محاسبه ظرفیت شبکه عصبی هاپفیلد و ارائه روش عملی افزایش حجم حافظه

محمدرضا عارف[۱] * (PH.D)، مجید سلیمانپور[۲] * * (M.SC.)

چکیده :

ظرفیت (حجم حافظه) مدل عصبی هاپفیلد ، به عنوان پارامتری مهم در بهره برداری از این شبکه ، بررسی شده است. دراین مقاله، باند بالایی ظرفیت شبکه هاپفیلد با روشی نو، از طریق مدلسازی آن با کانال شانون محاسبه می‌شود. به منظور دستیابی به حداکثر حافظه ممکن ، الگوریتم یادگیری شبکه بررسی می‌شود و به اثبات می‌رسد که ظرفیت شبکه در عمل محدود به ماکزیمم تعداد الگوهای آموزشی دو به دو متعامد است. سپس ضمن اثبات چند قضیه در خصوص کدهای متعامد، حجم حافظه عملی شبکه رادر حالت ورودی خش دارو بی‌خش ، با فرض کد گذاری مناسب روی الگوهای آموزشی ، بررسی کرده ، سرانجام با استفاده از شبیه‌سازی مدل هاپفیلد نتایج محاسبات نظری و میزان افزایش حجم حافظ را ارزیابی می‌کنیم .

۱- مقدمه

مدل عصبی هاپفیلد که در سال ۱۹۸۲ معرفی شد [۱،۲،۳]، به عنوان حافظهٔ با محتوای نشانی‌پذیر[۳] و راه حلی بر مسئله بهینه‌سازی قابل به کارگیری است . این شبکه از تعداد زیادی نورون با اتصالات متقابل انبوه تشکیل شده است. ورودی هر نورون ، برابر مجموع خروجیهای وزن شدهٔ سایر نورونهاست (شکل ۱) . وزن w^{ij} که مربوط به اتصال نورون j ام به نورون i ام است براساس قانون هب [۴] تنظیم می‌شود. بیان قانون هب چنین است که هربار که نورون j ام ، نورون i ام را تحریک کند، کارآیی نورون j ام در آتش کردن نورون i ام افزایش می‌یابد. در حقیقت بر اساس این قانون ، تغییرات قدرت اتصال میان دو نورون i و j تابعی از مقدار همبستگی متقابل مؤلفه‌های i ام و j ام درالگوی آموزشی شبکه‌است.اگرالگوی آموزشی شبکه $x^{(a)}$ باشد این تعییر وزن برابر خواهد بود با :

$$\Delta w_{ij}=kx_i^{(a)}x_j^{(a)} \qquad (1)$$

که در آن $x_i^{(a)}$ مؤلفه i ام الگوی آموزشی $x^{(a)}$ و k ضریبی ثابت و در اغلب موارد برابر یک است. مقدار وزن اتصال w_{ij} ، به این ترتیب برابر خواهد بود با مجموع تغییرات وزن نسبت به کلیه الگوهای آموزشی،که اگر الگوها n مؤلفه‌ای و به تعداد m باشند خواهیم داشت:

$$w_{ij}=\sum_{a=1}^{m} x_i^{(a)}x_j^{(a)} \quad , \quad W_{ii}=O \quad , 1\leq i,j\leq n \quad j\neq i \qquad (2)$$

با توجه به ساختار مدل هاپفیلد، سطح فعالیت نورون i ام برابر $\sum_j w_{ij}u_j$ است که u_j مؤلفه j ام بردار حالت u است و هرگاه این سطح فعالیت از یک سطح آستانه θ_i فراتر رود ، حالت خود را تغییر داده و مقدار جدید خود را بر اساس رابطهٔ زیر به دست می‌آورد:

---

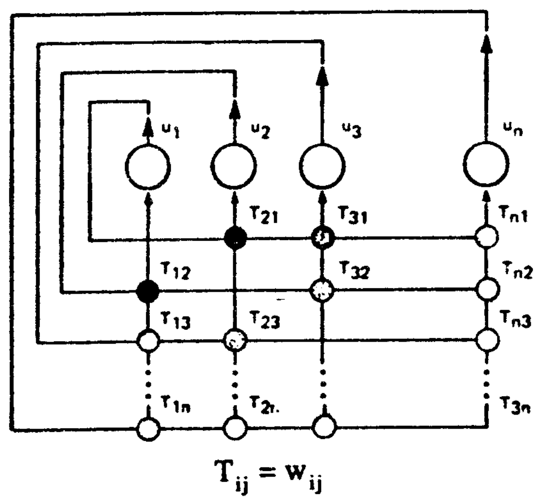
* - عضو هیئت علمی دانشگاه صنعتی اصفهان

* * - عضو هیئت علمی دانشگاه امام حسین (ع)

۳- Content - Addressable Memory

$$u_i=f\left[\sum_{\substack{j=1\\ j\neq i}}^{n} w_{ij}u_j-\theta_1\right] \qquad (۳)$$

f رابطهٔ غیرخطی ورودی - خروجی یا تابع تصمیم گیری نورون است. برحسب اینکه تابع f از نوع هلالی[1] یا تقریب پله‌ای آن باشد، دو نوع مدل هاپفیلد تحت عناوین پیوسته و دودویی، معرفی می‌شود.



شکل (۱)- شبکه هاپفیلد با مقادیر وزنهای w_{ij} که به صورت دوایر نمایش داده شده‌اند و رنگ آنها مقدار وزن را نشان می‌دهد.

دراین مقاله ، مدل دودویی هاپفیلد درنظر گرفته شده است که وجوه مشخصه آن دو مقدار بودن ورودیها (۱ و ۱-) و دو حالتی بودن نورونها (روشن یا خاموش) است. تغییر حالات شبکه به این ترتیب است که پس‌از اعمال ورودی x، درلحظه اولیه، خروجی نورون i برابر مؤلفه iام x یعنی x_i است. سپس مقدار $u_i(t+1)$ یعنی حالت نورون iام در زمان $t+1$ از رابطه زیر به دست می‌آید:

$$u_i(t+1)=f\left[\sum_{j=1}^{n} w_{ij}u_j(t)\right] \quad , 1\leq i\leq n \qquad (۴)$$

تابع f در رابطه (۴)، تقریب غیر خطی پله‌ای یا محدودکننده سخت از تابع هلالی ورودی - خروجی نورون بیولوژیک، دارای دو مقدار ۱ + و ۱- است. مقدار سطح آستانه θ_i در رابطه (۳)، دراین مدل برابر صفر فرض شده است. مقدار w_{ij} از رابطه (۲) به دست می‌آید. به این ترتیب تغییر حالت نورون iام براساس قاعده زیر انجام خواهد شد:

$$\begin{array}{lcl} \sum_j w_{ij}u_j>0 & & u_i=1 \\ & \Rightarrow & \\ \sum_j w_{ij}u_j<0 & & u_i=-1 \end{array} \qquad (۵)$$

رابطه برگشتی (۴) آنقدر تکرار خواهد شد تا خروجی تمام نورونها ثابت بماند ، یعنی :

$$u_i(t+1)=u_i(t)=f\left[\sum_{j=1}^{n} w_{ij}u_j(t)\right] \qquad (۶)$$

برای تضمین همگرایی شبکه به سمت پایداری و تثبیت خروجی تمام نورونها، کافی است شرط $w_{ij}=w_{ji}$ برقرار باشد [۱]. اگر $x^{(a)}$ یکی از حالات پایدار محلی در شبکه باشد و شبکه در نقطهٔ x در مجاورت $x^{(a)}$ شروع به کار کند، یعنی $x=x^{(a)}+\Delta$ با گذشت زمان شبکه به پایداری محلی $x^{(a)}$ نزدیک و در،آن نقطه پایدار خواهد شد. در این صورت بردار $x^{(a)}$ را می‌توانیم اطلاعات ذخیره شده در

1- Sigmoid

شبکه در نظر بگیریم که نقطه شروع یا حالت اولیه سیستم α اطلاعات ناقص و مغشوش $x^{(\alpha)}$ بوده و شبکه اطلاعات کامل را بازیابی یا آدرس کرده است. از این رو شبکه هاپفیلد می‌تواند به عنوان یک حافظه محتوای‌نشانی‌پذیر یا حافظهٔ تداعی [1] به کار گرفته شود.

