

## ЭКСПЕРИМЕНТАЛЬНАЯ СРЕДА: СИМУЛЯТОР IBSIM-EDGE

*Евдокимова И.А., Андриалович И.В.*

*Белорусский государственный университет информатики и радиоэлектроники,  
г. Минск, Республика Беларусь*

*Научный руководитель: Лихачевский Д.В. – к.т.н., доцент, доцент кафедры ПИКС*

**Аннотация.** Подход TPRUDF был разработан на основе симулятора «IBSim-Edge» [1], обеспечивая суть среды ИВ в реальном времени. Симулятор IBSim-Edge моделирует распространение и обработку потоковых записей данных, генерируемых устройствами ИВ в среде граничных вычислений [2]. Он фиксирует поведение гетерогенной инфраструктуры Интернет вещей и периферийных вычислений и позволяет пользователю тестировать предлагаемую инфраструктуру, наборы данных и алгоритмы [3], а не просто моделировать потребление ресурсов в умных пространствах [4].

**Ключевые слова:** симулятор, слияние, Интернет вещи, обработка данных, MEL, DAG, SID LID, наборы данных

**Введение.** Симулятор IBSim-Edge предлагает множество функций, в том числе неоднородность устройств ИВ для поддержки различных доменов, состав набора данных, приложения ИВ и ресурсы обработки. Он представляет приложение ИВ в виде направленного ациклического графа (DAG) устройств ИВ, генерирующих микроэлементы (MEL). Каждый MEL – это абстрактный компонент приложения ИВ, который полностью инкапсулирует сервисы и данные. Единый MEL создается устройством ИВ и передается на облачный сервер для обработки.

**Основная часть.** Можно улучшить симулятор в классе MEL, добавив следующие дополнительные метаданные, которые идентифицируют MEL для поддержки целей обработки TPRUDF:

1 Идентификатор источника (SID): идентификатор устройства ИВ, который создает набор данных.

2 Идентификатор местоположения (LID): идентификатор пограничного устройства, которое передает MEL на сервер.

3 Размер обработки.

4 Время генерации (GT): время, когда MEL генерируется устройством ИВ.

Каждое устройство в приложении ИВ создает определяемый пользователем набор данных в форме MEL. MEL передаются на разные пограничные устройства (перед передачей в облачный центр обработки данных) для минимальной обработки. Затем каждое граничное устройство передает свои MEL в облачный центр обработки данных, поддерживая высокую скорость передачи данных ИВ и предотвращая узкие места в обработке. Следовательно, симулятор IBSim-Edge моделирует непрерывно генерируемые массивные MEL разнородных устройств ИВ с их метаданными для обработки на облачном сервере.

Рисунок 1 представляет модифицированную системную архитектуру IBSim-Edge после ее интеграции со слоями TPRUDF.

Слой в TPRUDF «Ресурсы вывода данных, используемые для слияния пространственно-временных данных» добавляется поверх слоя «IBSim-Edge Orchestrator», в котором MEL анализируются из файла DAG.

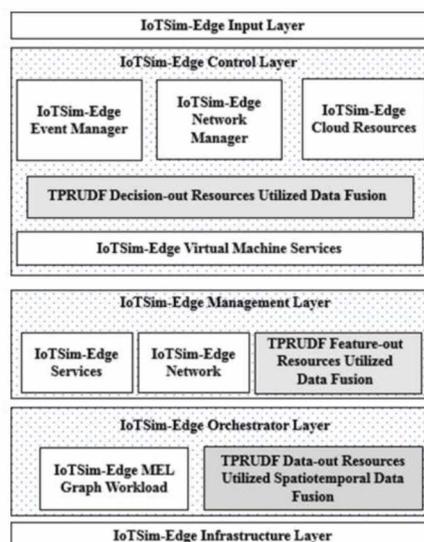


Рисунок 1 – Системная архитектура механизма (TPRUDF), интегрированного с симулятором IBSim-Edge

Этот слой работает на каждом отдельном пограничном устройстве на уровне «IBSim-Edge Management», где выполняется задача извлечения признаков. В конце добавляется на слой «Управление IBSim-Edge», где задачи принятия решений выполняются с использованием облачного сервера. Тем не менее, симулятор не обрабатывает представленные функции данных ИВ, продемонстрированные ранее, также не рассматривает какие-либо подходы к использованию ресурсов для обработки MEL. Следовательно, интеграция уровней подхода TPRUDF с симулятором IBSim-Edge предложит комплексную среду обработки ИВ и эффективное использование облачных ресурсов.

