

В. Ю. Осипов, доктор техн. наук, проф., e-mail: osipov_vasily@mail.ru,
Санкт-Петербургский институт информатики и автоматизации РАН

Ассоциативно-пространственная адресация к памяти рекуррентных нейронных сетей

В интересах расширения когнитивных возможностей рекуррентных нейронных сетей исследуются подходы к наделению их ассоциативно-пространственной адресацией к памяти. Наличие дополнительной пространственной адресации к памяти в рекуррентных нейронных сетях позволяет им управлять ассоциативными взаимодействиями обрабатываемых сигналов в зависимости от текущих состояний слоев. Предлагается новый метод такой ассоциативно-пространственной адресации, снижающий сложность ее реализации. Он базируется на усовершенствованной модели рекуррентной нейронной сети с управляемыми синапсами. Приводятся результаты моделирования, подтверждающие целесообразность и справедливость предлагаемых решений.

Ключевые слова: рекуррентная нейронная сеть, ассоциативно-пространственная адресация к памяти, когнитивная обработка информации

Введение

Основами всей современной когнитивной обработки информации в рекуррентных нейронных сетях (РНС) выступают ассоциативное запоминание и ассоциативный вызов из памяти сигналов. Однако этого, недостаточно для наделения РНС широкими когнитивными возможностями [1, 2]. Желательно, чтобы РНС имели, помимо ассоциативной, также пространственную адресацию к памяти в зависимости от текущих состояний слоев сети.

В последние годы некоторые результаты в этом направлении совершенствования РНС достигнуты. Они опираются на известные положения по управлению синапсами [3–5] и пространственной обработкой сигналов в нейронных сетях [6–10].

В работах [11–13] предложены методы когнитивной обработки информации в двухслойной РНС с управляемыми синапсами. Нейронам такой сети свойственны три состояния: ожидания, возбуждения и временной невосприимчивости после возбуждения. Каждый нейрон одного слоя в общем случае связан синапсами со всеми нейронами другого слоя. Нейроны одного слоя связаны друг с другом не имеют. Время задержки единичных образов в образуемых контурах этой сети меньше времени невосприимчивости нейронов после возбуждения. Подаваемые в РНС сигналы (в общем случае цветные динамические воздействия) предварительно раскладывают на составляющие в базисе, согласованном с входным слоем. При этом каждую из них преобразуют в последовательность единичных образов (импульсов) с частотой повторения как функцией от амплитуды составляющей. Сигналы в сети представляют в виде последовательных совокупностей единичных образов (СЕО) в соответствии с заданными правилами. При передаче СЕО от слоя к слою осуществляют их пространственные сдвиги. За счет цикличности обработки этих СЕО с простран-

ственными сдвигами каждый слой РНС разделяется на конечное число равных полей. Сеть наделяется прозрачной логической структурой. Результаты распознавания запоминают на элементах сети. Причем за счет частичного отражения СЕО от принимающих слоев частично стирают устаревшую информацию с синапсов [14, 15]. В итоге обрабатываемые СЕО продвигаются вдоль слоев сети в заданных направлениях. Они успешно ассоциируются друг с другом и с запомненными сигналами с учетом обратных результатов распознавания. Приоритетность коротких связей между нейронами взаимодействующих слоев обеспечивает однозначное соответствие между входом и выходом РНС. В качестве результатов обработки используют последовательные СЕО на выходном слое после обратного преобразования их в исходные сигналы. Пример такой РНС с простой спиральной структурой слоев показан на рис. 1.

Через воздействие на синапсы, в зависимости от текущих состояний слоев, в работах [16, 17] реко-

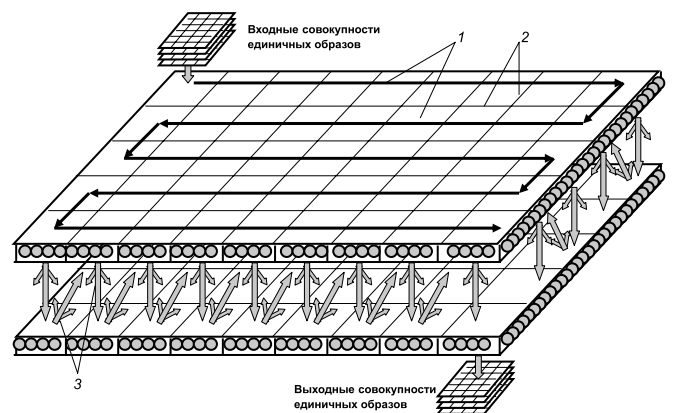


Рис. 1. Рекуррентная нейронная сеть со спиральной структурой слоев:

1, 3 — направления продвижения СЕО вдоль слоев и между ними; 2 — линии разбивки слоев на логические поля

мендовано изменять формы и поворачивать поперечные сечения расходящихся и сходящихся единичных образов в РНС. Это позволило наделить РНС не только ассоциативной, но и пространственной адресацией к ее памяти. На основе ее были предложены решения по формированию в РНС трех сигнальных систем [18, 19], обеспечивающих глубокую обработку информации и взаимодействие с внешним миром.

