

Искусственный интеллект: от мемристора до применения систем

Стратегический диалог и секция научно-технической конференции по теме искусственного интеллекта на форуме «Микроэлектроника 2019»

Ю. Ковалевский

В 2019 году научно-технические доклады по теме нейроморфных вычислений и искусственного интеллекта (ИИ), представленные на конференции «ЭКБ и микроэлектронные модули» в рамках Международного форума «Микроэлектроника 2019», прошедшего в Алуште с 30 сентября по 5 октября 2019 года, были впервые объединены в отдельную секцию. Это позволило специалистам по микроэлектронным технологиям, процессорной архитектуре, программным алгоритмам, объединенных этой специфической темой, познакомиться с новейшими достижениями и – главное – существующими требованиями и проблемами в смежных областях. Для успеха в создании отечественных систем ИИ такое общение – необходимое условие. Но не достаточное: ключевым аспектом также является понимание перспектив применения разработок, рынка и потребностей заказчиков. Решению этой задачи был призван помочь прошедший на форуме стратегический диалог с участием как разработчиков, так и потребителей решений с применением ИИ.

СТРАТЕГИЧЕСКИЙ ДИАЛОГ «ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ЗАКАЗЧИКИ И РЫНКИ ДЛЯ ТЕХНИЧЕСКИХ РЕШЕНИЙ НА ОСНОВЕ РАЗРАБОТОК РОССИЙСКИХ ДИЗАЙН-ЦЕНТРОВ В СФЕРЕ ИИ»

Модераторами стратегического диалога выступили **А. А. Алясев, заместитель генерального директора по науке IVA Technologies**, и **А. Э. Мохнаткин, заместитель генерального директора ЗАО НТЦ «Модуль»**.

В рамках дискуссии слово было предоставлено представителям компаний и организаций, применяющих ИИ, с тем чтобы они поделились своим опытом внедрения интеллектуальных систем и видением дальнейшего развития этого рынка с учетом будущих требований к локализации производства электронной аппаратуры и увеличению доли российской ЭКБ в применяемых решениях, а также представителям российских компаний – разработчиков аппаратных средств для ИИ, которые рассказали о своих решениях, обозначили проблемы, которые они испытывают в отношении их развития и внедрения, и высказали свои пожелания и предложения для преодоления этих проблем.

Первым выступлением со стороны заказчиков стала презентация **Н. Б. Волковой (ГКУ «Инфогород»)**. Она

охарактеризовала Москву как крупнейший рынок, а органы исполнительной власти города – как крупнейших заказчиков систем на базе ИИ. По ее словам, на портале госуслуг Москвы в настоящее время представлено более 340 услуг в цифровой форме, которыми пользуется около 8 млн москвичей. В 2018 году Москва заняла первое место в рейтинге городов по развитию электронного правительства, составленном ООН.

В 2018 году была разработана концепция «Умный город – 2030». При этом было опрошено несколько тысяч экспертов, в разработку концепции включилось 5 тыс. горожан.

В рамках концепции рассматривается шесть направлений: «Человеческий и социальный капитал», «Городская среда», «Цифровая мобильность», «Городская экономика», «Безопасность и экология», «Цифровое правительство». По всем направлениям имеются крупные системы, среди которых многим известны Единая медицинская информационно-аналитическая система города Москвы (ЕМИАС), Московская электронная школа (МЭШ) и др. Н. Б. Волкова отметила, что в большей части этих продуктов уже начинается использование систем на базе машинного обучения, и привела ряд примеров такого использования.

Также прозвучала информация о том, что сейчас департаменту поручена проработка вопросов, связанных с развитием применения ИИ, в частности в правовой сфере для упрощения соответствующих экспериментов. Н. Б. Волкова предложила направлять в департамент свои предложения по снижению нормативно-правовых барьеров для внедрения отечественных разработок в области ИИ.



