

СТАТИСТИЧЕСКИЙ МЕТОД ВЫДЕЛЕНИЯ ВРЕМЕННЫХ ПАТТЕРНОВ ИЗ ЕСТЕСТВЕННЫХ ЯЗЫКОВ

М. Н. Калимолдаев, А. А. Пак*, С. С. Нарынов*

Институт проблем информатики и управления
Министерства образования и науки Республики Казахстан,
Республика Казахстан, 050010, Алма-Ата

*ТОО Alem Research, Республика Казахстан, 050010, Алма-Ата

УДК 004.8.032.26

Рассматривается алгоритм рекурсивной самоорганизующейся карты (Recursive Self Organizing Map, RSOM) для выделения статистически значимых временных паттернов из символического потока контекстно-зависимых грамматик. Главной идеей метода является, предложенное Thomas Voegtlin в 2002, объединение классического алгоритма Теuvo Kohonen и неявного представления времени в виде комбинации текущего входа и контекстного отклика сети. В статье приведены результаты экспериментов над текстами на казахском и английском языках.

Ключевые слова: обработка естественных языков, нейронные сети, самоорганизация.

In the paper the algorithm of recursive self-organizing map (RSOM) is examined in the application of the extraction of statistically significant temporal patterns from the symbolic data flow of context dependent grammars. The main idea of the method is the union of classical algorithm of Teuvo Kohonen and implicit representation of time in the form of the combination of input and network's contextual response. Originally the idea was proposed by Thomas Voegtlin, in 2002. In the paper there are experiments' results for kazakh and english languages.

Key words: natural language processing, neural networks, self-organization.

Введение. На сегодняшний день суммарный объем символической информации, аккумулированный в Интернете, превысил все знания человечества, накопленные до создания глобальной информационной среды. Современные методы индексации текстов не справляются в полной мере с задачей быстрого поиска информации. Проблема индексации накопленного знания является нетривиальной задачей. Размер текстов, накопленный в сети, составляет приблизительно 8000 терабайт, среднегодовой объем электронной почты — 10000 терабайт. Очевидно, что в таком большом потоке данных выборка документов, полученная с помощью методов полнотекстового поиска, может содержать много нерелевантных или слабореlevantных результатов из-за синонимического представления запроса, опечаток, грамматических неточностей, сленга и др. [1]. Предложенная Tim Berners-Lee в 2001 году идея семантической паутины была призвана решить вышеобозначенные проблемы, однако в 2006 г. автор идеи признал, что реализация в силу ряда причин невозможна. Принцип семантической паутины основан на оперировании смыслами взамен слов. Иначе говоря, вычислительная машина должна быть способна обобщать слова до некоторых понятий. На сегодняшний день эта задача является отчасти решенной при помощи методов корпусной лингвистики. Однако недостатком подобного подхода является колоссальный объем работы, связанный с разметкой текстов и поддержанием в актуальном состоянии

морфологии, синтаксиса и семантики конкретного языка. Без учета грамматических ошибок, сленга и др.

В данной статье будет рассмотрен метод, позволяющий в автоматическом режиме выделять темпоральные паттерны и конструкции естественного языка. Метод основан на нейросетевом подходе самоорганизации и обучении без учителя. На сегодняшний день было разработано несколько способов в нейрокомпьютинге для представления времени, большинство из них связаны с задачами распознавания речи или автоматической обработки языка. Условно можно разделить представление времени на явное, где время дополняет размерность пространства и коэффициент временной задержки используется для оценки мер процесса внутри окна наблюдения [2]. И неявное, когда время выражено опосредовано через логику представления изучаемого феномена [3], что характерно для рекуррентных сетей или так называемых „Leaky integrator“ процессов. В общем случае явное представление времени сверхчувствительно и ограничено к временным деформациям, напротив неявное является более устойчивым и универсальным. Действительным ограничением квазиинтегративных нейронов является тот факт, что информация с долговременной зависимостью обычно затухает по экспоненциальному закону. Однако рекуррентное представление необязательно должно иметь эти ограничения [4], по крайней мере до тех пор, пока не будет использован градиентный спуск.

