Математика и её приложения: ЖИМО. 2011. Вып. 1 (8). С. 145 – 150.

УДК 512.54

С. И. Хашин[1]

Распределение длин циклов несимметричных сетей Хопфилда

Ключевые слова: распознавание образов, нейронные сети, сети Хопфилда.

В работе исследуются свойства нейронных сетей Хопфилда. Доказано, что многие их свойства сводятся к случаю однородных сетей. Доказываются свойства однородных сетей Хопфилда, касающиеся распределения длин циклов.

Keywords: image recognition, new neural networks, Hopfield network.

We investigate properties of neural Hopfild networks; many of these properties are reduced to properties of homogeneous networks. Also we prove the properties of homogeneous Hopfild networks concerning distributions of cycles lengths.

1. Введение

При булевом сжатии файлов [1] код отдельного буфера содержит три поля: поле принадлежности, поле кратности, поле порядка и еще одно общее для всех буферов поле. В работе [2] исследовалась возможность использования методов распознавания образов для задачи разбиения файла на буферы, как составной части общей задачи сжатия файла методами булевой алгебры. Там же были рассмотрены основные методы распознавания образов, которые могут быть применены в этом случае. Один из важных методов распознавания образов — построение сетей Хопфилда, можно попытаться использовать в нашей задаче.

В настоящей работе рассматриваются некоторые варианты этого метода и его обобщения, могущие быть полезными при решении задачи сжатия файла методами булевой алгебры.

2. Сети Хопфилда

Определение 1. *Нейроном* $W = \{w_0, \ldots, w_n\}$ будем называть функцию переменных $x_1, \ldots, x_n$ вида

$$W(x_1, \ldots, x_n) = F(w_0 + w_1 x_1 + \cdots + w_n x_n),$$

где w_i — внутренние параметры нейрона (весовые коэффициенты), а передаточная функция F определена так:

$$F(t) = \begin{cases} -1 \text{ при } t < 0, \\ +1 \text{ при } t \geq 0. \end{cases}$$

[1]Ивановский государственный университет; E-mail: `khash2@mail.ru`. Работа выполнена при финансовой поддержке РФФИ (проект 10-07-00350а).

В дальнейшем будем предполагать, что переменные x_i так же принимают значения ± 1. Таким образом, нейрон можно рассматривать как некоторую булеву функцию от n переменных.

Замечание 1. Булева функция b от переменных $x_1, \ldots, x_n$ может быть реализована с помощью нейрона тогда и только тогда, когда множества

$$B_0 = \{(x_1, \ldots, x_n) : \ b(x_1, \ldots, x_n) = 0\}$$

и

$$B_1 = \{(x_1, \ldots, x_n) : \ b(x_1, \ldots, x_n) = 1\}$$

линейно разделимы в $\mathbb{R}^n$ (при этом считаем, что $0 \in \mathbb{Z}_2$ соответствует $0 \in \mathbb{R}$, а единица из $\mathbb{Z}_2$ — единице из $\mathbb{R}$).

Определение 2. Два нейрона будем называть *эквивалентными*, если соответствующие булевы функции совпадают.

Определение 3. Нейрон будем называть *однородным*, если $w_0 = 0$.

Из произвольного нейрона можно получить однородный, если добавить еще одну переменную x_0:

$$W'(x_0, x_1, \ldots, x_n) = F(w_0 x_0 + w_1 x_1 + \cdots + w_n x_n).$$

Обозначим через A_n множество векторов из $\mathbb{R}^n$, $x = (x_1, \ldots, x_n)$, у которых каждая из координат равна ± 1. Мощность множества A_n равна 2^n.

Определение 4. *Базовым* вектором $B(W)$ однородного нейрона W назовем сумму всех векторов x из A_n, для которых $W(x) = 1$.

Другими словами, если

$$W(x_1, \ldots, x_n) = F(w_1 x_1 + \cdots + w_n x_n),$$

то

$$B(W) = \sum_{x \in A_+} x, \text{ где } A_+ = \{x \in A_n : \ w_1 x_1 + \cdots + w_n x_n \geq 0\}.$$

Очевидно, что сумма всех векторов, для которых $W(x) = 0$, равна $-B(W)$.

Из определения следует, что нейроны W_1 и W_2 эквивалентны тогда и только тогда, когда $B(W_1) = B(W_2)$.

Пример 1. На плоскости, т. е. при $n = 2$, базисным вектором нейрона могут быть лишь векторы $((0, \pm 2), (\pm 2, 0), (\pm 1, \pm 1)$, всего 8 различных векторов.

Пример 2. В случае $n = 3$ базисными векторами нейрона могут быть лишь следующие векторы:

$$(0, 0, \pm 4), (0, \pm 4, 0), (\pm 4, 0, 0),$$
$$(\pm 2, \pm 2, 0), (\pm 2, 0, \pm 2), (0, \pm 2, \pm 2),$$
$$\frac{4}{3}(\pm 2, \pm 1, \pm 1),\ \frac{4}{3}(\pm 1, \pm 2, \pm 1),\ \frac{4}{3}(\pm 1, \pm 1, \pm 2),$$
$$(\pm 1, \pm 1, \pm 1),$$

всего 40 различных векторов.

