

УДК 519.876.5

В.М. Булойчик, В.Д. Гришко

НЕЙРОСЕТЕВОЙ МЕТОД ПОИСКА ЭКСТРЕМУМОВ ДИСКРЕТНО ЗАДАННОЙ ДВУМЕРНОЙ ФУНКЦИИ

Рассматривается подход к использованию нейросетевых технологий при решении аналитических задач с данными в виде значений дискретно заданной функции. Дан пример использования описанного метода на примере электронной карты местности.

Введение

В последние десятилетия в мире бурно развивается новая прикладная область математики – нейроматематика, которая ориентирована на описание процессов функционирования искусственных нейронных сетей (ИНС). Актуальность исследований в этом направлении подтверждается необходимостью решения с помощью ИНС массы прикладных задач: автоматизации процессов распознавания образов, адаптивного управления сложными процессами, прогнозирования событий, создания экспертных систем, организации ассоциативной памяти и многих других. В настоящее время не представляется возможным создавать универсальные мощные сети, комплексно решающие широкий круг подобных задач. Это вынуждает разрабатывать специализированные ИНС, ориентированные на решение отдельных прикладных задач и функционирующие по различным алгоритмам. В настоящей статье приводится синтез и математическое описание специализированных ИНС, с помощью которых решается задача поиска экстремумов дискретно заданной функции, а также приводятся результаты их экспериментальных исследований с помощью имитационной модели.

Многие числовые данные из систематизированных источников, например электронной карты местности (ЭКМ), можно рассматривать как значения дискретно заданной функции. Одной из наиболее часто решаемых на практике задач при обработке таких данных является поиск глобального и локальных экстремумов функции, представляющей собой аппроксимированную земную поверхность. Такая задача имеет место при выборе в зоне аэропорта позиции для средств обеспечения посадки самолетов, при поиске места для бурения скважин к водонесущим и другим слоям, выявлении зон с наиболее напряженной экологической ситуацией и т. п. В военном деле, когда всесторонний учет тактических свойств местности может сыграть решающую роль в бою, важность быстрого решения подобной задачи не вызывает сомнения.

Очевидно, что данная задача имеет большую размерность и применение традиционных способов ее решения, ориентированных на последовательную реализацию, приводит к значительным затратам вычислительных ресурсов ЭВМ и часто не обеспечивает получение решения в требуемом ритме развития обстановки. В связи с этим представляет интерес использование подхода, обеспечивающего распараллеливание задачи и основанного на применении современных технологий обработки информации, в частности нейросетевой технологии, наиболее приемлемой для решения подобных комбинаторных задач [1].

1. Основные теоретические положения

Основу предлагаемого подхода составляет использование искусственной нейронной сети (ИНС) Хопфилда. Кратко рассмотрим ее возможности.

Полносвязная сеть Хопфилда представляет собой динамическую систему. Все состояния нейронов такой сети характеризуют некоторое фазовое пространство, а каждому отдельному состоянию сети соответствует точка в этом фазовом пространстве. Структура соединений нейронной сети и значения ее весовых коэффициентов определяют очертания некоторой поверхности, которую принято называть функцией вычислительной энергии [2-4].

При дискретном представлении такой сети изменения состояний нейронов описываются

следующим соотношением:

$$y_i(t+1) = F(u_i) = F\left(\sum_j w_{ij} y_j(t) - \theta_i\right), \quad (1)$$

$$y_i(0) \quad i, j = \overline{1, n},$$

где $y_i(t+1)$ – значение i -го нейронного элемента в момент времени $t+1$;

t – дискретно изменяющееся время;

$F(\bullet)$ – активационная функция нейронов;

θ_i – пороговое значение i -го нейрона;

h_i – входное значение i -го нейронного элемента;

n – количество нейронов в сети.

При возбуждении сети (т. е. при подаче хотя бы на один ее i -й вход значения $h_i > 0$) она производит вычисления, следуя по пути, уменьшающему вычислительную энергию. Релаксационный процесс изменения состояний нейронов можно интерпретировать как движение точки в фазовом пространстве в устойчивое положение, определяемое минимумом вычислительной энергии.

