



Известия высших учебных заведений. Прикладная нелинейная динамика. 2024. Т. 32, № 5
Izvestiya Vysshikh Uchebnykh Zavedeniy. Applied Nonlinear Dynamics. 2024;32(5)

Научная статья
УДК 530.182

DOI: 10.18500/0869-6632-003121
EDN: MJFDNA

Импульсный бинарный нейрон — детектор причинно-следственных связей*

М. В. Киселев¹✉, Д. А. Ларионов^{1,2}, А. М. Урусов^{1,3}

¹Чувашский государственный университет имени И. Н. Ульянова, Чебоксары, Россия

²Частное учреждение «Цифрум», Госкорпорация «Росатом», Москва, Россия

³Акционерное общество «Газпромбанк», Москва, Россия

E-mail: ✉mkiselev1@yandex.ru, denis.larionov@gmail.com, imurusov@mail.ru

Поступила в редакцию 31.10.2023, принята к публикации 11.04.2024,
опубликована онлайн 10.09.2024, опубликована 30.09.2024

Аннотация. Цель. Распознавание причинно-следственных связей является фундаментальной функцией нейронных сетей, обучающихся целенаправленному поведению, осуществляющих планирование действий и формирующих модели динамики внешнего мира. Эта функциональность особенно важна для реализации обучения с подкреплением. В контексте импульсных нейронных сетей события представлены в виде импульсов (спайков), испускаемых нейронами сети или входными узлами. Обнаружение причинно-следственных связей между этими событиями является необходимым для эффективной реализации обучения с подкреплением. **Методы.** В данной работе представлен новый подход к распознаванию причинно-следственных связей с использованием импульсного бинарного нейрона. Этот подход основан на специально разработанных простых и эффективных правилах синаптической пластичности. При этом учитываются временные аспекты обнаруженных причинно-следственных связей, а также то, что спайковые сигналы могут иметь вид как одиночных импульсов, так и плотных последовательностей импульсов (всплесков), как это наблюдается в биологическом мозге. Кроме того, в данном исследовании уделяется большое внимание вопросу эффективной реализации предложенных моделей на современных и перспективных нейропроцессорах. **Результаты.** В сравнении с точными методами машинного обучения, такими как алгоритмы деревьев решений и сверточные нейронные сети, наш нейрон демонстрирует удовлетворительную точность, несмотря на свою простоту. **Заключение.** В данной работе представлена архитектура импульсной нейронной сети, включающая нейроны описываемого типа, которая может эффективно применяться в более сложных информационных окружениях, что делает ее перспективным кандидатом для реализации обучения с подкреплением в импульсных нейронных сетях.

Ключевые слова: импульсные нейронные сети, бинарный нейрон, зависящая от времени спайков пластичность, модулированная дофамином пластичность, анти-Хеббовская пластичность, обучение с подкреплением, нейроморфная аппаратура.

Для цитирования: Киселев М. В., Ларионов Д. А., Урусов А. М. Импульсный бинарный нейрон — детектор причинно-следственных связей // Известия вузов. ПНД. 2024. Т. 32, № 5. С. 589–605. DOI: 10.18500/0869-6632-003121. EDN: MJFDNA

Статья опубликована на условиях Creative Commons Attribution License (CC-BY 4.0).

*Работа публикуется по материалам доклада, сделанного на конференции «Нейроинформатика — 2023».

A spiking binary neuron — detector of causal links*

M. V. Kiselev¹✉, D. A. Larionov^{1,2}, A. M. Urusov^{1,3}

¹I. N. Ulianov Chuvash State University, Cheboksary, Russia

²Cifrum, subsidiary of Rosatom, Moscow, Russia

³Gazprombank (JSC), Moscow, Russia

E-mail: ✉mkiselev1@yandex.ru, denis.larionov@gmail.com, imurusov@mail.ru

Received 31.10.2023, accepted 11.04.2024, available online 10.09.2024, published 30.09.2024

Abstract. *Purpose.* Causal relationship recognition is a fundamental operation in neural networks aimed at learning behavior, action planning, and inferring external world dynamics. This operation is particularly crucial for reinforcement learning (RL). In the context of spiking neural networks (SNNs), events are represented as spikes emitted by network neurons or input nodes. Detecting causal relationships within these events is essential for effective RL implementation. *Methods.* This research paper presents a novel approach to realize causal relationship recognition using a simple spiking binary neuron. The proposed method leverages specially designed synaptic plasticity rules, which are both straightforward and efficient. Notably, our approach accounts for the temporal aspects of detected causal links and accommodates the representation of spiking signals as single spikes or tight spike sequences (bursts), as observed in biological brains. Furthermore, this study places a strong emphasis on the hardware-friendliness of the proposed models, ensuring their efficient implementation on modern and future neuroprocessors. *Results.* Being compared with precise machine learning techniques, such as decision tree algorithms and convolutional neural networks, our neuron demonstrates satisfactory accuracy despite its simplicity. *Conclusion.* We introduce a multi-neuron structure capable of operating in more complex environments with enhanced accuracy, making it a promising candidate for the advancement of RL applications in SNNs.

Keywords: spiking neural network, binary neuron, spike timing dependent plasticity, dopamine-modulated plasticity, anti-Hebbian plasticity, reinforcement learning, neuromorphic hardware.

For citation: Kiselev MV, Larionov DA, Urusov AM. A spiking binary neuron — detector of causal links. *Izvestiya VUZ. Applied Nonlinear Dynamics.* 2024;32(5):589–605. DOI: 10.18500/0869-6632-003121

This is an open access article distributed under the terms of Creative Commons Attribution License (CC-BY 4.0).

Введение

Если мы стремимся создать интеллектуальную систему на основе нейронных сетей, способную формировать адаптивное поведение для достижения определенных целей, необходимо наделять ее способностью идентифицировать и фиксировать в ее структуре причинно-следственные связи между событиями, происходящими как внутри сети, так и во внешней среде. Эти связи могут включать в себя последовательности однородных паттернов, образующих единый пространственно-временной паттерн, команды, генерируемые сетью, и реакции внешней среды на них или события, предшествующие вознаграждению, и само вознаграждение. Поэтому способность различать причины и следствия должна рассматриваться как базовая функциональность нейросетевых структур или отдельных нейронов. В большинстве сценариев обучаемая сеть не имеет доступа к априорным знаниям, описывающим причинно-следственные связи в ее окружении, — она должна выводить их из наблюдаемых временных закономерностей появления различных событий, исходя из предположения, что если событие В часто наблюдается в определенном временном интервале после события А, то А служит причиной, а В — следствием.

В данном исследовании рассматривается реализация описанной выше функциональности в рамках импульсной нейронной сети (ИмНС), а если более точно, одного конкретного нейрона как части сети. В контексте ИмНС информация кодируется последовательностями спайков,

*The paper presents materials of a talk given at the conference “Neuroinformatics — 2023”.

что делает необходимым формулирование задачи в этих терминах. Формализуем задачу относительно одного конкретного нейрона (назовем его детектором причинно-следственных связей). Этот нейрон принимает сигналы в виде спайков от пресинаптических нейронов. Будем считать, что срабатывание некоторого заранее неизвестного набора пресинаптических нейронов соответствует событию А (причина), тогда как активация некоторого другого нейрона соответствует событию В (следствие). Предположим, что событие В почти всегда происходит в течение времени T_p после события А, тогда задача нейрона-детектора состоит в том, чтобы срабатывать каждый раз при выявлении активности набора нейронов, соответствующих наступлению события А. Подобно другим задачам обучения нейросетей, будем решать эту с помощью изменения синаптических весов нейрона-детектора (синаптической пластичности). Важно отметить особенность, которая существенно ограничивает применимость большинства существующих на сегодняшний день моделей синаптической пластичности, основанных на относительной задержке между отдельными пре- и постсинаптическими спайками: нейрон-детектор и пресинаптические нейроны могут генерировать как единичные спайки, так и их плотно сгруппированные и продолжительные последовательности, так что невозможно сказать, какой из спайков был раньше или позже — пре- или пост-синаптический. Описанная ниже модель синаптической пластичности учитывает это обстоятельство.