Однако в известных работах [16—19] зависимости параметров ассоциативно-пространственного обращения к памяти от текущих состояний слоев РНС оставлены без должного внимания. Кроме этого, реализация дополнительной пространственной адресации к памяти в известной РНС осуществляется путем воздействий на синапсы со стороны блока управления. В интересах этого блок управления считывает текущие состояния слоев и вырабатывает соответствующие решения. Однако технически реализовать это управление в РНС больших размеров довольно сложно.

Требуется уточнить модель РНС с рассматриваемыми свойствами, развить методы ассоциативно-пространственной адресации к памяти (АПАП) и взаимодействия сигналов, снизить сложность реализации.

Модели РНС с АПАП

Согласно модели [14—19] различают прямые $w_{ij}(t)$ и обратные $w_{ij}^*(t)$ веса (проводимости) синапсов, которые определяются как

$$\begin{aligned} w_{ij}(t) &= k_{ij}(t) \cdot \beta(r_{ij}(t)) \cdot \eta_{ij}(r_{ij}(t)); \\ w_{ij}^*(t) &= k_{ij}^*(t) \cdot \beta(r_{ij}(t)) \cdot \eta_{ij}(r_{ij}(t)), \end{aligned} \quad (1)$$

где $k_{ij}(t) = 1 - \exp(-\gamma \cdot g_{ij}(t))$, $k_{ij}^*(t) = 1 - \exp(-\gamma(g_{ij}(t) - g_0))$ — прямой и обратный весовые коэффициенты; γ — постоянный коэффициент; $g_{ij}(t)$ — число запомненных воздействий единичных образов на синапс на момент времени t ; $(g_{ij}(t) - g_0)$ — число запомненных воздействий единичных образов, которые могут быть стерты с синапса, $g_{ij}(t) \geq g_0$; $\beta(r_{ij}(t))$ — функция ослабления расходящихся единичных образов; $\eta_{ij}(r_{ij}(t))$ — функция ослабления сходящихся единичных образов; $r_{ij}(t)$ — удаленность связываемых через синапсы нейронов. При формализации запоминания и стирания результатов воздействий на синапсы единичных образов величина $g_{ij}(t)$ равна

$$g_{ij}(t) = g_{ij}(t - \Delta t) \pm \Delta g^{(*)}(I_{ij}^{(*)}(t)), \quad (2)$$

$\Delta g(I_{ij}(t))$, $\Delta g^*(I_{ij}^*(t))$ — прирост и снижение $g_{ij}(t)$ в зависимости от проходящего через синапсы тока $I_{ij}(t)$ в прямом и тока $I_{ij}^*(t)$ в обратном направлениях. В частном случае $\Delta g(I_{ij}(t)) = A I_{ij}(t)$, а $\Delta g^*(I_{ij}^*(t)) = D \cdot I_{ij}^*(t) \leq (g_{ij}(t - \Delta t) - g_0)$, где A , D — константы.

Каждая из связей (синапсов) имеет свои значения функций ослабления $\beta_{ij}(r_{ij})$, $\eta_{ij}(r_{ij})$, соответственно, расходящихся и сходящихся единичных образов.

При возбуждении i -го нейрона на его выходе формируется заряд Q_i . Переносимый часть этого заряда прямой ток $I_{ij}(t)$, протекающий через ij -й синапс, выражается как $I_{ij}(t) = U_i(t)w_{ij}(t)$, где U_i — потенциал на выходе i -го нейрона передающего слоя. Полагается, что на входе возбуждаемого j -го нейрона принимающего слоя потенциал близок к нулю. Это допущение характерно для всех известных импульсных нейронных сетей [3, 4].

Суммарному заряду, переносимому от i -го нейрона к возбуждаемым j -м нейронам за отводимое время T , свойственно условие

$$\sum_{j=1}^n \int_0^T I_{ij}(t) dt = Q_i = \text{const.}$$

Изменение в расходящемся пучке любого $I_{ij}(t)$ через функцию $\beta(r_{ij}(t))$, от которой зависит $w_{ij}(t)$, приводит к пространственному перераспределению переносимых зарядов. При этом относительные токи $\delta I_{ij}(t)$ через синапсы расходящегося пучка равны

$$\delta I_{ij}(t) = I_{ij}(t) / \sum_{j=1}^n I_{ij}(t) = w_{ij}(t) / \sum_{j=1}^n w_{ij}(t). \quad (3)$$

Суммарный ток $I_{j\Sigma}(t)$ на входе возбуждаемого j -го нейрона без учета частичного отражения сигналов равен $I_{j\Sigma}(t) = \sum_{i=1}^n U_i(t)w_{ij}(t)$. Для того, чтобы j -й нейрон возбудился, этот ток должен перенести некоторый конечный заряд

$$Q_{jR} = \sum_{i=1}^n \int_0^T U_i(t)w_{ij}(t) dt.$$

В формировании заряда Q_{jR} участвуют токи $I_{ij}(t)$ от различных нейронов. Изменение любого из них через функцию $\eta_{ij}(r_{ij}(t))$ в составе $w_{ij}(t)$ приводит к перераспределению вкладов переносимых зарядов в Q_{jR} .