Такое свойство сети Хопфилда на практике часто используется при решении задач ассоциативного распознавания [5]. В этом случае впадины и долины в сети соответствуют наборам информации, которую хранит сеть. Если процесс возбуждения начинается с приближенной или неполной информации (т. е. сеть возбуждена в области притяжения некоторого i -го локального минимума вычислительной энергии, соответствующего определенной информации), то он следует по пути, который ведет к ближайшей впадине (т. е. к этому i -му локальному минимуму). Траектория движения точки может характеризовать процесс «вспоминания» хранимой сетью информации.

Этот же процесс может быть интерпретирован и как процесс поиска локального минимума функции вычислительной энергии, т. е. как решение задачи поиска наименьшего (наибольшего) значения функции, формируемой сетью. Более детально рассмотрим это свойство сети Хопфилда.

Пусть структура соединений нейронов в сети и значения их весовых коэффициентов определяют функцию вычислительной энергии, имеющую вид плоской поверхности, что соответствует состоянию нейронов, когда $\forall i \theta_i = 0$, а все весовые коэффициенты одинаковы и отсутствуют обратные связи нейронов на себя, т. е. $\forall ij w_{ij} = a$, $w_{ii} = 0$. Кроме того, функция активации должна быть возрастающей, т. е.

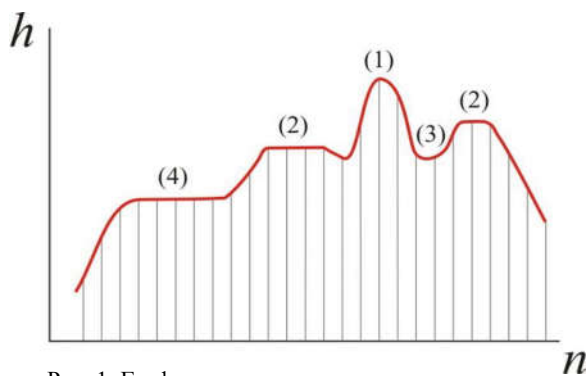


Рис. 1. Графическое двумерное представление входной информации

$$\frac{dF(u)}{du} > 0. \quad (2)$$

При подаче на вход сети некоторой дискретно аппроксимированной функции $h(i)$, например функции, изображенной на рис. 1, степень возбужденности отдельных нейронов, а также вычислительная энергия всей сети будут определяться только уровнем входных сигналов.

В связи с выполнением условия (2) релаксационный процесс изменения состояний нейронов, следуя в направлении минимизации вычислительной энергии, приведет к тому, что

возбужденным останется только тот нейрон, который соответствует глобальному экстремуму функции $h(i)$. Полное математическое описание процесса поиска глобального экстремума функции с помощью дискретной синхронной сети Хопфилда приведено в [6].

Рассмотренная нейронная сеть решает задачу поиска глобального экстремума. Однако на практике, при представлении непрерывной функции в дискретном виде, зачастую имеет место

наличие нечеткого глобального экстремума, т. е. «протяженность» его вершины превышает величину дискретизации, либо в результате дискретизации появляется несколько экстремумов функции, значения которых являются одинаковыми на заданном интервале.

Из-за наличия на входе сети хотя бы двух равных максимальных (минимальных) значений итерации будут продолжаться до перехода всех нейронов в нулевое состояние. Данный факт объясняется тем, что первоначально весовые коэффициенты межнейронных связей установлены в единичное (равное) состояние и сеть, пытаясь «вспомнить» находящуюся в данных связях информацию, стремится к минимуму вычислительной энергии, соответствующему нулевому состоянию всех нейронов сети, т. е. сеть не обеспечит решение поставленной задачи.

Такая ситуация характерна и для процесса поиска локальных экстремумов с помощью сети Хопфилда. Кроме того, применение рассмотренного метода для селекции локальных экстремумов связано с проблемой «краевого эффекта» – специфическим взаимодействием соседних нейронов на краях исследуемой зоны и в ее внутреннем пространстве, что приводит к искажению значений состояния краевых нейронов, а в последующем – и всей сети в целом.

Рассмотрим варианты построения ИНС, обеспечивающие устранение этих негативных моментов.