Многочисленные исследования демонстрируют, как ИмНС могут выявлять причинно-следственные связи между различными событиями. Однако данное исследование предлагает уникальное сочетание трех отличительных атрибутов.

1. Причинно-следственные связи между событиями распознаются единственным нейроном-детектором, а не сетью.
2. Учитывается временной аспект причинно-следственных связей, при котором события-причины происходят за определенное время до соответствующих им событий-последствий.
3. Используются специально разработанные локальные правила синаптической пластичности для обучения.

В большинстве работ, имеющих на сегодняшний день, данная задача рассматривается вне временного аспекта — для такого типа задач часто используется термин Байесовский вывод (Bayesian inference). Эта задача близко связана с обучением с учителем, где сеть должна определить наиболее достоверные факторы, независимо или в сочетании указывающие на принадлежность определенного объекта к целевому классу. В статьях [1–3] представлены примеры исследований, в которых применяются различные подходы для решения таких задач. Следует отметить, что обычно в задачах обучения с учителем время явно не учитывается. Даже когда объектом обучения являются временные ряды, каждый временной ряд рассматривается как единая сущность, относимая к некоторому целевому классу. В отличие от этого, данное исследование более тесно связано с обучением с подкреплением, где все сигналы (входные сигналы, команды сети и вознаграждение/наказание) существуют в непрерывном времени и временные интервалы имеют большое значение.

Наш подход может быть важен с точки зрения концепции свободной энергии Фристана [4], поскольку понимание, что является «ожидаемым событием», важно для количественной оценки величины «неожиданности» событий в терминах свободной энергии. Настоящее исследование заостряет внимание на распознавании причинно-следственных связей в контексте времени, что имеет особое значение для динамических систем.

Следует отметить, что рассматриваемая исследовательская задача перекликается с еще одной обширной областью машинного обучения — прогнозированием временных рядов. Подходы прогнозирования временных рядов направлены на предсказание будущих значений определенных переменных (дискретных или непрерывных) на основе их текущих и недавних значений, а также, возможно, значений других связанных переменных. Естественно, если возможно определить

причинно-следственные связи между значениями определенных переменных или состояний объекта и значениями определенных параметров в будущем — это дает инструмент для прогнозирования будущих значений. Однако наша основная цель отличается от традиционного прогнозирования временных рядов, поскольку мы не сосредоточены на предсказании конкретного точного значения определенной переменной в определенный момент времени. Вместо этого наша цель — вывести причинно-следственные правила, которые указывают на то, что после события А ожидается, что событие В произойдет в интервале времени длиной T_p . Следовательно, задача может быть более точно характеризована как прогнозирование будущих событий, а не временных рядов.

Примечательно, что на сегодняшний день относительно немного применений ИмНС для решения подобных задач. Один из подходов, описанный в [5], использует систему NeuCube [6], которая основана на так называемой машине с жидким состоянием (Liquid State Machine, LSM) [7]. LSM — это большая хаотичная непластичная ИмНС, разработанная для преобразования комбинации временных рядов и статических (или медленно изменяющихся) параметров в многомерное представление в виде частот генерации спайков нейронами внутри LSM. Благодаря большому количеству нейронов в LSM, представления различных пространственно-временных паттернов в форме нейронной активности LSM с высокой вероятностью являются линейно сепарабельными. Следовательно, задачи классификации, связанные с такими представлениями, могут быть эффективно решены с помощью простых линейных классификаторов. Как описано в [5], были продемонстрированы несколько примеров применения NeuCube для прогнозирования редких событий. Один конкретный пример — прогнозирование инсультов, рассматривается более подробно в [8]. Хотя подход на основе LSM продемонстрировал успех в широком диапазоне задач, у него есть заметный недостаток в том, что для достижения эффективности LSM она должна быть большой и, следовательно, требовать значительных вычислительных мощностей. В отличие от этого, подход, предложенный в данном исследовании, эффективно решает аналогичную задачу, используя всего один нейрон, что является более ресурсоэффективным решением.

В статье [9] показано, как специальные структуры ИмНС могут использоваться для получения графа причинно-следственных связей, но опять же, без учета временного аспекта, как упоминалось ранее.

Наконец, существует еще одно направление исследований ИмНС, тесно связанное с нашим исследованием. Как описано ниже, для решения задачи выявления причинно-следственных связей используется комбинация двух моделей синаптической пластичности, обычно называемых, хотя и в очень приблизительном смысле, Хеббовской и дофаминовой пластичностью. Хеббовский принцип пластичности, когда он применяется к пластичности импульсных нейронов, часто обозначается как модель STDP (Spike Timing Dependent Plasticity) [10], а дофаминовая пластичность обычно связана с эффектами, связанными с вознаграждением. Коллективное терминологическое обозначение для этих объединенных моделей пластичности — R-STDP (Reward Spike Timing Dependent Plasticity). Различные модели R-STDP изучались в многочисленных работах (например, [11–14]), некоторые из которых уже были протестированы в реальных приложениях [15]. В настоящее время нет единого мнения о том, как лучше объединять эти два типа синаптической пластичности, так что спектр рассматриваемых моделей остается весьма широк. Кроме того, в отличие от нашего подхода, где вознаграждение имеет форму спайкового сигнала, эти модели обычно представляют вознаграждение как глобальную вещественную переменную. Насколько нам известно, ни одна из предшествующих работ не использовала подобные правила синаптической пластичности для цели обнаружения причинно-следственных связей.

Кроме того, цель данного исследования заключалась в создании достаточно простой модели пластичности для ее эффективной реализации на современных и будущих нейропроцессорах. Согласно работе [16], существует явный тренд в появлении и развитии программных и аппаратных систем, которые основаны не на преобразовании традиционных сверточных нейронных

сетей в форму ИмНС, а на использовании «нативных» моделей импульсных нейронов и правил пластичности, открывая возможность для непрерывного обучения ИмНС.

В следующем разделе будет подробно описана наша инновационная модель синаптической пластичности, которая объединяет Хеббовскую (фактически анти-Хеббовскую) и дофаминовую пластичность. После этого будет рассмотрено применение модели к задаче предсказания вознаграждений в обучении с подкреплением (reinforcement learning, RL), используя в качестве примера задачу предсказания вознаграждений в задаче «Пинг-понг». В заключение мы опишем наше видение того, как нейроны такого рода могут формировать сетевые структуры, способные выводить сложные графы причинно-следственных связей, необходимые для построения моделей внешнего мира в RL. Далее, будут оценены преимущества и ограничения данного подхода и определен план будущих исследований.

1. Методы и материалы

В данном исследовании рассматривается процесс обучения одного импульсного нейрона в контексте выявления причинно-следственных связей между событиями. Этот нейрон (рис. 1) связан с группой пресинаптических нейронов, образующих множество C , чья активность представляет различные события. Будем интерпретировать эти события как потенциальные триггеры для другого события, которое назовем «целевым событием». Это целевое событие соответствует спайку от отдельного пресинаптического нейрона, обозначаемого как S , который не является частью множества C . Момент времени j -й генерации спайка i -м пресинаптическим нейроном из множества C обозначим как t_{ij} . Моменты времени, когда нейрон S генерирует спайк, будем обозначать как t_j^S . Мы говорим, что некоторое событие является причиной целевого события, если целевое событие часто наблюдается не позднее времени T_p после этого события. Как уже упоминалось, возможные причины целевого события определяются специфической активностью пресинаптических нейронов, которую должен попытаться распознать обучаемый нейрон-детектор.