Обратный ток $I_{ij}^*(t)$ через ij -й синапс выражается как

$$I_{ij}^*(t) = I_{j\Sigma}(t)K_j(t)w_{ij}^*(t) / \sum_{i=1}^n w_{ij}^*(t),$$

$K_j(t) = I_{j\Sigma}^*(t)/I_{j\Sigma}(t)$ — коэффициент отражения единичных образов от принимающих нейронов; $I_{j\Sigma}^*(t)$ — суммарный ток, отражаемый от j -го нейрона.

Результирующий ток, протекающий через входное сопротивление j -го нейрона, равен $I_{j\text{вх}}(t) = I_{j\Sigma}(t) - I_{j\Sigma}^*(t) = I_{j\Sigma}(t)(1 - K_j(t))$. В случаях, когда ток $I_{j\text{вх}}(t)$ или соответствующий накопленный заряд больше порогового значения, j -й нейрон возбуждается. На его выходе формируется единичный образ (импульс), и нейрон переходит в состояние невосприимчивости. В состоянии невосприимчиво-

сти каждый нейрон находится одинаковое время, которое больше, чем время задержки единичных образов в двухслойных контурах сети.

Изменяя $r_{ij}(t)$ и другие параметры в функциях $\beta(r_{ij}(t))$, $\eta_{ij}(r_{ij}(t))$ в зависимости от текущих состояний слоев РНС, можно осуществлять пространственные сдвиги обрабатываемых сигналов, формировать новые и поворачивать формы поперечных сечений расходящихся и сходящихся единичных образов. В результате обеспечивается возможность ассоциативно-пространственной адресации к памяти и соответствующего взаимодействия сигналов в РНС. С физической точки зрения это достижимо прямым изменением проводимостей синапсов. Такие ассоциативно-пространственная адресация к памяти и взаимодействие сигналов справедливы, если потенциалы на входах возбуждаемых нейронов невелики и практически не зависят от потенциалов на их выходах.

В случаях, когда потенциалы на входах нейронов могут быть сравнимы с потенциалами на их выходах, ситуация существенно меняется. Еще в работе [20] замечено, что в биологических нейронах при их возбуждении токи могут протекать встречно по дендритам и формировать на выходе их значительные потенциалы.

С учетом этого в предлагаемой модели РНС с ассоциативно-пространственной адресацией к памяти (АПАП) прямой ток $I_{ij}(t)$, протекающий через ij -й синапс, должен быть равен $I_{ij}(t) = (U_i(t) - U_{jR}(t))w_{ij}(t)$, где $U_{jR}(t)$ — потенциал на входе j -го нейрона. В этом случае

$$Q_i = \sum_{j=1}^n \int_0^T (U_i(t) - U_{jR}(t))w_{ij}(t)dt = \text{const.} \quad (4)$$

Относительное значение $\delta I_{ij1}(t)$ прямого тока через ij -й синапс в расходящемся пучке единичных образов имеет вид:

$$\begin{aligned} \delta I_{ij1}(t) &= \\ &= (U_i(t) - U_{jR}(t))w_{ij}(t) / \sum_{j=1}^n (U_i(t) - U_{jR}(t))w_{ij}(t). \end{aligned} \quad (5)$$

В отличие от выражения (3), свойственного всем известным моделям импульсных нейронных сетей, отношение (5) учитывает различие потенциалов на синапсах в расходящемся пучке. Аналогично можно получить выражение для относительных токов $\delta I_{ij2}(t)$ применительно к сходящемуся пучку. Оно также зависит от различия потенциалов, приходящихся на отдельные синапсы.

Из анализа выражений (3)—(5) видно, что при изменениях $\Delta U_{ij}(t) = U_{iT}(t) - U_{jR}(t)$ в зависимости от текущих состояний слоев в сети происходит пространственное перераспределение переносимых зарядов в расходящихся и сходящихся пучках. Изменение $\Delta U_{ij}(t)$ равносильно приросту или уменьшению проводимостей синапсов.

Реальные значения $\Delta g(I_{ij}(t))$, $\Delta g^*(I_{ij}^*(t))$ в выражении (2) и изменения проводимостей (весов) (1)

синапсов должны в этом случае определяться в зависимости от перераспределяемых токов в расходящихся и сходящихся пучках единичных образов.

Учет этого эффекта, ранее не принимаемого во внимание, позволяет по-новому взглянуть на пространственную обработку сигналов в РНС. Как в биологических нейронах, так и в их электрических моделях потенциалы на входах могут быть сравнимы с потенциалами на выходах. В результате РНС способна в широких пределах осуществлять пространственное перераспределение переносимых зарядов в расходящихся и сходящихся пучках единичных образов.