۲- تعریف ظرفیت اطلاعاتی در حافظه‌ها

ظرفیت اطلاعاتی حافظه ها در حالت کلی، به صورت زیر تعریف می‌شود[۵]:

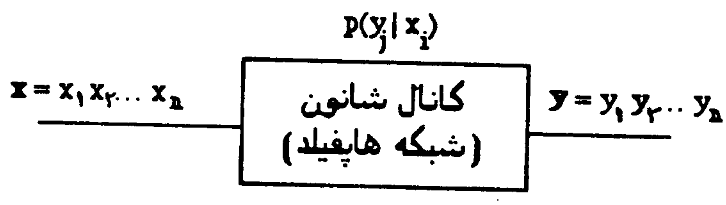
ظرفیت اطلاعاتی یک حافظه عبارت است از لگاریتم تعداد کل حالات مختلف و قابل تشخیص از یکدیگر در آن حافظه. بنابراین مهمترین پارامتر در تعیین ظرفیت یک حافظه تعیین حالات مختلف و متمایز در آن حافظه است. مثلاً در یک حافظه RAM با M خط آدرس و یک خط داده (حافظه ۱*M)، تعداد ۲^M خانه حافظه وجود دارد. هر خانه از حافظه با تعداد M بیت آدرس مشخص شده و شامل یک بیت داده می‌باشد و مجموعا ۲^Mبیت داده در حافظه قابل ذخیره‌سازی است. تعداد حالات مختلف و متمایز در چنین حافظه‌ای برابر با $۲^{۲^M}$ است که براساس تعریف کلی ظرفیت اطلاعاتی حافظه‌ها ، ظرفیت این حافظه برابر خواهد بود با :

$$C=\log_{۲}۲^{۲^M}=۲^M \qquad (۷)$$

در شبکه هاپفیلد برحسب اینکه حالتهای انتقالی یا حالتهای پایدار را به عنوان حالات متمایز حافظه درنظر بگیریم ، ظرفیتهای متفاوتی به دست خواهیم آورد. لیکن به دلیل اینکه حالتهای پایدار شبکه ، جنبه کاربردی مطلوب‌تری دارند ، ظرفیت شبکه هاپفیلد را برابر لگاریتم حالات پایدار تعریف می‌کنیم .

۳- محاسبه باند بالایی ظرفیت شبکه هاپفیلد

دراین مقاله ، ظرفیت شبکه هاپفیلد ، از طریق مدلسازی آن با کانال شانون محاسبه می‌شود. از دیدگاه نظریهٔ اطلاعات، شبکه هاپفیلد (یا به طورکلی هر شبکه عصبی) را می‌توان یک کانال انتقال اطلاعات فرض کرد که ورودی آن بردار n مؤلفه‌ای $x=x_۱x_۲\cdots x_n$ متعلق به فضای ۲^n بردار n مؤلفه‌ای و خروجی آن بردار n مؤلفه‌ای $y=y_۱y_۲\cdots y_n$ متعلق به فضای محدود به الگوهای آموزش داده شده به شبکه $(x^{(s)}, s=۱,\cdots,m)$اند. با این تعبیر، ظرفیت شبکه را می‌توان از طریق محاسبه ظرفیت کانال محاسبه کرد(شکل ۲).



شکل (۲)- شبکه هاپفیلد به عنوان مدل کانال شانون

تابع انتروپی ورودی H(X)، بنابر تعریف مبین مقدار متوسط ابهام و عدم قطعیتی است که نسبت به هر نماد ورودی X وجود دارد و یا برابر مقدار متوسط اطلاعات موجود در هر نماد ورودی است که برحسب بیت بر نماد اندازه‌گیری می‌شود. همچنین تابع انتروپی شرطی H(X|Y)، مقدار متوسط اطلاعات نسبت به ورودی X رابا داشتن خروجی Y بیان می‌کند. چون همواره $H(X|Y)\le H(X)$ است ، داشتن خروجی Y، میزان ابهام و عدم‌قطعیت‌نسبت‌به X راکاهش می‌دهد.یعنی ورودی X با عبور از کانال ، بخشی از محتوای اطلاعاتی خود را از دست می‌دهد. از اینرو مقدار تابع انتروپی شرطی H(X|Y) به مقدار اطلاعات گم شده در کانال نیز تعبیر شده است [۶]. بنابراین مقدار متوسط اطلاعات در خروجی نسبت به ورودی برابر $H(X)-H(X|Y)$ است که اطلاعات متقابل

1- associative memory

کانال، $I(X;Y)$ نامیده می‌شود و با اطلاعات موجود در ورودی نسبت به خروجی که با بیان مشابه برابر $H(Y)-H(Y|X)$ است، مساوی خواهد بود. ظرفیت کانال بنابه تعریف [۷و۸] ماکزیمم مقدار ممکن اطلاعات متقابل است. این مقدار ماکزیمم اطلاعات متقابل، زمانی حاصل می‌شود که منبع اطلاعات باکانال منطبق باشد یعنی ورودی X از منبعی تولید شود که توزیع احتمال آن ، مقدار اطلاعات متقابل را ماکزیمم کند:

$$C \triangleq \max_{p(x)} [H(X)-H(X|Y)] = \max_{p(x)} [H(Y)-H(Y|X)] \qquad (۸)$$

شبکه هاپفیلد در حالت ایدئال بایستی بتواند هر ورودی X راکه شکل خش‌دار یا ناقصی از یکی از حافظه‌های شبکه مثل $x^{(n)}$ است، به درستی تشخیص دهد و $x^{(s)}$ را در خروجی به دست آورد.به عبارت دیگراگرمجموعه $\{x^{(s)}\}$ شامل تمام بردارهای ورودی متناظر با حافظه $x^{(s)}$ باشد، برای هر عنصر از این مجموعه ، بایستی بااحتمال یک ، خروجی به دست آمده، $x^{(s)}$ باشد. شبکه هاپفیلد به عنوان یک کانال شانون در شکل (۲) نمایش داده شده است . ماتریس انتقال این کانال دارای عناصر صفر و یک است ودر هر سطر آن یک عنصر یک و بقیه عناصر صفر است . کانالهایی که دارای این مشخصات باشند، کانال غیر تصادفی[1] نامیده می‌شوند. در این کانالها با داشتن ورودی هیچ‌گونه ابهامی در مورد خروجی وجود ندارد . بنابراین :

