Red Multivaluada Estocástica para Optimización. Aplicación al Problema del MaxCut

Domingo López-Rodríguez[1], Enrique Mérida-Casermeiro[1] & Juan M. Ortiz-de-Lazcano-Lobato[2]

[1] Departamento de Matemática Aplicada
[2] Departamento de Lenguajes y Ciencias de la Computación
E.T.S. Ingeniería Informática
Universidad de Málaga
email: [1]{dlopez,merida}@ctima.uma.es
[2]jmortiz@lcc.uma.es

MAEB, Febrero 2007

UNIVERSIDAD DE MÁLAGA

Índice

UNIVERSIDAD DE MÁLAGA

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

UNIVERSIDAD DE MÁLAGA

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

- En el caso general, NP-completo.

UNIVERSIDAD DE MÁLAGA

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

- En el caso general, NP-completo.
- En el caso de que el grafo sea plano, el problema pertenece a P (se puede hallar una solución en tiempo polinómico).

UNIVERSIDAD DE MÁLAGA

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

- En el caso general, NP-completo.
- En el caso de que el grafo sea plano, el problema pertenece a P (se puede hallar una solución en tiempo polinómico).

Aplicaciones:

- Clasificación y agrupamiento de datos,

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

- En el caso general, NP-completo.
- En el caso de que el grafo sea plano, el problema pertenece a P (se puede hallar una solución en tiempo polinómico).

Aplicaciones:

- Clasificación y agrupamiento de datos,
- Reconocimiento de patrones,

El Problema Clásico de MaxCut

Dado un grafo no dirigido $G = (V, E)$ ponderado en los ejes por $C = (c_{i,j})$ con $c_{i,j} \geq 0$,
El objetivo es obtener una partición de V en dos (K) subconjuntos de forma que se maximice el coste total de los ejes cuyos extremos están en distintos conjuntos de la partición.

- En el caso general, NP-completo.
- En el caso de que el grafo sea plano, el problema pertenece a P (se puede hallar una solución en tiempo polinómico).

Aplicaciones:

- Clasificación y agrupamiento de datos,
- Reconocimiento de patrones,
- Diseño de redes de comunicaciones. . .

UNIVERSIDAD DE MÁLAGA

Trabajo Previo para 2-MaxCut

- En 1997, Alberti et al. presentaron una red tipo Hopfield para MaxCut.

Trabajo Previo para 2-MaxCut

- En 1997, Alberti et al. presentaron una red tipo Hopfield para MaxCut.
- En 1998, Bertoni et al. publicaron otra red con mejores resultados.

Trabajo Previo para 2-MaxCut

- En 1997, Alberti et al. presentaron una red tipo Hopfield para MaxCut.
- En 1998, Bertoni et al. publicaron otra red con mejores resultados.
- Takefuyi y sus colaboradores desarrollaron en 1996 el modelo neuronal **maximum** que obtenía buenos resultados sobre un amplio conjunto de problemas de optimización combinatorial.

Trabajo Previo para 2-MaxCut

- En 1997, Alberti et al. presentaron una red tipo Hopfield para MaxCut.
- En 1998, Bertoni et al. publicaron otra red con mejores resultados.
- Takefuyi y sus colaboradores desarrollaron en 1996 el modelo neuronal **maximum** que obtenía buenos resultados sobre un amplio conjunto de problemas de optimización combinatorial.
- En 2001, Galán Marín et al. propusieron el modelo OCHOM, que obtiene soluciones de forma más eficiente que 'maximum'. Es muy versátil y converge rápidamente a una solución válida sin necesidad de ajustar ningún parámetro.

UNIVERSIDAD DE MÁLAGA

Trabajo Previo para 2-MaxCut

- En 1997, Alberti et al. presentaron una red tipo Hopfield para MaxCut.
- En 1998, Bertoni et al. publicaron otra red con mejores resultados.
- Takefuyi y sus colaboradores desarrollaron en 1996 el modelo neuronal **maximum** que obtenía buenos resultados sobre un amplio conjunto de problemas de optimización combinatorial.
- En 2001, Galán Marín et al. propusieron el modelo OCHOM, que obtiene soluciones de forma más eficiente que 'maximum'. Es muy versátil y converge rápidamente a una solución válida sin necesidad de ajustar ningún parámetro.
- En 2004, Wang et al. propusieron una variante estocástica de OCHOM.

UNIVERSIDAD DE MÁLAGA

Objetivo de Este Trabajo

- En 2005, propusimos el modelo neuronal **MREM** para resolver el problema de MaxCut. Además, permite hacer una K-partición del grafo como generalización natural de MaxCut.