Несмотря на это, необходимость в блоке управления, как в работах [16—19], не отпадает. Однако по отношению к синапсам в предлагаемом решении функции этого блока могут быть сведены лишь к неоперативной корректировке через них пространственных параметров расходящихся и сходящихся единичных образов. Формально через блок управления можно существенно изменять структуру РНС в процессе ее функционирования, корректировать базовые функции ослабления расходящихся и сходящихся единичных образов, варьировать порогом возбуждения нейронов и внутренним временем сети.

Базовые функции ослабления сигналов в РНС с АПАП

В качестве базовых функций ослабления сигналов в ряде случаев могут выступать известные радиальные функции Гаусса, Габора, прямая и обратная мультикватратичные и др. [2, 3, 7]. В частности, для расходящихся пучков применима модифицированная функция Габора, определенная только в положительной области. Ее выражение без учета зависимости параметров от времени имеет вид:

$$\beta_{ij}(x_{ij}, y_{ij}) = \exp\left(-\frac{X_{ij}^2 + \gamma^2 Y_{ij}^2}{2\sigma^2}\right) \cos^2\left(\frac{2\pi}{\lambda} X_{ij}\right), \quad (6)$$

где $X_{ij} = x_{ij}\cos\varphi + y_{ij}\sin\varphi + c$; $Y_{ij}(t) = -x_{ij}\sin\varphi + y_{ij}\cos\varphi + f$; φ — угол поворота; c, f — параметры сдвигов по осям X, Y ; $\gamma = \sigma_x/\sigma_y$; σ_x, σ_y — средние квадратические отклонения по осям X, Y ; $\sigma_x = \sigma$; λ — пространственная длина волны. Отличие (6) от классической функции Габора в том, что $\cos(\cdot)$ в правой части возведен в квадрат. Неплохо ведет себя в качестве $\beta(r_{ij}(t))$ также обратная мультикватратичная функция. Эти же функции применимы также в качестве $\eta_{ij}(r_{ij}(t))$.

Задание для РНС базовых функций ослабления единичных образов в расходящихся и сходящихся пучках с конкретными параметрами позволяет разделить сеть не только определенной структурой, но и своеобразной "наследственной" информацией. Именно на основе ее сеть будет обучаться и действовать. Однако однозначных подходов к определению

нию целесообразных базовых функций ослабления для такой РНС пока нет. Формы их могут зависеть от реализуемых в РНС пространственных сдвигов и направлений продвижения СЕО вдоль слоев, количества и типов ассоциируемых в сети сигналов, их реальной взаимосвязи и других факторов.

Опираясь на результаты работ [16–19] и учитывая, что основные ассоциативные взаимодействия единичных образов реализуются не внутри нейросетевых каналов, по которым продвигаются СЕО, а между каналами, можно определиться с исходной ориентацией этих функций в пространстве слоев сети. Целесообразно ориентировать базовые функции ослабления в плоскости слоев перпендикулярно нейросетевым каналам. Причем для функции (6) провалы должны приходиться на каналы с такими же направлениями продвижения СЕО вдоль них, что и у канала, содержащего нейрон, порождающий расходящийся пучок единичных образов. Это позволяет не только обеспечить хорошее ассоциативное взаимодействие сигналов в сети, но и расширить возможности ее памяти, снизить избыточность запоминания информации.

В ряде случаев для определения таких целесообразных функций сначала можно задаться одной из известных радиальных функций. Затем следует приравнять в $w_{ij}(t) = k_{ij}(t) \cdot \beta_{ij}(t) \cdot \eta_{ij}(t)$ все весовые коэффициенты $k_{ij}(t)$ единице. Применяя известные методы [3, 4], можно обучить сеть решению некоторых опорных задач, варьируя только значениями функций ослабления $\beta_{ij}(t)$ расходящихся и $\eta_{ij}(t)$

сходящихся единичных образов. Закрепляя найденные $\beta_{ij}(t)$, $\eta_{ij}(t)$ и снимая ограничения с $k_{ij}(t)$, позволяя этим коэффициентам изменяться по рассмотренным выше правилам, получаем РНС с "наследственной информацией", способную к самообучению.

Результаты ассоциативно-пространственной адресации к памяти РНС

Для того чтобы показать, как могут перераспределяться относительные токи и переносимые заряды в расходящихся пучках единичных образов, сделаем допущения. Пусть $k_{ij}(t)$ и $\eta_{ij}(r_{ij}(t))$ в выражениях (1) принимают значения единицы. Тогда вместо $w_{ij}(t)$ в выражениях (3), (5) можно подставить $\beta(r_{ij}(t))$. Используем в качестве $\beta(r_{ij}(t))$ модифицированную функцию Габора (6). В результате согласно (3) и (5) получим относительные распределения токов в пространстве. Примеры таких сглаженных распределений представлены на рис. 2, а–г.