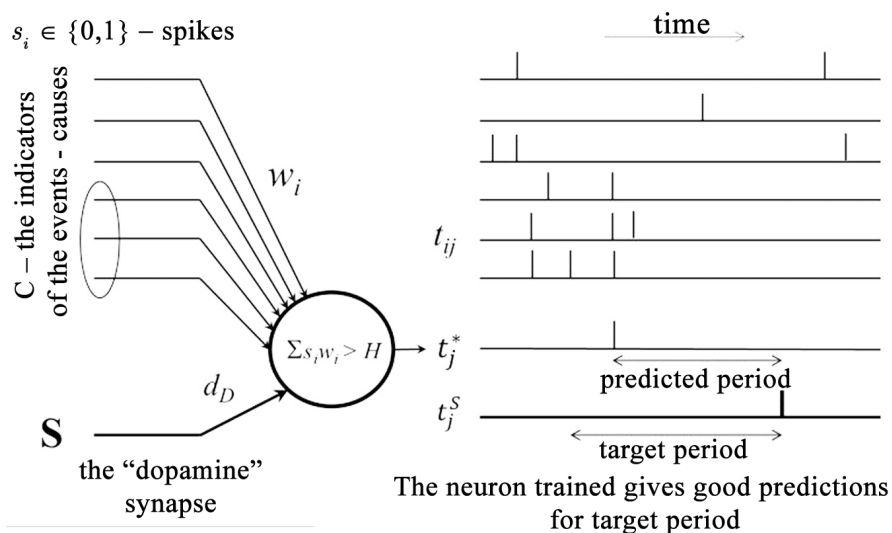


Рис. 1. Схематичная модель бинарного импульсного нейрона и временная диаграмма его входных и выходных спайков (для обученного состояния)

Fig. 1. The schematic model of binary spiking neuron and the temporal diagram of its pre- and postsynaptic spikes (for the trained state)

T_p — это временная константа, фиксирующая временной масштаб конкретной задачи. Предполагается, что целевые события редкие — это означает, что T_p намного меньше минимального значения интервалов между спайками $t_j^S - t_{j-1}^S$. Это единственное важное предположение — без него наша задача поиска причинно-следственных связей, по-видимому, теряет смысл.

Введем также понятие «целевой период», охватывающее временной интервал длиной T_p , предшествующий каждому t_j^S . Обучаемый нейрон должен пометить целевые периоды своей активностью (спайками, испускаемыми им в моменты времени t_j^*). Если он научится делать это с достаточной точностью, это означает, что он успешно распознает причинно-следственную связь между конкретным событием (активностью пресинаптических нейронов, которая вызывает срабатывание нейрона) и целевым событием. Для оценки точности этого распознавания введем понятие «предсказанный период». Каждый предсказанный период начинается в момент времени t_j^* и заканчивается либо через время T_p от этого момента, либо в один из моментов t_j^S — в зависимости от того, что произойдет раньше. Общая продолжительность времени T_{err} , в течение которого целевые периоды и предсказанные периоды не пересекаются, служит естественной метрикой для измерения неточности предсказания целевого события. Целью обучаемого нейрона является максимизация метрики, представленной формулой

$$R = 1 - \frac{T_{err}}{T_{tar}}, \quad (1)$$

где T_{tar} обозначает общую продолжительность целевых периодов.

В данном исследовании используется самая простая модель нейрона, называемая «бинарным нейроном». Этот нейрон работает в дискретном времени. В каждый квант времени он получает спайки через свои пластичные синапсы с весами w_i . Предполагается, что величина этих квантов — порядка возможного времени рассогласования входных спайков из множества S , индицирующих событие-причину, так что значительная часть этих спайков приходит в рамках одного кванта. Если сумма весов синапсов, получивших спайки в данный квант, больше порогового значения H , то нейрон генерирует спайк. Выбор столь простой модели делает наш результат общим — фактически он не зависит от конкретной модели нейрона. После соответствующей дискретизации времени любая модель импульсного нейрона может быть приближена бинарным нейроном, который сохраняет основное свойство — нейрон срабатывает, когда несколько сильных возбуждающих синапсов получают спайки в течение короткого временного периода. Чтобы сделать веса w_i безразмерными, мы устанавливаем $H = 1$.

1.1. Общая идея метода и правила синаптической пластичности, используемые в исследовании. Предполагается, что информация о потенциальных событиях-причинах, приводящих к целевому событию, закодирована в спайках, исходящих от пресинаптических нейронов из множества S . Синапсы, отвечающие за передачу этих импульсов, являются пластичными, и их синаптические веса должны корректироваться таким образом, чтобы заставить постсинаптический нейрон-детектор срабатывать во время целевого периода.

Активность обучаемого постсинаптического нейрона-детектора и корректировка его синаптических весов должны быть связаны следующим образом.

- a. Необученный нейрон должен быть неактивен — механизм пластичности должен усиливать те синапсы, которые заставляли бы нейрон генерировать спайки в правильное время. По этой причине мы устанавливаем веса всех пластичных синапсов равными 0 в начале обучения.
- b. Если нейрон генерирует спайк в неправильное время (вне целевых периодов), то синапсы, которые помогли ему сгенерировать ложный спайк, должны быть подавлены.
- c. Если нейрон генерирует спайк в правильное время, то с его синаптическими весами ничего не должно происходить, в противном случае их изменение может вывести его из обученного состояния.

Это достигается благодаря специфическому выбору правил синаптической пластичности. Важно, что свойства пластичных синапсов полностью отличны от единственного синапса, через который подключен пресинаптический нейрон S . Мы называем его «дофаминовым» синапсом, потому что спайки, поступающие на него, контролируют пластичность всех остальных синапсов.

Принципы А, В и С, указанные выше, выполняются благодаря комбинации двух правил пластичности.

1. Дофаминовая пластичность. Каждый раз, когда обучаемый нейрон получает спайк от нейрона S , все пластичные синапсы, получившие спайки в течение времени T_p перед этим «дофаминовым» спайком, усиливаются.
2. Анти-Хеббовская пластичность. Все синапсы, способствующие срабатыванию нейрона, подавляются.

Очевидно, что в случае баланса между дофаминовой и анти-Хеббовской пластичностью условия А, В и С удовлетворяются, обеспечивая успешное обучение и функционирование нейрона.

1.2. Модель синаптической пластичности в деталях. Аналогично нашим предыдущим исследованиям [17, 18], используемые в данной работе правила синаптической пластичности являются аддитивными и применяются к переменной, называемой «синаптический ресурс», обозначаемой как W , а не непосредственно к синаптическому весу, обозначаемому как w . Функциональная зависимость между W и w выражается формулой

$$w = w_{\min} + \frac{(w_{\max} - w_{\min}) \cdot \max(W, 0)}{w_{\max} - w_{\min} + \max(W, 0)}, \quad (2)$$

где w_{\min} и w_{\max} — константы. Очевидно, что значения w пробегают диапазон $[w_{\min}, w_{\max})$, когда W изменяется от $-\infty$ до $+\infty$. В данном исследовании $w_{\min} < 0$, а $w_{\max} > 0$, так что синаптическая пластичность может сделать возбуждающий синапс тормозным и наоборот.

Как уже упоминалось ранее, модель синаптической пластичности состоит из двух отдельных и независимых компонент. Они описаны в подразделах 1.2.1 и 1.2.2.