$$H(Y|X)=H(x^{(s)}|x)=0 \qquad (۹)$$

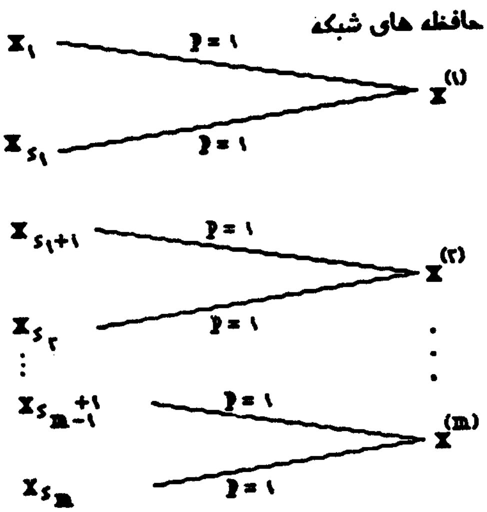
وظرفیت‌کانال(شبکه هاپفیلد)از رابطه زیربه دست می‌آید:

$$C=\max_{p(x)} [H(Y)] = \max_{p(x)} [H(x^{(s)})] \qquad (۱۰)$$

ازطرف‌دیگر،حالات پایدار خروجی شبکه هاپفیلد از میان 2^n حالت ممکن انتخاب می‌شود که برای توزیع احتمال یکنواخت خروجی ، ظرفیت برابر لگاریتم آن خواهد بود:

$$C=\log_2 2^n = n \qquad (۱۱)$$

نتیجه به دست آمده، به عنوان باند بالایی ظرفیت حالات پایدار شبکه هاپفیلد ، مقدار محاسبه شده در [۵] را تایید می‌کند. روش محاسبه ظرفیت شبکه‌های عصبی ، از طریق تشبیه آنها به کانالهای مخابراتی معادل، روش جدیدی است که بر اساس رفتار ورودی - خروجی شبکه و مستقل از پارامترهای داخلی آن ، قادر به حل مسئله است و از این رو می‌تواند روشی عام تلقی شود.



شکل (۳)- شبکه هاپفیلد به عنوان یک کانال غیر تصادفی

۴- بهبود حجم حافظه عملی شبکه هاپفیلد

ظرفیت عملی و قابل دسترس شبکه هاپفیلد بر اساس آزمایشهای خود وی [۱] که در آنها الگوهای آموزشی تصادفی انتخاب شدند.برابر $0/15n$ اعلام شده است. این مقدار به صورت آماری اندازه‌گیری شده و فرض شده است که حداقل در ۵۰ درصد اوقات ، خروجی به دست آمده با حافظه اصلی ،کمتر از $0/05n$ مؤلفه اختلاف داشته باشد . در محاسبات مک‌الیس و همکارانش [۹]، مقدار ظرفیت برای پاسخ صحیح در تمام اوقات برابر $n/4\ln(n)$ و برای پاسخ صحیح در اکثر اوقات برابر

۱- deterministic channel

$n/2\ln(n)$ به دست آمد که برای n مشابه با آزمایش هاپفیلد به ترتیب برابر $0/05n$ و $0/1n$ می‌باشند. البته نتایج به دست آمده در [۹]، عملاً شبیه‌سازی و آزمایش نشده است. در این بخش الگوریتم شبکه هاپفیلد را مجدداً بررسی و راههای بهبود حجم حافظه را جستجو خواهیم کرد. سپس نتایج به دست آمده را با شبیه‌سازی، ارزیابی خواهیم نمود.

۴-۱ مروری بر الگوریتم شبکه هاپفیلد

در بخش مقدمه، شبکه هاپفیلد را معرفی کردیم. در این بخش به تجزیه و تحلیل بیشتر الگوریتم شبکه هاپفیلد درمدل دودویی می‌پردازیم. دریک شبکه با n نورون و m الگوی آموزشی $[x^{(s)}=x_{s1}\cdots x_{sn}, s=1,\cdots,m]$ با مؤلفه‌های ۱ و ۱-، بردار خروجی y برای ورودی x به صورت زیر محاسبه می‌شود:

$$y=\mathrm{sgn}[w_x],\quad W=\sum_{i=1}^{m} W_s \qquad \begin{vmatrix} x_{s1}^2 & x_{s1}x_{s2} & \cdots & x_{s1}x_{sn} \\ x_{s2}x_{s1} & x_{s2}^2 & \cdots & x_{s2}x_{sn} \\ \vdots & & & \\ x_{sn}x_{s1} & x_{sn}x_{s2} & \cdots & x_{sn}^2 \end{vmatrix} \tag{۱۲}$$

$$x_s\ W_s=x^{(s)}(x^{(s)T}= \begin{vmatrix} x_{s1} \\ x_{s2} \\ \vdots \\ x_{sn} \end{vmatrix} [x_{s1}x_{s2}\ \cdots\ x_m]= \begin{vmatrix} x_{s1}^2 & x_{s1}x_{s2} & \cdots & x_{s1}x_{sn} \\ x_{s2} & x_{s2}^2 & \cdots & x_{s2}x_m \\ \vdots & \vdots & & \vdots \\ x_{sn}x_{s1} & x_{sn}x_{s2} & \cdots & x_{sn}^2 \end{vmatrix} \tag{۱۳}$$

ماتریس W_s ماتریس همبستگی متقابل مؤلفه‌های الگوی آموزشی sام نامیده می‌شود. عناصر قطری این ماتریس برابر ۱ و بقیه، حاصل ضربهای دو به دو مؤلفه‌ها یا همبستگی متقابل میان آنهاست. در این صورت خروجی به طریق زیر محاسبه می‌شود:

$$y= \mathrm{sgn}[w_x]= \mathrm{sgn}[W_1+W_2+\cdots+W_m]x=\mathrm{sgn}[W_1x+W_2x+\cdots+W_mx] \tag{۱۴}$$

$$W_sx= \begin{vmatrix} x_{s1}x_{s1} & x_{s1}x_{s2} & \cdots & x_{s1}x_{m1} \\ x_{s2}x_{s1} & x_{s2}x_{s2} & \cdots & x_{s2}x_m \\ \vdots & \vdots & & \vdots \\ x_{sn}x_{s1} & x_{sn}x_{s2} & \cdots & x_{sn}x_{sn} \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{vmatrix}$$

$$= \begin{vmatrix} x_{s1} & x_{s1}x_1\ x_{s1} & x_{s2}\ x_2 & \cdots & x_{s1}x_{sn}x_n \\ x_{s2} & x_{s1}x_1\ x_{s2} & x_{s2}\ x_2 & \cdots & x_{s2}x_{sn}\ x_n \\ \vdots & \vdots & & & \vdots \\ x_{sn} & x_{s1}x_1\ x_{sn} & x_{s2}\ x_2 & \cdots & x_{sn}x_{sn}x_n \end{vmatrix} =(x_{s1}x_1+x_{s2}x_2+\cdots+x_{sn}\ x_n)+ \begin{vmatrix} x_{s1} \\ x_{s2} \\ \vdots \\ x_{sn} \end{vmatrix}$$

$$= \left[\sum_{i=1}^{n} x_{si}x_i\right]x^{(s)} \tag{۱۵}$$

بنابراین رابطهٔ (۱۴) را می‌توان به این صورت بسط داد:

$$y=\mathrm{sgn}[w_x]=\mathrm{sgn}\Big[\sum_{s=1}^{m} W_s x\Big]=\mathrm{sgn}\Big\{\sum_{s=1}^{m}\Big[\sum_{j=1}^{n} x_{si}x_i\Big]\ x^{(s)}\Big\}$$

$$=\mathrm{sgn}\Big[\sum_{s=1}^{m} a_s x^{(s)}\Big],\ a_s=\sum_{i=1}^{n} x_{si}x_i \qquad (۱۶)$$