А. С. Кочуров, исполнительный директор ООО «Союзинфотех», поделился успешным опытом разработки системы на базе ИИ для обеспечения безопасности детей. Он отметил, что система была разработана на основе отечественных программных решений и с использованием в том числе аппаратных средств НТЦ «Модуль», что позволило реализовать функционал, который было бы крайне сложно создать на зарубежных решениях, поскольку российские разработчики аппаратных средств были очень плотно задействованы в этом процессе. Особенностью системы является отслеживание местонахождения детей не на основе датчиков или геолокации, а с помощью распознавания видеоданных с камер. Сейчас система активно внедряется в различных регионах России.

Вопрос применения отечественных аппаратных платформ крупными отечественными телекоммуникационными операторами волнует многих в электронной и радиоэлектронной отрасли, поскольку эти компании теоретически могли бы обеспечить значительный рынок для российской аппаратуры и, как следствие, ЭКБ. Однако **Э. А. Гайфутдинов, директор Центра по управлению технологическими партнерствами ПАО «Ростелеком»**, в своем выступлении отметил, что данная компания подходит к внедрению систем ИИ в свою инфраструктуру очень аккуратно, поскольку пока не видит законченных решений, позволяющих решить некую задачу так, чтобы это привело к экономии или другим очевидным положительным результатам. Сейчас внимание обращено на задачу анализа собираемых данных о состоянии сетей с целью прогнозирования возникновения неисправностей; ведется поиск решений, которые могли бы с ней справиться. В качестве перспективного направления Э. А. Гайфутдинов отметил распределенные вычисления, а именно принятие части решений в системах, критичных к задержкам, как можно ближе к базовым станциям или даже к абонентским устройствам, однако это – вопрос будущего.

Отвечая на вопросы участников мероприятия, Э. А. Гайфутдинов сказал, что на данный момент сложно деконструировать решения, которые могли бы быть интересны ПАО «Ростелеком», до компонентной базы.

Применение тех или иных аппаратных средств в области ИИ невозможно без наличия разработанного для них программного обеспечения, поэтому очень важным вопросом является то, какие аппаратные платформы выбирают для создания аппаратно-программных комплексов разработчики ПО. Этот вопрос осветил в своем выступлении **А. А. Цессарский, генеральный директор IVA Cognitive**. Он указал на высокие требования к вычислительной мощности аппаратуры в области систем машинного зрения и распознавания лиц, которая является основной специализацией компании. По этой причине в настоящее время разработки обычно ведутся на основе аппаратуры фирмы Nvidia. Однако компания видит ряд причин для более широкого использования отечественной ЭКБ и аппаратуры, среди которых А. А. Цессарский выделил три основных.

Первая причина – экономическая целесообразность, учитывая высокую стоимость зарубежных решений, которая часто превышает ожидания конечных заказчиков систем ИИ на рынке.

В качестве второй причины А. А. Цессарский назвал риски, связанные с возможными ограничениями поставок из-за рубежа и в целом с работой с единственным, базальтернативным поставщиком.

Третья причина заключается в том, что с отечественными компаниями работать проще: они ближе, более открыты к диалогу. Это может позволить не просто разрабатывать ПО на имеющейся аппаратуре, но воздействовать на аппаратную часть и во взаимодействии с ее разработчиком создавать лучше скомбинированные аппаратно-программные комплексы, способные стать уникальными решениями с большим потенциалом.

Говоря про возможности, имеющиеся у отечественных ИС для ИИ на рынке, А. А. Цессарский разделил потенциальный рынок на два основных направления.

Примером первого, серверного применения, служит обработка данных, поступающих с городских видеокамер. Даже построение ЦОД для интеллектуальной обработки видеoinформации в реальном времени требует весьма значительного бюджета, а если

к этому добавить дальнейшую аналитику на базе ИИ, могут потребоваться очень большие инвестиции. По мнению А. А. Цессарского, в плотном сотрудничестве с разработчиками аппаратуры и ЭКБ компания могла бы предложить решения с большей функциональностью при том же бюджете.

