

Aplicación neuronal del enfriamiento funcional discreto al problema del viajante

Domingo López Rodríguez, Enrique Mérida Casermeiro
Departamento de Matemática Aplicada
ETS Ingeniería Informática
Universidad de Málaga
Campus de Teatinos, s/n 29071 Málaga
{dlopez,merida}@ctima.uma.es

Resumen

El objetivo de este trabajo es presentar una nueva técnica de optimización discreta que permite reducir el número de mínimos locales en la función de energía de la red, consiguiendo así mejorar sensiblemente la calidad de las soluciones obtenidas. Debido a su generalidad, esta técnica puede ser aplicada a la mayoría de las redes neuronales discretas. Proponemos sus bases teóricas, así como mostramos su aplicación en el ámbito de las redes neuronales recurrentes, para ello utilizamos como banco de pruebas el problema del viajante de comercio por ser el problema de optimización combinatoria más conocido y el más comúnmente empleado como test para medir la eficacia de las técnicas algorítmicas aplicadas a este tipo de problemas. Hemos comparado esta técnica con métodos de probada eficacia, consiguiendo mejorar la calidad media de las soluciones de forma notable, logrando en la mayor parte de las simulaciones realizadas soluciones muy cercanas al óptimo global del problema.

1. Introducción

El problema del viajante de comercio (TSP, por sus siglas en inglés), es uno de los más conocidos y estudiados problemas de optimización combinatoria. La lista de aplicaciones reales de este problema es muy extensa, abarcando desde la elección automática de ru-

tas para robots, hasta la localización de agujeros en las placas de circuitos impresos [15], pasando por la revisión de turbinas de gas, la elección de la sucesión de trabajos para una máquina, o el análisis de la estructura de cristales [3], entre otros.

Aparte de la gran cantidad de aplicaciones del TSP, otra característica que posee es su intratabilidad computacional.

Podemos definir este problema de la siguiente forma: dadas N ciudades $\{X_1, \dots, X_N\}$ (puntos en el plano) y las distancias $d_{i,j}$ entre cada par de ciudades X_i y X_j , encontrar el recorrido cerrado de menor longitud que pase una única vez por cada ciudad. A pesar de la sencillez de esta definición, este problema es uno de los máximos representantes de la clase de complejidad de problemas NP-completos, lo que indica el grado de dificultad en su resolución. Por tanto, resultan interesantes algoritmos que obtengan buenas aproximaciones al óptimo en un tiempo reducido.

Por eso, además de los métodos ya clásicos de Optimización y de Investigación Operativa, se han desarrollado métodos basados en algoritmos genéticos [13], enfriamiento estadístico [1], búsqueda tabú [4], y redes neuronales [14].

Al trabajar con redes neuronales, el interés radica en conseguir una formulación adecuada del problema, de forma que la resolución del mismo aparezca de forma natural como un problema de minimización de una función de energía.

La primera red neuronal utilizada para resolver problemas de optimización combinatoria fue el modelo de Hopfield analógico, presentado por Hopfield y Tank en 1985 [5], precisamente para este problema.

Este modelo, superior al discreto en cuanto a posibilidad de escapar de mínimos locales, posee algunas deficiencias a la hora de resolver este problema, que ya fueron mencionadas en [17], como es la necesidad de ajustar *a priori* determinados parámetros que se utilizan en la función de energía, problema que comparte con el modelo discreto.

Otras estrategias para resolver el TSP están basadas en la red auto-organizada de Kohonen [6]. La más eficaz de entre todas ellas es la llamada KNIES [2], que incorpora de forma explícita algunas características estadísticas a la red original.

Recientemente, Mérida et al. [7, 8], presentaron un modelo de red multivaluada, MREM, generalización de la bipolar de Hopfield y de otras multivaluadas [9] con la que se mejoran los resultados obtenidos por KNIES, con la ventaja de que no se necesita ajustar ningún parámetro, contrariamente a lo que ocurre con KNIES, y que ha sido utilizada con mucho éxito para resolver otros problemas de optimización combinatoria [10, 11, 12].

Su dinámica computacional está basada en una técnica heurística que detectaba los posibles cruces existentes en el recorrido e intentaba deshacerlos, minimizando así su distancia total.