Следующая теорема показывает, что эти векторы однозначно описывают все классы неэквивалентных ненулевых нейронов от двух и трех переменных.

Теорема 1. *Пусть нейрон W задается функцией*

$$W(x_1, \ldots, x_n) = F(w_1 x_1 + \cdots + w_n x_n),$$

и его базовый вектор равен $B(W) = (b_1, \ldots, b_n)$. Тогда нейрон W эквивалентен нейрону W' с функцией

$$W'(x_1, \ldots, x_n) = F(b_1 x_1 + \cdots + b_n x_n).$$

Доказательство. Для данного нейрона W выделим из множества векторов A_n подмножество

$$A_+ = \{x = (x_1, \ldots, x_n) \in A_n : w_1 x_1 + \cdots + w_n x_n \geq 0\},$$

т. е. те векторы, сумма которых и образует вектор $B = B(W)$. Нам надо доказать, что

$$\forall\ x \in A_+ : (x, B) \geq 0.$$

Предположим противное, т. е. что для некоторого $x \in A_+$ скалярное произведение $(x, B) < 0$. Тогда вместе с каждым вектором $x \in A_+$, симметричный ему относительно оси (O, B) вектор x' также принадлежит A_+, при этом

$$x' = -x + 2\frac{(x, B)}{(B, B)} B.$$

Но тогда скалярное произведение векторов $x + x'$ и $W = (w_1, \ldots, w_n)$ будет равно

$$2\frac{(x, B)}{(B, B)}(B, W),$$

что меньше 0, так как

$$(B, W) = \sum_{x \in A_+} (x, W) > 0,$$

а $(x, B) < 0$ по предположению. С другой стороны,

$$(x + x', W) = (x, W) + (x', W) \geq 0.$$

Имеем противоречие. Таким образом, для всех $x \in A_+$ справедливо неравенство $(x, B) \geq 0$, что и требовалось доказать. ■

Определение 5. *Реальной размерностью* однородного нейрона будем называть количество ненулевых компонент его базисного вектора.

Теорема 2. *У базисного вектора однородного n-мерного нейрона с реальной размерностью $k \leq n$ все ненулевые координаты равны по абсолютной величине 2^{n-k}.*

Доказательство получается непосредственным вычислением и индукцией по размерности нейрона n и его реальной размерности.

Эта теорема дает полное описание всех базисных векторов однородных нейронов. Например, в случае размерности 4 — это векторы

$$(\pm 16, 0, 0, 0),\ (\pm 8, \pm 8, 0, 0),\ (\pm 2, \pm 2, \pm 2, 0),\ (\pm 1, \pm 1, \pm 1, \pm 1)$$

и векторы, получающиеся из них перестановками координат.

Следствие. *Любой однородный нейрон эквивалентен нейрону с коэффициентами* $(0, \pm 1)$[1].

Доказательство. Из теоремы 2 следует, что каждый базовый вектор пропорционален вектору с координатами $(0, \pm 1)$, а при умножении коэффициентов на положительную константу однородный нейрон не изменяется. ■

Нейронная сеть Хопфилда [5] состоит из $n > 1$ нейронов.

$$\begin{array}{l} w_1(x_1, \ldots, x_n) = F(a_{10} + a_{11}x_1 + \cdots + a_{1n}x_n), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ w_n(x_1, \ldots, x_n) = F(a_{n0} + a_{n1}x_1 + \cdots + a_{nn}x_n). \end{array} \tag{1}$$

Коэффициенты всех нейронов можно объединить в одну матрицу $A = (a_{ij})$ размера $(n+1)\times n$, содержащую внутренние параметры всех нейронов и называемую *порождающей* матрицей сети. Квадратную часть этой матрицы, без столбца свободных коэффициентов будем называть *основной* матрицей и обозначать $\tilde{A}$.

Выход каждого нейрона связан со входами всех нейронов. В начале мы имеем некоторый вектор $(x_1, \ldots, x_n)$ состояния нейронной сети. Не ограничивая общности, можно считать, что $x_i \in \{\pm 1\}$. Одна итерация сети заключается в вычислении значений $(x'_1, \ldots, x'_n)$ по формулам (1):

$$\begin{array}{ll} x'_1 & = F(a_{10} + a_{11}x_1 + \cdots + a_{1n}x_n), \\ \ldots & \\ x'_n & = F(a_{n0} + a_{n1}x_1 + \cdots + a_{nn}x_n). \end{array} \tag{2}$$

Из произвольного начального состояния мы будем повторять пересчет по формулам (2). Так как общее количество возможных состояний сети конечно, то через некоторое количество итераций состояния начнут повторяться через несколько шагов. Если следующее состояние окажется равным предыдущему, то мы получим неподвижную точку в пространстве состояний. Если же состояния будут повторяться через $k \geq 2$ шагов, то мы получим цикл длины k.