1.2.1. Анти-Хеббовская пластичность. Стандартная модель STDP [10] утверждает, что спайки, поступающие за короткое время до срабатывания постсинаптического нейрона, усиливают синапсы, которые их получают. Эта концепция соответствует принципу Дональда Хебба, который утверждает, что синаптическая пластичность должна отражать причинно-следственные связи между срабатываниями нейронов — синапсы, ответственные за генерацию спайков постсинаптическим нейроном, должны быть усилены. Этот принцип был подтвержден множеством нейрофизиологических наблюдений. Однако глубокие исследования пластичности в биологических нейронах показали, что в природе существуют и несколько альтернативных моделей синаптической пластичности [19, 20]. Кроме того, в различных организмах были обнаружены примеры правил пластичности, действующих в направлении, противоположном Хеббовскому принципу (анти-Хеббовская пластичность) [21]. Это позволяет заключить, что различные виды синаптической пластичности подходят для решения разных задач. Кроме того, стандартная модель STDP теряет смысл в случае (который довольно распространен в биологическом мозге), когда мы имеем дело не с одиночными пре- и постсинаптическими спайками, а с их плотно сгруппированными во времени последовательностями. В этом случае бессмысленно говорить о конкретной последовательности прихода пресинаптического и постсинаптического спайков, потому что есть множество постсинаптических спайков в непосредственной близости до и после момента прихода конкретного пресинаптического спайка.

По этой причине нами был разработан новый вариант модели анти-Хеббовской пластичности, рассмотренный ниже.

Как уже упоминалось, изменения весов в стандартной модели STDP связаны с одиночными пре- и постсинаптическими спайками. Однако в случае плотных последовательностей спайков эти правила теряют свою применимость. В нашей модели синаптическая пластичность связана с последовательностями постсинаптических спайков, а не с отдельными спайками. Назовем эти последовательности «плотными спайковыми последовательностями» (англ. tight spike sequences, TSS). Принимая постоянную ISI_{\max} (ISI — inter-spike interval) как меру «плотности» TSS, определим TSS как последовательность спайков, соответствующую следующим критериям.

1. Отсутствуют спайки в течение времени ISI_{\max} до первого спайка в TSS.
2. Интервалы между всеми соседними спайками в TSS не превышают ISI_{\max} .
3. Отсутствуют спайки в течение времени ISI_{\max} после последнего спайка в TSS.

В данной работе ISI_{\max} устанавливается равным T_p .

Предлагаемая модель анти-Хеббовской пластичности задается следующими правилами.

1. Ресурс любого синапса может изменяться не более одного раза в течение одной TSS. Здесь и далее понятие TSS относится к постсинаптическим спайкам.
2. Изменяются ресурсы только тех синапсов, которые получают хотя бы один спайк в течение TSS.

Все синаптические ресурсы изменяются (уменьшаются) на одно и то же значение d_H , независимо от точного времени пресинаптических импульсов.

1.2.2. Дофаминовая пластичность. Описываемый нейрон-детектор имеет синапс (соединенный с нейроном S), модулирующий пластичность остальных синапсов. Когда он получает спайк, синаптические ресурсы всех пластичных синапсов, получивших хотя бы один пресинаптический спайк в течение временного интервала T_p до этого спайка, изменяются (увеличиваются) на одно и то же значение d_D .

Подчеркнем, что дофаминовая пластичность и анти-Хеббовская пластичность — это два совершенно логически независимых механизма. Хотя их сочетание как раз и дает желаемый эффект обучения нейрона.

1.2.3. Стабильность нейрона. В нашей модели пластичность синапсов d_H и d_D не является постоянной. В начале обучения значения d_H и d_D должны быть достаточно большими. Однако для уже обученного нейрона, который постоянно делает точные предсказания, они должны стремиться к нулю. Эта адаптация важна для предотвращения дальнейших изменений синаптических весов нейрона, которые могут нарушить его обученное состояние. Чтобы учесть эту адаптивную функцию, в состояние нейрона вводится дополнительный компонент s , называемый «стабильностью». Значения пластичности синапсов экспоненциально уменьшаются до нуля при росте значения стабильности в соответствии с уравнениями

$$d_H = \overline{d}_H \cdot \min(2^{-s}, 1), \quad d_D = \overline{d}_D \cdot \min(2^{-s}, 1). \quad (3)$$

Здесь \overline{d}_H и \overline{d}_D являются константами модели нейрона. Чтобы сбалансировать анти-Хеббовскую и дофаминовую пластичность (что необходимо для выполнения условия С из подраздела 2.1, мы устанавливаем $\overline{d}_H = \overline{d}_D$. Значение стабильности нейрона изменяется в двух случаях.

1. Уменьшается на константу d_S при каждом TSS.
2. Меняется на величину $d_s \cdot \max\left(2 - \frac{|t_{TSS} - ISI_{\max}|}{ISI_{\max}}, -1\right)$ при приходе дофаминового спайка.

Здесь t_{TSS} — временной интервал между началом последнего TSS и дофаминовым спайком.

Очевидно, что если TSS начался ровно ISI_{\max} ($= T_p$) времени назад до дофаминового спайка (то есть целевого события), то увеличение стабильности нейрона будет максимальным и равным d_S — если учесть его уменьшение на d_S в соответствии с правилом 1. Это соответствует наиболее точному предсказанию целевого события и служит показателем того, что нейрон обучен.

Напротив, если дофаминовый спайк происходит в момент, когда нейрон находился в неактивном состоянии в течение продолжительного времени — это признак недостаточной обученности нейрона, в результате чего его стабильность уменьшается на d_S , чтобы облегчить дальнейшее обучение.

1.3. Тестовая задача — найти причину получения вознаграждения в игре «Пинг-понг» из набора тестов ATARI. Описанный подход обладает большим потенциалом в области обучения с подкреплением (RL). В то время как обучение с учителем может рассматриваться как определение причинно-следственных связей между предикторами в качестве причин и целевым значением в качестве следствия, задачи RL охватывают более широкий спектр определения причинно-следственных связей, явно включая элемент времени. Сигналы вознаграждения могут поступать редко и, возможно, со значительной задержкой по отношению к состояниям мира или действиям агента, которые они оценивают. Чтобы преодолеть проблемы недостаточной частоты сигналов оценки, необходимо использовать механизм промежуточных целей, который также основан на выводе причинно-следственных связей.

Более того, ключевой момент для реализации наиболее развитого варианта RL, известного как RL, основанное на моделях (model-based RL), заключается в создании агентом внутренней модели динамики внешнего мира и реакций мира на действия агента. Механизм создания модели неизбежно включает нахождение сети причинно-следственных связей между изменениями состояния мира и действиями агента. Таким образом, вполне разумно утверждать, что вывод причинно-следственных связей является одной из основных операций в RL.

Исходя из этого, для проверки возможностей модели нейрона была выбрана одна из задач RL из часто используемого тестового набора игр ATARI [22]. Эта задача связана с компьютерной игрой «Пинг-понг», где мяч перемещается внутри квадратной области, отскакивая от ее стен. Область имеет только три стены. Вместо левой стены — ракетка, которая движется в вертикальном направлении по левой границе этой квадратной области. Ракеткой управляет агент, который может двигать ее вверх и вниз. Когда мяч попадает в ракетку и отскакивает назад, агент получает сигнал вознаграждения. Если мяч пересекает левую границу без попадания в ракетку, то агент получает сигнал наказания, и мяч возвращается в случайную точку средней вертикальной линии квадратной области, получая случайное направление движения и скорость, и игра продолжается. Используя полученные сигналы вознаграждения/наказания, агент должен понять, что его цель — отразить мяч и обучиться этому.

В нашем примере сеть (фактически один нейрон) должна решить первую задачу — понять, какие условия приводят к получению вознаграждения в ближайшем будущем.

Информация, поступающая на пластичные синапсы нейрона, в этой задаче включает текущие положения мяча и ракетки, скорость мяча. В то время как в окончательной формулировке этой задачи нейронная сеть должна работать с первичной растровой информацией (изображением на экране), предмет данной статьи лежит не в области компьютерного зрения, а состоит в нахождении причинно-следственных связей. Поэтому мы предполагаем, что входные слои нейросети уже обработали первичные растровые данные и преобразовали их в спайковое представление, подающееся на пластичные синапсы нейрона — детектора причинно-следственных связей.

