2006 International Joint Conference on Neural Networks
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

# A Columnar Competitive Model with Simulated Annealing for Solving Combinatorial Optimization Problems

Eu Jin Teoh, *Student Member, IEEE*, Huajin Tang, *Member, IEEE*, and Kay Chen Tan, *Member, IEEE*

***Abstract*— One of the major drawbacks of the Hopfield network is that when it is applied to certain polytopes of combinatorial problems, such as the traveling salesman problem (TSP), the obtained solutions are often invalid, requiring numerous trial-and-error setting of the network parameters thus resulting in low-computation efficiency. With this in mind, this article presents a columnar competitive model (CCM) which incorporates a winner-takes-all (WTA) learning rule for solving the TSP. Theoretical analysis for the convergence of the CCM shows that the competitive computational neural network guarantees the convergence of the network to valid states and avoids the tedious procedure of determining the penalty parameters. In addition, its intrinsic competitive learning mechanism enables a fast and effective evolving of the network. Simulation results illustrate that the competitive model offers more and better valid solutions as compared to the original Hopfield network.**

***Index Terms*— Competitive learning, simulated annealing, combinatorial optimization, traveling salesman problem.**

## I. INTRODUCTION

From the view of mathematical programming, the TSP can be described as a quadratic 0-1 programming problem with linear constraints,

$$\text{minimize} \quad E^{obj}(\mathbf{v}) \tag{1}$$

$$\text{subject to} \quad S_i = \sum_{x=1}^{n} v_{xi} = 1 \quad \forall i \in \{1, \cdots, n\}, \tag{2}$$

$$S_x = \sum_{i=1}^{n} v_{xi} = 1 \quad \forall x \in \{1, \cdots, n\}, \tag{3}$$

and a redundant constraint $S = \sum_x \sum_i v_{x,i} = n$, where $v_{xi} \in \{0, 1\}$ and $E^{obj}$ is the total tour length described by a valid 0-1 solution $\mathbf{v}$. The energy function to be minimized in the network is

$$E(\mathbf{v}) = E^{obj}(\mathbf{v}) + E^{cns}(\mathbf{v}), \tag{4}$$

where $E^{cns}$ is the constraints described by (2) and (3).

In his seminal work, Hopfield [11] solved the TSP, as a specific instance of combinatorial optimization problems using a highly-interconnected network of nonlinear analog neurons. However, the convergence of network to valid states and preferably quality ones depended heavily on the values of penalty terms in the energy function. It required careful setting of these parameters in order to obtain valid quality solutions, which is often a difficult and trial-and-error procedure.

It has been a continuing research effort to improve the performance of Hopfield network since its origination [2]. The behavior of Hopfield network was analyzed based on the eigenvalues of connection matrix [1] and the parameter settings for TSP was derived. The local minimum escape (LME) algorithm [12] was proposed to improve the local minimum by combining the network disturbing technique with the Hopfield network's local minima search property. Most recently, a parameter setting procedure based on the stability conditions of the energy function was presented [3]. Although these methods have been successful to some extent for improving the quality and validity of solutions, spurious states were often generated. Moreover, the existing methods require a large volume of computational resources, which restricts their practical applications.

## II. COLUMNAR COMPETITIVE MODEL

The dynamics of Hopfield networks can be described by a system of differential equations and the activation function is a hyperbolic tangent. Let $\mathbf{v}, \mathbf{i^b}$ be the vectors of neuron activities and biases, and $\mathbf{W}$ be the connection matrix, then the energy function of the Hopfield network for the high-gain limit expression exists such that

$$E = -\frac{1}{2}\mathbf{v}^T \mathbf{W} \mathbf{v} - (\mathbf{i^b})^T \mathbf{v}. \tag{5}$$

Hopfield has shown that the network will converge to local minima of energy function (5) if $\mathbf{W}$ is symmetric [11].

In this section, the columnar competitive model (CCM) which is constructed by incorporating WTA into the network in column-wise is introduced. Competitive learning by winner-takes-all (WTA) has been recognized to play an important role in many areas of computational neural networks, such as feature discovery and pattern classification [4], [9], [10], [6]. Nevertheless, the potential of WTA as a means of eliminating all spurious states is seen due to its intrinsic competitive nature that can elegantly reduce the number of penalty terms, and hence the constraints of the network for optimization. The WTA mechanism can be described as: given a set of $n$ neurons, the input to each neuron is calculated and the neuron with the maximum input value is declared the winner. The winner's output is set to '1' while the remaining neurons will have their values set to '0'.

The neurons of the CCM are evolved based on the WTA learning rule. The competitive model handles the columnar

Eu Jin Teoh and Kay Chen Tan are with the Department of Electrical and Computer Engineering, National University of Singapore, 4 Engineering Drive 3, Singapore 117576 (email: eletankc@nus.edu.sg).