مشاهده می‌شود که مقدار خروجی ، پس‌از اجرای یک مرحله محاسبه، برابر با علامت مجموع الگوهای آموزشی وزن داده شده است . وزن هر الگو، a_s را مقدار حاصل ضرب داخلی الگو در بردار ورودی شبکه تعیین می‌کند. از این رو اگر وزن مربوط به الگوی صحیح، ماکزیمم و وزنهای دیگر ، می‌نیمم و در حالت ایدئال صفر باشند، شبکه با اجرای یک مرحله محاسبه ، به جواب صحیح دست می‌یابد. بنابراین با فرض اینکه الگوی ورودی x مربوط به حافظه $x^{(j)}$ باشد، کافی است:

$$\sum_{i=1}^{n} x_{si}x_i \to o \qquad \forall_s \neq j \qquad (۱۷)$$

در بهترین حالت که ورودی بی‌خش و یکی از حافظه‌های اصلی شبکه باشد، $x=x^{(j)}$ لازم می‌آید که :

$$\sum_{i=1}^{n} x_{si}x_{ji}=o \qquad \forall_s \neq j \qquad (۱۸)$$

شرط برقراری رابطه (۱۸) این است که $x^{(j)}$ برای تمام مقادیر $s\neq j$ بر $x^{(s)}$ عمود باشد.(قطر ماتریس W غیرصفر فرض می‌شود) و برای اینکه این رابطه برای سایر حافظه‌ها، به عنوان ورودی شبکه ، صادق باشد، در حالت کلی بایستی برای تمام مقادیر $i\perp j$ شرط $x^{(i)}\perp x^{(j)}$ برقرار باشد.به‌این‌ترتیب‌می‌توان گفت که مسئله حجم حافظه قابل دسترس‌شبکه تبدیل به مسئله ماکزیمم تعداد الگوهای n مؤلفه‌ای‌که‌دوبه‌دونسبت به هم متعامد باشند، خواهد شد.

قضیهٔ ۱- دوبردار $x^{(i)}$ و $x^{(j)}$ با n مؤلفه ۱ و ۱- نسبت به هم عمودند، اگر و تنها اگر فاصله همینگ میان دوبردار برابر $\frac{n}{۲}$ باشد.

اثبات- شرط لازم: فرض می‌کنیم d فاصله همینگ میان دوبردار و a=n-d تعداد مؤلفه‌های مشابه آنها باشد، شرط تعامد دو بردار ، صفر بودن حاصل ضرب داخلی آنهاست:

$$\sum_{k=1}^{n} x_{ik}x_{jk}=o \qquad (۱۹)$$

با توجه به فرض قضیه داریم:

$$\sum_{k=1}^{n} x_{sk}x_{jk}=a(1)+d(-1)=a-d=o \Rightarrow a=d \qquad (۲۰)$$

و چون n=a+d است:

$$n=a+d=۲d \Rightarrow d=\frac{n}{۲} \qquad (۲۱)$$

شرط کافی: اگر $d=\frac{n}{۲}$ باشد، a نیز برابر $\frac{n}{۲}$ خواهد بود و بنابراین خواهیم داشت :

$$\sum_{k=1}^{n} x_{ik}x_{jk}=a(1)+d(-1)=\frac{n}{2}-\frac{n}{2}=0 \qquad (22)$$

قضیه ۲- فرم کلی کدهای متعامد با مؤلفه‌های ۱و۱- به صورت $(4k,2k)$[1] است.

اثبات- برای اثبات قضیه کافی است نشان دهیم که مقدار فاصله در کدهای متعامد بایستی حتماً زوج باشد. در این صورت چون لازم است $n=2d$ باشد. بنابراین n فقط می‌تواند از مضارب صحیح ۴ باشد.

فرض می‌کنیم x_1 و x_2 کلمات یک کد متعامد و هر دو بر x_0 یک کلمه دلخواه دیگر از کد، عمود باشند بدون اینکه در جامعیت مسئله تأثیری داشته باشد، فرض می‌کنیم تمام مؤلفه‌های x_0 یک باشند. در این صورت دو بردار x_1 و x_2 دارای $\frac{n}{2}$ مؤلفه ۱ و $\frac{n}{2}$ مولفه ۱- خواهند بود. برای اینکه x_1 و x_2 نیز برهم عمود باشند، لازم خواهد بود که این دو دارای $d=a+b=\frac{n}{2}$ مؤلفه مشابه (a تعداد ۱های مشابه و b تعداد ۱-های مشابه) و $d=a'+b'=\frac{n}{2}$ مؤلفه مخالف (a' تعداد ۱های مخالف و b' تعداد ۱-های مخالف در کد x_1)، باشند. از طرفی چون تعداد مؤلفه‌های ۱ و ۱- بایستی مساوی و برابر d باشند، $d=a+a'=b+b'$، نتیجه می‌شود که $a=b'$ و $b=a'$. اما $a+b'$ تعداد مؤلفه‌های ۱ بردار x_2 و $a'+b$ تعداد مؤلفه‌های ۱- آن‌اند و برابر d اند لذا بایستی $d=2a=2b$ و زوج باشد.

قضیه ۳- اگر $m\perp$ تعداد الگوهای n مؤلفه‌ای دو به دو متعامد باشد، با شرایط $n=2^k$، حداکثر تعداد این الگوها برابر n است، $m\perp_{max}=n$.

اثبات- بر اساس نتیجه قضیه ۱، ماکزیمم تعداد الگوهای n مؤلفه‌ای که دو به دو متعامد باشند، برابر خواهد بود با ماکزیمم تعداد الگوهای به طول n و به فاصله $\frac{n}{2}$ و نه $\frac{n}{2} \geq d$. کافی است ثابت کنیم که ماکزیمم تعداد الگوهای به طول n و با فاصله $d=\frac{n}{2}$، برابر n است یعنی $A(2^k,2^{k-1})=2^k$[2].

کد مرتبه اول رید - مولر [۱۰]، با مشخصه $(2^k-2,2^k,2^{k-1})$، دارای ویژگیهای مهم و مفید متحدالفاصله بودن و تعامد (با نمادهای صفر و یک) است. اگر به این کد، یک بیت بررسی توازن اضافه کرده مؤلفه‌های صفر را به ۱- تبدیل کنیم، تعداد 2^k کلمه کد به طول 2^k و به فاصله 2^{k-1} خواهیم داشت که با توجه به نتیجه قضیه ۱، دو به دو متعامدند. حال باید اثبات شود که این تعداد، ماکزیمم مقدار ممکن است.

با اضافه کردن کلمات مکمل و یک بیت بررسی توازن، از روی کد $(2^k-1,2^k,2^{k-1})$ می‌توان کد بهینه $(2^k,2^{k+1},2^{k-1})$ را ساخت [۱۰]، یعنی ماکزیمم تعداد کلمات به طول 2^k و فاصله 2^{k-1}، برابر 2^{k+1} است. با توجه به فرض قضیه باید ثابت کنیم که تنها نیمی از این تعداد می‌توانند متعامد باشند. با توجه به اینکه نیمی از 2^{k+1} کلمه کد، همان کلمات کد $(2^k-1,2^k,2^{k-1})$ و نیم دیگر مکمل آن است و یک بیت بررسی توازن به هر دو دسته اضافه شده است، در کد جدید، فاصله‌های n میان کلمات مکمل وجود دارد و متحدالفاصله نیستند. برای حفظ شرط تعامد لازم است تمام کلمات مکمل حذف شوند به این ترتیب نیمی از کلمات باقی می‌ماند که ماکزیمم تعداد ممکن خواهد بود و مشخصه $(2^k,2^k,2^{k-1})$ را خواهند داشت.

