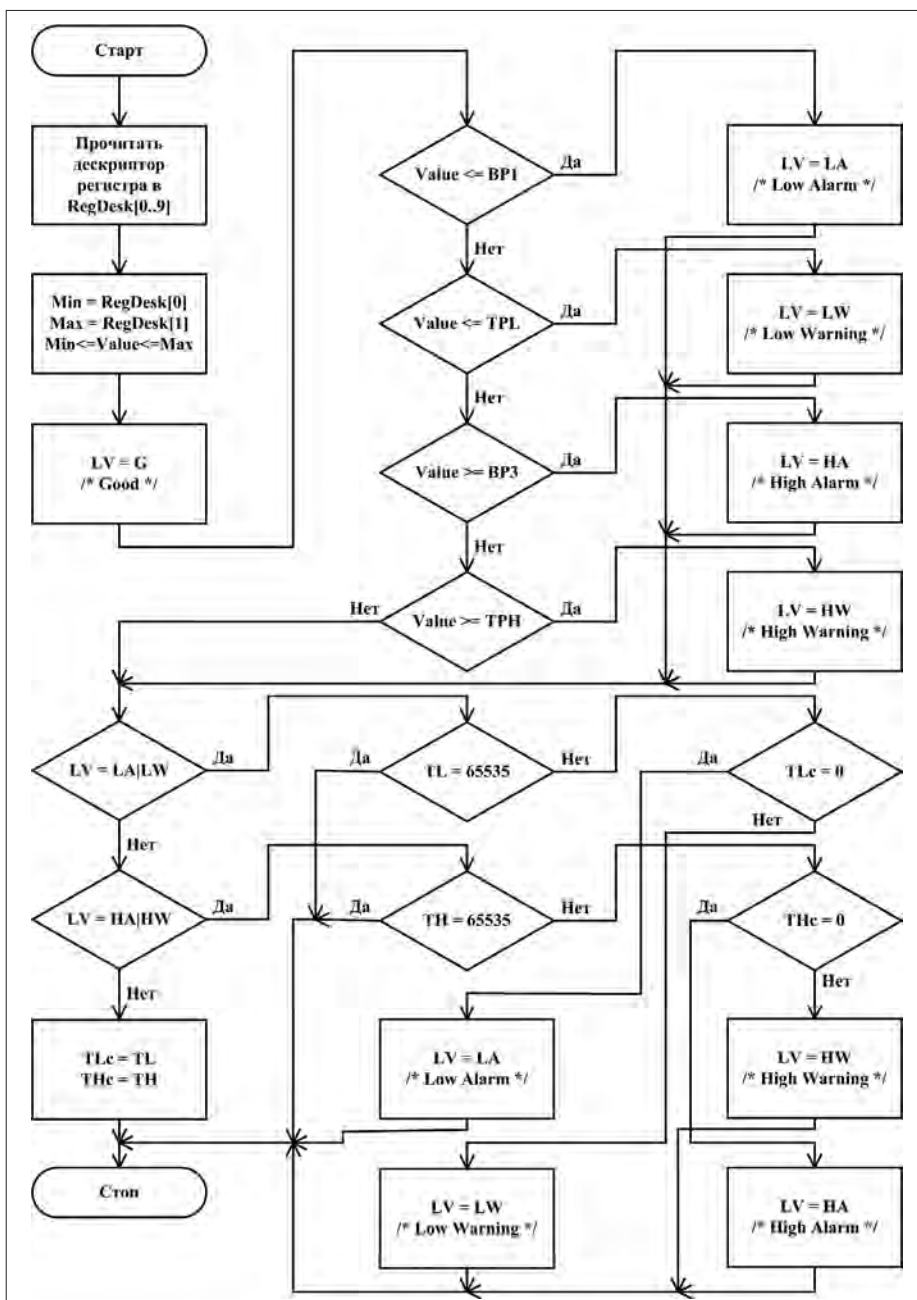


Рисунок 3 – Алгоритм масштабирования/фаззификации

для исходных значений -5.1149 и $+35.8439$ с учетом их умножения на 100. Параметр 5 показывает значение регистра Value, отправляемое Edge/Sensing Gateway и принимаемое Edge Hub, для контроля его целостности при передаче. Эти данные в алгоритме на рис. 3 обозначены как Value, и на их основании вычисляется лингвистическая переменная LV, на которую указывает параметр 6.