Входные узлы (источники спайков) подразделяются на следующие секции.

1. Координата X мяча. Состоит из 30 узлов, отображающих горизонтальное положение мяча. Горизонтальное измерение разбито на 30 равных непересекающихся интервалов. Когда мяч находится в интервале i , i -й узел генерирует спайки с частотой 300 Гц. Чтобы установить пространственные и временные масштабы, мы предполагаем, что размер квадратной области составляет 10×10 см (так что координаты границ равны ± 5 см), а дискретный шаг эмуляции времени составляет 1 мс.

2. Координата Y мяча. Состоит из 30 узлов, фиксирующих вертикальное положение мяча. Аналогично X , но для вертикальной оси.
3. Компонента X скорости мяча. Состоит из 9 узлов, соответствующих горизонтальной скорости мяча. Когда мяч вбрасывается в середине квадратной области, его скорость устанавливается равной случайному значению из диапазона $[10, 33.3]$ см/с. Его исходное направление движения также случайно, но выбирается таким образом, чтобы абсолютное значение его компоненты X не была меньше 10 см/с. Весь диапазон возможных значений компоненты скорости X мяча разбит на 9 интервалов таким образом, чтобы вероятности обнаружить мяч в случайный момент времени в каждом из этих интервалов были примерно равны. Пока компонента X скорости мяча находится в каком-то интервале, соответствующий входной узел генерирует спайки с частотой 300 Гц.
4. Компонента Y скорости мяча. Состоит из 9 узлов, фиксирующих вертикальную скорость мяча. Та же логика, что и для компоненты X скорости.
5. Координата Y ракетки. Состоит из 30 узлов, фиксирующих вертикальное положение ракетки. Аналогично Y координате мяча. Размер ракетки составляет 1.8 см, поэтому ракетка занимает чуть больше 5 вертикальных интервалов.
6. Относительное положение мяча и ракетки в ближней зоне. Состоит из 25 узлов, соответствующих положениям мяча, близким к ракетке. Квадратное поле зрения размером 3×3 см перемещается вместе с ракеткой, так что центр ракетки всегда находится в центре левой границы этого поля. Поле зрения разбито на 5×5 квадратных зон. Когда мяч находится в какой-то зоне, соответствующий входной узел генерирует спайки с частотой 300 Гц.

Таким образом, всего имеется 133 входных узла, передающих свои спайки обучаемому нейрону. Цель нейрона — распознать условия, приводящие к получению сигнала вознаграждения в течение 100 мс, и, соответственно, мы устанавливаем $T_p = 100$ мс.

1.4. Выбор параметров нейрона с использованием генетического алгоритма. Хотя модель нейрона кажется относительно простой, она включает несколько параметров, требующих настройки. Их четыре.

- Максимальное изменение синаптического ресурса $\overline{d_H}$. Этот параметр контролирует скорость обучения. Низкие значения делают обучение медленным, высокие могут сделать его нестабильным.
- Минимальное значение веса синапса w_{\min} . Оно отрицательное.
- Максимальное значение веса синапса w_{\max} .
- Скорость изменения стабильности d_S .

Оптимальные значения $\overline{d_H}$ и d_S определяются силой причинно-следственных связей — в случае слабого детерминизма или высокого уровня шума большие значения этих параметров приведут к нестабильности обучения. w_{\min} и w_{\max} должны быть выбраны на основе среднего потока входных спайков — количества входных узлов и средней частоты спайков на узел.

Хотя общие принципы для установки этих параметров достаточно ясны, было принято решение найти их оптимальные значения, используя генетический алгоритм и очень широкие диапазоны поиска: $[0.03, 1]$ для $\overline{d_H}$; $[0.003, 1]$ для w_{\min} ; $[0.03, 1]$ для w_{\max} ; $[0.003, 3]$ для d_S . Для установки их случайных значений использовалось лог-равномерное распределение. Размер популяции был выбран равным 300; уровень элитарности составлял 0.1; вероятность мутации на каждую хромосому была равна 0.5. Критерием оптимизации был R (1). Он измерялся за последние 600 секунд 2000-секундной записи игры в пинг-понг, где ракетка двигалась хаотично. Общее количество вознаграждений составило 951. Генетический алгоритм завершался, когда три последовательных поколения не показывали увеличения R .

2. Результаты и обсуждение

Лучший результат, достигнутый в ходе наших экспериментов, был получен в 17-м поколении генетического алгоритма. Значение R составило 0.553. Оптимальные значения параметров: $\overline{d_H} = 0.056$; $w_{\min} = -0.017$ (нулевое значение слегка снижает R); $w_{\max} = 0.48$ (то есть для срабатывания требуется не менее трех входных спайков); $d_S = 0.23$.

Учитывая значительную неопределенность взаимосвязи между текущим состоянием мира и получением вознаграждения из-за дискретного описания мира и хаотичного движения ракетки, полученное значение R представляется удовлетворительным. Чтобы объективно оценить этот результат, было произведено сравнение нашей техники с традиционными методами машинного обучения. Чтобы обеспечить корректное сравнение, было выбрано два алгоритма машинного обучения абсолютно разной природы: дерево решений и сверточная нейронная сеть. Все алгоритмы обучались на тех же двоичных сигнальных данных от входных узлов; каждый шаг (квант времени) эмуляции служил отдельным обучающим примером. В качестве целевой переменной использовалось булевское значение, индицирующее нахождение в целевом периоде. Алгоритмы машинного обучения применялись к тем же данным, что и наш нейрон (первые 1 400 000 шагов), и создавали модели оценки вероятности того, что текущий шаг принадлежит к целевому периоду. Кванты времени, когда возвращаемое моделью значение превышало некоторый порог (чьё значение определялось с помощью оптимизационной процедуры), рассматривались аналогично срабатыванию нейрона. По тем же правилам, описанным в начале раздела 2, на последних 600 000 шагах определялись предсказанные периоды и рассчитывалось значение R . Оптимальное значение порога находилось из условия максимизации значения R на первых 1 400 000 шагах. Алгоритм дерева решений использовал критерий расщепления на взаимной информации. Сверточная сеть включала 2 сверточных слоя ReLU.

Максимальное значение R , полученное деревом решений, равно 0.742, сверточная сеть дала результат $R = 0.731$. Наблюдаемая близость результатов, показанных очень разными методами, доказывает корректность нашего подхода к определению теоретического предела для значения R в текущей задаче. Таким образом, оценка этого предела, равная 0.75, кажется реалистичной.

Хотя результат, достигнутый описанным единичным нейроном (74% от теоретического максимума), можно рассматривать как скромный, мы считаем нашу модель успешной. Рассматриваемая модель нейрона очень проста — ей соответствует предиктивная модель, содержащая только 133 степени свободы. В отличие от этого, например, модель дерева решений включает 51 уровень и 403 нетерминальных узла. По сути, функция нашего нейрона аналогична конъюнкции логических значений, соответствующих нескольким его сильнейшим синапсам (см. ниже). Совпадение этих факторов рассматривается как причина целевого события. Однако очевидно, что получению вознаграждения в нашем примере могут предшествовать несколько существенно разных условий. Поэтому хорошая предсказательная модель — это скорее дизъюнкция нескольких конъюнкций. Таким образом, весьма вероятно, что сеть, состоящая из описываемых нейронов, могла бы давать гораздо более точный прогноз вознаграждения, и мы рассмотрим возможную структуру такой сети ниже.