Objetivo de Este Trabajo

- En 2005, propusimos el modelo neuronal **MREM** para resolver el problema de MaxCut. Además, permite hacer una K-partición del grafo como generalización natural de MaxCut.
- En este trabajo presentamos una versión estocástica de MREM (**sMREM**) y su aplicación a MaxCut.

UNIVERSIDAD DE MÁLAGA

El modelo de Hopfield Discreto

Este modelo clásico se caracteriza por:

- El estado de cada neurona viene caracterizado por su salida, pudiendo ser binaria ($\{0,1\}$) o bipolar ($\{-1,1\}$).

El modelo de Hopfield Discreto

Este modelo clásico se caracteriza por:

- El estado de cada neurona viene caracterizado por su salida, pudiendo ser binaria ($\{0, 1\}$) o bipolar ($\{-1, 1\}$).
- El estado de la red está determinado por la salida de sus N neuronas, es decir, $\vec{S}(t) \in \{0, 1\}^N$, o $\{-1, 1\}^N$).

UNIVERSIDAD DE MÁLAGA

El modelo de Hopfield Discreto

Este modelo clásico se caracteriza por:

- El estado de cada neurona viene caracterizado por su salida, pudiendo ser binaria ($\{0, 1\}$) o bipolar ($\{-1, 1\}$).
- El estado de la red está determinado por la salida de sus N neuronas, es decir, $\vec{S}(t) \in \{0, 1\}^N$, o $\{-1, 1\}^N$).
- Todas las neuronas están conectadas entre sí. La conexión entre (i, j) entre las neuronas i y j tiene asociado un peso sináptico $w_{i,j}$.

UNIVERSIDAD DE MÁLAGA

El modelo de Hopfield Discreto

Este modelo clásico se caracteriza por:

- El estado de cada neurona viene caracterizado por su salida, pudiendo ser binaria ($\{0,1\}$) o bipolar ($\{-1,1\}$).
- El estado de la red está determinado por la salida de sus N neuronas, es decir, $\vec{S}(t) \in \{0,1\}^N$, o $\{-1,1\}^N$).
- Todas las neuronas están conectadas entre sí. La conexión entre (i,j) entre las neuronas i y j tiene asociado un peso sináptico $w_{i,j}$.
- Cada neurona tiene un umbral (θ_i).

UNIVERSIDAD DE MÁLAGA

El modelo de Hopfield Discreto

Este modelo clásico se caracteriza por:

- El estado de cada neurona viene caracterizado por su salida, pudiendo ser binaria ($\{0, 1\}$) o bipolar ($\{-1, 1\}$).
- El estado de la red está determinado por la salida de sus N neuronas, es decir, $\vec{S}(t) \in \{0, 1\}^N$, o $\{-1, 1\}^N$).
- Todas las neuronas están conectadas entre sí. La conexión entre (i, j) entre las neuronas i y j tiene asociado un peso sináptico $w_{i,j}$.
- Cada neurona tiene un umbral (θ_i).
- Existe una **función de energía**:

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

Modelo **RE**currente **M**ultivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.

Modelo REcurrente Multivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.
- El estado de la neurona i está determinado por su salida s_i.

UNIVERSIDAD DE MÁLAGA

Modelo REcurrente Multivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.
- El estado de la neurona i está determinado por su salida s_i.
- El estado de la red viene caracterizado por el **vector de estado** $\vec{S} = (s_1, s_2, \ldots, s_N) \in \mathcal{M}^N$. ($N$ es el número de neuronas)

Modelo REcurrente Multivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.
- El estado de la neurona i está determinado por su salida s_i.
- El estado de la red viene caracterizado por el **vector de estado** $\vec{S} = (s_1, s_2, \ldots, s_N) \in \mathcal{M}^N$. ($N$ es el número de neuronas)
- Existe una **función de energía** $E : \mathcal{M}^N \to \mathbb{R}$

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

UNIVERSIDAD DE MÁLAGA

Modelo REcurrente Multivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.
- El estado de la neurona i está determinado por su salida s_i.
- El estado de la red viene caracterizado por el **vector de estado** $\vec{S} = (s_1, s_2, \ldots, s_N) \in \mathcal{M}^N$. ($N$ es el número de neuronas)
- Existe una **función de energía** $E : \mathcal{M}^N \to \mathbb{R}$

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

  - $W = (w_{i,j})$ es la matriz de pesos sinápticos.

Modelo REcurrente Multivaluado (MREM)

- La salida de la neurona puede tomar cualquier valor en un **conjunto discreto** $\mathcal{M}$.
- El estado de la neurona i está determinado por su salida s_i.
- El estado de la red viene caracterizado por el **vector de estado** $\vec{S} = (s_1, s_2, \ldots, s_N) \in \mathcal{M}^N$. ($N$ es el número de neuronas)
- Existe una **función de energía** $E : \mathcal{M}^N \to \mathbb{R}$

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

  - $W = (w_{i,j})$ es la matriz de pesos sinápticos.
  - f es una función de similitud.