На рис. 2, а показано распределение согласно (3) для (6) без учета индивидуальной разности потенциалов на синапсах в расходящемся пучке. Рисунки, 2, б–г отражают возможные изменения исходного распределения (рис. 2, а) с таким учетом. Из анализа этих распределений видно, что если потенциалы на входах нейронов принимающего слоя могут быть сравнимы с потенциалами на выходах возбужденных нейронов, то сеть способна в широких пределах изменять адресацию к памяти в зависимости от текущих состояний слоев. Это под-

тверждают существенные изменения (рис. 2, б–г) исходной формы (рис. 2, а) пространственного распределения относительных токов. Причем обеспечивается возможность адресации к памяти не только по направлению, но и по конкретным областям, даже по отдельным элементам. Это справедливо не только для изменения распределения относительной базовой функции (6), но и других. При таком ассоциативно-адресном обращении к памяти сохраняются все возможности селективного запоминания и вызова сигналов, а также стирания устаревшей информации за счет частичного отражения единичных образов от принимающих слоев. Обеспечивается устойчивое функционирование РНС при обработке потоков сигналов в реальном времени. Предельное количество информации C в битах, запоминаемое на синапсах такой сети, равно

$$C = -\frac{n^2}{2} \ln \left(1 - \frac{2W_{\max}}{n^2 \beta \bar{\eta}} \right), \quad (7)$$

где n — число нейронов в каждом из двух слоев сети; W_{\max} — максимальный суммарный вес синапсов сети, при ко-

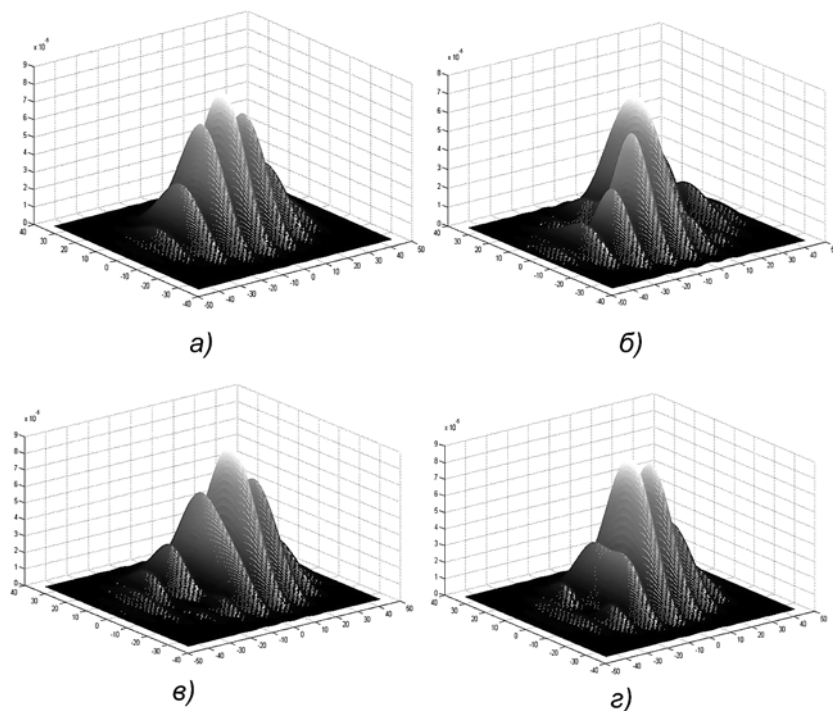


Рис. 2. Примеры пространственного распределения относительных токов в расходящихся пучках единичных образов: а — распределение без учета индивидуальной разности потенциалов на синапсах; б–г — распределения с таким учетом

тором число запоминаемых новых единичных образов равно числу стираемых устаревших образов; $\bar{\beta}$, $\bar{\eta}$ — средние значения функций $\beta(r_{ij}(t))$, $\eta_{ij}(r_{ij}(t))$. Выражение (7) можно получить из анализа выражений (1), (2). Оценка (7) емкости памяти РНС не противоречит результатам работ [2, 21, 22]. В то же время она расширяет взгляды на предельные возможности памяти РНС с учетом стирания устаревшей информации.

Такое обращение к памяти и, соответственно, управление ассоциативным взаимодействием сигналов в РНС позволяет реализовывать условные переходы от обработки одних сигналов к другим.

В случае чистого запоминания входных СЕО они проходят по РНС без искажений, оставляя в ней след о своем появлении в виде изменения весов синапсов с учетом текущих состояний слоев. При запоминании СЕО в РНС со спиральной структурой слоев сети (см. рис. 1) каждая сигнальная конструкция (СЕО, их группа) наиболее прочно связывается с другими, когда при продвижении вдоль слоев расстояние между ними становится минимальным. В итоге устанавливаются прочные связи различных сигнальных конструкций друг с другом. Когда одновременно с запоминанием входных СЕО осуществляется вызов из памяти связанных сигналов, также учитываются текущие состояния слоев. Причем в зависимости от этих состояний, определяющих текущие значения функций ослабления расходящихся и сходящихся единичных образов, возможно формирование различных последовательностей, вызываемых из памяти сигналов. Так можно вызывать из памяти РНС сигнальные конструкции перед или после вызывающего сигнала, дополняя его. Например, пусть РНС помнит сигнальную конструкцию "нейроны связаны синапсами", где каждой букве соответствует одна или несколько СЕО. Тогда в зависимости от текущих состояний ее слоев одним и тем же сигналом "связаны" из памяти РНС могут быть вызваны: "нейроны" или (и) "синапсами", другие связанные сигналы. Также сеть может не отреагировать на входной сигнал. Причем группировка этих конструкций может осуществляться из различных, но связанных между собой, запомненных сигналов.