Рассмотрим процесс обучения и его результаты несколько более подробно. Динамика активности нейрона, его стабильности и общего изменения весов (суммы абсолютных значений изменений весов каждого синапса) представлены на рис. 2. Поскольку исходные веса всех пластичных синапсов были равны нулю, сначала работала только дофаминовая пластичность. В течение первых 100 секунд нейрон не генерировал спайков. После 250 секунд частота генерации спайков стабилизировалась. Из-за механизма стабилизации весов (как видно из рис. 2, стабильность

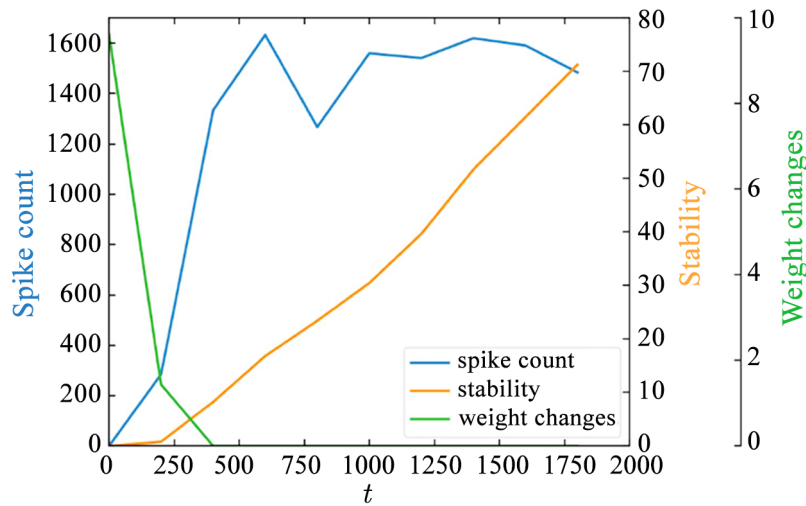


Рис. 2. Динамика частоты спайков, стабильности и изменения веса обучающегося нейрона (цвет онлайн)

Fig. 2. The time course of firing frequency, stability and weight changes of the learning neuron (color online)

нейрона почти линейно растет) синаптические веса не менялись после 400 секунд. В целом процесс обучения занял 400 секунд вместо запланированных 1400 секунд.

Результаты обучения представлены на рис. 3, который отражает значения синаптических ресурсов обучающегося нейрона на 2000-й секунде. Самый левый график соответствует 30 входным узлам, кодирующим координату X мяча. Вертикальная ось всех графиков, кроме самого правого, отражает значение синаптического ресурса. Второй график соответствует 30 входным узлам, кодирующим координату Y мяча (синяя линия) и ракетки (оранжевая линия). Следующие два графика представляют 9+9 входных узлов, кодирующих горизонтальную и вертикальную составляющую скорости мяча. На правом графике показана цветовая кодировка значений синаптических ресурсов 25 входных узлов, соответствующих местоположению мяча внутри сетки 5×5, которая движется с ракеткой. Распределение значений синаптических ресурсов на этих графиках выглядит разумным и соответствует ожиданиям.

Итак, можно сделать заключение, что представленный нейрон способен обнаруживать причинно-следственные связи между наблюдаемыми событиями. Однако, как уже обсуждалось, даже при достаточной длительности обучения, одиночный нейрон не может точно предсказать наступление целевого события в реальных задачах. Это связано с тем, что целевое событие может быть вызвано несколькими совершенно разными по своей природе событиями. Для более точных прогнозов требуется сеть нейронов, способных распознавать причинно-следственные связи и учитывать несколько значимых причин.

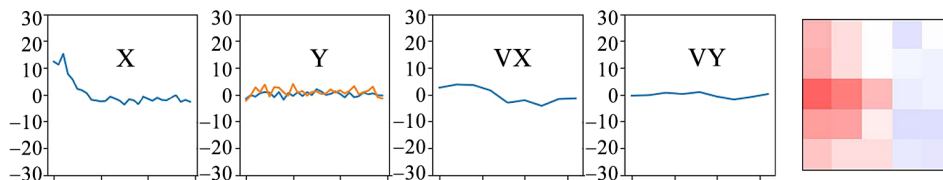


Рис. 3. Синаптические ресурсы обученного нейрона (цвет онлайн)

Fig. 3. The synaptic resource of the neuron trained (color online)

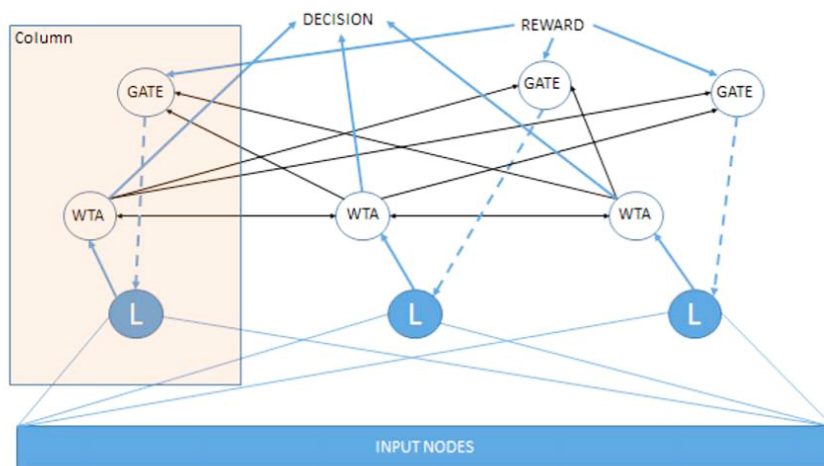


Рис. 4. Возможная нейросетевая структура, состоящая из нескольких нейронов, распознающих причинно-следственные связи, для предсказания целевого события, возникающего в результате разных причин. Синие стрелки — возбуждающие связи, черные — блокирующие, пунктирные — дофаминовые (цвет онлайн)

Fig. 4. The proposed structure of an SNN, which includes several neurons recognizing causal relationships for predictions of target events caused by several other events. The blue arrows depict excitatory connections, the black arrows — blocking connections, the dashed lines — dopamine connections (color online)

Можно предположить возможную архитектуру такой сети (рис. 4). В этой сети нейроны, распознающие различные причины целевого события (синие круги L), входят в колончатую структуру, где каждая колонка соответствует отдельной причине. Эти колонки конкурируют за распознавание событий-причин за счет латеральных блокирующих связей между их нейронами по принципу «победитель забирает все» (“winner-takes-all”, WTA). «Победивший» (сработавший первым) нейрон блокирует не только другие нейроны WTA, но и нейроны GATE в других колонках. Если один нейрон L генерирует спайк, то другие нейроны L не должны этого делать, потому что нейроны L должны распознавать различные причинные связи. Если какой-то нейрон L генерирует спайк после победителя, то он не получит сигнала вознаграждения, потому что этот сигнал не пройдет через GATE, заблокированный победителем. Синапсы, вызвавшие его спайк, будут подавлены анти-Хеббовской пластичностью, чтобы эта ситуация не повторилась в следующий раз. Если победивший нейрон правильно генерировал спайк, то он будет вознагражден, так как его GATE не заблокирован. На наш взгляд, такой архитектурный подход имеет потенциал для распознавания сложных сетей причинно-следственных связей, что и планируется проверить в дальнейших исследованиях.

Заключение

Неоспоримо, что способность распознавать причинно-следственные связи в динамических потоках данных является фундаментальной функциональностью для любой самообучающейся системы, работающей в реальном мире. В данном исследовании продемонстрирована возможность реализации этой критически важной функции на уровне отдельного нейрона благодаря предложенному сочетанию анти-Хеббовской и дофаминовой пластичности. Учитывая ключевую роль таких механизмов в контексте реализации обучения с подкреплением в ИмНС, описанная модель была проверена на простой, но нетривиальной задаче обучения с подкреплением — компьютерной игре «Пинг-понг» из набора тестов ATARI. Полученные результаты,

включая оценку теоретического верхнего предела точности прогнозирования в этой задаче, подтверждают эффективность предложенной модели нейрона для определения причинно-следственных связей.