В ряде научных работ были предложены различные подходы интегрирования представления времени в самоорганизующуюся карту Кохонена (Self-Organizing Map, SOM). Явное представление было рассмотрено в работе [5, 6], методы, основанные на латеральных и рекуррентных связях, были предложены в [7, 8], техники, базирующиеся на квазиинтеграторе [9, 10], также были изучены различные комбинации вышеозначенных принципов в работах [11, 12, 13]. Основным лейтмотивом перечисленных работ является идея обобщения процесса самоорганизации на время. Однако не совсем ясно, каким образом самоорганизация и время могут быть совмещены. Ошибка квантования или искажения в алгоритме SOM с учетом времени не может быть использована напрямую.

В данной статье будет рассмотрен алгоритм рекуррентного модифицированного SOM, где обратная связь будет использована для представления времени. Представление времени неявное и самореферентное в том смысле, что карта учится классифицировать собственные выходы. Однако основная идея заключается в добавлении в классический алгоритм SOM рекуррентных связей при сохранении оригинального принципа самоорганизации. Для того чтобы оценить сходимость сети, было предложено обобщение ошибки квантования временных рядов [14].

Самоорганизация и самореферентность. Широко известным примером рекуррентной нейронной сети является архитектура Simple Recurrent Network (SRN), предложенная Jeff Elman в 1990 году. SRN — это расширение классического Multi Layered Perceptron (MLP), а именно добавлен еще один скрытый слой, которой хранит отклики сети с заданным временным лагом в качестве дополнения к входному вектору. Алгоритм обратного распространения применяется к входному вектору и его копии с лагом. Поскольку SRN обучается на своих прошлых активностях, то его представление на скрытом слое является самореферентным. Самореферентные связи прямого и обратного распространения являются гомогенными. Нет никакой разницы между ними в уравнениях, описывающих активность сети, или правилах обучения. Далее будет показан подход, основанный на данном принципе применительно к SOM, иными словами, будут добавлены рекуррентные связи к оригинальной архитектуре SOM. Следует отметить, что уже существуют рабо-

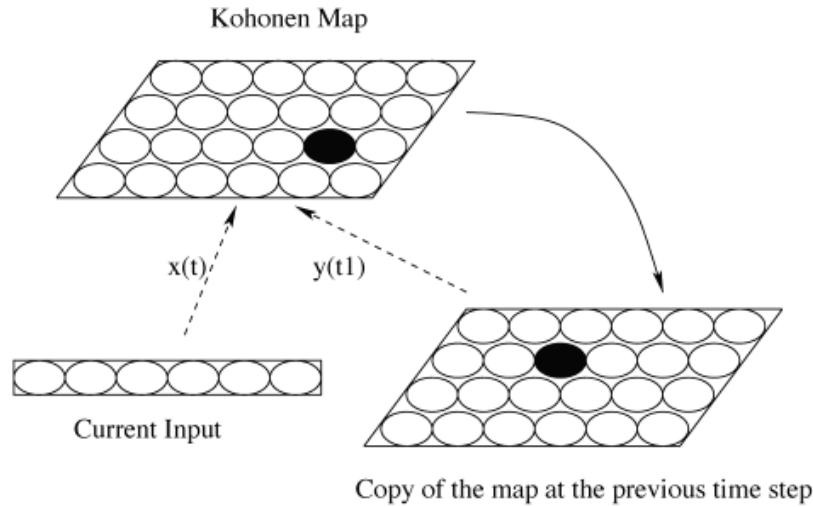


Рис. 1. Архитектура рекурсивного СОКа. Оригинальный алгоритм карты рекурсивно применяется к входному вектору $\mathbf{x}(t)$ и отклику сети, полученному на предыдущем шаге $\mathbf{y}(t-1)$. Пунктирные стрелки обозначают связи обучения. Сплошная стрелка означает связь один к одному. Сеть учится ассоциировать текущий вход к предыдущему контексту. Взято из работы [14]

ты [7, 8, 12], в которых рекуррентные связи были интегрированы в карту Кохонена, но в этих работах прямая и обратная связь не гомогенны. Архитектура, представленная в данной статье, использует активность сети с задержкой и гомогенность прямых обратных связей (см. рис. 1). Текущий входной вектор и предыдущий отклик сети рассматриваются как один входной вектор классического SOM алгоритма. Иначе говоря, сеть учится кластеризовать пару „входной вектор \mathbf{x} и $\mathbf{y}(t-1)$ контекст“. Длинные последовательности сеть заучивает итеративно, используя более короткие последовательности в качестве описательных признаков.