El objetivo de este trabajo es presentar una técnica adicional a este modelo MREM para poder escapar de ciertos mínimos locales, y mejorar así su eficacia en problemas de difícil resolución, como el que estamos tratando. Comenzaremos, por tanto, por describir el modelo MREM para este problema, para después explicar de forma detallada la técnica que hemos llamado ‘Enfriamiento Funcional’.

2. El modelo multivaluado MREM

El modelo neuronal MREM consiste en una serie de neuronas multivaluadas, donde el estado de la i -ésima neurona queda caracteriza-

do por su salida s_i , pudiendo tomar cualquier valor de un conjunto finito, que denotaremos por \mathcal{M} . Este conjunto de posibles salidas no tiene por qué ser un conjunto numérico, pudiendo estar formado por elementos tales como: ‘rojo’, ‘verde’ y ‘azul’, o ‘verdadero’ y ‘falso’.

El estado completo de la red queda completamente determinado por el vector de estados $\vec{S} = (s_1, s_2, \dots, s_N) \in \mathcal{M}^N$, donde N es el número de neuronas de la red.

Asociado a cada estado de la red, se define una función de energía

$$E(\vec{S}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{i,j} f(s_i, s_j) \quad (1)$$

donde $W = (w_{i,j})$ es una matriz de tamaño $N \times N$ que representa la conexión entre las distintas neuronas ($w_{i,j}$ es el peso que ejerce la neurona j sobre la i) y $f: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ es una función de similitud, es decir, $f(s_i, s_j)$ representa una medida de la similitud existente entre las salidas de las neuronas i y j .

El objetivo de la red es minimizar la función de energía descrita antes. Para ello, partiendo de un estado aleatorio \vec{S}_0 , en cada iteración cambiará el estado actual de la red, \vec{S} , por otro estado \vec{S}' (definido por la dinámica computacional) siempre y cuando se verifique que $E(\vec{S}') < E(\vec{S})$. Si no se cumple esta condición, la red deja de iterar porque ha llegado a un mínimo local (para la dinámica en cuestión).

En principio, este modelo admite múltiples dinámicas (pudiendo ser tanto secuenciales como paralelas). La elección de la dinámica más adecuada vendrá determinada por el problema a que se enfrente la red.

3. MREM para resolver el problema del viajante

Para resolver este problema con la red MREM, hemos de poder identificar un estado de la red con una solución al problema, y, por otra parte, asociaremos la distancia total de un recorrido, con la función de energía de la red.

En cuanto a la primera identificación, sabemos que una solución al TSP se puede representar por una permutación del conjunto

$\{1, \dots, N\}$, siendo N el número de ciudades. Por tanto, la red estará compuesta por N neuronas, cada una de las cuales tomará un valor en el conjunto $\mathcal{M} = \{1, \dots, N\}$, de forma que un vector de estado $\vec{S} = (s_1, \dots, s_N)$ se corresponda con una permutación de $\{1, \dots, N\}$ (éstos son los estados factibles), significando $s_i = k$ que la ciudad k -ésima será visitada en el lugar i -ésimo.

Por otra parte, si, en la definición de la función de energía dada por la expresión (1), hacemos $f(x, y) = -2d_{x,y}$ y

$$w_{i,j} = \begin{cases} 1 & (j = i + 1) \vee \\ & \vee ((i = N) \wedge (j = 1)) \\ 0 & \text{en otro caso} \end{cases}$$

llegamos a que $E(\vec{S}) = \sum_{i=1}^{N-1} d_{s_i, s_{i+1}} + d_{s_N, s_1}$ es la longitud total del recorrido que representa el estado \vec{S} .

La dinámica computacional se basa en partir de un estado factible y actualizar la red de forma que siempre se mantenga dentro del conjunto de estados factibles, realizando en cada momento actualizaciones 2-óptimo sobre el estado actual, es decir, se estudian secuencialmente todos los pares de neuronas p, q con $p > q + 1$, verificando si existe cruce entre los segmentos $[s_p, s_{p+1}]$ y $[s_q, s_{q+1}]$. En ese caso se verifica:

$$d_{s_p, s_{p+1}} + d_{s_q, s_{q+1}} < d_{s_p, s_q} + d_{s_{p+1}, s_{q+1}}$$

y se invierte el sentido del arco entre las ciudades s_{p+1} y s_q .

Como mejora, se le ha añadido la técnica 3-óptimo, consistente en descomponer el recorrido en tres arcos que podrán ser recombinados de todas las formas posibles, eligiendo como estado siguiente de la red aquél que más reduzca la función de energía. Para más detalle puede consultarse [7, 8].