2. Построение ИНС для поиска глобального экстремума

Для нахождения нечеткого глобального экстремума в полносвязную ИНС предлагается ввести «решающий» нейрон, выделенный на рис. 2. Его задачей будет селекция такого состояния всей сети, когда отсутствует приращение количества «обнуленных» нейронов, а текущие состояния активных нейронов будут одинаковыми (начиная с этого момента будет происходить автокомпенсация одинаково возбужденных нейронов). Поэтому функция активации решающего нейрона должна иметь вид

$$\varphi(t+1) = \begin{cases} 1 & \text{при } m(t+1)=m(t) \cap \forall i, j (y_i(t+1)=y_j(t+1)) \text{ при } (y_i(t+1) \neq 0) \cap (y_j(t+1) \neq 0), \\ 0 & \text{в противном случае,} \end{cases} \quad (3)$$

где $m(t) = \sum_{i=1}^n i, i = \overline{1, n}, j = \overline{1, n}.$

Для первоначального рассмотрения примем, что функция активации всех нейронов (за исключением решающего элемента) является линейно возрастающей. До установления устойчивого состояния итерационные вычисления в сети производятся по правилу

$$y_j(t+1) = F[y_j(t) - \varepsilon \sum_{i \neq j} y_i(t)], 1 \leq j \leq n, 1 \leq i \leq n, \varepsilon < \frac{1}{n}, t = 0, 1, \dots, \quad (4)$$

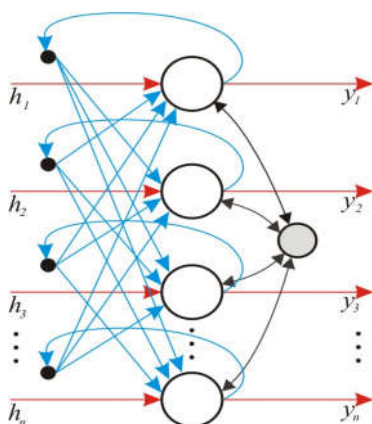


Рис. 2. Фрагмент полносвязной ИНС с решающим элементом

где ε – коэффициент, определяющий «скорость» изменения состояния сети.

Начальные состояния нейронов такой сети соответствуют входным значениям исследуемой дискретной функции. После нескольких итераций сеть придет в устойчивое состояние: активным будет один или несколько нейронов, первоначальные состояния которых были максимальными.

Таким образом, после введения дополнительного нейрона с предложенной функцией активации (3) процесс вычислений закончится не при отсутствии изменений состояния всех нейронов сети (как это было в «классическом» варианте сети), а при отсутствии приращения количества «обнуленных» нейронов и эквивалентном состоянии активных нейронов сети.

При подаче на вход рассматриваемой ИНС значений дискретной функции \mathbf{H} с ее помощью может быть найден глобальный экстремум с точностью до величины интервала дискретизации и параметра ε . Количество итераций, необходимых для поиска глобального экстремума, можно сократить за счет нелинейного преобразования входного сигнала:

$$h_i^*(0) = h_i^k(0), k = 2, 3, \dots, i = \overline{1, n}. \quad (5)$$

Это обеспечит усиление влияния нейронов с большими значениями входных сигналов на «более слабых конкурентов» и, соответственно, наискорейшее приведение сети в устойчивое состояние. Значение коэффициента k выбирается на основе априорной информации о входной функции и позволяет в значительных пределах менять параметр ε , обеспечивая тем самым наискорейшую стабилизацию сети.

За счет параллельной обработки информации в такой сети время итерций τ_u будет значительно меньше времени t_m решения задачи поиска глобального экстремума любым известным последовательным методом, т. е. $\tau_u \ll t_m$. Кроме того, к явным достоинствам описанного метода следует отнести его нечувствительность к объему входных данных – время решения задач для 10, 100, 1000 и т. д. входных сигналов практически не изменяется, что резко контрастирует с более чем экспоненциальным ростом времени решения задачи при использовании последовательных методов [4, 6].

3. Построение ИНС для поиска локальных экстремумов

С целью решения задачи поиска локальных экстремумов (для определенности – локальных максимумов) сформулируем *необходимое условие* существования локального максимума: точка может являться локальным максимумом, если значения точек ее ближайшего окружения не превышают ее собственного значения. Очевидно, что нахождение и анализ знака дифференциала функции в этих точках обеспечит проверку данного условия.