Второе направление – ИС для граничных вычислений, встраиваемые в устройства на последней миле, например в «умные» видеокамеры. Применение таких решений часто оправдано, в частности при ограниченной пропускной способности канала связи до ЦОД.

Было отмечено, что компания уже движется в сторону разработки отечественных программно-аппаратных комплексов в сотрудничестве с IVA Technologies. Также рассматриваются ИС компании НТЦ «Модуль». А. А. Цессарский пригласил и других разработчиков и производителей ЭКБ для решений на базе ИИ к диалогу с целью создания систем для обоих обозначенных секторов рынка.

Участники мероприятия задали А. А. Цессарскому ряд вопросов, один из которых касался причин применения компанией платформ Nvidia в настоящее время. Одной из важнейших таких причин было названо то, что с этими платформами удобно работать благодаря сформировавшемуся вокруг них сообществу. Подобных сообществ вокруг российских решений пока нет.

А. А. Тихонов, исполнительный директор АО «Лаборатория Касперского», сделал акцент на использовании ИИ для обеспечения информационной безопасности. Он поддержал тезис о том, что ИИ находит применение в различных частях сети, причем не только в ЦОД и на периферийных устройствах, но и по всей цепочке между ними. Он отметил, что угроз сейчас становится больше, а сами они – динамичнее, и приходится их «ловить на лету», с чем системы на базе машинного обучения могут справиться.

Учитывая повсеместное проникновение ИИ, рынок для микроэлектроники в данном направлении представляется очень большим. А. А. Тихонов выразил готовность содействовать внедрению таких российских решений в сетевом, серверном оборудовании и т. п.



Тему кибербезопасности продолжил **А. М. Федюлин, директор Центра разработки ПО АО «КТ – Беспилотные системы»**, который рассказал об особенностях применения ИИ в беспилотных летательных аппаратах, в частности указав, что ИИ играет важную роль в бортовых системах, поскольку многие решения должны приниматься непосредственно на борту для обеспечения автономности БПЛА.

Со стороны разработчика ИС для задач ИИ выступил **А. В. Черников, заместитель начальника отделения ЗАО НТЦ «Модуль»**. Он рассказал о видении компании в отношении развития ее аппаратных платформ для ИИ, согласно которому основное внимание должно уделяться не жестким аппаратным решениям, реализующим только один наперед заданный тип нейросетей, а программируемым платформам, позволяющим достаточно гибко настраивать вычислительный граф нейросети и, как следствие, решать различные задачи. Причина такой расстановки приоритетов в том, что архитектуры нейросетей постоянно развиваются, меняются форматы и рядность входных данных, и нейросетей, которые могли бы решать все виды задач, не существует.

При этом было отмечено, что полностью исключать из рассмотрения жесткую реализацию нельзя, поскольку такие аппаратные платформы демонстрируют более высокую производительность в решении распространенных в настоящее время задач, таких как распознавание образов.

Также А. В. Черников присоединился к уже озвученному ранее тезису о необходимости создания активного сообщества пользователей российских аппаратных платформ, подобного сообществу вокруг платформ Nvidia, что является критически важным условием развития отечественных систем ИИ.

От еще одного разработчика аппаратных средств ИИ выступил **Е. В. Терентьев, заместитель генерального директора IVA Technologies**. Рассказав о направлениях деятельности компании, основными из которых являются встраиваемые системы, серверные решения и IP-блоки, он сообщил, что в 2018 году компания создала IP-ядро ускорителя тензорных вычислений, которое постоянно совершенствуется и, несмотря на короткую историю существования, уже получило признание у крупных зарубежных заказчиков.

Е. В. Терентьев посетовал на то, что, несмотря на имеющиеся отечественные решения мирового уровня, существуют прецеденты, когда российские крупные компании предпочитают зарубежные продукты. Он отметил, что отечественные разработчики часто ставят в сложные условия, когда заказчик требует готовое решение, которое сравнивается с импортным аналогом, и при этом никаких гарантий закупки даже в случае превосходства аналога по характеристикам нет. В решении этой проблемы, по мнению Е. В. Терентьева, может помочь создаваемый консорциум

по направлению ИИ, в рамках которого можно будет наладить диалог между разработчиками и заказчиками, сформировать кооперационные цепочки и создать инструменты для обеспечения гарантированного заказа на решения, которые нужны заказчикам, удовлетворяют их заявленным требованиям и разработаны в установленные сроки.