4. Enfriamiento funcional

En esta sección, comenzaremos por presentar de forma rigurosa este método de optimización, más allá de su posible aplicación neuronal, proporcionando los teoremas y resultados básicos que aseguran su convergencia, aunque no incluyendo sus demostraciones,

pues su longitud excedería el tamaño a que está restringido este trabajo.

4.1. Fundamentos teóricos

Supongamos que queremos minimizar una función F que toma valores reales y que está definida sobre un conjunto discreto V .

Construimos una sucesión $\{F_n\}_{n \geq 1}$ de funciones definidas sobre el mismo espacio que F , y que verifiquen las dos propiedades siguientes:

1. $F_n \rightarrow F$, puntualmente, es decir, para todo $x \in V$, se tiene que $\lim_{n \rightarrow \infty} F_n(x) = F(x)$.
2. $\sum_{n=1}^{\infty} \|F_n - F_{n+1}\|_{\infty} < \infty$

donde $\|g\|_{\infty} = \max_{x \in V} |g(x)|$.

Supongamos que, para minimizar F_n , partimos de un punto $x_1^{(n)}$ y que, mediante un método iterativo, construimos una sucesión $x_k^{(n)}$ de forma que $F_n(x_k^{(n)}) \geq F_n(x_{k+1}^{(n)})$, y con $\lim_{k \rightarrow \infty} x_k^{(n)} = x_*^{(n)}$.

Esta forma de construir la sucesión $\{x_k^{(n)}\}$ nos permite asegurar la convergencia del método, pues se tiene el siguiente resultado, que también nos define la estrategia a seguir, es decir, que para minimizar F_{n+1} , tomamos como punto inicial el mínimo obtenido para F_n , $x_*^{(n)}$:

Teorema 4.1 *Bajo las hipótesis anteriores sobre F_n , si ponemos $x_1^{(n+1)} = x_*^{(n)}$, con $x_1^{(1)}$ aleatorio, entonces la sucesión $F_n(x_*^{(n)})$ es convergente.*

Corolario 4.2 *Existe un punto x_* tal que*

$$F(x_*) = \lim_{n \rightarrow \infty} F_n(x_*^{(n)}) = \lim_{n \rightarrow \infty} F(x_*^{(n)})$$

Corolario 4.3 *Si $x_*^{(n)}$ es mínimo global de F_n para todo $n \in \mathbb{N}$, entonces x_* es mínimo global de F .*

Más aun, se puede demostrar que, en el límite, estamos minimizando la función F :

Lema 4.4 *Dados $x, x' \in V$ con $F(x) < F(x')$, existe n_0 tal que si $n \geq n_0$ entonces $F_n(x) < F_n(x')$.*

Corolario 4.5 Existe un $n_0 \in \mathbb{N}$ tal que para todo $x, x' \in V$ con $F(x) < F(x')$, se verifica $F_n(x) < F_n(x')$ para todo $n \geq n_0$.

Proposición 4.6 Existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$, entonces $F_n(x) \leq F_n(x')$ implica que se tiene $F(x) \leq F(x')$.

Demostración.

Sea $\varepsilon = \min\{F(x) - F(x') : x, x' \in V \text{ y } F(x) - F(x') > 0\} > 0$. Como V es un conjunto discreto (finito), ε está bien definido. Por esta definición, si x, x' son tales que $F(x) - F(x') < \varepsilon$, entonces se tendrá que $F(x) - F(x') \leq 0$.

Sea entonces $0 < \varepsilon' < \varepsilon$. Como $F_n \rightarrow F$ puntualmente, dados x, x' , existirá un natural $n_{(x,x')}$ (dependiendo de x, x') tal que si $n \geq n_{(x,x')}$, entonces se tiene

$$|F(x) - F_n(x)| < \frac{\varepsilon'}{2}$$

$$|F(x') - F_n(x')| < \frac{\varepsilon'}{2}$$

en particular, $[F(x) - F(x')] - [F_n(x) - F_n(x')] < \varepsilon'$, luego $F(x) - F(x') < \varepsilon' + F_n(x) - F_n(x')$.