برای سایر مقادیر n، $n=4t\neq 2^k$، نیز می‌توان نشان داد که همواره رابطه $m\perp_{max}=n$ برقرار است. با مراجعه به جدول بهترین کدهای شناخته شده [۱۰] در محدوده $n\leq 512$ و $d\leq 30$، ملاحظه می‌شود که کدهای $(4t,8t,2t)$ تنها توسط کد هادامارد ساخته می‌شوند. با توجه به مشخصات ماتریس هادامارد و کد حاصل از آن، به سادگی می‌توان دید که نیمی از کلمات کد $(4t,8t,2t)$، مکمل نیم دیگر آن‌اند و اگر کلمات مکمل را حذف کنیم کد حاصل

۱- هر کد به طول n و فاصله حداقل d را که دارای تعداد کلمات m باشد به صورت (n,m,d) و یا (n,d) نمایش می‌دهیم.

۲- برای هر کد به طول n و فاصله حداقل d، $A(n,d)$ عبارت از تعداد ماکزیمم کلمات کد است.

(۴t،۴t،۲t) متعامد خواهد بود.

بنابراین به طور کلی می‌توان پذیرفت که: $m\perp_{max}=n, \forall n=4t$.

این نتیجه جالب، با توجه به تأثیر تعامد الگوهای آموزشی در بهبود ظرفیت عملی شبکه هاپفیلد، تأییدی بر امکان دستیابی به باند بالایی ظرفیت این شبکه است.

۲-۴- حجم حافظه عملی شبکه هاپفیلد

تا اینجا، تأثیر مثبت و بهبود دهنده الگوهای آموزشی متعامد، بر ظرفیت شبکه هاپفیلد را دیدیم. همچنین تعداد و نحوهٔ ساختن این الگوهای متعامد را مشخص کردیم حال به بررسی کمی ظرفیت عملی شبکه و میزان نزدیکی آن به مقدار حدی n می‌پردازیم.

حجم حافظه عملی در شرایط بی‌خش - فرض می‌کنیم ورودی شبکه یکی از m الگوی متعامد آموزش داده شده به شبکه واز حافظه‌های اصلی آن باشد ($x^{(j)}$) خروجی شبکه پس‌از یک مرحله محاسبه، با استفاده از رابطه (۱۶) برابر است با:

$$y=sgn(Wx^{(j)})=sgn\left[\left(\sum_i x_i^{(1)}x^{(j)}\right)x^{(1)}+\cdots+\left(\sum_i x_i^{(j)}x_i^{(j)}\right)x^{(j)}+\cdots+\left(\sum_i x_i^{(m)}x_i^{(j)}\right)x^{(m)}\right] \qquad (۲۳)$$

چون m الگوی آموزشی، دوبدو متعامدند ضرایب $\sum_i x_i^{(j)}x_i^{(j)}$ در رابطه(۲۳)،به ازای‌تمام مقادیر $k\neq j$ برابر صفر است. اگر $k=j$ باشد، این ضریب برابر n خواهد بود و علامت $Wx^{(j)}$ و در نتیجه خروجی y برابر $x^{(j)}$ است. در این صورت تعداد الگوهای آموزشی شبکه، برابر ماکزیمم تعداد الگوهای متعامد، یعنی n می‌تواند باشد. این نتیجه مهم ،بدین معنی است که توانسته‌ایم به ظرفیت حدی نظری، به صورت عملی دستیابی پیدا کنیم، به طوری که شبکه همواره با انجام یک مرحله اجرای الگوریتم در حالت همزمانی [1] و با اجرای حداکثر n مرحله در حالت ناهمزمانی [2] به درستی الگوی ورودی را از میان n الگوی آموزشی بازیابی خواهد کرد.

حجم حافظه عملی در شرایط خش‌دار - بدون اینکه در جامعیت مسئله تأثیری داشته باشد، فرض می‌کنیم حافظه $x^{(j)}$ دارای مؤلفه‌های مشابه ۱+ و ورودی x، شکل خش‌دار آن، دارای تعداد Pn خطا باشد ($0<P<\frac{1}{2}$). دراین صورت ورودی دارای تعداد Pn مؤلفه ۱- خواهد بود. اگر y_i مؤلفه iام خروجی باشد، مقدار آن از رابطه زیر به دست می‌آید:

$$y_i=sgn(Wx)_i=sgn[(W_1x)_i+\cdots+(W_jx)_i+\cdots+(W_mx)_i] \qquad (۲۴)$$

جمله $(W_jx)_i$، حاصلضرب سطر iام ماتریس $W_j=x^{(j)}\cdot(x^{(j)})^T$ در ورودی x است. با توجه به فرض‌های بالا، مقدار آن برابر $n-2Pn$ است. سایر جملات رابطه (۲۴)، نیز حاصل ضرب سطر iام ماتریس W_k، $k\neq j$، در بردار ورودی xاند. با توجه به تعامد الگوهای $x^{(k)}$، سطرهای ماتریس W_k شامل $\frac{n}{2}$ مؤلفه ۱+ و $\frac{n}{2}$ مؤلفه ۱- اند و حاصل ضرب آنها در بردار x مقادیری صحیح بین ۲Pn+ و ۲Pn- خواهند داشت. اگر این مقادیر را متغیرهای تصادفی با توزیع احتمال یکنواخت فرض کنیم، در رابطه (۲۴)، تعداد ۱-m مقدار تصادفی خواهیم داشت که میانگین آنها برابر خواهد بود:

۱- همزمانی (Synchronous) :حالتی است که درآن تمام نورونها تغییر همزمان دارند.

۲- حالت ناهمزمانی (Asynchronous) : حالتی است که در آن در هر لحظه تنها یک نورون ، بطور تصادفی تغییر نماید.

$$E[(W_k x)_i] = \int_{-۲Pn}^{۲Pn} vp(v)dv = \int_{-۲Pn}^{۲Pn} \frac{v}{۴Pn} dv = o \quad (۲۵)$$

بنابراین می‌توان گفت که :

$$\bar{y}_i = n - ۲Pn > o \quad \forall P \quad (۲۶)$$

احتمال پاسخ صحیح در خروجی پس‌از اجرای یک مرحله محاسبه خروجی، برابر احتمال این است که مؤلفه y_i مثبت باشد. با توجه به رابطه (۲۶) ، می‌توان نتیجه‌گیری کرد که شبکه در ازای ورودیهای خش‌دار واقع در کره همینگ به شعاع Pn و مرکز یکی‌از حافظه‌ها، با شرط $O<P<\frac{۱}{۲}$، برای تعداد m=n الگوی آموزشی متعامد، حداقل در ۵۰درصداوقات دارای پاسخ صحیح خواهدبود.

۳-۴- شبیه‌سازی کامپیوتری

به منظور بررسی عملی نتایج محاسبات نظری، با استفاده از یک دستگاه کامپیوتر شخصی IBM/AT ، برنامه شبیه‌سازی مدل هاپفیلد به زبان توربو پاسکال ۵، نوشته شد. البته یک برنامهٔ شبیه‌سازی به نام PDP[۱۱] به زبان C در دسترس بود که به علت لزوم تغییرات قابل توجه در برنامه ، نوشتن برنامه مستقل ترجیح داده شد. ویژگیهای برنامه شبیه‌سازی شبکه هاپفیلد که تسهیلات لازم جهت انجام آزمایشها را از طریق یک ادیتور ساده فراهم می‌کند،به قرار زیر است.

