DETERMINATION OF PARAMETERS IN RELAXATION-SEARCH NEURAL NETWORKS FOR OPTIMIZATION PROBLEMS

Gursel Serpen
Electrical Engineering & Computer Science Department
The University of Toledo, Toledo OH 43606 USA

David L. Livingston
Director of Engineering Division
VWC College
Ranoke, VA

Azadeh Parvin
The Civil Engineering Department
The University of Toledo, Toledo OH 43606 USA

ICNN 97 - Houston, TX
HIGHLIGHTS

♦ A procedure

* Defines constraint weight parameters of the Hopfield network in order to establish the solutions of a given optimization problem as stable equilibrium points in the state space of the network dynamics *

♦ Demonstration

* Application of the methodology is demonstrated on a well known benchmark problem, the Traveling Salesman Problem. *

♦ Simulation Results

* Indicate that the proposed bounds on the constraint weight parameters establish the solutions as stable points and consequently, the Hopfield network consistently converges to a solution after each relaxation.*
INTRODUCTION

◊ Single-layer relaxation-type recurrent neural networks: Hopfield, Mean-Field Annealing, and Boltzmann Machine

Uncertainty in determining values for the parameters of the network

Convergence to INVALID solutions

◊ A procedure to determine values for the parameters of a single-layer recurrent neural network: specifically discrete-dynamics Hopfield network.

It establishes solutions of a given optimization problem as stable equilibrium points in the state space of the network dynamics.
Let \( s_i \) represent a node output where \( s_i = 0,1 \) for \( i = 1,2,\ldots,N \) and \( N \) is the number of network nodes. Then, the equation given by

\[
E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} s_i s_j - \sum_{i=1}^{N} b_i s_i + \sum_{i=1}^{N} \Theta_i s_i, \ i \neq j
\]

is the Liapunov function whose local minima are the final states of the network with node dynamics defined by

\[
s_{i}^{k+1} = \begin{cases} 
0 & \text{if } net_{i}^{k} < \Theta, \\
1 & \text{if } net_{i}^{k} > \Theta, \text{ and} \\
\text{s}_{i}^{k} & \text{if } net_{i}^{k} = \Theta, \ i = 1,2,\ldots,N ,
\end{cases}
\]

where \( k \) is a discrete time index, and

\[
net_{i}^{k} = \sum_{j=1}^{N} w_{ij} s_j^{k} + b_i
\]

with \( i \neq j \) and \( \Theta_i \) is the threshold of node \( s_i \).
The weight term is defined by

\[ w_{ij} = \sum_{\varphi=1}^{Z} g_{\varphi} \delta_{ij}^{\varphi} d_{ij}^{\varphi}, \]

where \( Z \) is the number of constraints.

Given the set of constraints \( C_{\varphi} \in \{C_1, C_2, \ldots, C_Z\} \),

\( g_{\varphi} \in R^+ \) if the hypotheses nodes \( s_i \) and \( s_j \) represent for \( C_{\varphi} \) are mutually supporting and

\( g_{\varphi} \in R^- \) if the same hypotheses are mutually conflicting.

The term \( \delta_{ij}^{\varphi} \) is equal to 1 if the two hypotheses represented by nodes \( s_i \) and \( s_j \) are related under \( C_{\varphi} \) and is equal to 0 otherwise.

The \( d_{ij}^{\varphi} \) term is equal to 1 for all \( i \) and \( j \) under a hard constraint and is a pre-defined cost for a soft constraint, which is typically associated with a cost term in optimization problems.
TRAVELLING SALESPERSON PROBLEM

Consider the following energy function proposed by Hopfield to map this problem to the network topology,

\[
E(s) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} g_r \delta^r_{ij} s_i s_j - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} g_c \delta^c_{ij} s_i s_j - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} g_\eta d_{ij} \delta^\eta_{ij} s_i s_j,
\]

where

\[
\delta^r_{ij} = 1 \text{ if } \text{row}(i) = \text{row}(j) \text{ or } \delta^r_{ij} = 0, \text{ otherwise;}
\]

\[
\delta^c_{ij} = 1 \text{ if } \text{col}(i) = \text{col}(j) \text{ or } \delta^c_{ij} = 0, \text{ otherwise;}
\]

\[
\delta^\eta_{ij} = 1 \text{ if } \left[|\text{col}(i) - \text{col}(j)| = 1 \land \text{row}(i) \neq \text{row}(j)\right] \text{ or } \delta^\eta_{ij} = 0, \text{ otherwise;}
\]

\(i \neq j; g_r, g_c, g_\eta \in R^-, d_{ij} \) is the distance between cities \text{row}(i) and \text{row}(j), and superscripts/subscripts \( r, c, \gamma \) and \( \eta \) stand for row, column, global and distance inhibitions, respectively.

Functions \text{row}(i) and \text{row}(j) return the row and column location of nodes \( i \) and \( j \), respectively.
The row and column inhibition energy terms are minimum if, at most, one node is active for each row and column.