**Экспериментальная методология и наборы данных.** В этом пункте представлена экспериментальная методология, использующая наборы данных и связанную облачную среду, а также конфигурации моделирования для проверки эффективности слоев TPRUDF. Эксперименты проводились на машине с ядром i7, 2,70 ГГц, 1 Тб на жестком диске и 8 ГБ ОЗУ. Для настройки пограничной среды три разных размера обработки пограничных устройств настраиваются для имитации различных приложений ИВ следующим образом: (Ed1, 50 КБ), (Ed2, 100 КБ) и (Ed3, 500 КБ). Чтобы смоделировать облачную среду для экспериментов, настроен один облачный центр обработки данных с 1000 хостами и 30 виртуальными машинами (VM). Каждый хост имеет 64 ядра, где каждое ядро имеет 1000 MIPS и 144 000 МБ ОЗУ. Несколько виртуальных машин были настроены на 2000 MIPS, 8192 МБ ОЗУ и пропускную способность 1000 МБ/с:

1 Наборы данных для трех тематических исследований: производительность TPRUDF оценивается с использованием трех наборов данных для трех реальных тематических исследований из разных доменов ИВ с различными форматами и сценариями обработки.

В последнее время ИВ играет жизненно важную роль в интеллектуальных средах, особенно в прогнозировании погоды. Одной из интеллектуальных систем прогнозирования погоды является прогнозирование солнечного ветра, целью которого является прогнозирование времени прихода быстрых потоков солнечного ветра на Землю с использованием измерений солнечного ветра в режиме реального времени, таких как: давление, направление ветра, видимость неба, температура, местоположение, категория данных, ощущение и дата/время. Первый набор данных (D1), который мы исследуем для оценки TPRUDF, связан с такой системой [5]. Размер этого набора данных составляет 1,3 МБ и представлен 820 записями в формате JSON.

С точки зрения умного дома успех умных домов в основном зависит от принятия людьми этой концепции в контексте их повседневной жизни. Таким образом, одна из систем умного дома представляет собой систему рекомендаций по услугам умного дома, которая фиксирует деятельность каждого и предлагает определенные услуги, адаптированные для каждого человека в зависимости от его повседневного образа жизни. Система анализирует связь между услугами умного дома и одиннадцатью атрибутами пользовательского опыта, такими как: активное время суток, время активных и приключенческих игр, будильники, книги по искусству, дополненная реальность, езда на велосипеде, бокс, общение и готовка. Второй набор данных (D2) относится к такой системе [5]. Размер этого набора данных составляет 8,17 МБ и представлен 12 687 записями в формате CSV.

Наконец, для интеллектуального транспорта предлагается система прогнозирования загруженности дорог на основе ИВ с использованием интеллектуальных датчиков, установленных на трассе и прилегающих улицах, для регистрации количества пешеходов, велосипедистов, автомобилей и других транспортных средств. Третий набор данных (D3) связан с этой системой [68]. Размер этого набора данных составляет 56,86 МБ, он представлен 52 814 записями в формате CSV и определяется десятью атрибутами, такими как: скорость, направление, автомобиль, пешеход, велосипедист, мотоцикл, автобус, дата/время и местоположение. Детали трех использованных наборов данных для трех принятых практических примеров приведены в Таблице 1.

Таблица 1 – Краткое описание трех используемых общедоступных реальных наборов данных

Набор данных	домен Интернет вещей	Размер в МБ	Формат данных	Количество записей	Количество атрибутов
D1	Умная среда	1,3	JSON	820	8
D2	Умные дома	8,17	CSV	5687	11
D3	Умный транспорт	5686	CSV	32814	10

2 Конфигурации. Перед конфигурацией моделирования было настроено три приложения ИВ (App1, App2, App3), связанные с наборами данных трех принятых реальных примеров (D1, D2, D3) и спецификациями обработки трех периферийных устройств (Ed1, Ed2, Ed3) [68]. Каждое приложение содержит несколько устройств ИВ, которые отвечают за моделирование подключенного набора данных, генерируя MEL каждую секунду.

App1 имеет 10 устройств ИВ (SID<sub>1</sub>, SID<sub>2</sub>, ..., SID<sub>10</sub>), App2 имеет 20 ((SID<sub>1</sub>, SID<sub>2</sub>, ..., SID<sub>20</sub>),) ИВ-устройства, а App3 имеет 40 ИВ устройств (SID<sub>1</sub>, SID<sub>2</sub>, ..., SID<sub>40</sub>). Каждый отдельный MEL имеет фиксированный размер обработки равный 10 КБ для хранения данных подмножества из исходного набора данных. Например, App1 имеет 131 MEL для D1, App2 – 817 MEL для D2, а App 3 – 5686 MEL для D3. Мы настроили фиксированный размер выборки равный 6, 20 и 70 для выборки MEL в App1, App2 и App3 с использованием STDF Редуктора данных ИВ для достижения максимального уровня точности и интервала обновления 1000 мс для результатов выборки STDF. Эти конфигурации считываются симулятором IBSim-Edge на этапе инициализации для моделирования данного приложения ИВ.