الف: امکان معرفی الگوهای آموزشی تعریف شده در یک دفتر راهنمای معین،

ب:محاسبه جدول فاصله همینگ الگوهای آموزشی بر اساس الگوریتم خاصی که طراحی و اثبات شده است [۱۲]،

ج- امکان معرفی بردار حالت اولیه یا ورودی تعریف شده در یک دفتر راهنمای معین ،

د- امکان ایجاد خطا به تعداد دلخواه و به صورت تصادفی در ورودی تعیین شده ،

هـ - انتخاب همزمانی و ناهمزمانی شبکه هاپفیلد

و- تعیین میزان خطای مجاز در پایداری شبکه،

ز- انتخاب نحوه نمایش خروجیهای مرحله‌ای،

ح- تعیین نزدیکترین حافظه به پایداری شبکه و میزان فاصله همینگ با آن ،

ط- محاسبه تعداد مراحل طی شده تا رسیدن به پایداری ،

ی- تولید الگوهای تصادفی به طول معین با استفاده از مولد اعداد تصادفی درکامپیوتر.

با استفاده از این برنامه آزمایشهای مختلفی صورت گرفت که نتایج آن به شرح زیر است:

۱-۳-۴- شبیه‌سازی در حالت عمومی

این آزمایش به منظور بررسی و مشاهدهٔ رفتار عمومی شبکه هاپفیلد انجام گرفت . در این آزمایش تعدادی الگو که نمایشگر اعداد و الفبای انگلیسی بودند، به طول ۱۲۰ بیت به شبکه آموزش داده شدند. ورودیهای مختلف شبکه از طریق ایجاد تعداد دلخواهی خطا در الگوهای آموزش داده شده ، شناخته شدند. آزمایشهای مکرر با شرط تقارن و قطر صفر در ماتریس وزنها، نتایج کیفی زیر را به دنبال داشته است:

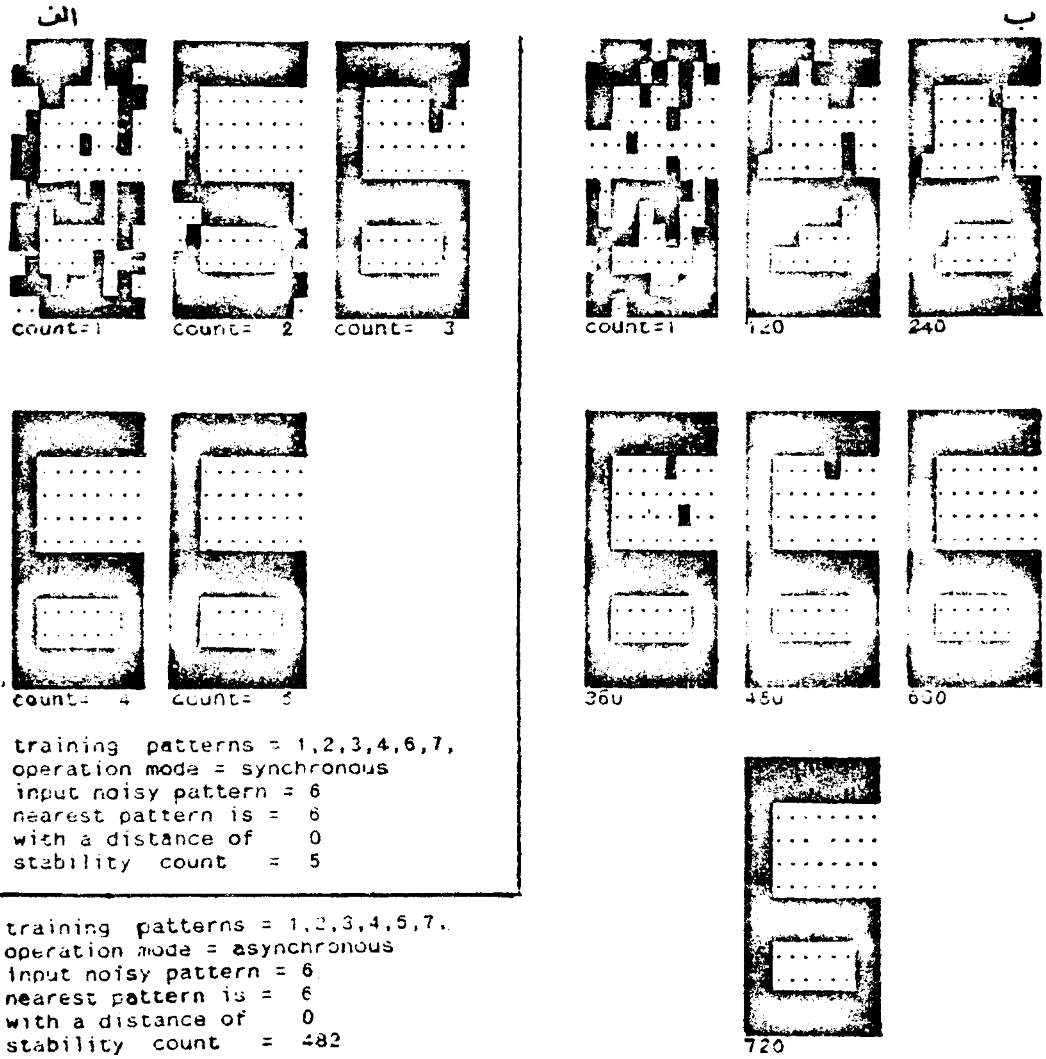
الف- افزایش تعداد الگوهای آموزشی بیش‌از یک حد معین ، موجب بی‌ثباتی حافظه‌ها و پاسخ مبهم و تعریف نشده می‌شود. حتی اگر ورودی خش‌دار نبوده و یکی از الگوهای آموزشی باشد.

ب- افزایش خش در ورودی بیش‌از یک حد معین ، موجب بروز خطا در خروجی است حتی اگر تعداد الگوهای آموزشی به اندازه کافی کوچک باشد.

نتیجه یک آزمایش انجام شده با ۶ الگوی آموزشی که اعداد 1 ، 2 ، 3 ، 4 ،6 و 7 را نمایش می‌دادند ،در ازای ورودی الگوی 6 با ۳۰بیت خطا (۲۵درصد)، در دو حالت همزمانی (شکل الف- ۴) و ناهمزمانی (شکل ب-۴) نمایش داده شده است . همانگونه که در شکل‌ها مشاهده می‌شود ، شبکه درهمزمانی طی ۵ مرحله و در ناهمزمانی طی ۴۸۲ مرحله اجرای الگوریتم ، به درستی الگوی

خش‌دار ورودی را با زیابی کرده است.

ج- جهت بردار خطا و خش در الگوی ورودی نیز درنتیجه خروجی مؤثر است. آزمایشها نشان می‌دهند که برای تعداد الگوی معین m به طول n و حداقل فاصله همینگ d،پاسخ شبکه برای ورودیهای خش‌دار با تعداد خطای ثابت (قدر مطلق خطا ثابت)ولی به صورت تصادفی (جهت بردار خطا متغیر)، ثابت و قابل پیش‌بینی نیست. بنابراین بهتر است که آزمایشها به صورت آماری انجام شود و تأثیر جهت بردار خطا در نتایج حاصل منظور شود. به این منظور تابع احتمال ظرفیت را تعریف می‌کنیم.