Supongamos que tenemos $F_n(x) \leq F_n(x')$, es decir, $F_n(x) - F_n(x') \leq 0$, entonces, sustituyendo en la expresión anterior, nos queda $F(x) - F(x') \leq \varepsilon' < \varepsilon$, y por lo comentado al principio de esta demostración, se tiene $F(x) - F(x') \leq 0$, esto es, $F(x) \leq F(x')$.

Tomando por tanto $n_0 = \max\{n_{(x,x')} : x, x' \in V\}$, ya se tiene el resultado que buscábamos. \square

Por tanto, no será necesario en la práctica el construir toda la sucesión F_n , sino que bastará con minimizar hasta cierta aproximación F_{n_0} de F . Pero este resultado además implica el siguiente:

Corolario 4.7 Si $x_*^{(n)}$ es mínimo local de F_n para todo n mayor que un cierto $n_0 \in \mathbb{N}$, entonces x_* es mínimo local de F .

Nótese que todos los resultados de convergencia hasta ahora nos hablan de la convergencia de $\{F_n(x_*^{(n)})\}$, para llegar a la conclusión de que $F(x_*)$ es mínimo local de F .

Lo que no es necesariamente cierto, bajo estas condiciones tan generales, es que se tenga $x_*^{(n)} \rightarrow x_*$. Veámoslo con un ejemplo:

Ejemplo 4.1 Sea $M \in \mathbb{N}$ y definimos $F : \{0, \frac{1}{M}, \dots, \frac{M-1}{M}, 1\} \rightarrow \mathbb{R}$ dada por $F(x) = 0$ para todo $x \in V$.

La sucesión F_n vendrá definida de la siguiente forma:

- Si n es impar, F_n será la función cuya gráfica es el segmento de extremos $(0, \frac{1}{n})$ y $(1, \frac{1}{n+1})$ evaluado en los puntos de V . Su expresión analítica es: $F_n(x) = \frac{1}{n} - \frac{1}{n(n+1)}x$. Es convexa y su mínimo global se alcanza en $x_*^{(n)} = 1$.
- Si n es par, F_n será la función cuya gráfica es el segmento de extremos $(0, \frac{1}{n+1})$ y $(1, \frac{1}{n})$, evaluada en V . Su expresión analítica es: $F_n(x) = \frac{1}{n+1} + \frac{1}{n(n+1)}x$. Es convexa y su mínimo global se alcanza en $x_*^{(n)} = 0$.

Evidentemente, F_n converge puntualmente a F . Veamos que también verifica la propiedad $\sum_{n=1}^{\infty} \|F_n - F_{n+1}\|_{\infty} < \infty$.

Se puede comprobar fácilmente que es $\|F_n - F_{n+1}\|_{\infty} = \frac{2}{n(n+2)}$ para todo n , y por tanto $\sum_{n=1}^{\infty} \|F_n - F_{n+1}\|_{\infty} = \sum_{n=1}^{\infty} \frac{2}{n(n+2)} < \infty$.

Además, es $F_n(x_*^{(n)}) = \frac{1}{n+1} \rightarrow 0$, que es el valor mínimo global de F , pero $\{x_*^{(n)}\} = \{1, 0, 1, 0, 1, 0, \dots\}$ no es una sucesión convergente.

Sin embargo, podemos dar unas condiciones suficientes bastante generales que impliquen la convergencia $x_*^{(n)} \rightarrow x_*$. Llamemos Ω_F al conjunto de los mínimos locales de F .

Proposición 4.8 Si todos los mínimos locales de F son estrictos, entonces $x_*^{(n)}$ es convergente, y su límite verifica todos los resultados anteriores.

Demostración.

Para cada $x \in V$, definimos un entorno \mathcal{N}_x de x tal que, para buscar el mínimo partiendo del punto x , se han de comprobar todos los $\hat{x} \in \mathcal{N}_x$, y elegir como siguiente punto aquel que tenga menor valor de la función objetivo.

Supongamos $\Omega_F = \{\xi_1, \dots, \xi_l\}$. Si sólo existe un j tal que $F(\xi_j) = F(x_*)$, ya lo tendríamos, pues sería $x_*^{(n)} \rightarrow x_* = \xi_j$.