Пространство состояний сети Хопфилда будем отождествлять с n-мерным векторным пространством над полем из двух элементов $\mathbb{Z}_2$. Итерации сети H задают некоторое (нелинейное) отображение $\varphi_H : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$.

[1] Напомним, что мы рассматриваем нейроны только при исходных векторах с координатами ± 1.

В работе [4] доказано, что если матрица $\tilde{A}$ симметрична и положительно определена, то сеть Хопфилда не содержит циклов длины большей 1, то есть содержит только неподвижные точки. Обычно при распознавании образов используют именно такие сети Хопфилда.

3. Однородные сети Хопфилда

Будем называть сеть Хопфилда *однородной*, если $a_{i0} = 0$ для всех i.

Пространство состояний сети Хопфилда H_1 от n переменных $(x_1, \ldots, x_n)$ естественно вкладывается в пространство состояний сети H_2 от $n+1$ переменных $(x_0, x_1, \ldots, x_n)$:

$$I : (x_1, \ldots, x_n) \to (1, x_1, \ldots, x_n).$$

Определение 6. Будем говорить, что отображение φ пространства состояний сети H_1 в состояния сети H_2 является *вложением* сети H_1 в сеть H_2, если для любого вектора $v = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$ выполнено:

$$\varphi_{H_2}(Iv) = I\varphi_{H_1}(v),$$

т. е. если сеть H_2 действует на векторах $(1, x_1, \ldots, x_n)$ так же, как сеть H_1 на векторах $(x_1, \ldots, x_n)$.

Теорема 3. *Произвольная сеть Хопфилда вкладывается в однородную сеть.*

Доказательство. Пусть матрица сети H_1 имеет вид:

$$\begin{pmatrix} a_{10} & a_{11} & \ldots & a_{1n} \\ & & \ldots & \\ a_{n0} & a_{n1} & \ldots & a_{nn} \end{pmatrix}.$$

Рассмотрим однородную сеть Хопфилда H_2 с матрицей

$$\begin{pmatrix} S & 0 & \ldots & 0 \\ a_{10} & a_{11} & \ldots & a_{1n} \\ & & \ldots & \\ a_{n0} & a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

где S — некоторая положительная константа, такая, что для всех j

$$S \geq \sum_{i=1}^{n} |a_{ji}|.$$

Из этого условия следует, что для любого вектора $w = (x_0, x_1, \ldots, x_n)$ первая координата вектора $\varphi_{H_2} v$ всегда равна 1. Поэтому

$$\varphi_{H_2}(Iv) = I\varphi_{H_1}(v),$$

т. е. сеть H_1 вкладывается в H_2. ■

Теорема 4. *Произвольная симметричная сеть Хопфилда вкладывается в однородную симметричную сеть.*

Доказательство . Если матрица (3) сети H_1 симметрична, построим расширенную симметричную матрицу

$$\begin{pmatrix} S_2 & a_{10} & \dots & a_{n0} \\ a_{10} & a_{11} & \dots & a_{1n} \\ & & \dots & \\ a_{n0} & a_{n1} & \dots & a_{nn} \end{pmatrix},$$

где S — некоторая положительная константа, такая, что для всех j

$$S \geq 2 \sum_{i=1}^{n} |a_{ji}|.$$

Из этого условия следует, что для любого вектора $w = (x_0, x_1, \dots, x_n)$ первая координата вектора $\varphi_{H_3} v$ соответствующей сети Хопфилда H_3 всегда равна 1. Поэтому

$$\varphi_{H_3}(Iv) = I\varphi_{H_1}(v),$$

т. е. сеть H_1 вкладывается в H_3. ■

Список литературы

1. *Толстопятов А. А.* О возможности использования булевых уравнений для сжатия файлов // Вестник ИвГУ. – 2003. – Вып. 3. – С. 82–84.

2. *Толстопятов А. А.* Возможные подходы к разбиению файла на буферы при булевом сжатии // Математика и ее приложения: Журн. Иванов. матем. об-ва. – 2009. – Вып. 1 (6). – С. 129–138.

3. *Хашин С. И.* Экспериментальное исследование несимметричных сетей Хопфилда // Математика и ее приложения: Журн. Иванов. матем. об-ва. – 2010.– Вып. 1(7). – С. 109–114.

4. *Cohen M. A., Grossberg S. G.* Absolute stability of global pattern formation and parallel memory storage by compatitive neural networks // IEEE Transactions on Systems: Man and Cybernetics. – 1983. – V. 13. – P. 815–826.

5. *Hopfield J. J.* Neural Networks and Physical Systems with Emergent Collective Computational Abilities // Proc. Nat. Academy of Sciencies, USA. – Vol. 79. – 1982. – P. 2554–2558.

Поступила в редакцию 20.12.2011