شکل (۴)- پاسخ شبکه هاپفیلد به ورودی عدد b با ۲۵درصد خطا در حالتهای : الف- همزمانی ب- ناهمزمانی

تابع احتمال ظرفیت - اگر برای هر تعداد الگوی آموزشی (m) تعداد N آزمایش ،برای ورودیهای مختلف واقع در کره همینگ Pn)Pn در طول آزمایش ثابت نگه داشته می‌شود) انجام شود و درتعداد k آزمایش، پاسخ شبکه صحیح باشد، احتمال ظرفیت برای هر مقدار m برابر خواهد بود با :

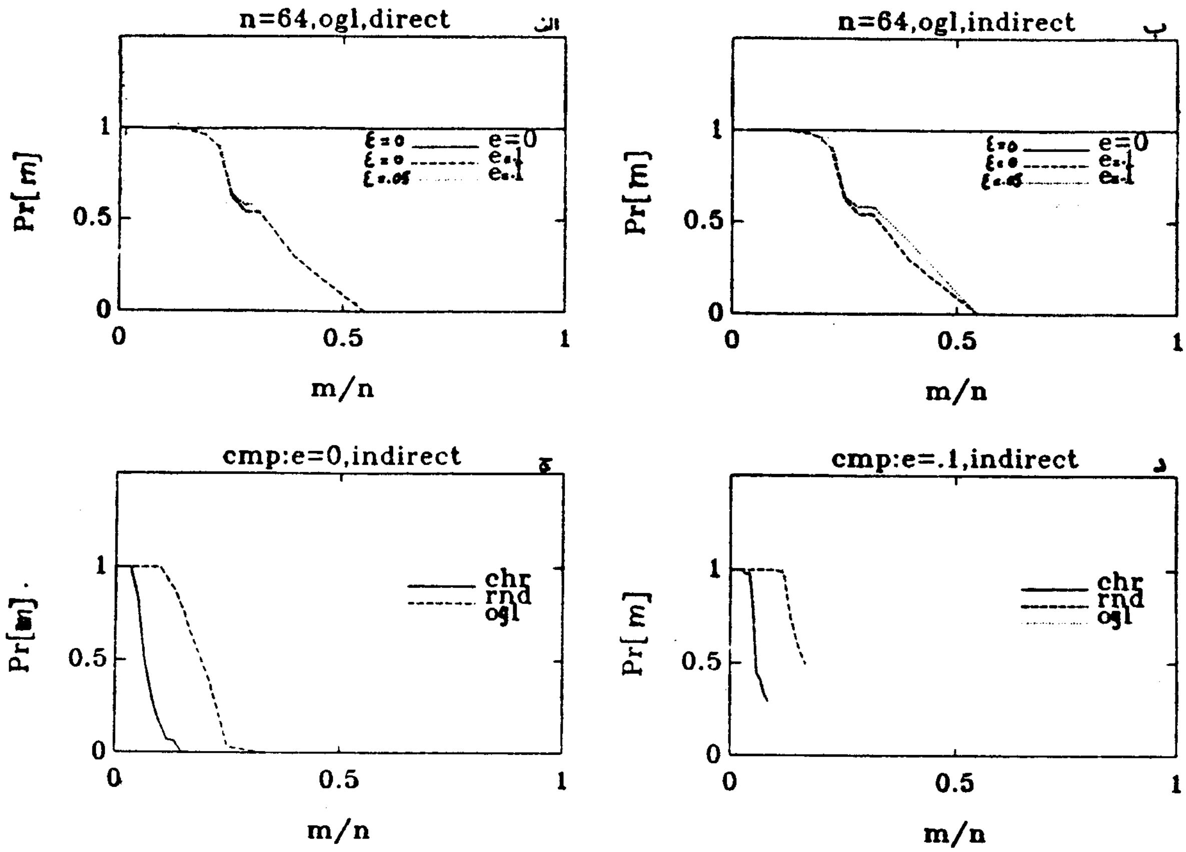
$$P=\frac{k}{N} \qquad (۲۷)$$

اگر منحنی P را برای مقادیر مختلف m رسم کنیم ، منحنی حاصل ، تابع احتمال ظرفیت شبکه خواهد بود. البته برای هر تابع احتمال ظرفیت ، پاسخ صحیح شبکه ، می‌تواند به طور نسبی و دلخواه تعریف شود. مثلاً می‌توان مقدار خطا

در خروجی (ε) یا تعداد مؤلفه‌های غلط در پاسخ شبکه را صفر و یا مقداری کوچک فرض کرد . دراین صورت اگر ε=o فرض شود،تعدادآزمایشهای موفق kبراساس تعداد دفعات پاسخ صددرصد صحیح شبکه، شمرده می‌شود و در غیر این صورت ،پاسخ‌هایی که به تعداد ε و کمتر مؤلفه خطا داشته باشند، درنظر گرفته خواهند شد. به همین ترتیب ،می‌توان تابع احتمال ظرفیت را در حالتهای خاصی مانند همگرایی مستقیم به دست آورد.

۲-۳-۴- شبیه‌سازی حالت بهبود یافته

دراین مرحله با روشی مشابه ، منحنی احتمال ظرفیت برای ۶۴ الگوی متعامد به طول ۶۴ بیت در سه حالت مختلف به دست آورده شد. ابتدا با الگوریتم ارائه شده در [۱۲]کد متعامد (۳۲،۶۴،۶۴) تولید شد و پس‌از آموزش آن به شبکه ، برای ورودی بی‌خش و خش‌دار ۱۰ درصد ، منحنیهای احتمال ظرفیت در دو حالت همگرایی مستقیم و غیر مستقیم ،به دست آمد. همچنین برای مقایسه نتایج این بخش با نتایج تجربی هاپفیلد ،منحنی احتمال ظرفیت برای ε=۰/۰۵ یعنی مجاز بودن خطا در ۵ درصد از مؤلفه‌های خروجی ،نیز ترسیم شد. نتایج شبیه سازی انجام شده به شرح زیرند:
(شکل ۵):



شکل (۵)- منحنیهای احتمال ظرفیت و مقایسه آنها، الف- احتمال ظرفیت برای الگوهای متعامد با همگرایی مستقیم، ب- احتمال ظرفیت برای الگوهای متعامد با همگرایی غیر مستقیم ، ج- مقایسه توابع احتمال ظرفیت برای الگوهای مختلف و ورودی بی‌خش ، د- شرایط مشابه ج برای ورودی خش‌دار.

الف- مقایسه شکلهای (۵-الف) و (۵-ب) نشان می‌دهد که پاسخ شبکه با الگوهای آموزشی متعامد ، در همزمانی و ناهمزمانی یکسان است. این بدین معنی است که شبکه تا مرز ظرفیت خود، همواره بطور مستقیم به سمت پاسخ درست همگرا می‌شود. یعنی درهمزمانی و ورودی بدون خش در دو مرحله و با ورودی خش‌دار در سه مرحله به پاسخ درست می‌رسد. درنتیجه شبکه بسیار سریع پاسخ می دهد.

ب- نتیجه آزمایشها در حالت ورودی بی‌خش، محاسبات نظری را کاملاً تأیید می‌کنند (شکلهای ۵ - الف و ب).

ج- برای ورودی خش‌دار ۱۰ درصد و $m=0/24n$، با احتمال ۸۵ درصد ، پاسخ خروجی کاملاً صحیح $(\varepsilon=0)$ و با احتمال ۹۳ درصد، کمتر از ۰/۰۵ مؤلفه‌ها دچار خطا خواهند بود $(\varepsilon=0/05)$ (شکلهای ۵- الف و ب).