Допустимы различные варианты структуры РНС, позволяющей переставлять сигнальные конструкции в предложениях с использованием ассоциативно-пространственной адресации к памяти. В частности, помимо структуры, показанной на рис. 1, возможен вариант, приведенный на рис. 3. Основным выходом в такой РНС выступает выход 2, а не выход 1.

В соответствии с этой структурой (рис. 3) сеть может многократно прокручивать внутри себя входную или вызванную из памяти последовательность СЕО и формировать из нее новые последовательности с переставленными сигнальными конструкциями относительно исходного предложения.

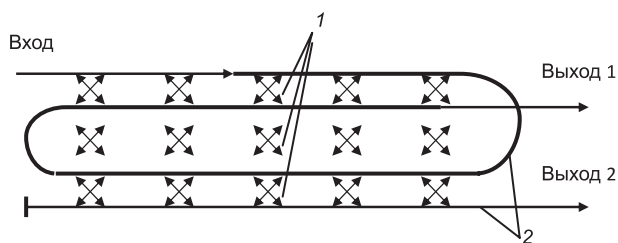


Рис. 3. Структура рекуррентной нейронной сети на уровне нейросетевых каналов, реализующая перестановку сигнальных конструкций в предложениях:

1 — основные ассоциативные взаимодействия сигналов в РНС; 2 — нейросетевые каналы в РНС, по которым продвигаются совокупности единичных образов

Ослабляя одни и усиливая другие связи (1, рис. 3) в зависимости от текущих состояний слоев, РНС может на выходе 2 формировать различные порядки вызываемых из памяти сигналов. Например, пусть РНС держит в своей оперативной памяти на самих нейронах сигнальную конструкцию "веса синапсов это их проводимости". Тогда из нее РНС может сформировать "проводимости синапсов это их веса" и другие конструкции.

Заключение

Таким образом, опираясь на полученные результаты исследования, есть основания утверждать, что не только предлагаемым искусственным, но и биологическим РНС свойственна не просто ассоциативная, а ассоциативно-пространственная адресация к памяти. Рекуррентные сети в зависимости от текущих состояний слоев способны осуществлять пространственное перераспределение переносимых зарядов в расходящихся и сходящихся пучках единичных образов. Имеет место сочетание ассоциативного и пространственного методов обращения к памяти РНС. Такой подход позволяет снять многие ограничения, препятствующие пониманию процессов когнитивной обработки информации в РНС. За счет такой адресации предоставляется возможность оперативно регулировать загрузку памяти второй сигнальной системы РНС входным потоком. Этим обеспечивается глубокая обработка информации в РНС. Также можно оперативно связывать вторую сигнальную систему РНС с третьей и управлять выходом сети, выдавать не все результаты обработки, а только "осознанные".

При реализации синапсов РНС на мемристорах и использовании электрических моделей нейронов с существенной зависимостью потенциалов на входах от выходных уровней можно снизить сложность создания таких сетей. Рекомендуется в известных моделях РНС учитывать текущие потенциалы на входах возбуждаемых нейронов и осуществлять пространственное перераспределение переносимых зарядов в зависимости от состояний слоев.

Предложенные решения развивают методы когнитивной обработки информации в РНС в части

ассоциативно-пространственной адресации к памяти и снижения сложности реализации. Они могут быть использованы при разработке перспективных нейросетевых когнитивных машин с высокоскоростной параллельной обработкой различных сигналов в условиях высокой неопределенности ситуаций и событий.