UNIVERSIDAD DE MÁLAGA

MREM

Dinámica: La red evoluciona de forma que decrezca el valor de la función de energía. Son posibles diversas dinámicas. Se elegirá la más adecuada al problema que se esté estudiando.

UNIVERSIDAD DE MÁLAGA

Table of contents

UNIVERSIDAD DE MÁLAGA

Implementación del Problema K-MaxCut con MREM.

Si identificamos la función de coste de K-MaxCut:

$$\sum_{\mathbf{v_i \in A_m, v_j \in A_n, i > j, m \neq n}} \mathbf{c_{i,j}}$$

con la función de energía de MREM:

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

obtenemos $w_{i,j} = c_{i,j}$ y $f(x,y) = \delta_{x,y}$ (delta de Kronecker) $\{$Maximizar el coste de los ejes cortados por la partición$\}$ $\equiv$ $\{$Minimizar el coste de los ejes cuyos extremos están en el mismo conjunto de la partición$\}$

Table of contents

UNIVERSIDAD DE MÁLAGA

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.
2. Repetir hasta que no haya cambios en el vector de estado:

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.
2. Repetir hasta que no haya cambios en el vector de estado:
   - El scheduling (controlador) selecciona un valor $d \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor\}$.

UNIVERSIDAD DE MÁLAGA

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.
2. Repetir hasta que no haya cambios en el vector de estado:
    - El scheduling (controlador) selecciona un valor $d \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor\}$.
    - En paralelo, cada neurona p estudia todas las posibilidades de actualizarse ella misma junto con la neurona $q = (p + d) \mod (N)$, con $0 < q \leq N$.

UNIVERSIDAD DE MÁLAGA

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.
2. Repetir hasta que no haya cambios en el vector de estado:
    - El scheduling (controlador) selecciona un valor $d \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor\}$.
    - En paralelo, cada neurona p estudia todas las posibilidades de actualizarse ella misma junto con la neurona $q = (p + d) \mod (N)$, con $0 < q \leq N$.
    - De esta forma, p calcula la matriz de potencial U_p con $K \times K$ componentes.

UNIVERSIDAD DE MÁLAGA

Dinámica Best-2 (sólo 2 neuronas se actualizan)

1. El estado inicial se elige aleatoriamente.
2. Repetir hasta que no haya cambios en el vector de estado:
   - El scheduling (controlador) selecciona un valor $d \in \{1, \ldots, \lfloor \frac{N}{2} \rfloor\}$.
   - En paralelo, cada neurona p estudia todas las posibilidades de actualizarse ella misma junto con la neurona $q = (p + d) \mod (N)$, con $0 < q \leq N$.
   - De esta forma, p calcula la matriz de potencial U_p con $K \times K$ componentes.
   - $(U_p)_{(i,j)}$ se asocia al decremento de energía correspondiente a la actualización $s_p = i$, $s_q = j$.

Dinámica Best-2 (cont.)

- Usaremos la siguiente expresión:
  $U_p = -\Delta E = \sum_{i=1}^{N} (\Delta_{i,p} + \Delta_{i,q}) - \Delta_{p,q}$ donde
  $\Delta_{i,j} = w_{i,j}\left(f(s_i, s_j) - f(s_i', s_j')\right)$, $s_i(t) = s_i$, $s_i(t+1) = s_i'$.

Dinámica Best-2 (cont.)

- Usaremos la siguiente expresión:
  $U_p = -\Delta E = \sum_{i=1}^{N} (\Delta_{i,p} + \Delta_{i,q}) - \Delta_{p,q}$ donde
  $\Delta_{i,j} = w_{i,j} \left(f(s_i, s_j) - f(s_i', s_j') \right)$, $s_i(t) = s_i$, $s_i(t+1) = s_i'$.
- La neurona p calcula $\alpha(p) = \text{máx}_{i,j} U_p$, el máximo potencial, asociado a $\vec{\tilde{S}}_{p=i;q=j} \equiv \vec{S}_p$.

UNIVERSIDAD DE MÁLAGA

Dinámica Best-2 (cont.)

- Usaremos la siguiente expresión:
$U_p = -\Delta E = \sum_{i=1}^{N} (\Delta_{i,p} + \Delta_{i,q}) - \Delta_{p,q}$ donde
$\Delta_{i,j} = w_{i,j}\left(f(s_i, s_j) - f(s_i', s_j')\right)$, $s_i(t) = s_i$, $s_i(t+1) = s_i'$.
- La neurona p calcula $\alpha(p) = \max_{i,j} U_p$, el máximo potencial, asociado a $\vec{\tilde{S}}_{p=i;q=j} \equiv \vec{S}_p$.
- El controlador selecciona el siguiente estado de la red:
$\vec{S}(t+1) = \vec{\tilde{S}}_p$ para el cual $p = \arg\max_{\vec{\alpha}}$.