Обратим внимание на то, что для нахождения локальных экстремумов однослойной ИНС уже недостаточно. Для решения такой задачи предлагается ИНС, представленная на рис. 3.

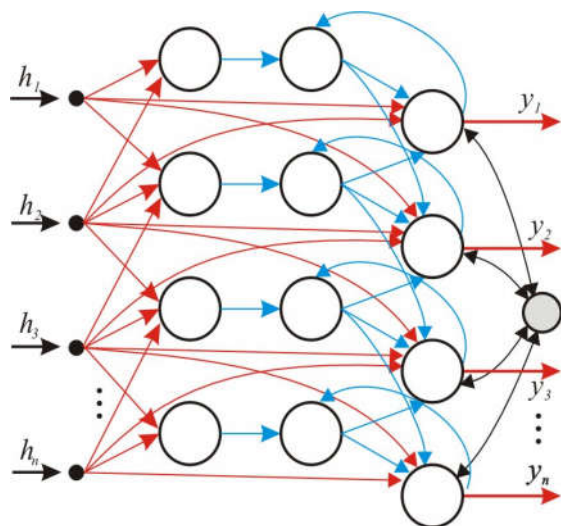


Рис. 3. Фрагмент ИНС для нахождения локальных экстремумов

Как видно из рисунка, каждый нейрон такой сети связан с другими нейронами своего и соседних слоев посредством перекрестных и обратных связей. При этом текущее состояние нейронов зависит не только от их предыдущего состояния, но и от состояния связанных с ним нейронов. А поскольку радиус связей предлагается принять равным единице, то синтезированную сеть можно отнести к клеточной нейронной сети [7].

Принцип функционирования предлагаемой ИНС состоит в следующем: первый слой нейронов из всей совокупности входных сигналов выделяет «кандидатов» на значение локального максимума, т. е. реализует вычисление и анализ знака дифференциала функции в каждой дискретной точке и передает выработанную информацию во второй слой.

Функция активации нейронов первого слоя является пороговой. Ее значение зависит от текущего состояния каждого нейрона и нейронов, его окружающих, т. е.

$$F_i^{(1)}(h) = \begin{cases} 0 & \text{при } h_j > h_i \\ 1 & \text{в противном случае,} \end{cases} \quad (6)$$

где $i = \overline{1, n}$, $j = \overline{i-1, i+1}$.

Второй и третий слои посредством нескольких итераций производят окончательный выбор значений входной функции, соответствующих локальным максимумам (см. рис. 1, области 1 и 2). Решающий элемент определяет момент окончания поиска требуемых значений и, соответственно, выдачи результатов работы сети.

Весовые коэффициенты всех связей в рассматриваемой ИНС принимаются равными единице.

Функция активации третьего слоя нейронов также является пороговой и зависит от входных значений, подаваемых на ИНС, и состояния нейронов второго слоя:

$$F_i^{(3)}(t+1) = \begin{cases} 0 & \text{при } y_i^{(2)}(t) = 0 \cup y_j^{(2)}(t) = 0 \cup h_j > h_i, \\ 1 & \text{в противном случае,} \end{cases} \quad (7)$$

где $i = \overline{1, n}$, $j = \overline{i-1, i+1}$.

Состояние нейронов второго слоя в момент времени $t=0$ определяется состоянием нейронов первого слоя, т. е.

$$y_i^{(2)}(0) = y_i^{(1)}.$$

В течение последующих итераций за счет обратной связи значения состояний нейронов второго слоя устанавливаются равными

$$y_i^{(2)}(t+1) = y_i^{(3)}(t+1).$$

Моментом окончания итераций в сети является поступление сигнала на нейроны третьего слоя от нейрона решающего элемента.

Функция активации решающего нейрона имеет вид

$$F_p(t+1) = \begin{cases} 0 & \text{при } y_i^{(3)}(t+1) \neq y_i^{(3)}(t) \text{ для } \forall i = \overline{1, n}, \\ 1 & \text{в противном случае.} \end{cases} \quad (8)$$

Такое функционирование синтезированной сети обеспечивает нахождение всех локальных максимумов входного сигнала. При этом ИНС придет в устойчивое состояние, при котором активными будут нейроны третьего слоя, соответствующие локальным максимумам входного сигнала.