Наряду с этим, очевидно, что описанный в работе нейрон сам по себе недостаточно сложен, чтобы описывать причинно-следственные связи между многими различными факторами. В этом смысле имеется прямая аналогия с традиционными нейронными сетями. Известно, что уже единичный нейрон обладает ограниченной возможностью аппроксимировать узкий класс многомерных нелинейных функций, однако для аппроксимации широкого класса функций требуется многослойная сеть таких нейронов. Поэтому в работе представлена возможная архитектура ИмНС, в которой используются нейроны описанного типа, которая должна быть способна выводить комплексы причинно-следственных связей непосредственно из сырых данных. В последующих исследованиях планируется тщательно проверить и настроить эту ИмНС структуру. Планируется также расширить возможности этой ИмНС для учета временных аспектов причинно-следственных связей, переходя от вопроса «Каковы возможные последствия данных событий?» к более сложному — «Когда эти последствия могут проявиться?». Полагаем, что данное исследование является существенным шагом в развитии импульсных нейронных сетей, которые могут моделировать и понимать сложную причинно-следственную динамику реального мира.

Список литературы

1. *Moreno-Bote R., Drugowitsch J.* Causal inference and explaining away in a spiking network // *Scientific Reports*. 2015. Vol. 5. P. 17531. DOI: 10.1038/srep17531.
2. *Lansdell B. J., Kording K. P.* Neural spiking for causal inference and learning // *PLoS Computational Biology*. 2023. Vol. 19, no. 4. P. e1011005. DOI: 10.1371/journal.pcbi.1011005.
3. *Skatchkovsky N., Jang O., Simeone O.* Bayesian continual learning via spiking neural networks // *Frontiers in Computational Neuroscience*. 2022. Vol. 16. P. 1037976. DOI: 10.3389/fncom.2022.1037976.
4. *Friston K.* The history of the future of the bayesian brain // *Neuroimage*. 2012. Vol. 62, no. 2. P. 1230–1233. DOI: 10.1016/j.neuroimage.2011.10.004.
5. *Kasabov N., Scott N. M., Tu E., Marks S., Sengupta N., Capecchi E., Othman M., Doborjeh M. G., Murlu N., Hartono R., Espinosa-Ramos J. I., Zhou L., Alvi F. B., Wang G., Taylor D., Feigin V., Gulyaev S., Mahmoud M., Hou Z.-G., Yang J.* Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: Design methodology and selected applications // *Neural Networks*. 2016. Vol. 78. P. 1–14. DOI: 10.1016/j.neunet.2015.09.011.
6. *Kasabov N. K.* NeuCube: a spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data // *Neural Networks*. 2014. Vol. 52. P. 62–76. DOI: 10.1016/j.neunet.2014.01.006.
7. *Schliebs S., Fiasché M., Kasabov N.* Constructing robust liquid state machines to process highly variable data streams // *International Conference on Artificial Neural Networks ICANN 2012: Artificial Neural Networks and Machine Learning — ICANN*. 2012. P. 604–611. DOI: 10.1007/978-3-642-33269-2_76.
8. *Doborjeh M., Doborjeh Z., Merkin A., Krishnamurthi R., Enayatollahi R., Feigin V., Kasabov N.* Personalized spiking neural network models of clinical and environmental factors to predict stroke // *Cognitive Computation*. 2022. Vol. 14. P. 2187–2202. DOI: 10.1007/s12559-021-09975-x.
9. *Fernando C.* From blickets to synapses: Inferring temporal causal networks by observation // *Cognitive Science*. 2013. Vol. 37, no. 8. P. 1426–1470. DOI: 10.1111/cogs.12073.
10. *Markram H., Gerstner W., Sjöström P. J.* A history of spike-timing-dependent plasticity // *Frontiers in Synaptic Neuroscience*. 2011. Vol. 3. P. 4. DOI: 10.3389/fnsyn.2011.00004.

11. *Kerr R. R., Grayden D. B., Thomas D. A., Gilson M., Burkitt A. N.* Coexistence of reward and unsupervised learning during the operant conditioning of neural firing rates // *PLoS ONE*. 2014. Vol. 9, no. 1. P. e87123. DOI: 10.1371/journal.pone.0087123.
12. *Yuan M., Wu X., Yan R., Tang H.* Reinforcement learning in spiking neural networks with stochastic and deterministic synapses // *Neural Computation*. 2019. Vol. 31, no. 12. P. 2368–2389. DOI: 10.1162/neco_a_01238.
13. *Mozafari M., Ganjtabesh M., Nowzari-Dalini A., Thorpe S. J., Masquelier T.* Bio-inspired digit recognition using reward-modulated Spike-Timing-Dependent Plasticity in deep convolutional networks // *Pattern Recognition*. 2019. Vol. 94. P. 87–95. DOI: 10.1016/j.patcog.2019.05.015.
14. *Frémaux N., Sprekeler H., Gerstner W.* Functional requirements for reward-modulated Spike-Timing-Dependent Plasticity // *The Journal of Neuroscience*. 2010. Vol. 30, no. 40. P. 13326–13337. DOI: 10.1523/JNEUROSCI.6249-09.2010.
15. *Juarez-Lora A., Ponce-Ponce V. H., Sossa H., Rubio-Espino E.* R-STDP spiking neural network architecture for motion control on a changing friction joint robotic arm // *Frontiers in Neurobotics*. 2022. Vol. 16. P. 904017. DOI: 10.3389/fnbot.2022.904017.
16. *Ivanov D., Chezhegov A., Kiselev M., Grunin A., Larionov D.* Neuromorphic artificial intelligence systems // *Frontiers in Neuroscience*. 2022. Vol. 16. P. 959626. DOI: 10.3389/fnins.2022.959626.
17. *Kiselev M., Ivanov A., Ivanov D.* Approximating conductance-based synapses by current-based synapses // *Advances in Neural Computation, Machine Learning, and Cognitive Research IV. Neuroinformatics 2020. Studies in Computational Intelligence*. 2020. Vol. 925. P. 394–402. DOI: 10.1007/978-3-030-60577-3_47.
18. *Kiselev M. V.* A synaptic plasticity rule providing a unified approach to supervised and unsupervised learning // *Proceedings of International Joint Conference on Neural Networks*. 2017. P. 3806–3813. DOI: 10.1109/IJCNN.2017.7966336.
19. *Ho V. M., Lee J. A., Martin K. C.* The cell biology of synaptic plasticity // *Science*. 2011. Vol. 334, no. 6056. P. 623–628. DOI: 10.1126/science.1209236.
20. *Citri A., Malenka R. C.* Synaptic plasticity: Multiple forms, functions, and mechanisms // *Neuropsychopharmacology Reviews*. 2008. Vol. 33. P. 18–41. DOI: 10.1038/sj.npp.1301559.
21. *Roberts P. D., Leen T. K.* Anti-hebbian spike-timing-dependent plasticity and adaptive sensory processing // *Frontiers in Computational Neuroscience*. 2010. Vol. 4. P. 156. DOI: 10.3389/fncom.2010.00156.
22. *Jiajun F.* A review for deep reinforcement learning in atari: Benchmarks, challenges, and solutions // *ArXiv:abs/2112.04145*. 2022.

References

1. *Moreno-Bote R, Drugowitsch J.* Causal inference and explaining away in a spiking network. *Scientific Reports*. 2015;5:17531. DOI: 10.1038/srep17531.
2. *Lansdell BJ, Kording KP.* Neural spiking for causal inference and learning. *PLoS Computational Biology*. 2023;19(4):e1011005. DOI: 10.1371/journal.pcbi.1011005.
3. *Skatchkovsky N, Jang O, Simeone O.* Bayesian continual learning via spiking neural networks. *Frontiers in Computational Neuroscience*. 2022;16:1037976. DOI: 10.3389/fncom.2022.1037976.
4. *Friston K.* The history of the future of the bayesian brain. *Neuroimage*. 2012;62(2):1230–1233. DOI: 10.1016/j.neuroimage.2011.10.004.
5. *Kasabov N, Scott NM, Tu E, Marks S, Sengupta N, Capecchi E, Othman M, Doborjeh MG, Murli N, Hartono R, Espinosa-Ramos JI, Zhou L, Alvi FB, Wang G, Taylor D, Feigin V, Gulyaev S, Mahmoud M, Hou Z-G, Yang J.* Evolving spatio-temporal data machines based on the NeuCube neuromorphic framework: Design methodology and selected applications. *Neural Networks*. 2016;78:1–14. DOI: 10.1016/j.neunet.2015.09.011.