UNIVERSIDAD DE MÁLAGA

MREM Estocástica

Tiene la misma arquitectura que MREM y está basada en la misma función de energía. Pero ahora el siguiente estado de la red dependerá de una sucesión decreciente $\{T_n\}$, de igual forma que el Enfriamiento Estadístico (Simulated Annealing).

Su dinámica consiste en:

- Se genera aleatoriamente el estado inicial $S_1^{(1)}$.

UNIVERSIDAD DE MÁLAGA

MREM Estocástica

Tiene la misma arquitectura que MREM y está basada en la misma función de energía. Pero ahora el siguiente estado de la red dependerá de una sucesión decreciente $\{T_n\}$, de igual forma que el Enfriamiento Estadístico (Simulated Annealing).
Su dinámica consiste en:

- Se genera aleatoriamente el estado inicial $S_1^{(1)}$.
- Dado un estado $\vec{S}$, se elige otro estado $\vec{S'}$ dentro de un entorno $\mathcal{N}_{\vec{S}}$ de $\vec{S}$.

MREM Estocástica

Tiene la misma arquitectura que MREM y está basada en la misma función de energía. Pero ahora el siguiente estado de la red dependerá de una sucesión decreciente $\{T_n\}$, de igual forma que el Enfriamiento Estadístico (Simulated Annealing).

Su dinámica consiste en:

- Se genera aleatoriamente el estado inicial $S_1^{(1)}$.
- Dado un estado $\vec{S}$, se elige otro estado $\vec{S'}$ dentro de un entorno $\mathcal{N}_{\vec{S}}$ de $\vec{S}$.
- Se calcula el incremento de energía $\Delta E = E(\vec{S'}) - E(\vec{S})$.

UNIVERSIDAD DE MÁLAGA

MREM Estocástica

Tiene la misma arquitectura que MREM y está basada en la misma función de energía. Pero ahora el siguiente estado de la red dependerá de una sucesión decreciente $\{T_n\}$, de igual forma que el Enfriamiento Estadístico (Simulated Annealing).
Su dinámica consiste en:

- Se genera aleatoriamente el estado inicial $S_1^{(1)}$.
- Dado un estado $\vec{S}$, se elige otro estado $\vec{S'}$ dentro de un entorno $\mathcal{N}_{\vec{S}}$ de $\vec{S}$.
- Se calcula el incremento de energía $\Delta E = E(\vec{S'}) - E(\vec{S})$.
- La red aceptará el nuevo estado $\vec{S'}$ con probabilidad $\mathbb{P}(\Delta E)$, dependiente de T_n.

UNIVERSIDAD DE MÁLAGA

MREM Estocástica

Tiene la misma arquitectura que MREM y está basada en la misma función de energía. Pero ahora el siguiente estado de la red dependerá de una sucesión decreciente $\{T_n\}$, de igual forma que el Enfriamiento Estadístico (Simulated Annealing).
Su dinámica consiste en:

- Se genera aleatoriamente el estado inicial $S_1^{(1)}$.
- Dado un estado $\vec{S}$, se elige otro estado $\vec{S'}$ dentro de un entorno $\mathcal{N}_{\vec{S}}$ de $\vec{S}$.
- Se calcula el incremento de energía $\Delta E = E(\vec{S'}) - E(\vec{S})$.
- La red aceptará el nuevo estado $\vec{S'}$ con probabilidad $\mathbb{P}(\Delta E)$, dependiente de T_n.
- Este proceso se repite M veces por cada valor de T_n.

Table of contents

Condiciones de Convergencia

Estas dos condiciones garantizan la convergencia a estados de mínima energía:

1. $$\mathbb{P}\left(\mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}) > \mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m}})\right) = \mathbf{0}, \forall \mathbf{m} \in \{\mathbf{1}, \mathbf{2}, \ldots, \mathbf{M} - \mathbf{1}\}$$

Condiciones de Convergencia

Estas dos condiciones garantizan la convergencia a estados de mínima energía:

1.
$$\mathbb{P}\left(\mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}) > \mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m}})\right) = \mathbf{0}, \forall \mathbf{m} \in \{\mathbf{1}, \mathbf{2}, \ldots, \mathbf{M} - \mathbf{1}\}$$

2. La probabilidad de aceptación de $\vec{\mathbf{S}}'$ como $\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}$ es de la forma:
$$\mathbb{P}(\mathbf{\Delta E}) = \begin{cases} \mathbf{1}, & \text{si } \mathbf{\Delta E} < \mathbf{0} \\ \mathbf{g_n}(\mathbf{\Delta E}) < \mathbf{1}, & \text{si } \mathbf{\Delta E} \geq \mathbf{0} \end{cases}$$