В классическом алгоритме SOM каждый i нейрон сравнивает свой вектор весов w_i , ($1 < i < N$) с входным вектором $\mathbf{x}(t)$, где t — это переменная времени. Обычно используется евклидова норма для определения победителя. Победитель определяется соревнованием среди N нейронов по наименьшей ошибке $E_i = \|\mathbf{w}_i - \mathbf{x}(t)\|$. По правилу обучения веса нейрона-победителя должно быть модифицировано согласно

$$\Delta \mathbf{w}_i = \gamma h_{ik}(\mathbf{x}(t) - \mathbf{w}_i), \quad (1)$$

где i — индекс нейрона-победителя, γ — функция скорости обучения, h_{ik} — функция соседства. Традиционно используют так называемую „мексиканскую шляпу“ или гауссиану, см. рис. 2. Существенным различием между ними является тот факт, что „мексиканская шляпа“ позволяет больше специализировать отклик нейрона-победителя.

В данной статье используется такая же нотация. Дополнительно, пусть $\mathbf{y}(t)$ — вектор активностей карты в момент времени t . Каждый i нейрон карты ($1 < i < N$) имеет пару весовых векторов, \mathbf{w}_i^x и \mathbf{w}_i^y , которые сопоставляются входному вектору $\mathbf{x}(t)$ и вектору активностей карты за предыдущую итерацию $\mathbf{y}(t-1)$ соответственно. Для того чтобы определить нейрон-победитель, предлагается объединить ошибки квантования, соответствующие $\mathbf{x}(t)$ и $\mathbf{y}(t-1)$. Поскольку прямые и обратные связи являются гомогенными, то суммарная ошибка квантования выглядит следующим образом

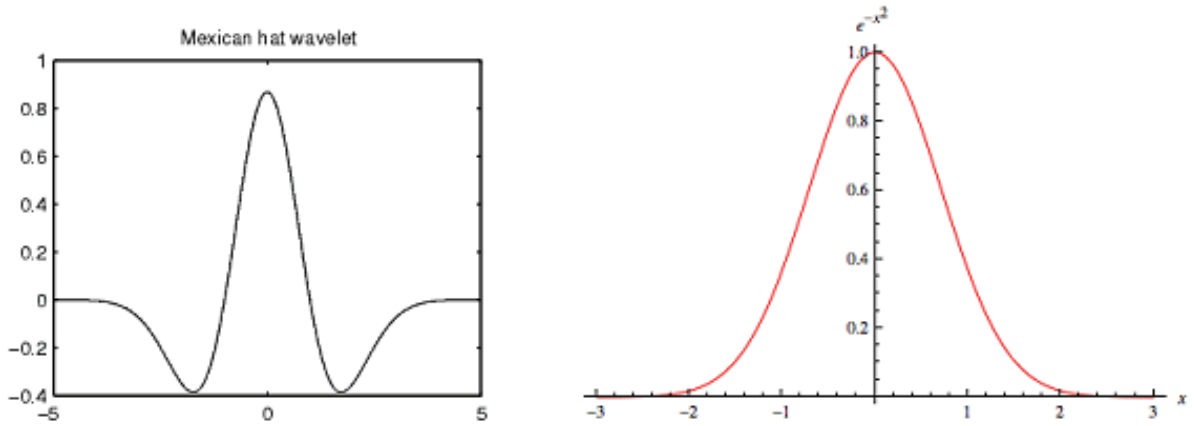


Рис. 2. Вейвлет „мексиканская шляпа“ (слева). Взят с сайта [15]. Гауссиана (Справа). Используются в качестве функции соседства между нейронами SOM.

Рис. взят с сайта [16]

$$E_i = \alpha \|\mathbf{x}(t) - \mathbf{w}_i^x\|^2 + \beta \|\mathbf{y}(t-1) - \mathbf{w}_i^y\|^2, \quad (2)$$

где $\alpha > 0$ — вес ошибки текущего входа, $\beta > 0$ — вес ошибки контекста. Индекс нейрона-победителя будет определен по формуле $k = \arg \min \{E_i\}_{1 < i < N}$, тогда правила эволюции прямых и рекуррентных весов будут иметь вид:

$$\Delta \mathbf{w}_i^x = \gamma h_{ik}(\mathbf{x}(t) - \mathbf{w}_i^x), \quad (3)$$

$$\Delta \mathbf{w}_i^y = \gamma h_{ik}(\mathbf{y}(t-1) - \mathbf{w}_i^y), \quad (4)$$

где γ — коэффициент скорости обучения и h_{ik} — функция соседства. Следует отметить, что это правила обучения оригинального SOM, примененные к $\mathbf{x}(t)$ и $\mathbf{y}(t-1)$.