Список литературы

1. **Haikonen Pentti O. A.** The Role of Associative Processing in Cognitive Computing // *Cognitive Computing*. 2009. N. 1. P. 42–49.
2. **Palm G.** Neural associative memories and sparse coding // *Neural Networks*. 2013. N. 37. P. 165–171.
3. **Haykin S.** *Neural Networks and Learning Machines*. Third Edition. New-York: Prentice Hall, 2008. 936 p.
4. **Galushkin A. I.** *Neural Networks Theory*. Berlin Heidelberg: Springer-Verlag, 2007. 396 p.
5. **Liaw J.-S., Berger T. W.** Dynamic Synapse for Signal Processing in Neural Networks. US Patent 6643627 B2, Nov. 4, 2003.
6. **Cover T. M.** Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern recognition // *IEEE Transactions on Electronic Computers*. 1965. June. P. 326–334.
7. **Jeong S., Lee M.** Adaptive object recognition model using incremental feature representation and hierarchical classification // *Neural Networks*. 2012. N. 25. P. 130–140.
8. **Chartier S., Giguere G., Langlois D.** A new bidirectional heteroassociative memory encompassing correlational, competitive and topological properties // *Neural Networks*. 2009. N. 22. P. 568–578.
9. **Fukushima K.** Artificial vision by multi-layered neural networks: Neocognitron and its advances // *Neural Networks*. 2013. N. 37. P. 103–119.
10. **Kohonen T.** Essentials of the self-organizing map // *Neural Networks*. 2013. N. 37. P. 52–65.
11. **Осипов В. Ю., Осипова М. В.** Способ и устройство интеллектуальной обработки информации в нейронной сети. Патент на изобретение RU 2413304, опубл. 27.02.2011, бюл. № 6.
12. **Осипов В. Ю.** Способ интеллектуальной обработки информации в нейронной сети. Патент на изобретение RU 2427914, опубл. 27.08.2011, бюл. № 24.
13. **Осипов В. Ю.** Рекуррентная нейронная сеть со структурой слоев в виде двойной спирали // *Информационные технологии*. 2014. № 7. С. 56–60.
14. **Осипов В. Ю.** Стирание устаревшей информации в ассоциативных интеллектуальных системах // *Мехатроника, автоматизация, управление*. 2012. № 3. С. 16–20.
15. **Осипов В. Ю.** Пределы памяти рекуррентных нейронных сетей со стиранием устаревшей информации // *Научный вестник Новосибирского государственного технического университета*. 2014. № 3 (56). С. 115–122.
16. **Осипов В. Ю.** Метод управления синапсами в рекуррентной нейронной сети // *Информационные технологии*. 2013. № 7. С. 61–66.
17. **Осипов В. Ю.** Способ интеллектуальной обработки информации в нейронной сети. Патент на изобретение RU 2514931, опубл. 10.05.2014, бюл. № 13.
18. **Осипов В. Ю.** Рекуррентная нейронная сеть с двумя сигнальными системами // *Информационно-управляющие системы*. 2013. № 4. С. 8–15.
19. **Осипов В. Ю.** Ассоциативная интеллектуальная машина с тремя сигнальными системами // *Информационно-управляющие системы*. 2014. № 5. С. 12–17.
20. **Alberts B., Bray P., Lewis J., Raff M., Roberts K., Watson J. D.** (Eds.). *Molecular Biology of The Cell*. Second Edition. New York: Garland Publishing, Inc. 1989. 1219 p.
21. **Palm G.** Information capacity in recurrent McCulloch-Pitts networks with sparsely coded memory states // *Network*. 1992. N. 3. P. 177–186.
22. **Bosch H., Kurfess F.** Information storage capacity of incompletely connected associative memories // *Neural Networks*. 1998. N. 11 (5). P. 869–876.

V. Yu. Osipov, Professor, e-mail: osipov_vasily@mail.ru,
Saint-Petersburg Institute for Informatics and Automation of RAS

Associative and Spatial Addressing to Memory of Recurrent Neural Networks

Approaches to enhance the cognitive capabilities of recurrent neural networks are investigated. In the interests it is proposed to empower such a networks of associative and spatial addressing to memory. The additional spatial addressing to memory of recurrent neural networks allows them to manage the associative interaction of processed signals. The improved model of recurrent neural network with associative and spatial addressing to memory is disclosed. It is shown that a memory addressing can be easily implemented in the electrical models recurrent neural networks. Examples of spatial redistribution of the currents in such networks, depending on the current state of interacting layers are shown. The proposed solutions allow to remove a number of constraints to understanding the processes of cognitive information processing in recurrent neural networks. These solutions can be used to develop advanced cognitive neural machines with high-speed parallel processing of different signals.

Keywords: recurrent neural network, associative and spatial addressing, cognitive information processing

References

1. **Haikonen Pentti O. A.** The Role of Associative Processing in Cognitive Computing, *Cognitive Computing*, 2009, no. 1, pp. 42–49.
2. **Palm G.** Neural associative memories and sparse coding, *Neural Networks*, 2013, no. 37, pp. 165–171.
3. **Haykin S.** *Neural Networks and Learning Machines*. Third Edition. New-York, Prentice Hall, 2008, 936 p.
4. **Galushkin A. I.** *Neural Networks Theory*. Berlin Heidelberg, Springer-Verlag, 2007, 396 p.
5. **Liaw J.-S., Berger T. W.** *Dynamic Synapse for Signal Processing in Neural Networks*. US Patent 6643627 B2, Nov. 4, 2003.
6. **Cover T. M.** Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern recognition, *IEEE Transactions on Electronic Computers*, June, 1965, pp. 326–334.
7. **Jeong S., Lee M.** Adaptive object recognition model using incremental feature representation and hierarchical classification, *Neural Networks*, 2012, no. 25, pp. 130–140.
8. **Chartier S., Giguere G., Langlois D.** A new bidirectional heteroassociative memory encompassing correlational, competitive and topological properties, *Neural Networks*, 2009, no. 22, pp. 568–578.

9. **Fukushima K.** Artificial vision by multi-layered neural networks: Neocognitron and its advances, *Neural Networks*, 2013, no. 37, pp. 103–119.

10. **Kohonen T.** Essentials of the self-organizing map, *Neural Networks*, 2013, no. 37, pp. 52–65.