3 Исходный сценарий IBSim-Edge. Исходный сценарий обработки симулятора IBSim-Edge начинается со всех устройств ИВ в группе обеспечения доступности баз данных, которые генерируют MEL на основе присоединенных наборов данных, а затем передают смесь сгенерированных MEL на граничные устройства, которые, в свою очередь, передают их для обработки в облачном центре обработки данных. Этот сценарий повторяется каждую секунду.

4 IBSim-Edge после интеграции TPRUDF.

5 После интеграции TPRUDF диспетчер данных на основе ИВ на уровне «Ресурсы вывода данных, используемые для слияния пространственно-временных данных» развертывается на всех граничных устройствах, которые каждую секунду получают MEL, для авторизации, чтения, обслуживания, группировки и выборки включенных второстепенных наборов данных ИВ из одинаковой SID и передачи их для обработки на сервер.

6 На стороне облачного центра, используя менеджер объединения данных на основе ИВ в том же уровне, вместо случайной обработки MEL в облачном центре обработки данных в исходном сценарии, все выборочные MEL каждого SID группируются в соответствии с одним и тем же LID (идентификатор пограничного устройства), в котором они передаются. Затем дерево сгруппированных MEL на LID строится для каждого SID для агрегирования самых свежих MEL (т. е. MEL с минимальным GT) для каждого местоположения.

7 После этого уровень объединения данных об использовании ресурсов для извлечения объектов получает агрегированные MEL для каждого SID и формулирует свои наборы данных в  $Z \times X$  матрице для уменьшения размера их наборов данных и вывода  $Y \times X$  матрицы путем удаления коррелированных функций через PCA.

**Заключение.** Наконец, слияние данных, используемых для принятия решения, получает  $K \times X$  некоррелированные матрицы функций и доступные виртуальные машины в облачном центре обработки данных, чтобы рекомендовать наиболее оптимальные виртуальные машины для их обработки с использованием двух методов использования ресурсов: GA и PSO, и выбрать лучшие результирующие виртуальные машины из обоих методов, используя метод слияния логики. Этот сценарий повторяется каждую секунду для всех устройств SID в приложении Интернет вещей.

### Список литературы

1. Jha DN, Alwase K. I, Areeb Alshoshan XH, Naha RK. SKSG Battula, IoT-Sim-Edge/ DN Jha, K. Alwaseel, XH Areeb Alshoshan, RK Naha, SKSG Battula, // среда моделирования для моделирования поведения Интернета вещей и сред граничных вычислений, Software. Практика. – Эксп. 50 – 2020 – pp 844–867.
2. Пасрича С., Аюб Р., Кишиневский М., Мандал С., Ограс У. Обзор по управлению энергопотреблением для мобильных устройств и устройств/ С. Пасрича, Р. Аюб, М. Кишиневский, С. Мандал, У. Ограс // IoT, IEEE Des. Тест. – №37– 2020 – pp 7–24.
3. С. Мессауд, А. Брадай, С. Хашиш, Р. Бухари, П. Тран, А. Куанг, О. Ахмед, М. Атри, Обзор машинного обучения в Интернете вещей: алгоритмы, стратегии и приложения, Интернет вещей – 12–2020 – pp 100-314.
4. Мусса С.. Многоагентный симулятор для управления ресурсами в интеллектуальных пространствах/ С. Мусса, Г. А. Ага//Материалы совместных конференций по всеобъемлющим вычислениям (JCPCC), IEEE – 2009 – 765–770 стр.
5. iot datasets . – [Электронный ресурс]. URL: <https://data.world/datasets/iot>. – Дата доступа: 23.02.2022.

UDC 004.021

## EXPERIMENTAL ENVIRONMENT: SIMULATION AND IVSIM-EDGE

I.A. Evdokimova, I.V. Andryalovich

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*D.V. Likhachevsky– Cand. of Sci., associate professor, associate professor of the department of ICSD*

**Annotation.** The TPRUDF approach was developed based on the simulator «IVSim-Edge» simulator [1], providing the essence of a real-time IV environment. The IVSim-Edge simulator models the propagation and processing of streaming data records generated by UI devices in a boundary computing environment [2]. It captures the behavior of heterogeneous Internet of Things and edge computing infrastructure and allows the user to test the proposed infrastructure, datasets and algorithms [3], rather than just simulating resource consumption in smart spaces [4].

**Keywords:** simulator, fusion, Internet of Things, data processing, MEL, DAG, SID, LID, dataset