د- آزمایش مشابه با آزمایش هاپفیلد [1] ، یعنی با احتمال بالای پنجاه درصد، کمتر از ۰/۰۵ مؤلفه‌های خروجی خطا داشته باشند، برای الگوهای متعامد نتیجه‌ای برابر $0/37n$ به دنبال دارد. در حالیکه آزمایش هاپفیلد برای الگوهای تصادفی ، $0/15n$ را تأیید می‌کند، بنابراین الگوهای متعامد ، در شرایط مشابه، ۲/۵ برابر ظرفیت شبکه هاپفیلد را افزایش داده‌اند.

هـ- برای همه حالتهای بالا آزمایشها، درشرایطی که قطر ماتریس وزنها، صفر باشد نیز انجام گرفت که دراکثر اوقات، نتیجه مشابه و بعضاً دارای کیفیتی نازل‌تر بود.

و- مقایسه ظرفیت شبکه هاپفیلد برای ورودی بی‌خش در سه حالت الگوهای متعامد، الگوهای تصادفی و حروف ، گویای برتری کامل الگوهای متعامد است . در حالی که شبکه با احتمال صددرصد برای $m=n$ الگوی متعامد پاسخ درست می‌دهد. با همین احتمال برای الگوهای تصادفی، $m=0/1n$ و برای حروف $m=0/03n$ است (شکل ۵ - ج).

ز- مقایسه ظرفیت شبکه هاپفیلد برای ورودی خش‌دار ۱۰ درصد خطا برای الگوهای مختلف نیز حاکی از کیفیت بهتر شبکه برای الگوهای متعامد است. دراین شرایط با احتمال ۹۰ درصد برای الگوهای متعامد $m=0/24n$ ، برای الگوهای تصادفی، $m=0/14n$ و برای حروف $m=0/05n$ به دست می‌آید. میزان بهبودی ظرفیت نسبت به حروف ۵ برابر و نسبت به الگوهای تصادفی بیش‌از ۷۰ درصد است (شکل ۵-د).

نکته باقیمانده ، تفاوت نتایج به دست آمده برای الگوهای متعامد در حالت ورودی خش‌دار با محاسبات انجام شده در بخش (۲ - ۴) است. تنها عاملی که به نظر می‌رسد مانع تطابق نتایج باشد، فرض توزیع یکنواخت برای مقادیر جملات رابطه (۲۴) برای مقادیر کوچک m است. و محتمل است که با افزایش آن در نتیجه بهبود حاصل شود.

۵- نتیجه‌گیری

در این مقاله ، تأثیر مثبت کدینگ الگوهای آموزشی در افزایش حجم حافظه شبکه عصبی هاپفیلد (مدل دودویی) به طریق نظری و عملی مطرح شد. ابتدا از دیدگاهی تازه، شبکه هاپفیلد به یک کانال شانون تشبیه شد و از این زاویه ظرفیت آن مورد بررسی قرار گرفت. نتیجه محاسبه ظرفیت ، مؤید محاسبات انجام شده در [5] بود، ضمن اینکه جامعیت و سادگی روش در خور توجه است. سپس با بررسی نظری الگوریتم یادگیری مدل هاپفیلد، این نتیجه حاصل شد که در صورت اعمال کدینگ مناسب با هدف متعامد ساختن الگوهای آموزشی شبکه ، پاسخ شبکه بسیار مطلوب‌تر از حالات دیگر، حتی حالت الگوهای تصادفی ، خواهد بود. لذا به عنوان راه‌حلی مؤثر بر افزایش حجم حافظه قابل دسترس مدنظر قرار گرفت. نتیجه نظریهٔ اعمال الگوهای متعامد ، در حالت ورودی بی‌خش ، بسیار جالب وبرابر حد نظری ظرفیت شبکه است. همچنین بررسیهای نظریه نشان می‌دهد که در شرایط خش‌دار نیز ، حجم حافظه به صورت محسوسی

افزایش خواهد یافت. شبیه‌سازی شبکه هاپفیلد و آموزش آن با الگوهای متعامد ،بررسیهای نظری را به طور قاطع تأیید کرده و نتایج زیر را در بر داشته است.

– دستیابی به حد ظرفیت شبکه هاپفیلد در حالت ورودیهای بی‌خش و بهبود آن به میزان ۱۰ برابر نسبت به الگوهای آموزشی تصادفی وبیش‌از ۳۰ برابر نسبت به الگوهای آموزشی غیر تصادفی ،

– افزایش حجم حافظه به میزان ۷۰٪ نسبت به الگوهای تصادفی و ۵ برابر نسبت به الگوهای غیر تصادفی در حالت ورودی خش‌دار،

– بهبود بخشیدن حجم حافظه به دست آمده توسط آقای هاپفیلد (ابداع کننده شبکه هاپفیلد) به میزان ۲/۵برابر.

مراجع

[۱]- J. J. Hopfield,«Neural networks and physical systems with emergent collective computational abilities,» Proceedings of the National Academy of Sciences 79, 1982.

[۲]- J. J. Hopfield,«Neurons with graded response have collective computati onal properties like those of two-state neurons,»Proceedings of National Academy of Sciences 81, 1984.

[۳]- J. J. Hopfield, D. W. Tank,« Computing with Neural Circuits: A Model,» Science, vol. 233,1986.

[۴]- D. O. Hebb,«The Organization of Behavior,» New York: Wiley, 1949.

[۵]- Y. S. Abu-Mostafa, J. M. Jacques, Information Capacity of the Hopfield Model, IEEE Transactions on Information Theory, Vol IT - 31, July 1985.

[۶]- K. S. Shanmugam, «Digital and Analog Communication Systems» John Wiley & Sons, 1979.

[۷]- R. Ash,«Information Theory,» Wiley, New York, 1964.

[۸]- R. G. Gallager,«Information Theory and Reliable Communication,»Wiley, New York, 1968.

[۹]- R. J. Mc Eliece, E. C. Posner, E. R. Rodemich, and S. S. Venkatesh, The Capacity of the Hopfield Associative Memory, IEEE T - IT, 33(4), 1987.

[۱۰]- F. J. Mac Williams, N. J. A. Sloane,«The Theory of Error Correcing Codes,»Elsevier, Science Publishers, North - Holland Mathematical Library, 6th Printing: 1988.

[۱۱]- D. E. Romelhart, J. L. Mc Clelland,«Parallel Distributed Processing : Explorations in the Microstructure of Cognition , » MIT Press, 1986.

[۱۲]- مجید سلیمانی‌پور « شبکه‌های عصبی و کاربرد تئوری اطلاعات» پایان‌نامه کارشناسی ارشد، دانشکده برق دانشگاه خواجه نصیرالدین طوسی ، خرداد ۱۳۷۰

The Capacity of the Hopfield Neural Network and A Practical Way To Increase its Memory.

M. R. AREF , M. SOLAIMANIPOUR

abstract

The capacity of the Hopfield model has been considered as an imortant parameter in using this model. In this paper, the Hopfield neural network is modeled as a Shannon Channel and an upperbound to its capacity is found.

For achieving maximum memory, we focus on the training algorithm of the network, and prove that the capacity of the network is bounded by the maximum number of the orthogonal training patterns. Then, the pratical memory of the network, for noiseless and noisy inputs, by appropiate coding of the training patterns is examined. Finally, the theoretical results and the increase rate of the memory is evaluated by simulation.