Supongamos que existen $\{\xi_{j_1}, \dots, \xi_{j_k}\} \subset \Omega_F$ tales que $F(\xi_{j_i}) = F(x_*)$ para todo i . Como son los ξ_j mínimos locales estrictos, se tiene que $F(\xi_j) < F(x)$ para todo $x \in \mathcal{N}_{\xi_j}$, para todo j . Por el Corolario 4.5, existe un n_0 tal que, si $n \geq n_0$, $F_n(\xi_j) < F_n(x)$ para todo $x \in \mathcal{N}_{\xi_j}$, para todo j .

Ahora bien, para algún $n_1 \geq n_0$, debe ser $x_*^{(n_1)} = \xi_{j_i}$ para algún i , pues en caso contrario, no podría tenerse $F_n(x_*^{(n)}) \rightarrow F(x_*)$ y $F(x_*)$ mínimo local de F .

Pero, por lo dicho antes, en la iteración $n_1 + 1$ se tendría $F_{n_1+1}(x_1^{(n_1+1)}) = F_{n_1+1}(x_*^{(n_1)}) = F_{n_1+1}(\xi_{j_i}) < F_{n_1+1}(x)$ para todo $x \in \mathcal{N}_{\xi_{j_i}}$, y así $x_*^{(n_1+1)} = \xi_{j_i}$. Y de forma sucesiva se demuestra que $x_*^{(n)} = \xi_{j_i}$ para todo $n \geq n_1$. Luego $x_*^{(n)} \rightarrow x_* = \xi_{j_i}$. \square

Como se puede comprobar, en el Ejemplo 4.1 fallaba una condición: los mínimos locales no son estrictos.

Luego, en la práctica, podremos asegurar la convergencia de $x_*^{(n)}$ a un mínimo local de F y, ya que la sucesión $F(x_*^{(n)})$ no es necesariamente decreciente, se puede esperar que esta técnica escape de ciertos mínimos locales, mejorando así la eficacia del algoritmo.

4.2. Aplicación del ‘Enfriamiento Funcional’ a las redes neuronales

De forma análoga a lo expuesto en la teoría del ‘Enfriamiento Funcional’, ya que la función a minimizar en el caso de las redes recurrentes es la función de energía E , la podemos identificar con F . La idea, por tanto, será buscar sucesiones E_n (análogas a F_n), que verifiquen las mismas restricciones que E definida en la expresión (1), pero más fáciles de minimizar que E . De hecho, lo ideal sería que cada E_{n+1} fuera un poco más difícil de minimizar que E_n .

Si bien esta técnica se puede aplicar sin problemas a la red discreta de Hopfield, resulta más útil basarse en la red MREM anteriormente descrita, por ser ésta más general que la

de Hopfield, y porque las soluciones de múltiples problemas, incluyendo TSP, están mejor representadas según el modelo de esta red multivaluada que en el de Hopfield.

Como sucesión de funciones de energía iteradas vamos a considerar las dadas por la expresión:

$$E_n(\vec{S}) = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{i,j}^{(n)} f^{(n)}(s_i, s_j) \quad (2)$$

donde $W^{(n)} = (w_{i,j}^{(n)})$ es una sucesión de matrices que verifica $W = \lim_{n \rightarrow \infty} W^{(n)}$, y $f^{(n)}: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ es una sucesión de funciones que tiende a f puntualmente. Evidentemente, se tiene que $E(\vec{S}) = \lim_{n \rightarrow \infty} E_n(\vec{S})$ para todo vector de estados $\vec{S} \in \mathcal{M}^N$.

La dinámica empleada para minimizar E_n se corresponde exactamente con la del modelo MREM mencionado antes.

Por tanto, para cada n , dado $\vec{S}_1^{(n)}$, tendremos una sucesión $\{\vec{S}_i^{(n)}\}_{i \geq 1}$ que converge a un cierto $\vec{S}_*^{(n)}$.

Teorema 4.9 Si definimos $\vec{S}_1^{(n+1)} = \vec{S}_*^{(n)}$, $\vec{S}_1^{(1)}$ aleatorio, y se verifican las siguientes condiciones:

1. $\sum_{n=1}^{\infty} \|W^{(n)} - W^{(n+1)}\|_{\infty} < \infty$
2. $\sum_{n=1}^{\infty} \|f^{(n)} - f^{(n+1)}\|_{\infty} < \infty$
3. $\|f^{(n)}\|_{\infty} < K$ para todo n

entonces las sucesiones $\{E_n(S_*^{(n)})\}_{n \geq 1}$ y $\{E(S_*^{(n)})\}_{n \geq 1}$ convergen a $E(\vec{S}_*)$ para cierto estado $\vec{S}_* = \lim_{n \rightarrow \infty} \vec{S}_*^{(n)}$.