UNIVERSIDAD DE MÁLAGA

Condiciones de Convergencia

Estas dos condiciones garantizan la convergencia a estados de mínima energía:

1. $$\mathbb{P}\left(\mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}) > \mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m}})\right) = \mathbf{0}, \forall \mathbf{m} \in \{\mathbf{1}, \mathbf{2}, \ldots, \mathbf{M} - \mathbf{1}\}$$

2. La probabilidad de aceptación de $\vec{\mathbf{S}}'$ como $\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}$ es de la forma:

$$\mathbb{P}(\mathbf{\Delta E}) = \begin{cases} \mathbf{1}, & \text{si } \mathbf{\Delta E} < \mathbf{0} \\ \mathbf{g_n(\Delta E)} < \mathbf{1}, & \text{si } \mathbf{\Delta E} \geq \mathbf{0} \end{cases}$$

donde $g_n(x) > 0 \, \forall x$ y depende de T_n.

UNIVERSIDAD DE MÁLAGA

Condiciones de Convergencia

Estas dos condiciones garantizan la convergencia a estados de mínima energía:

1. $$\mathbb{P}\left(\mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}) > \mathbf{E}(\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m}})\right) = \mathbf{0}, \forall \mathbf{m} \in \{\mathbf{1}, \mathbf{2}, \ldots, \mathbf{M} - \mathbf{1}\}$$
2. La probabilidad de aceptación de $\vec{\mathbf{S}}'$ como $\vec{\mathbf{S}}^{(\mathbf{n})}_{\mathbf{m+1}}$ es de la forma:
$$\mathbb{P}(\boldsymbol{\Delta}\mathbf{E}) = \begin{cases} \mathbf{1}, & \text{si } \boldsymbol{\Delta}\mathbf{E} < \mathbf{0} \\ \mathbf{g_n}(\boldsymbol{\Delta}\mathbf{E}) < \mathbf{1}, & \text{si } \boldsymbol{\Delta}\mathbf{E} \geq \mathbf{0} \end{cases}$$

donde $g_n(x) > 0\,\forall x$ y depende de T_n.

Se puede probar que si $\{g_n\}$ converge uniformemente a 0, entonces la segunda condición implica la primera.

Notas sobre Convergencia

Varios teoremas demuestran la convergencia del modelo MREM Estocástico, con el esquema de muestreo-aceptación aquí propuesto, a un mínimo local de la función de energía E.

Esta convergencia no depende de la tasa de convergencia de la sucesión $\{T_n\}$ y sólo depende de la convergencia de $\{||g_n||\}$ a 0.

Table of contents

UNIVERSIDAD DE MÁLAGA

Grafos

Para probar este método, se han generado 280 grafos aleatorios basados en los siguientes parámetros:

- El número de vértices en el grafo $N \in \{20, 50, 100, 150, 200\}$.

UNIVERSIDAD DE MÁLAGA

Grafos

Para probar este método, se han generado 280 grafos aleatorios basados en los siguientes parámetros:

- El número de vértices en el grafo $N \in \{20, 50, 100, 150, 200\}$.
- La densidad de ejes en el grafo $\rho \in \{0{,}05; 0{,}15; 0{,}25; 0{,}5; 0{,}75; 0{,}9\}$.

UNIVERSIDAD DE MÁLAGA

Grafos

Para probar este método, se han generado 280 grafos aleatorios basados en los siguientes parámetros:

- El número de vértices en el grafo $N \in \{20, 50, 100, 150, 200\}$.
- La densidad de ejes en el grafo $\rho \in \{0{,}05; 0{,}15; 0{,}25; 0{,}5; 0{,}75; 0{,}9\}$.
- Los costes para los ejes se eligieron aleatoriamente en $\{1, \ldots, 10\}$. ($\{1, \ldots, 5\}$ para los grafos de la Tabla 1)

UNIVERSIDAD DE MÁLAGA

Grafos

Para probar este método, se han generado 280 grafos aleatorios basados en los siguientes parámetros:

- El número de vértices en el grafo $N \in \{20, 50, 100, 150, 200\}$.
- La densidad de ejes en el grafo $\rho \in \{0{,}05; 0{,}15; 0{,}25; 0{,}5; 0{,}75; 0{,}9\}$.
- Los costes para los ejes se eligieron aleatoriamente en $\{1, \dots, 10\}$. ($\{1, \dots, 5\}$ para los grafos de la Tabla 1)

Para cada grafo se realizaron 10 ejecuciones independientes. Consideramos $K = 2$ (bipartición), $K = 3$ y $K = 4$.