The energy term associated with the constraint weight parameter $g_\gamma$ is minimum if exactly $M$ nodes are active within the overall network topology.

The energy term for the distance constraint is minimum if the solution has the minimum total distance.

Comparison of this energy function with the generic energy function yields the following weight matrix entries, external bias terms, and thresholds:

$$w_{ij} = g_r \delta^r_{ij} + g_c \delta^c_{ij} + g_\eta \delta^\eta_{ij} d_{ij} + g_\gamma,$$

$$b_i = \frac{1}{2} (1 - 2M) g_\gamma$$ and

$$\Theta_i = 0 \text{ for } i, j = 1, 2, \cdots, N; \ i \neq j$$
Consider the input to an inactive node within a solution \((\text{net}_i < \Theta_i)\).

- An inactive node receives inputs from two active nodes under the row and the column inhibition constraints.
- A total of \(M\) active nodes collectively contribute \(Mg_\gamma\) to the input of an inactive node due to global inhibitory interaction.
- An inactive node receives inputs from two active nodes in two neighboring columns under the inhibitory distance constraint.
- Note also that \(\Theta_i = 0\) and \(b_i = 0.5(1 - 2M)g_\gamma\) for \(i = 1, 2, \ldots, N\).

As a result, the inequality for an inactive node to be stable takes the form of

\[
|g_r| + |g_c| + 2g_\eta d_{\min} + \frac{1}{2}|g_\gamma| \geq 0.
\]

where \(d_{\min} \leq d_i\) for \(i = 1, 2, \ldots, N\).
An active node: \( \text{net}_i > \Theta_i \), for \( i = 1, 2, \ldots, N \).

* No nodes contribute to the input of the active node under the column and row inhibition constraints.
* Other active nodes contribute \((M-1)g\gamma\).
* Two active nodes within the neighboring columns under inhibitory distance constraint contribute.
* We also have \( \Theta_i = 0 \) and \( b_i = 0.5(1-2M)g\gamma \) for \( i = 1, 2, \ldots, N \).

Then the inequality for the active node becomes

\[
4 |g\eta| d_{\text{max}} \leq |g\gamma|
\]
The inequality for an inactive node to be stable takes the form of

\[ |g_r| + |g_c| + 2|g_\eta|d_{\min} + \frac{1}{2}|g_\gamma| \geq 0 \]

where \( d_{\min} \leq d_i \) for \( i = 1, 2, \ldots, N \).

This inequality is satisfied for any values of the constraint weight parameter magnitudes.

The inequality for the active node in a solution to be stable

\[ 4|g_\eta|d_{\max} \leq |g_\gamma| \]

This inequality is the only one which needs to be satisfied for solutions to be the stable points of the network dynamics.
SIMULATION RESULTS

* The TSP was mapped to discrete network dynamics for the problem sizes of 10, 20, 30, 40, and 50 cities and for three distinct operating points.

* The distances between cities are random variables uniform in the interval [0, 1].

* A total of 100 relaxations were realized for each problem size and operating point pair.

* The operating points employed in the experiments were chosen at the limiting points of the admissible subspace of the constraint weight parameter space. Specifically, small and large values for the difference given by

\[
\left| g_\gamma \right| - 4 \left| g_\eta \right| d_{\text{max}}
\]

were used as the test values for the operating points.
SIMULATION RESULTS
(cont’d)

<table>
<thead>
<tr>
<th>Operating Points</th>
<th>(g_r)</th>
<th>(g_c)</th>
<th>(g_\eta)</th>
<th>(g_\gamma)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-0.1</td>
<td>-1.0</td>
</tr>
<tr>
<td>2</td>
<td>-100.0</td>
<td>-100.0</td>
<td>-0.1</td>
<td>-1.0</td>
</tr>
<tr>
<td>3</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-0.1</td>
<td>-100.0</td>
</tr>
</tbody>
</table>

Table 1. Operating Point Definitions.

<table>
<thead>
<tr>
<th>Cities</th>
<th>Operating Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP 1</td>
<td>OP 2</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
</tr>
<tr>
<td>20</td>
<td>100%</td>
</tr>
<tr>
<td>30</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>100%</td>
</tr>
<tr>
<td>50</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2. Convergence Rate vs. Operating Point × TSP Size.

The results show the network converged to a solution after each relaxation in all test cases.
The quality of the solutions, the measure of which is the Normalized Total Distance (NTD), was observed for various problem sizes.

The NTD is computed by dividing the total distance of a solution by the number of cities. The problem sizes used in the tests are 10, 30, and 50 cities.

A total of 100 relaxations were performed for each problem size.

Figure. Quality of Solutions vs. Problem Size.
CONCLUSIONS

♦ Simulation study conducted on TSP for up to 50 cities indicated that the proposed bounds on constraint weight parameters led the Hopfield network towards a solution after each relaxation.

♦ The procedure essentially made the solution point set equal to stable point set for the TSP.

♦ The same procedure generally establishes the solution point set as a subset of the stable point set: some non-solution points become stable and the network might converge to a nonsolution in that case.