Demostración.

Veamos que se puede aplicar el Teorema 4.1 a este caso. Para ello, lo único que debemos verificar es que $\sum_{n=1}^{\infty} \|E_{n+1} - E_n\|_{\infty} < \infty$.

Dado $\vec{S} \in \mathcal{M}$, se puede comprobar que es

$$\begin{aligned} & |E_{n+1}(\vec{S}) - E_n(\vec{S})| \leq \\ & \leq \frac{N}{2} \left[\|W^{(n+1)}\|_{\infty} \|f^{(n+1)} - f^{(n)}\|_{\infty} + \right. \\ & \quad \left. + \|W^{(n+1)} - W^{(n)}\|_{\infty} \|f^{(n)}\|_{\infty} \right] \end{aligned}$$

Por tanto, tomando máximo, tenemos que:

$$\begin{aligned} & \|E_{n+1} - E_n\|_\infty \leq \\ & \leq \frac{N}{2} \left[\|W^{(n+1)}\|_\infty \|f^{(n+1)} - f^{(n)}\|_\infty + \right. \\ & \quad \left. + \|W^{(n+1)} - W^{(n)}\|_\infty \|f^{(n)}\|_\infty \right] \end{aligned}$$

Pero, como $W^{(n)} \rightarrow W$, se tiene que $\|W^{(n)}\|_\infty \rightarrow \|W\|_\infty$ y por lo tanto existe M tal que $\|W^{(n)}\|_\infty < M$ para todo n . Y, por la condición 3., se tiene $\|f^{(n)}\|_\infty < K$ para todo n . Luego

$$\begin{aligned} & \|E_{n+1} - E_n\|_\infty \leq \\ & \leq \frac{N}{2} \left[M \|f^{(n+1)} - f^{(n)}\|_\infty + \right. \\ & \quad \left. + K \|W^{(n+1)} - W^{(n)}\|_\infty \right] \end{aligned}$$

Sumando, para $n = 1, \dots$, se tiene (aplicando las propiedades 1. y 2.):

$$\begin{aligned} & \sum_{n=1}^{\infty} \|E_{n+1} - E_n\|_\infty \leq \\ & \leq \frac{N}{2} \left[M \sum_{n=1}^{\infty} \|f^{(n+1)} - f^{(n)}\|_\infty + \right. \\ & \quad \left. + K \sum_{n=1}^{\infty} \|W^{(n+1)} - W^{(n)}\|_\infty \right] < \infty \end{aligned}$$

Así que podemos aplicar el Teorema 4.1, y sus corolarios, de donde llegamos al resultado que queríamos demostrar. \square

Corolario 4.10 *El valor $E(\vec{S}_*)$ es un mínimo local de la función de energía E .*

Hemos de hacer notar que en la mayoría de los problemas prácticos, esta condición se va a cumplir, pues en realidad no estaremos utilizando sucesiones infinitas de matrices ($\{W^{(n)}\}_{n \geq 1}$) y/o funciones ($\{f^{(n)}\}_{n \geq 1}$), sino un subconjunto finito de ellas, $\{W^{(1)}, W^{(2)}, \dots, W^{(L)} = W\}$ y $\{f^{(1)}, f^{(2)}, \dots, f^{(L)} = f\}$.

5. Resultados experimentales

Para este problema concreto vamos a considerar sólo un número finito de funciones de energía E_n , ($\{E_1, E_2, \dots, E_{n_{approx}+1} = E\}$), donde las correspondientes matrices de pesos son todas iguales a W ($W^{(n)} = W$) y las funciones de similitud $f^{(n)}(x, y)$ están definidas de la siguiente forma:

- $f^{(n)}(x, y) = f(x, y) = -2d_{x,y}$ cuando $m \leq d_{i,j} \leq m + \frac{(k-1)(M-m)}{n_{approx}}$.
- $-2M$ en caso contrario.

donde m y M son respectivamente los valores mínimo y máximo que toman los $d_{i,j}$.

Esto significa que, en las primeras iteraciones, el modelo va a intentar ordenar aquellas ciudades que estén cerca unas de otras, sin tener en cuenta las que estén a más distancia, induciendo una serie de ordenaciones parciales correspondientes a estas ciudades, para así pasar a calcular unas ordenaciones menos locales, teniendo en cuenta al final también las ciudades que estén más alejadas.