В заключение мероприятия, которое вышло за пределы отведенного на него времени, модераторы отметили, что поднятая тема – очень широкая и многоплановая, и предложили продолжить обсуждение за рамками данного диалога.

СЕКЦИЯ «НЕЙРОМОРФНЫЕ ВЫЧИСЛЕНИЯ. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ»

Научные и практические технические вопросы направления ИИ обсуждались на отдельной секции конференции «ЭКБ и микроэлектронные модули» в рамках форума, на которой были представлены 21 презентация, три стендовых и один заочный доклад.

Модераторами секции выступили **член-корр. РАН, д. ф.-м. н. Б. В. Крыжановский, д. т. н., проф. Е. С. Горнев и к. т. н. О. А. Тельминов.**

Многие доклады касались мемристорных структур – перспективного направления как в создании энергонезависимой памяти высокой плотности и быстродействия, так и в аппаратной реализации нейронных сетей. По этой теме были представлены доклады коллективов авторов из **НИЦ «Курчатовский институт», МФТИ, ВГТУ, МГУ имени М. В. Ломоносова, АО «НИИМЭ», ИП СО РАН** и др. Ряд докладов был посвящен непосредственно материалам и физической реализации мемристорных структур. Одна из таких работ, в которой рассматривались мемристорное переключение и манипулирование атомами кислорода на поверхности MoO_2 , была подготовлена коллективом авторов из **ИФТТ РАН и School of Physics, Trinity College Dublin (Ирландия)**. Мемристорные структуры на основе $SiO_x (x \sim 1)$ послужили темой доклада **МФТИ** и **АО «НИИМЭ»**. Другой материал – селенид висмута – исследовался в работе коллектива авторов из **ИФТТ РАН и ИПТМ РАН.**



В другую группу можно объединить доклады, в которых рассматривались общие вопросы архитектур и принципов построения нейросетей, а также их применения. В частности, коллектив авторов из **НИЦ «Курчатовский институт»** и **МФТИ** рассмотрел программные и аппаратные реализации основных компонент спайковых и формальных нейронных сетей. Среди конкретных областей применений нейросетей и машинного обучения были рассмотрены задачи вычислительной литографии, криптографии (**АО «НИИМЭ»**) и оценки соответствия эталонам многокомпонентных жидкостей (**ИПФ РАН, ННГУ имени Н. И. Лобачевского и ФГБНУ СКФНЦСВВ**).

Также прозвучали доклады, посвященные конкретной ЭКБ и аппаратным решениям для реализации нейросетей, в частности нейроморфному процессору «Алтай» (**ООО «Мотив НТ»**) и процессорной СБИС 1879ВМ8Я (**НТЦ «Модуль»**). Методы достижения максимальной эффективности платформы прототипирования высокопроизводительных СнК на задачах ИИ были представлены коллективом авторов из **НПЦ «Элвис»**.

4 октября, подводя итоги работы секции, ее модератор **Б. В. Крыжановский** сказал, что до сих пор в России разработчики ПО и аппаратуры для ИИ существовали по отдельности. На данном мероприятии люди, которые занимаются разработкой нейропроцессоров, познакомились с «софтовиками» и узнали их проблемы, а «софтовики» – в каком виде необходимо создавать программы, чтобы их можно было имплементировать на отечественных микропроцессорах.

Модератор секции высоко оценил уровень докладов, отметив, что слабых выступлений были единицы. Также радостным моментом было названо то, что среди докладчиков было много молодежи.

«Дебют состоялся», – заключил Б. В. Крыжановский.

Фото предоставлены организаторами форума «Микроэлектроника 2019» <http://microelectronica.pro/>