Заключение. Разработанный на основе нейронечеткой модели нейро-нечеткий классификатор обеспечивает детектирование точечных и временных аномалий с настраиваемой длиной скользящего окна для данных от разных источников, что устраняет чувствительность предложенной модели к размеру окна при сохранении возможности обнаружения различных типов аномалий.

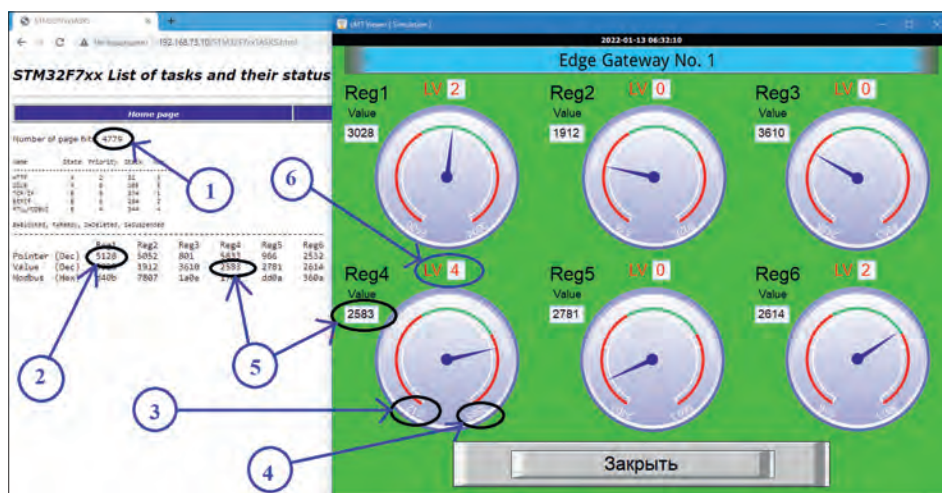


Рисунок 4 – Экран инструментальной ПЭВМ

Предложенное решение позволяет анализировать потоки данных телеметрии в распределенных телекоммуникационных системах с числом контролируемых параметров до 8000 на один Edge Hub с максимальным числом устройств Edge/Sensing Gateway до 62 и с числом регистров на каждом из них до 32.

Нейро-нечеткий классификатор входит в состав разработанного программного пакета «Библиотека управляющих программ для устройства Edge Hub».

ЛИТЕРАТУРА

- 5 Things to Know About the IoT Platforms Market. [Electronic resource]. – Mode of access: <https://iot-analytics.com/5-things-to-know-about-iot-platforms-market/>. – Date of access: 09.01.2022.
- Вишняков, В. А. Платформы для разработки сетей «интернета вещей» / В. А. Вишняков // «Вестник связи». – 2021. – № 2. – С. 62–65.
- Chandola, V., Banerjee, A., Kumar, V. Anomaly detection: A survey, ACM Computing Surveys, vol. 41(3), 2009, pp. 1–58.
- 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.
- The Industrial Internet of Things. Volume G1: Reference Architecture. [Electronic resource]. Mode of access: www.iiconsortium.org/IIC_PUB_G1_V1.80_2017-01-31.pdf. - Date of access: 10.01.2022.
- Радишевская, Т. А., Радишевский, В. А. Архитектурная модель слоя зондирования Промышленного Интернета Вещей // Проблемы инфокоммуникаций. – 2021 г. – № 1 (13). – С. 69–75.
- Рутковская, Д., Пилиньский, М., Рутковский, Л. Нейронные сети, генетические алгоритмы и нечеткие системы / пер. с польск. И. Д. Рудинского. М.: Горячая линия – Телеком, 2006. 452 с.
- Synthetic Control Chart Time Series Data Set. [Electronic resource]. – Mode of access: <https://archive.ics.uci.edu/ml/datasets/synthetic+control+chart+time+series>. – Date of access: 12.01.2022.

The article presents a neuro-fuzzy model for detecting anomalies in telemetry flows of distributed telecommunication systems. Signs of the state of the equipment are processed by an artificial neural fuzzy network of direct propagation using the mathematical apparatus of fuzzy sets. The basis of the model is a scaler/fuzzifier that determines input linguistic terms from data peaks or troughs and/or from uptrends or downtrends with different data rate thresholds. Adjustable sliding window length for data from different sources eliminates the sensitivity of the proposed model to the size of the window when detecting different types of anomalies. The neuro-fuzzy model is implemented as a library of subroutines in the EasyBuilder Pro development environment for cMT-SVR devices.

Keywords: neuro-fuzzy model; anomaly detection; scaler; fuzzifier; linguistic term.

Получено 31.03.2022.