Таким образом, в модели Recurrent SOM осуществляется обратная связь. Однако для того чтобы завершить определение модели, необходимо связать активность сети $\mathbf{y}(t)$ ко входу и синаптическим весам. Пусть компонента y_i вектора \mathbf{y} — это активность i нейрона, зависящая от квадрата ошибки квантования. Однако использование ошибки квантования напрямую $y_i = E_i$ может привести к нестабильности обучения. В таком случае предлагается использовать функцию активации $y_i = F(E_i)$. Принцип выбора функции активации будет разъяснен в следующем разделе.

Следует отметить что параметры α , β могут быть исключены из формулы $E_i = \|\mathbf{x}(t) - \mathbf{w}_i^x\|^2 + \|\mathbf{y}(t-1) - \mathbf{w}_i^y\|^2$ и $y_i = \sqrt{\frac{\beta}{\alpha}} F(\alpha E_i)$. Поэтому прямые и обратные связи являются гомогенными. Мы включили параметры α , β в E_i для удобства анализа сходимости и робастности алгоритма. В любом случае α , β могут быть рассмотрены как часть функции активации.

Стабильность и функция активации. Идея использовать в традиционном SOM алгоритме обратные связи была уже предложена [17, 18]. Однако серьезной проблемой была нестабильность в обучении [18]. Иными словами, при изменении многомерного представления обучающих примеров на карте, вектор обратных связей не успевал подстроиться, что препятствовало сходимости и стабильности обучения, что подтверждено в работе [17].

В данном разделе будет представлен аппроксимационный анализ стабильности рекуррентного SOM. Несмотря на приближительный характер анализа, будет дано представление о выборе трансферной функции.

Рассмотрим малое возмущение отклика сети $\delta\mathbf{y}(t)$ притом, что все остальные переменные и параметры модели остаются неизменными. Возмущение $\delta\mathbf{y}(t)$ будет иметь два эффекта. Во-первых, все последующие отклики $\mathbf{y}(t+n)$, $n > 0$ будут искажены рекуррентным распространением возмущения. Поставим в соответствие „будущим“ искажениям вектор $\delta\mathbf{y}(t+n)$.

Во-вторых, возмущение приведет к изменению распределения вектора $\mathbf{y}(t)$, что в свою очередь индуцирует долговременную адаптацию синаптических весов. В первом приближении будем рассматривать только кратковременную эволюцию. При условии достаточно малой величины коэффициента обучения можно пренебречь адаптацией весов.

Стабильность на кратковременной динамике зависит от эволюции $\delta\mathbf{y}(t+n)$. Если $\|\delta\mathbf{y}(t+n)\| \rightarrow 0$ при $n \rightarrow \infty$, то возмущение затухнет.

Более того, если $\sum_{n>0} \|\delta\mathbf{y}(t+n)\| < +\infty$, тогда распределение \mathbf{y} останется неизменным, и адаптации весов не произойдет. Также покажем, что эффект возмущения мал, если само возмущение мало. Можно выразить это утверждение, используя нотацию Ландау O :

$$\sum_{n>0} \|\delta\mathbf{y}(t+n)\| = O(\|\delta\mathbf{y}(t)\|). \quad (5)$$

Рассмотрим эффекты адаптации весов на стабильность обучения. Строго, стабильность весов должна быть рассмотрена, только когда коэффициент обучения стремится к нулю. Однако мы рассматриваем стабильность в более широком смысле: положим, что коэффициент обучения — малая константа и ожидаемые значения синаптических весов и их вариации можно сделать достаточно малыми правильным выбором коэффициента обучения. Для упрощения представления положим, что веса сходятся, когда используется традиционный алгоритм SOM и распределение входного вектора — константа [19].

