

Отзыв

на автореферат диссертации Бахтина Вадима Вячеславовича на тему «Метод синтеза нейросетевых устройств для реализации режима fog computing», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.2 – Вычислительные системы и их элементы.

«Туманные вычисления» — разновидность облачных вычислений, когда обработка данных происходит не в едином дата-центре, а в распределенной сети устройств. По оценкам Market Research Future, к 2030 году объем рынка fog computing должен достигнуть планки в \$343 млн. Неудивительно, что необходимы решения, позволяющие развернуть подобные вычислительные процессы в формате р2р.

С другой стороны, сейчас уже сложно представить какую-либо область деятельности человека, в которой не использовались бы искусственные нейронные сети. Однако для значительного повышения нейросетевых вычислений требуются сложные высокопроизводительные вычислительные системы.

Решению этих двух актуальных проблем и посвящена диссертационная работа Бахтина В.В., исследующего вопросы эффективного синтеза распределённых высокопроизводительных нейросетевых устройств для реализации туманных вычислений.

Новизна научных результатов диссертационного исследования состоит в том, что:

1. Разработана математическая модель искусственной нейронной сети для синтеза нейросетевых устройств, ориентированных на туманные вычисления. Она отличается от существующих тем, что с ее помощью возможно балансировать размеры декомпозированных блоков нейронной сети в зависимости от характеристик физических устройств, входящих в вычислительный каскад. Это позволяет учитывать требуемую загрузку вычислительных узлов при распределении блоков нейронной сети между различными устройствами.

2. Разработан метод синтеза вычислительных устройств для реализации искусственных нейронных сетей, ориентированных на туманные вычисления. Он отличается от существующих тем, что учитывает такие параметры, как мощность объединяемых устройств, объем передаваемых данных, пропорциональность блоков нейронной сети по слоям или по нейронам, а также имеет возможность реализации диагностики и реконфигурации. Всё это позволяет выбрать оптимальный вариант конфигурации нейронной сети по заранее выбранным метрикам и продолжать работу даже в случае отказа или

сбоя части устройств в каскаде.

3. Разработан алгоритм декомпозиции монолитной нейронной сети на каскад блоков блочной нейронной сети, адаптированной для туманных вычислений. Он отличается от существующих тем, что предлагает способ унификации хранения данных в памяти монолитной нейронной сети и блоков в блочной нейронной сети. Это позволяет проводить многократную декомпозицию распределённой системы в глубину, если потребуется, например, разбиение отдельного блока еще на несколько компонентов.

4. Разработан алгоритм выбора оптимального варианта декомпозиции нейронной сети для реализации её на распределённых вычислительных устройствах. Он отличается от существующих тем, основан на многокритериальной оптимизации путем нахождения Парето-оптимальных вариантов. Это позволяет находить оптимальную декомпозицию монолитной нейронной сети сразу по нескольким важным для вычислительной системы параметрам.

Теоретическое значение работы состоит в том, что разработанные модель, метод и алгоритмы имеют хороший уровень проработки, который обеспечивает их применимость для широкого круга задач, связанных с реализацией искусственных нейронных сетей в распределённых вычислительных системах.

Практическая значимость работы заключается в том, что предложенные метод и алгоритмы могут использоваться при разработке и эксплуатации вычислительных систем, состоящих из большого количества компонентов и используемых в составе сложных вычислительных комплексов.

В качестве **замечаний** к автореферату можно отметить следующие пункты:

1. В автореферате недостаточно подробно рассмотрен вопрос обучения распределённой нейросети и особенностей формирования обучающей выборки с учётом сложноструктурированной архитектуры ИНС.

2. В автореферате отсутствуют обоснования для вывода формул оценки среднего времени отклика пульта на входные сигналы и оценки времени выполнения нейросетевых вычислений. Почему в сумму вошли только эти слагаемые? Отражают ли они все этапы работы системы при оценке её производительности?

3. Не все переменные в формулах, представленных в автореферате, расшифрованы в тексте автореферата, что затрудняет понимание сути этих формул.

Несмотря на указанные замечания, считаю, что диссертационная работа выполнена на достаточно высоком уровне. Представленные результаты и выводы не противоречат результатам подобных российских и зарубежных

исследований, подтверждаются апробацией на практике, докладами на научных конференциях, статьями в журналах и т.п.

Считаю, что работа Бахтина В.В. соответствует требованиям п. 9 «Положения о присуждении ученых степеней» ВАК, утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. №842, а соискатель Бахтин Вадим Вячеславович заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.2 – Вычислительные системы и их элементы.

Я, Мельцов Василий Юрьевич, даю своё согласие на включение своих персональных данных в документы, связанные с работой диссертационного совета, и их дальнейшую обработку.

Кандидат технических наук, доцент,
доцент кафедры электронных
вычислительных машин
Вятского государственного университета,
руководитель СКБ «Современные
микропроцессорные системы»

В.Ю.

/ В. Ю. Мельцов /

«05» июня 2023 г.



венноручную подпись
Мельцова В.Ю. заверяю

Начальник управления по работе
с персоналом

Михайленко Е.Н.

Мельцов Василий Юрьевич

Кандидатская диссертация защищена по специальности: 05.13.13 –
Вычислительные машины, комплексы, системы и сети.

E-mail: meltsov@vyatsu.ru

Федеральное государственное бюджетное образовательное учреждения
высшего образования «Вятский государственный университет» (ФГБОУ ВО
«ВятГУ»)

610000, Кировская область, г. Киров, ул. Московская, д.36,
телефон: +7 (8332) 742-421 .