Los resultados de aplicar esta técnica a los problemas típicos de la librería TSPLIB, disponible en internet y creada por Reinelt [16], aplicando directamente *3-óptimo* para minimizar E (es decir, utilizando directamente MREM para minimizar E), o a través de la sucesión de E_n , vienen en la tabla 1, donde hemos considerado dos valores para n_{approx} : $n_{approx} = 10$ y $n_{approx} = 20$, sobre un total de 50 ejecuciones de cada uno de los métodos para cada uno de los problemas.

La elección del valor de n_{approx} viene dado por el tamaño del problema. Para problemas de hasta tamaño 100, un valor de 10 es adecuado, pero para problemas mayores, $n_{approx} = 10$ presenta una pega: las transiciones entre E_n y E_{n+1} no son tan 'suaves' como las que se deben imponer para obtener los resultados del teorema 4.9, y por tanto se aumentó su valor.

Se puede comprobar cómo, sobre todo en las dos últimas columnas, la técnica aquí propuesta ha conseguido mejorar en gran manera los resultados presentados por la red MREM original, y los de KNIES, que ya habían sido su-

Cuadro 1: Resultados comparativos (% error sobre el óptimo) de los modelos KNIES, MREM y ‘Enfriamiento Funcional’ tras realizar 50 ejecuciones.

Problema	Ópt	KNIES	MREM		Red Prop. ($n_{approx} = 10$)		Red Prop. ($n_{approx} = 20$)	
		Min	Min	Media	Min	Media	Min	Media
eil51	426	2.86	0.23	2.43	0	2.27	0	1.81
st70	675	1.51	0	1.89	0	1.79	0	1.79
eil76	538	4.98	1.3	3.43	0.18	2.62	0.56	2.64
rd100	7910	2.09	0	3.02	0	1.03	0	1.23
eil101	629	4.66	1.43	3.51	0.32	2.67	0	2.67
lin105	14379	1.29	0	1.71	0	1.18	0	0.73
pr107	44303	0.42	0.15	0.82	0	0.96	0	0.78
pr124	59030	0.08	0	1.23	0	1.39	0	0.80
bier127	118282	2.76	0.42	2.06	0.18	2.24	0.02	1.92
pr136	96772	4.53	0.37	3.79	0.90	3.31	0.15	2.43
pr152	73682	0.97	0.6	1.45	0	1.61	0.18	1.87

perados por MREM, de forma que en la mayoría de las simulaciones, el resultado medio dado por esta nueva técnica es mejor que el de MREM, mostrando así la capacidad para escapar de mínimos locales. Esto también se puede constatar si se comparan los resultados óptimos de estos tres métodos, ya que los de MREM y KNIES se ven ampliamente mejorados por los del ‘Enfriamiento Funcional’.

Hay que hacer notar que el tiempo de cómputo de esta red con ‘Enfriamiento Funcional’ con respecto del modelo MREM original, no se multiplica por n_{approx} , pues en las últimas aproximaciones, en las que ya n está cerca de $n_{approx} + 1$, el cambio en la función de energía aproximada, al pasar de E_n a E_{n+1} , es muy pequeño (pues así lo es el paso de $f^{(n)}$ a $f^{(n+1)}$), y el estado de la red estará muy próximo a un mínimo local, así que la red casi no iterará a partir de un cierto n_0 . De hecho, para $n_{approx} = 10$, el ‘Enfriamiento Funcional’ sólo tarda entre 2 y 6 veces el tiempo consumido por MREM (dependiendo del problema particular, no de su tamaño), consiguiendo una media de 4.03 veces el tiempo de MREM. Para $n_{approx} = 20$, el tiempo de cómputo de la nueva técnica se sitúa entre 3 y 8 veces el del modelo MREM original, con una media de sólo 6.41.

6. Conclusiones

En este trabajo hemos presentado una nueva técnica general de optimización que permite evitar diversos mínimos locales, mejorando así sensiblemente la calidad de las soluciones obtenidas por diversos algoritmos.

Hemos propuesto sus bases teóricas, en base a unos resultados bastante generales en cuanto a su aplicabilidad, y que podrían sentar las bases de una teoría más general de optimización. Además, hemos presentado su traducción neuronal, indicando cómo podría ser su utilización en este ámbito.