Полученный результат имеет важное практическое значение для решения ряда военно-прикладных задач на основе ЭКМ. В частности, описанный выше алгоритм используется в программном комплексе для предварительной селекции участков местности, предназначенных для размещения радиоэлектронных средств.

4. Результаты экспериментальных исследований

С целью проверки эффективности предложенного метода авторами были проведены экспериментальные исследования с использованием имитационной модели рассмотренной ИНС. В качестве входных данных для ИНС использованы реальные значения относительных высот

земной поверхности. На рис. 4 представлен фрагмент аппроксимированной земной поверхности, полученный с использованием матрицы высот ЭКМ.

Оценка проводилась по отношению к известным методам поиска экстремумов. А поскольку земную поверхность, как правило, невозможно описать аналитической зависимостью, то для решения задач поиска возвышающихся участков местности и впадин на практике наиболее часто применяются методы статистического направленного перебора. В основном такие методы обеспечивают сокращение времени поиска на 30-50% относительно времени, затрачиваемого на полный перебор [6]. Исходя из этих соображений, для сравнительной оценки предложенного метода и существующих последовательных методов решения сформулированной задачи в качестве базового принят метод полного перебора, который можно охарактеризовать количеством последовательно выполняемых операций.

Эксперимент проводился при следующих условиях:

- входные данные (высоты) изменялись в диапазоне 722 – 1397 м;
- размер анализируемой поверхности ограничен прямоугольником 2 500 x 3 000 м;
- шаг дискретизации – 100 м.

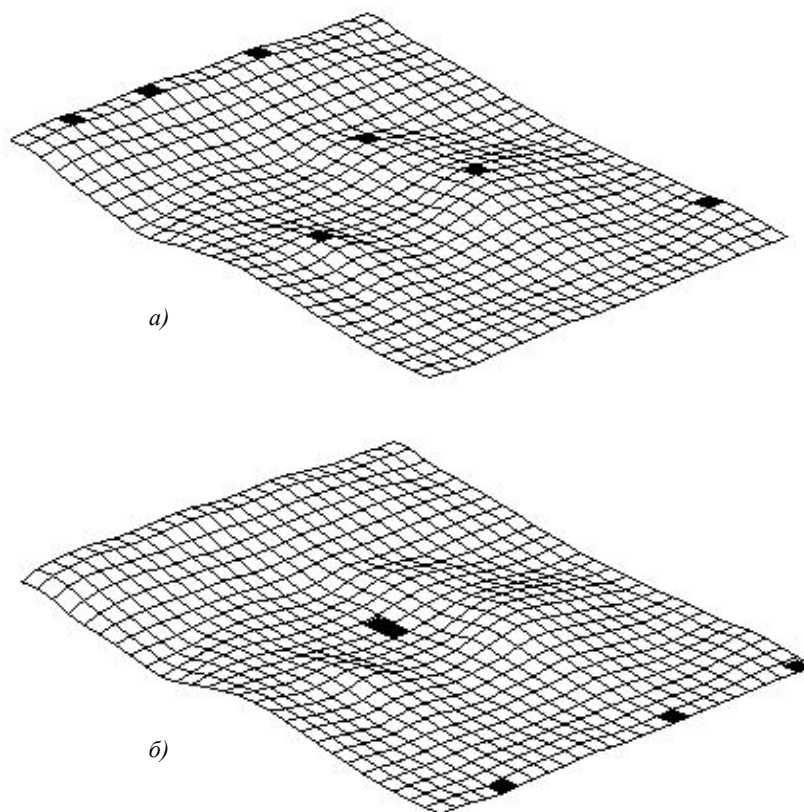


Рис. 4. Результат поиска локальных экстремумов на электронной карте местности:
а) максимумов; б) минимумов

Соответственно, каждый слой синтезированной ИНС представляет собой решетку 25 x 30 нейронов, а вся ИНС содержит 2 251 нейрон (три слоя по 750 нейронов в каждом и один решающий нейрон на всю ИНС). Ограничения на размеры исследуемой области были установлены исключительно из соображений наглядности результатов эксперимента.