6. Kasabov NK. NeuCube: a spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data. *Neural Networks*. 2014;52:62–76. DOI: 10.1016/j.neunet.2014.01.006.
7. Schliebs S, Fiasché M, Kasabov N. Constructing robust liquid state machines to process highly variable data streams. *International Conference on Artificial Neural Networks ICANN 2012: Artificial Neural Networks and Machine Learning – ICANN*. 2012;604–611. DOI: 10.1007/978-3-642-33269-2_76.
8. Doborjeh M, Doborjeh Z, Merkin A, Krishnamurthi R, Enayatollahi R, Feigin V, Kasabov N. Personalized spiking neural network models of clinical and environmental factors to predict stroke. *Cognitive Computation*. 2022;14:2187–2202. DOI: 10.1007/s12559-021-09975-x.
9. Fernando C. From blickets to synapses: Inferring temporal causal networks by observation. *Cognitive Science*. 2013;37(8):1426–1470. DOI: 10.1111/cogs.12073.
10. Markram H, Gerstner W, Sjöström PJ. A history of spike-timing-dependent plasticity. *Frontiers in Synaptic Neuroscience*. 2011;3:4. DOI: 10.3389/fnsyn.2011.00004.
11. Kerr RR, Grayden DB, Thomas DA, Gilson M, Burkitt AN. Coexistence of reward and unsupervised learning during the operant conditioning of neural firing rates. *PLoS ONE*. 2014;9(1): e87123. DOI: 10.1371/journal.pone.0087123.
12. Yuan M, Wu X, Yan R, Tang H. Reinforcement learning in spiking neural networks with stochastic and deterministic synapses. *Neural Computation*. 2019;31(12):2368–2389. DOI: 10.1162/neco_a_01238.
13. Mozafari M, Ganjtabesh M, Nowzari-Dalini A, Thorpe SJ, Masquelier T. Bio-inspired digit recognition using reward-modulated Spike-Timing-Dependent Plasticity in deep convolutional networks. *Pattern Recognition*. 2019;94:87–95. DOI: 10.1016/j.patcog.2019.05.015.
14. Frémaux N, Sprekeler H, Gerstner W. Functional requirements for reward-modulated Spike-Timing-Dependent Plasticity. *The Journal of Neuroscience*. 2010;30(40):13326–13337. DOI: 10.1523/JNEUROSCI.6249-09.2010.
15. Juarez-Lora A, Ponce-Ponce VH, Sossa H, Rubio-Espino E. R-STDP spiking neural network architecture for motion control on a changing friction joint robotic arm. *Frontiers in Neurorobotics*. 2022;16:904017. DOI: 10.3389/fnbot.2022.904017.
16. Ivanov D, Chezhegov A, Kiselev M, Grunin A, Larionov D. Neuromorphic artificial intelligence systems. *Frontiers in Neuroscience*. 2022;16:959626. DOI: 10.3389/fnins.2022.959626.
17. Kiselev M, Ivanov A, Ivanov D. Approximating conductance-based synapses by current-based synapses. *Advances in Neural Computation, Machine Learning, and Cognitive Research IV. Neuroinformatics 2020. Studies in Computational Intelligence*. 2020;925:394–402. DOI: 10.1007/978-3-030-60577-3_47.
18. Kiselev MV. A Synaptic Plasticity rule providing a unified approach to supervised and unsupervised learning. *Proceedings of International Joint Conference on Neural Networks*. 2017:3806–3813. DOI: 10.1109/IJCNN.2017.7966336.
19. Ho VM, Lee JA, Martin KC. The cell biology of synaptic plasticity. *Science*. 2011;334(6056): 623–628. DOI: 10.1126/science.1209236.
20. Citri A, Malenka RC. Synaptic plasticity: Multiple forms, functions, and mechanisms. *Neuropsychopharmacology Reviews*. 2008;33:18–41. DOI: 10.1038/sj.npp.1301559.
21. Roberts PD, Leen TK. Anti-hebbian spike-timing-dependent plasticity and adaptive sensory processing. *Frontiers in Computational Neuroscience*. 2010;4:156. DOI:10.3389/fncom.2010.00156.
22. Jiajun F. A review for deep reinforcement learning in atari: Benchmarks, challenges, and solutions. *ArXiv:abs/2112.04145*. 2022.

Киселев Михаил Витальевич — родился в Московской области (1966). Окончил с отличием физический факультет и аспирантуру МГУ им. М. В. Ломоносова. Кандидат технических наук (1993, МГТУ им. Н. Э. Баумана). Сооснователь и генеральный директор компании «Мегапьютер Интеллидженс» (1994–2020). В настоящее время — руководитель направления в частном учреждении «Цифрум» ГК «Росатом»; доцент факультета физики, прикладной математики и информационных технологий Чувашского государственного университета, руководитель лаборатории нейроморфных вычислений ЧГУ. Научные интересы: новые модели импульсных нейронов и синапсов, изучение их свойств, методы кодирования информации в нейронных сетях, модели механизмов кратковременной памяти, процессы самоорганизации хаотических импульсных нейронных сетей, моделирование больших импульсных нейронных сетей на нейроморфных массивно-параллельных компьютерах (GPU, SpiNNaker, Loihi), реализация механизмов обучения ИНС с учителем и без учителя в различных постановках (unsupervised learning, supervised learning, reinforcement learning). Опубликовал свыше 40 научных работ по этой тематике.



Россия, 428000 Чувашия, Чебоксары, Московский пр-т, 15
Чувашский государственный университет имени И. Н. Ульянова
E-mail: mkiselev1@yandex.ru
ORCID: 0000-0001-7403-6418
AuthorID (eLibrary.Ru): 734474
AuthorID (Scopus): 7005285769

Ларионов Денис Александрович — родился в Москве (1983). Окончил МФТИ (2006). В настоящее время обучается в аспирантуре Чувашского государственного университета им. Ульянова. Руководитель отдела искусственного интеллекта в компании ЧУ «Цифрум» (компания Госкорпорации «Росатом»). Эксперт фонда «Сколково» по искусственному интеллекту. Научный сотрудник МНИОИ им. П. А. Герцена. Область научных интересов: нейроморфные вычисления, предиктивная аналитика в промышленности, анализ медицинских данных. Автор нескольких научных и научно-популярных публикаций, научный редактор.



Россия, 115230 Москва, Варшавское шоссе, 46
Частное учреждение «Цифрум», Государственная корпорация «Росатом»
E-mail: denis.larionov@gmail.com
ORCID: 0000-0002-7437-2646
AuthorID (WoS): JDM-7863-2023
AuthorID (Scopus): 57417043900

Урусов Андрей Михайлович — родился в Чебоксарах (1993). Окончил магистратуру Чувашского государственного университета им. Ульянова (ЧГУ). В настоящее время обучается в аспирантуре ЧГУ. Работал в НПФ и ведущих банках страны (моделирование и количественная оценка рисков). Области научных интересов: нейроморфные вычисления, анализ данных, задачи оптимизации, количественные финансы, предиктивная аналитика в банковском деле, распределенные финансы. Автор нескольких научных публикаций.



Россия, 117420 Москва, ул. Наметкина, д. 16, корпус 1
Акционерное общество «Газпромбанк»
E-mail: imurusov@mail.ru
AuthorID (eLibrary.Ru): 864799