UNIVERSIDAD DE MÁLAGA

Table of contents

UNIVERSIDAD DE MÁLAGA

		MREM(best-2)		OCHOM		Wang	
N	ρ	Opt	Med	Opt	Med	Opt	Med
20	0.05	42.9	42.79	42.9	42.11	42.9	31.51
	0.25	49	48.71	48.7	47.95	49	35.06
	0.50	54.8	54.42	54.8	53.26	53.9	37.77
	0.75	60.7	60.02	60.7	59.06	60.7	40.88
	0.90	64.4	63.93	63.9	62.84	64.2	46.45
50	0.05	51.1	50.85	51.1	50.1	51	38.95
	0.25	86.8	85.63	86.1	83.59	86.7	51.98
	0.50	120.3	119.26	120.3	117.94	120.1	92.67
	0.75	154.3	152.73	153.5	150.78	154.3	109.52
	0.90	173.7	171.62	173.4	171.11	171.9	82.24
100	0.05	80.4	79.34	79	77	78.8	57.72
	0.25	207.1	203.79	207.4	201.52	202.5	143.26
	0.50	345.7	342.92	345	340.5	345	182.03
	0.75	475	471.45	475	468.29	473.6	333.03
	0.90	534	532.03	535.6	529.86	531.2	323.28
150	0.05	127.9	126.16	124.5	121.95	125.2	101.6
	0.25	386.5	382.89	384.3	377.76	380.2	267.08
	0.50	690.6	685.03	688.6	683.1	686.8	416.94
	0.75	990.1	985.56	990	985.92	990.1	599.64
	0.90	1157.1	1153.15	1154.3	1150.67	1148.8	694.86
200	0.05	194.7	192.32	192.9	187.15	192.6	163.5
	0.25	6075	6009	6020	5966.7	6052	2413.2
	0.50	11260	11194.7	11269	11206.7	11113	6667.8
	0.75	16650	16579.5	16599	16529.1	16498	13198.4
	0.90	19625	19520.8	19555	19439.5	19430	9715

Resultados comparativos para 2-MaxCut

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.
- Tomamos $M = 20$.

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.
- Tomamos $M = 20$.
- d se selecciona de forma decreciente desde $\lfloor \frac{N}{2} \rfloor$ a 1.

UNIVERSIDAD DE MÁLAGA

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.
- Tomamos $M = 20$.
- d se selecciona de forma decreciente desde $\lfloor \frac{N}{2} \rfloor$ a 1.
- La neurona p que se actualiza se selecciona al azar y $q = p + d \mod (N)$.

UNIVERSIDAD DE MÁLAGA

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.
- Tomamos $M = 20$.
- d se selecciona de forma decreciente desde $\lfloor \frac{N}{2} \rfloor$ a 1.
- La neurona p que se actualiza se selecciona al azar y $q = p + d \mod (N)$.
- Se selecciona un estado $\vec{\tilde{S}}_p$ con probabilidad proporcional a $\exp\left((U_p)_{i,j}\right)$ (U_p es la matriz con todos los posibles decrementos de energía si se actualizan p y q).

UNIVERSIDAD DE MÁLAGA

Parámetros de sMREM

- sMREM usa $0 \leq n \leq n_a = 5$ y $T_n = 1 - \frac{n}{n_a}$.
- Tomamos $M = 20$.
- d se selecciona de forma decreciente desde $\lfloor \frac{N}{2} \rfloor$ a 1.
- La neurona p que se actualiza se selecciona al azar y $q = p + d \mod (N)$.
- Se selecciona un estado $\vec{\tilde{S}}_p$ con probabilidad proporcional a $\exp\left((U_p)_{i,j}\right)$ (U_p es la matriz con todos los posibles decrementos de energía si se actualizan p y q).
- Si $\Delta E < 0$, se acepta el estado $\vec{\tilde{S}}_p$, de otra forma, ese estado se acepta con probabilidad