В результате моделирования рассмотренной ИНС после выполнения четырех параллельных вычислительных операций (ПВО) были выделены как участки локальных максимумов высот (рис. 4, а), так и локальных минимумов (рис. 4, б) анализируемой местности. Под ПВО в данном случае понимаются вычисления, производимые одновременно одним слоем ИНС. Результаты других экспериментов по поиску локальных и глобальных экстремумов высот приведены в табл. 1.

Таблица 1

Результаты работы модели ИНС по поиску экстремумов высот земной поверхности

Тип местности (высоты), м	Размер области, м x м	Количество операций				Методом полного перебора, ПО
		Нейросетевым методом				
		Локальные экстремумы		Глобальные экстремумы		
Макс., ПВО	Мин., ПВО	Макс., ПВО (ПО)	Мин., ПВО (ПО)			
Горная (637 – 2188)	3100 x 2900	6	11	3 (24)	2 (16)	2621
	13100 x 7200	10	19	3 (24)	3 (24)	35745
	15500 x 13200	10	15	3 (24)	2 (16)	184357
Холмистая (203 – 759)	3300 x 3100	16	12	2 (16)	2 (16)	1510
	13300 x 7800	16	22	2 (16)	2 (16)	24987
	15500 x 13100	18	14	2 (16)	2 (16)	148822
Равнина (150 – 168)	3600 x 3000	20	9	3 (24)	2 (16)	1225
	14500 x 9300	48	11	3 (24)	2 (16)	2621
	13700 x 13000	32	21	3 (24)	3 (24)	35745

В графе 7 табл. 1 приведены средние значения количества операций, выполненных при последовательном поиске глобального экстремума на выбранном участке ЭКМ. Для получения этих значений проводились серии из 15 экспериментов с различными начальными точками анализа.

В процессе функционирования формального нейрона можно выделить следующие операции: умножение, суммирование и нелинейное преобразование. Поскольку нелинейное преобразование может включать в себя не одну, а несколько последовательных операций, то в первом приближении можно полагать, что ПВО соответствует 4 – 8 операциям, выполняемым при последовательном решении задачи. Согласно сделанному предположению в графах 5 и 6 табл. 1 в скобках показано максимальное эквивалентное количество последовательных операций (ПО) для выполненных ПВО.

Заключение

Полученные в ходе экспериментов результаты свидетельствуют о работоспособности и высокой эффективности предложенных методов поиска как глобального, так и локальных экстремумов дискретно заданной функции, полностью подтверждают сделанное выше теоретическое заключение о нечувствительности метода к размерности решаемой задачи.

Рассмотренный нейросетевой подход к поиску глобального и локальных экстремумов функции одной переменной предлагается использовать для решения аналогичной задачи для функции двух и более переменных при решении различных прикладных научно-технических задач.

Список литературы

1. Питиенко А.А. Использование нейросетевых технологий при решении аналитических задач в ГИС // Методы нейроинформатики / Под ред. А.Н. Горбаня. – Красноярск: КГТУ, 1998. – 205 с.
2. Hopfield J.J. and Tank D.W. Neural computation of decisions in optimization problems // Biological Cybernetics. – V.52. – 1985.
3. Нейроинформатика / А.Н. Горбань, В.Л. Дунин-Барковский, А.Н. Кирдин и др – Новосибирск: Наука, Сибирское предприятие РАН, 1998. – 295 с.
4. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. – М.: Мир, 1992. – 237 с.
5. Калан Р. Основные концепции нейронных сетей. – М.: Издательский дом «Вильямс», 2001. – 287 с.
6. Булойчик В.М. Военно-прикладные вопросы математического моделирования. Ч.2. Математические методы, используемые при моделировании процессов функционирования систем. – Мн.: МО РБ, 2000. – 156 с.

7. Садыхов Р.Х., Фролов И.Г. Клеточные нейронные сети в задачах предварительной обработки изображений. – Мн.: Ин-т техн. кибернетики НАН Беларуси, 2002. – 30 с.

Поступила 28.01.04

Военная академия Республики Беларусь

Минск-57

e-mail: vladim_grishko@tut.by

V.M. Buloichik, V.D. Grishko

**NEURAL NETWORK METHOD OF EXTREMUM SEARCH
OF TWO DIMENSIONAL DISCRETE SET FUNCTION**

An approach is considered for using neural networks to solve analytical tasks with data in the form of discrete functions. An application example of this approach for electronic map processing is presented.