Положим, $\mathbf{W}^y = (\mathbf{w}_i^y)$, $1 < i < N$ — матрица весов рекуррентных связей. Рассмотрим малые изменения $\delta\mathbf{W}^y$ во время обучения. Как выше было сказано, мы рассматриваем матрицу \mathbf{W}^y , элементы которой являются ожиданием рекуррентных весов во время обучения, $\delta\mathbf{W}^y$ — это возмущение этих весов. $\delta\mathbf{W}^y$ индуцирует прямое изменение отклика $\mathbf{y}(t)$ из-за того, что в уравнении (2) изменится \mathbf{w}_i^y . В дополнение, к непрямому изменению $\mathbf{y}(t)$ может привести распространение возмущения отклика сети, как это было показано выше. Возмущение отклика $\delta\mathbf{y}(t)$ — результат суперпозиции двух компонент, прямой и косвенной. Прямая компонента может быть получена из уравнения (2) и будет иметь тот же порядок энергии, что и $\delta\mathbf{W}^y$, при непрерывной трансферной функции. Косвенная компонента — сумма величин, возникших в результате распространения возмущения на шагах $t-1$, $t-2$ и т. д. Однако, если условие (5) удовлетворено, то энергия косвенной компоненты будет того же порядка, что и прямой компоненты. Поэтому возмущения отклика сети $\delta\mathbf{y}(t)$ и $\delta\mathbf{W}^y$ синаптических весов будут одного порядка:

$$\|\delta\mathbf{y}(t)\| = O(\|\delta\mathbf{W}^y\|). \quad (6)$$

В обучении подобной модели возникает так называемая проблема „движущейся цели“, где цель — это набор синапсов, стабильных в период обучения. Под движущейся целью подразумевается изменение облака данных, на котором обучается RSOM из-за изменения рекуррентных связей, иными словами, сеть учится классифицировать собственный

отклик, который меняется во время обучения. Пусть $\mathbf{T} = (\mathbf{t}_i)$, $1 < i < N$ соответствует матрице „движущихся целей“, где \mathbf{t}_i — „цель“, к которой должен сойтись текущий вектор \mathbf{w}_i^y по совместному стационарному распределению $(\mathbf{x}(t), \mathbf{y}(t-1))$. Следовательно, вектор „цель“ в любой определенный момент времени является результатом суперпозиции векторов синапсов и статистического ожидания примеров обучения, мы предполагаем, что для совместного стационарного распределения $(\mathbf{x}(t), \mathbf{y}(t-1))$ ряды $\mathbf{x}(t)$ — эргодичны. Однако любое изменение синаптических весов также изменит и распределение $(\mathbf{x}(t), \mathbf{y}(t-1))$, и вектор \mathbf{t}_i изменится. Для того чтобы гарантировать успевания модели за вектором цели, мы должны наложить дополнительное условие:

$$\|\delta\mathbf{T}\| < \|\delta\mathbf{W}^y\|. \quad (7)$$

Из-за локальных латеральных связей на карте каждый вектор цели \mathbf{t}_i является барьером ожидания $\mathbf{y}(t-1)$, когда нейрон i становится победителем в момент времени t . Хотя отношения соседства усложняют анализ возмущений в векторе целей, разумно предположить следующее неравенство:

$$\|\delta\mathbf{T}\| \leq \sqrt{N}E[\|\delta\mathbf{y}\|], \quad (8)$$

где $\delta\mathbf{T}$ — вариация матрицы целей из-за возмущения распределения \mathbf{y} и $E[\cdot]$ — статистическое ожидание. Из неравенства (8) с учетом (7) получаем:

$$\sqrt{N}E[\|\delta\mathbf{y}\|] \leq \|\delta\mathbf{W}^y\|. \quad (9)$$

Таким образом, мы получили два условия стабильности. Условие (5) гарантирует стабильность нейронной динамики, и условие (9) — стабильность весов в период обучения.

В случае, если оба условия удовлетворяются, тогда ожидаемые значения весов сходятся во время обучения. Выбор функции активации в работе представлен следующим выражением:

$$y_i(t) = F(E_i) = \exp(-\alpha\|\mathbf{x}(t) - \mathbf{w}_i^x\|^2 - \beta\|\mathbf{y}(t-1) - \mathbf{w}_i^y\|^2). \quad (10)$$

Функция является непрерывной и ограниченной на отрезке $(0, 1]$. Отклик нейрона-победителя находится ближе к 1, отклик проигравшего ближе к 0. Интуитивно активность нейрона-победителя должна быть стабильной в случае возмущения перцептивных или синаптических полей. Гауссовская функция активации сглаживает выбросы. Таким же свойством обладает и отклик проигравших нейронов. [14] показал, что (10) удовлетворяет условиям стабильности (5) и (9).

На рис. 3 представлены численные результаты экспериментов по определению области стабильности для рекурсивной карты Кохонена с функцией активации (10).