Como banco de pruebas para esta nueva técnica, hemos escogido el famoso problema del viajante de comercio, debido a que constituye el principal banco de pruebas de los algoritmos de optimización combinatoria, presentando muchos métodos para su resolución aproximada, habiendo incluso algunos de gran potencia, como MREM o KNIES.

Mediante el uso de simulaciones, hemos dejado constancia de la mejora que supone el ‘Enfriamiento Funcional’ ya que es capaz de evitar caer en muchos mínimos locales y de reducir la probabilidad de caer en otros, mejorando la calidad de la solución media de forma muy notable, y logrando, en muchos casos, soluciones muy cercanas al óptimo del proble-

ma.

Referencias

- [1] Aarts, E.H.L., Korst, J.H.M. y Laarhoven, P.J.M., *A Quantitative Analysis of the Simulated Annealing Algorithm: A Case Study for the Traveling Salesman Problem*, J. Stats. Phys. **50**, 189-206, 1988.
- [2] Aras, N., Oomen, B.J., y Altinel, I.K., *The Kohonen Network Incorporating Explicit Statistics and its Application to the Travelling Salesman Problem*, Neural Networks, **12**, 1273-1284, 1999.
- [3] Bland, R.E., y Shallcross, D. F., *Large Traveling Salesman Problem Arising from Experiments in X-ray Crystallography: a Preliminary Report on Computation*, Technical Report No. 730, School of OR/IE, Cornell University, Ithaca, Nueva York, 1987.
- [4] Fiechter, C.N., *A Parallel Tabu Search Algorithm for Large Scale Traveling Salesman Problems*, Working Paper 90/1 Department of Mathematics, Ecole Polytechnique Federale de Lausanne, Suiza, 1990.
- [5] Hopfield, J.J. y Tank, D.W., *Neural Computation of Decisions in Optimization Problems*, Biological Cybernetics, **52**, 141-152, 1985.
- [6] Kohonen, T., *Self-organizing Maps*, Springer, 1995.
- [7] Mérida-Casermeiro, E., Galán-Marín, G. y Muñoz-Pérez, J., *An Efficient Multivalued Hopfield Network for the Travelling Salesman Problem*, Neural Processing Letters, **14**, 203-216, 2001.
- [8] Mérida-Casermeiro, E., Muñoz-Pérez, J., y Domínguez-Merino, E., *An N-parallel Multivalued Network: Applications to the Travelling Salesman Problem*, Computational Methods in Neural Modelling, Lecture Notes in Computer Science, **2686**, 406-41, 2003.
- [9] Mérida-Casermeiro, E. y Muñoz-Pérez, J., *MREM: An Associative Autonomous Recurrent Network*, Journal of Intelligent and Fuzzy Systems, **12** (3-4), 163-173, 2002.
- [10] Mérida-Casermeiro, E., Muñoz-Pérez, J. y Benítez-Rochel, R., *Neural Implementation of Dijkstra's Algorithm*, Lecture Notes in Computer Science **2686**, 342-349, 2003.
- [11] Mérida-Casermeiro, E. y López-Rodríguez, D., *Multivalued Neural Network for Graph MaxCut Problem*, in Proceedings of International Conference of Computational Methods in Science and Engineering, ICCMSE 2004, **1**, 375-378, 2004.
- [12] Mérida-Casermeiro, E. y López-Rodríguez, D., *Graph Partitioning via Recurrent Multivalued Neural Networks*, IWANN 2005, Lecture Notes in Computer Science (aceptado).
- [13] Potvin, J.V., *Genetic Algorithms for the Traveling Salesman Problem*, Annals of Operations Research **63**, 339-370, 1996.
- [14] Potvin, J.V., *The Traveling Salesman Problem: A Neural Network Perspective*, INFORMS Journal on Computing, **5**, 328-348, 1993.
- [15] Reinelt, G., *The Travelling Salesman. Computational Solutions for TSP Applications*, Springer, 1994.
- [16] Reinelt, G., *TSPLIB - A Travelling Salesman Problem Library*, ORSA Journal of Computing, **3**, 376-384, 1991.
- [17] Wilson, V. y Pawley, G.S., *On the Stability of the TSP Problem Algorithm of Hopfield and Tank*, Biological Cybernetics, **58**, 63-70, 1988.