$$g_n(\Delta E) = exp\left(-\frac{\Delta E}{T_n}\right)$$

N	ρ	MREM			sMREM		
		Opt	Med.	t	Opt	Med.	t
50	0.05	276.8	256.28	0.05	276.8	256.8	1.31
50	0.15	672.8	631.56	0.06	687.2	637.2	1.37
50	0.25	1013.2	970.84	0.06	1020.4	971.24	1.32
50	0.5	1778.8	1724.08	0.06	1774.4	1729.26	1.29
50	0.75	2663.6	2475.48	0.05	2646.4	2476.92	1.34
50	0.9	2941.8	2876.18	0.06	2949.2	2883.94	1.33
100	0.05	990.2	917.72	0.15	971.2	920.68	5.58
100	0.15	2384.4	2323.9	0.16	2408.6	2340.46	5.54
100	0.25	3719.2	3620.9	0.14	3720.8	3646.9	5.28
100	0.5	6711.6	6637.08	0.13	6745.6	6683.9	5.41
100	0.75	9816.2	9524.1	0.14	9813.4	9581.22	5.21
100	0.9	11348.8	11215.06	0.14	11434.4	11254.66	5.29
150	0.05	2009.8	1933.6	0.26	2020.2	1939.86	14.48
150	0.15	5136	5014.62	0.26	5120.4	5066.62	13.85
150	0.25	7990	7807.16	0.26	8068.2	7868.72	12.84
150	0.5	14701.4	14531.06	0.24	14818	14661.4	12.61
150	0.75	21126.2	20899.94	0.22	21308.6	21018.06	12.58
150	0.9	24926	24589.62	0.22	25056	24729.68	12.05
200	0.05	3411.4	3321.84	0.38	3443.8	3340.26	27.58
200	0.15	8765.4	8653.52	0.42	8873.6	8726.9	25.51
200	0.25	13741	13533.9	0.35	13959.4	13707.9	24.24
200	0.5	25750.8	25500.18	0.34	25967.2	25740.64	22.41
200	0.75	37038.6	36789.2	0.32	37217.8	37063.42	22.04
200	0.9	43584.8	43296.26	0.33	43854.8	43527.38	23.17

Resultados comparativos para 2-MaxCut

UNIVERSIDAD DE MÁLAGA

N	ρ	MREM			sMREM		
		Opt	Med.	t	Opt	Med.	t
50	0.05	303	276.4	0.1	302.6	276.4	2.13
50	0.15	828	791.88	0.15	833.8	793.04	2.43
50	0.25	1295.4	1228.38	0.14	1303	1228.1	2.44
50	0.5	2371.4	2270.16	0.15	2366.6	2271.5	2.44
50	0.75	3351.8	3245.92	0.14	3350.2	3249.74	2.5
50	0.9	3907	3738.98	0.15	3913.4	3749.02	2.45
100	0.05	1181.2	1084.34	0.32	1183	1085.46	10.6
100	0.15	3130.8	2989.16	0.36	3149.8	2992.54	10.71
100	0.25	4698.8	4645.64	0.36	4705.8	4651.4	10.45
100	0.5	8833.2	8684.56	0.32	8872.4	8719.46	9.36
100	0.75	12758.2	12491.74	0.34	12792.4	12508.1	9.63
100	0.9	14921.4	14695.6	0.33	14948.8	14729.94	9.76
150	0.05	2472	2390.54	0.63	2484.2	2388.3	27.51
150	0.15	6564.6	6434.82	0.66	6574.6	6453.42	29.81
150	0.25	10463.4	10176.84	0.57	10494.6	10203.08	26.53
150	0.5	19274.8	19057.42	0.53	19363.6	19153.9	23.67
150	0.75	27829.4	27599.96	0.55	27974.2	27741.5	22.26
150	0.9	32846.8	32498.44	0.63	32925.6	32611.28	22.96
200	0.05	4360	4212	0.97	4375.8	4216.9	52.46
200	0.15	11353	11146.58	1.01	11429.8	11165.36	53.82
200	0.25	17967.8	17716.1	0.94	18102.4	17808.84	46.85
200	0.5	33712.8	33479.64	0.93	33883.2	33675.92	43.66
200	0.75	49103	48584.72	0.95	49334.6	48808.88	44.27
200	0.9	57865.8	57533.02	0.89	58249.6	57743.24	44.85

Resultados comparativos para 3-MaxCut

N	ρ	MREM			sMREM		
		Opt	Med.	t	Opt	Med.	t
50	0.05	306	281	0.17	306	281	3.25
50	0.15	877.8	817.3	0.24	879	817.18	3.83
50	0.25	1372.8	1314.26	0.29	1372.2	1315.62	4.49
50	0.5	2611.4	2500.92	0.27	2606.6	2501.84	4.01
50	0.75	3691.8	3609.52	0.27	3685	3609.36	4.07
50	0.9	4313.4	4184.84	0.27	4314.6	4184.96	4.12
100	0.05	1190.8	1121.12	0.43	1191.4	1121.38	13.82
100	0.15	3236	3158.6	0.72	3240	3161.56	18.08
100	0.25	5250.2	5127.06	0.7	5259	5136.6	19.82
100	0.5	9803.6	9621.48	0.69	9817.8	9637.76	16.14
100	0.75	14083.4	13886.32	0.61	14082	13907.98	17.68
100	0.9	16640.2	16426.72	0.69	16651	16441.38	17.74
150	0.05	2590.2	2511.96	1	2596.6	2512.4	40.81
150	0.15	7126	6993.94	1.28	7097.8	6985.3	46.8
150	0.25	11466	11141.7	1.25	11499.8	11143.42	46.76
150	0.5	21439.4	21203.94	1.11	21513.4	21230.16	42.37
150	0.75	31254.8	30843.96	1.17	31291.6	30896.64	37.86
150	0.9	36698	36473.34	1.12	36749.2	36533.94	39.46
200	0.05	4503.4	4407.62	1.8	4517.8	4407.98	87.3
200	0.15	12469.6	12289.82	1.92	12495.4	12304.4	89.73
200	0.25	19997.2	19566.3	1.85	19990.4	19586.06	79.62
200	0.5	37621.4	37157.58	1.69	37764	37261.42	74.45
200	0.75	54646	54277.62	1.61	54855.2	54438.5	68.9
200	0.9	64819.6	64298.8	1.8	64950.2	64441.94	71.52