Эксперимент и параметры обучения. Для тестирования и отладки были проведены эксперименты, использовалась классическая функция соседства $h_{ik} = \exp(-\frac{d(i,k)^2}{\sigma^2})$. Размер карты составлял 28×28 для английского языка, 36×36 для казахского языка. Тексты, на которых обучалась сеть: „Brave New World“, Aldous Huxley; сборник исторических эссе о Казахстане, находящихся в сети в открытом доступе. Буквы были закодированы бинарным вектором, для казахского языка длина вектора составляла 6, и для английского 5 бит. В предобработке пунктуация была убрана из текста, исключения составляла тильда в английском языке в таких словах как don't, haven't и т.д. Сам текст был разбит на

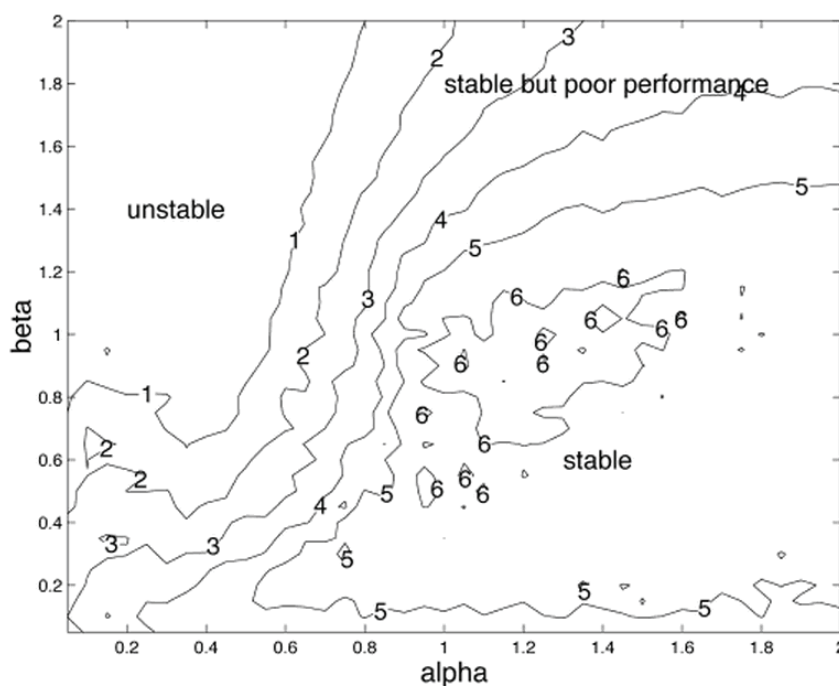


Рис. 3. Области стабильности рекуррентного SOM, обученного на бинарных последовательностях, для различных параметров α и β . На рис. представлена средняя длина рецептивного поля нейрона-победителя. Для каждого условия данные усреднялись по 8 экспериментам. Низкие значения указывают на плохую стабильность, большие значения указывают на хорошую стабильность. В верхней правой области наблюдается стабильность при плохой сходимости. Рис. взят из [14]

слова, разделителем служил пробел. Между словами контекст сети обнулялся. Коэффициент обучения также как и в оригинальной статье был константой $\gamma = 0,1$. Функция соседства представляла собой гауссиану с шириной $\sigma = 1,8$ и масштабом $s = 5$. Параметры $\alpha = 3$ и $\beta = 0,7$. По ходу обучения нейроны втягивали в свое перцептивное поле образы различных последовательностей букв. Объем текстов для казахского и английского языков составил приблизительно 400 Кб. Длительность обучения для английского языка составила около 8 часов. Для казахского языка — около 15 часов, что объясняется большой размерностью карты и кодирования букв. Реализация алгоритма была сделана на языке JavaSE2. Использовались математические библиотеки линейной алгебры и аналитической геометрии Cern Colt 1.2.0 [20]. Для визуализации экспериментальных данных использовалась программа Rajek64 3.11 [21].

Результаты. На рис. 4 представлены временные паттерны, которые были заучены нейронной сетью на текстах Aldous Huxley. Сеть была настроена для выявления характерных дифтонгов в тексте. Однако, в силу рекуррентной природы последняя буква первого дифтонга является также первой буквой второго дифтонга. На рис. 4 видны статистически значимые комбинации дифтонгов, образующие слова английского языка: was, with, whip, and, had, his. Характерные окончания: -ing -ent -n't

Данные, представленные на рисунке, подвергались постобработке. Прежде всего, были удалены дубликаты с общим корнем и одинаковым образом, что снизило количество узлов с 784 до 348. Также связи между узлами были прорежены по уровню 0,3.