11. **Osipov V. Yu., Osipova M. V.** Sposob i ustrojstvo intellektual'noj obrabotki informacii v nejronnoj seti [Method and device of intellectual processing of information in neural network]. Patent RU 2413304, 2011-02-27.

12. **Osipov V. Ju.** Sposob intellektual'noj obrabotki informacii v nejronnoj seti [Method of intellectual information processing in neural network]. Patent RU 2427914, 2011-08-27.

13. **Osipov V. Yu.** Rekurrentnaja nejronnaja set' so strukturoj sloev v vide dvojnoj spirali [Recurrent Neural Network with Structure of Layers in the Form of the Double Spiral]. *Informacionnye tehnologii*, no. 7, 2014, pp. 56–60 (In Russian).

14. **Osipov V. Yu.** Stiranje ustarevshej informacii v asociativnyh intellektual'nyh sistemah [Erase Outdated Information in Associative Intelligent Systems], *Mehatronika, avtomatizacija, upravlenie*, 2012, no. 3, pp. 16–20 (In Russian).

15. **Osipov V. Yu.** Predely pamjati rekurrentnyh nejronnyh setej so stiranijem ustarevshej informacii [Limits memory recurrent neural networks with deletion of obsolete information] // *Nauchnyj vestnik NGTU*, 2014, vol. 56, no. 3, pp. 115–122, (In Russian).

16. **Osipov V. Yu.** Metod upravlenija sinapsami v rekurrentnoj nejronnoj seti [The Method of Control Synapses in Recurrent Neural Network], *Informacionnye tehnologii*, no. 7, 2013, pp. 61–66 (In Russian).

17. **Osipov V. Ju.** Sposob intellektual'noj obrabotki informacii v nejronnoj seti [Method for intelligent information processing in neural network]. Patent RU 2514931, 2014-05-10.

18. **Osipov V. Yu.** Rekurrentnaja nejronnaja set' s dvumja signal'nymi sistemami [The Recurrent Neural Network with Two Signal System] // *Informacionno-upravliaiushchie sistemy*, 2013, no. 4, pp. 8–15 (In Russian).

19. **Osipov V. Yu.** Associativnaja intellektual'naja mashina s tremja signal'nymi sistemami [Three Signaling Systems Associative Machine] // *Informacionno-upravliaiushchie sistemy*, 2014, no. 5, pp. 12–17 (In Russian).

20. **Alberts B., Bray D., Lewis J., Raff M., Roberts K., Watson J. D.** (Eds.). *Molecular Biology of The Cell*. Second Edition. Garland Publishing, Inc. New York, 1989. 1219 p.

21. **Palm G.** Information capacity in recurrent McCulloch-Pitts networks with sparsely coded memory states. *Network*, 1992, no. 3, pp. 177–186.

22. **Bosch H., Kurfess F.** Information storage capacity of incompletely connected associative memories. *Neural Networks*, 1998, no. 11 (5), pp. 869–876.

УДК 004.032.26

И. В. Лёзина, канд. техн. наук, доц. кафедры информационных систем и технологий, chuchyck@yandex.ru, Самарский государственный аэрокосмический университет имени академика С. П. Королёва (национальный исследовательский университет) (СГАУ), Самара

Сравнительный анализ погрешности прогноза биржевых индексов методами экспоненциального сглаживания и нейросетевого моделирования

Прогнозирование значений биржевых индексов является актуальной задачей управления экономикой. Получаемые прогнозные значения служат основой для принятия различных управленческих решений. Целью данной работы является сравнение погрешности прогноза реальных данных биржевых индексов Dow Jones и Hang Seng. Для получения прогноза использовали методы экспоненциального сглаживания и нейросетевую модель многослойного перцептрона. Анализ погрешности использованных методов показал более высокую эффективность модели многослойного перцептрона по сравнению с традиционными методами экспоненциального сглаживания.

Ключевые слова: прогнозирование биржевых индексов, экспоненциальное сглаживание, нейросетевое моделирование, многослойный перцептрон

Введение

Выбор метода прогноза зависит в первую очередь от вида имеющейся информации. Прогнозирование на основе временных рядов осуществляется в тех случаях, когда существует ряд значений индекса за предыдущие периоды и предполагается, что в будущем закон его развития сохранится. При этом задача прогнозирования фактически сводится к задаче экстраполяции. К методам экстраполяции относятся метод скользящего среднего, метод наименьших квадратов и метод экспоненциального сглаживания [1–5]. Метод скользящего среднего состоит в замене фактических значений временного ряда расчетными, имеющими значительно меньшие

колебания, чем исходные данные. При этом среднее рассчитывается по группам данных за определенный интервал времени, причем каждая последующая группа образуется со сдвигом на один период. В результате первоначальные колебания временного ряда сглаживаются. Данный метод применяется для краткосрочного прогнозирования. Недостаток метода заключается в том, что требуется много данных для расчета прогнозного значения показателя.

Сущность метода наименьших квадратов состоит в минимизации суммы квадратов отклонений между наблюдаемыми и расчетными значениями. Расчетные значения находят по подобранному уравнению регрессии. Чем меньше расстояние между фактическими значениями и расчетными, тем более