Resultados comparativos para 4-MaxCut

Comentarios sobre los Resultados Experimentales

- Galán, G. & Muñoz, J. propusieron el modelo OCHOM. (IEEE. Trans. Neural Network,2001)

Comentarios sobre los Resultados Experimentales

- Galán, G. & Muñoz, J. propusieron el modelo OCHOM. (IEEE. Trans. Neural Network,2001)
- Wang et al. desarrollaron una versión estocástica de OCHOM que la superaba, pero presenta peores resultados que MREM de media. (Neurocomputing, 2004)

UNIVERSIDAD DE MÁLAGA

Comentarios sobre los Resultados Experimentales

- Galán, G. & Muñoz, J. propusieron el modelo OCHOM. (IEEE. Trans. Neural Network,2001)
- Wang et al. desarrollaron una versión estocástica de OCHOM que la superaba, pero presenta peores resultados que MREM de media. (Neurocomputing, 2004)
- MREM con la dinámica best-2 es un modelo muy potente. (LSCS, 2004), (LNCS, 2005) & ESANN'06.

UNIVERSIDAD DE MÁLAGA

Comentarios sobre los Resultados Experimentales

- Galán, G. & Muñoz, J. propusieron el modelo OCHOM. (IEEE. Trans. Neural Network,2001)
- Wang et al. desarrollaron una versión estocástica de OCHOM que la superaba, pero presenta peores resultados que MREM de media. (Neurocomputing, 2004)
- MREM con la dinámica best-2 es un modelo muy potente. (LSCS, 2004), (LNCS, 2005) & ESANN'06.
- OCHOM y el modelo de Wang son también buenos algoritmos.

UNIVERSIDAD DE MÁLAGA

Table of contents

UNIVERSIDAD DE MÁLAGA

Conclusiones

- En este trabajo se ha mostrado el modelo MREM con la dinámica best-2.

UNIVERSIDAD DE MÁLAGA

Conclusiones

- En este trabajo se ha mostrado el modelo MREM con la dinámica best-2.
- Se ha desarrollado una versión estocástica de MREM con la intención de ayudar a MREM a evitar ciertos mínimos locales de la función de energía.

UNIVERSIDAD DE MÁLAGA

Conclusiones

- En este trabajo se ha mostrado el modelo MREM con la dinámica best-2.
- Se ha desarrollado una versión estocástica de MREM con la intención de ayudar a MREM a evitar ciertos mínimos locales de la función de energía.
- Se han probado las bases teóricas de sMREM para demostrar su convergencia a mínimos de la función objetivo.

UNIVERSIDAD DE MÁLAGA

Conclusiones

- En este trabajo se ha mostrado el modelo MREM con la dinámica best-2.
- Se ha desarrollado una versión estocástica de MREM con la intención de ayudar a MREM a evitar ciertos mínimos locales de la función de energía.
- Se han probado las bases teóricas de sMREM para demostrar su convergencia a mínimos de la función objetivo.
- El modelo sMREM ha sido probado conjuntamente con MREM con el conocido problema MaxCut. Estos experimentos demuestran que en la mayoría de los casos, sMREM supera a MREM en resultados, usando más tiempo de cálculo.

UNIVERSIDAD DE MÁLAGA

Table of contents

UNIVERSIDAD DE MÁLAGA

Trabajo Futuro

- Debemos probar sMREM con otros problemas.

UNIVERSIDAD DE MÁLAGA

Trabajo Futuro

- Debemos probar sMREM con otros problemas.
- Se pueden elegir otras funciones de probabilidad de aceptación que deben ser estudiadas.

Trabajo Futuro

- Debemos probar sMREM con otros problemas.
- Se pueden elegir otras funciones de probabilidad de aceptación que deben ser estudiadas.
- Otras dinámicas (con otro esquema de entornos) para MREM y sMREM pueden obtener mejores resultados para este problema.

UNIVERSIDAD DE MÁLAGA