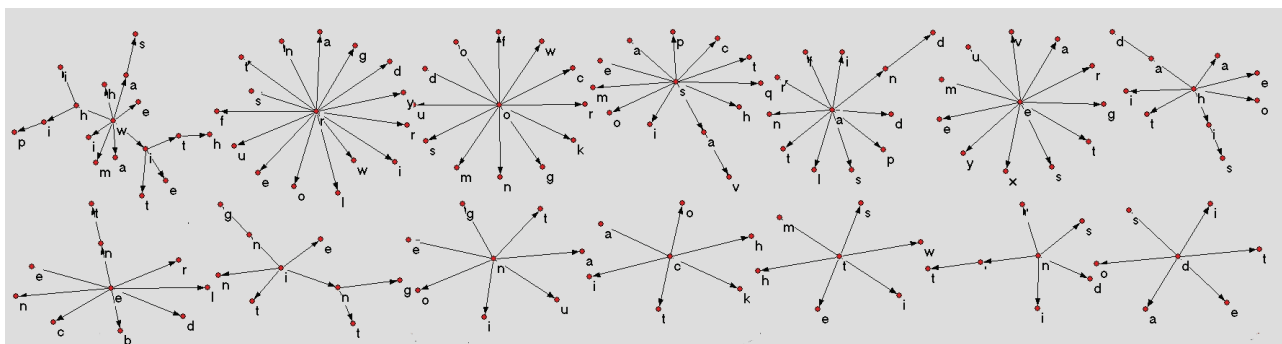


Рис. 4. Результаты обучения рекуррентного SOM на тексте английского языка

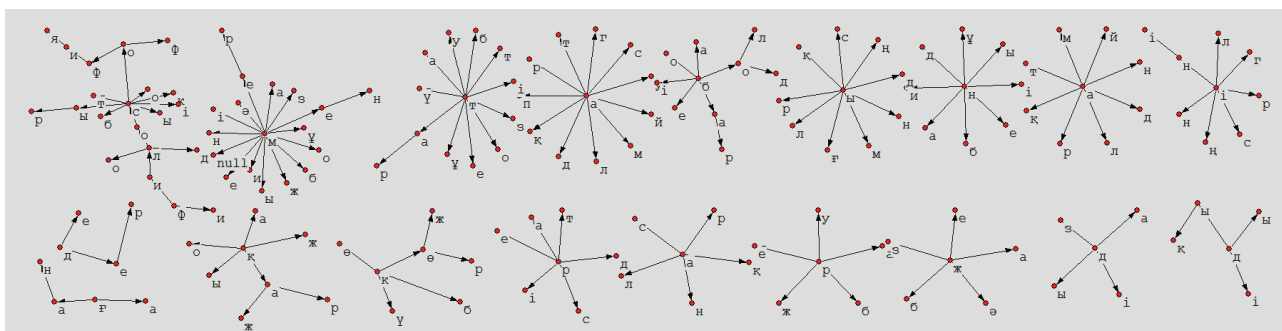


Рис. 5. Результаты обучения рекуррентного SOM на тексте казахского языка

На рис. 5 представлены временные паттерны казахского языка, характерные для исторических эссе, находящихся в свободном доступе. На изображение также видно, что сеть смогла выявить характерные сочетания, образующие целые слова: философ, философия, бар, іні, қар, Көр, бол. Характерные части слов: дық, дер, ған, тар, сыр.

Визуализация для казахского языка также была подвергнута фильтрации, редуцированию количества узлов с 1296 до 300 и прореживанию связей с уровнем 0,4.

Выводы. Сегодня существует большой ажиотаж вокруг обработки естественных языков. Это связано в первую очередь с глубокой востребованностью автоматических методов обработки больших объемов текстовых данных. Естественно, подобное возможно лишь при помощи методов машинного обучения. Здесь наблюдается тенденция в сторону более реалистичных когнитивных моделей, что связано с недавно появившимися эффективными методами „глубокого обучения“, подразумевающего выявление в потоках данных иерархии все более сложно устроенных признаков. Алгоритм рекуррентной карты является хорошим кандидатом для автоматического выявления сложных иерархий в потоке символов. Для развития требуется решить проблемы мертвых нейронов, дубликатов и скорости обучения.

Список литературы

1. SHUMSKY S. Selforganizing semantic networks / Neuroinformatics-2001 III All-Russian Scientific Conf. Lect.on Neuroinformatics. Moscow, 2001.
2. LANG, K. J., WAIBEL A. H., & HINTON, G. E. A time-delay neural network architecture for isolated word recognition. // Neural Networks. 1990. N 3, P. 23–43.

- 3 ELMAN, J. L. Finding structure in time. // Cognitive Science. 1990. N 14. P. 179–211.
- 4 BENGIO Y., SIMARD P., & FRASCONI P. Learning long-term dependencies with gradient descent is difficult // IEEE Transactions on Neural Networks. 1994. N 5(2). P. 157–166.
5. KANGAS J. On the analysis of pattern sequences by self-organizing maps // PhD Thesis, Helsinki University of Technology. 1994.
6. VESANTO J. Using the SOM and local models in time-series prediction / Proc. of the Workshop on Self-Organizing Maps'97 Espoo, Finland: Helsinki University of Technology. 1997. P. 209–214.
7. EULIANO N., & PRINCIPE J. Spatio-temporal self-organizing feature maps / Proc. of the Inter. Conf. on Neural Networks. 1996. P. 1900–1905.
8. HOEKSTRA A., & DROSSAERS M. An extended Kohonen feature map for sentence recognition / Proc. of the Inter. Conf. on Artificial Neural Networks. 1993. P. 404–407.
9. CHAPPELL G. J. & TAYLOR J. G. The temporal Kohonen map // Neural Networks. 1993. N 6. P. 441–445.
10. PRIVITERA C. M. & MORASSO P. The analysis of continuous temporal sequences by a map of sequential leaky integrators / Proc. of Intern. Conf. on Neural Networks. 1994. P. 3127–3130.
11. JAMES D. L. & MIKKULAINEN R. SARDNET: A self-organizing feature map for sequences // Advances in Neural Information Processing Systems. 1995. N 7. P. 577–584.
12. KOSKELA T., VARSTA M., HEIKKONEN J. & KASKI K.. Time series prediction using recurrent SOM with local linear models // Intern. J. of Knowledge-based Intelligent Eng. Sys. 1998. N 2(1). P. 60–68.
13. MOZAYYANI N., ALANOU V., DREYFUS J. & VAUCHER G. A spatiotemporal data coding applied to Kohonen maps / Proc. of the Intern. Conf. on Artificial Neural Networks. 1995. P. 75–79.
14. VOEGTLIN T. Recursive Self-Organizing Maps // Neural Networks. 2002. V. 15 N. 8–9. P. 979–992.
15. [electron. res.] <http://radio.feld.cvut.cz/matlab/toolbox/wavelet/mexihat.html>.
16. [electron. res.] <http://www5a.wolframalpha.com/Calculate/MSP/MSP6701h6a8290ga35bb8600003afghd1hgga79088?MSPStoreType=image/gif&s=61&w=304.&h=124.&cdf=RangeControl>
17. BRISCOE G. & CAELLI T. (1997). Learning temporal sequences in recurrent self-organising neural nets. In A. Sattar (Ed.), Advanced topics in artificial intelligence / Proc. of the 10th Australian Joint Conf. on Artificial Intell. Berlin: Springer. 1997. P. 427–435.
18. SCHOLTES J. C. Kohonen feature maps in natural language processing / Tech. Rep. CL-1991-01, Institute for Language, Logic and Information, University of Amsterdam. 1991.
19. VARSTA M., HEIKKONEN J. & MILLAN, J. D. R. Context-learning with the self-organizing map / Proc. of the Workshop on Self-Organizing Maps'97. Espoo, Finland: Helsinki University of Technology. 1997. P. 197–202.
20. КОНОНЕН Т. Self-Organizing Maps. Springer Verlag. 2001.
21. [electron. res.] <http://acs.lbl.gov/software/colt/>
22. [electron. res.] <http://pajek.imfm.si/doku.php>

Калимолдаев Максат Нурадилович — д-р физ.-мат. наук, проф., ген. дир., зав. лаб. математического моделирования и кибернетики Института проблем информатики и управления МОН РК, тел.: +7 (727) 272-37-11, e-mail: mnk@ipc.kz

Пак Александр Александрович — канд. техн. наук, ст. преп., рук. отд. науч. разраб. ТОО Alem Research, Республика Казахстан, тел.: +7 (701) 752-92-85, e-mail: aa.pak83@gmail.com
Нарынов Сергазы Сакенович — канд. техн. наук, ст. преп., ген. дир. ТОО Alem Research, Республика Казахстан, тел.: +7 (701) 723-01-62, e-mail: sergazy@gmail.com

Дата поступления — 18.09.2014