Huajin Tang is with ST Microelectronics Singapore, (email: huajin.tang@st.com).

constraints (2) elegantly, due to its competitive property of ensuring one '1' per column. As in two dimensional forms, $\mathbf{v} = \{v_{x,i}\}, \mathbf{i^b} = \{i^b_{x,i}\}$, where the subscript $x, i \in \{1, \ldots, n\}$ denotes the city index and the visit order, respectively. The strength of connection between neuron $(x, i)$ and neuron $(y, j)$ is denoted by $W_{xy,ij}$. Hence, the associated energy function can be written as

$$E(\mathbf{v}) = \frac{K}{2} \sum_x \sum_i (v_{x,i} \sum_{j \neq i} v_{x,j}) + \frac{1}{2} \sum_x \sum_{y \neq x} \sum_i d_{xy} v_{x,i} (v_{y,i+1} + v_{y,i-1}), \quad (6)$$

where $K > 0$ is a scaling parameter and $d_{xy}$ is the distance between cities $x$ and $y$. Comparing (5) and (6), the connection matrix and the external input of the network are computed as follows,

$$W_{xi,yj} = -\{K\delta_{xy}(1 - \delta_{ij}) + d_{xy}(\delta_{i,j+1} + \delta_{i,j-1})\} \quad (7)$$
$$\mathbf{i^b} = 0, \quad (8)$$

where $\delta_{i,j}$ is the Kronecker's delta.

The input to a neuron $(x, i)$ is calculated as

$$\begin{aligned} Net_{x,i} &= \sum_y \sum_j (W_{xi,yj} v_{yj}) + \mathbf{i^b} \\ &= -\sum_y d_{xy}(v_{y,i-1} + v_{y,i+1}) - K \sum_{j \neq i} v_{x,j} \quad (9) \end{aligned}$$

The WTA is applied based on a column-by-column basis, with the winner being the neuron with the largest input. The WTA updating rule is thus defined as

$$v_{x,i} = \begin{cases} 1, & \text{if } Net_{x,i} = \max\{Net_{1,i}, Net_{2,i}, \cdots, Net_{n,i}\} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

Hereafter, $v_{x,i}$ is evaluated by the above WTA rule. The algorithm of implementing the competitive model is summarized as follows:

**Competitive Model Algorithm**

1) Initialize the network, with each neuron having a small initial value $v_{x,i}$. A small random noise is added to break the initial network symmetry. Compute the $W$ matrix using (7).
2) Select a column (e.g., the first column). Compute the input $Net_{x,i}$ of each neuron in that column.
3) Apply WTA using (10), and update the output state of the neurons in that column.
4) Go to the next column, preferably the one immediately on the right for the convenience of computation. Repeat step 3 until the last column in the network is done. This constitutes the first epoch.
5) Go to step 2 until the network converges (i.e., the states of the network stop changes).

## III. Convergence of Competitive Model and Full Valid Solutions

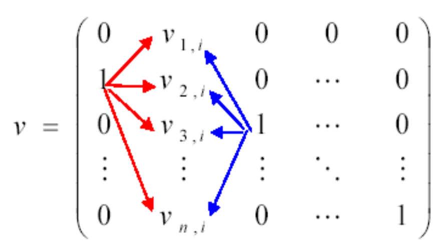
For the energy function (6) of CCM, the critical value of the penalty-term scaling parameter $K$ plays a predominant role in ensuring its convergence and driving the network to converge to valid states. Meanwhile, it is known that the stability of the original Hopfield networks is guaranteed by the well-known Lyapunov energy function. However, the dynamics of the CCM is so different from the Hopfield network, thus the stability of the CCM needs to be investigated.

In this section, our efforts are devoted to such two objectives, i.e., determining the critical value of $K$ that ensures the convergence of full valid solutions and proving the convergence of the competitive networks under the WTA updating rule. The following theorem is drawn.

*Theorem 1:* Let $K > 2d_{max} - d_{min}$, where $d_{max}$ and $d_{min}$ is the maximum and the minimum distance, respectively. Then the competitive model defined by (6)–(10) is always convergent to valid states.

*Proof:* The proof is composed of two parts based on the two objectives respectively.

**(i)** While the WTA ensures that there can only be one '1' per column, it does not guarantee the same for each row. Thise responsibility lies in the parameter $K$ of the penalty term. Without a loss of generality, it is assumed that after some updating the network reaches the following state, which is given by



Let the first row be an all-zero row. The input to each neuron in the $i$-th column is computed by

$$\begin{aligned} Net_{1,i} &= -(d_{12} + d_{13}), \\ Net_{2,i} &= -(K + d_{23}), \\ Net_{3,i} &= -(K + d_{23}), \\ Net_{n,i} &= -(K + d_{2n} + d_{3n}). \end{aligned}$$

For valid state, $v_{1,i}$ occupying the all-zero row will have to be the winner, i.e., $Net_{1,i} = \max_{x=1,\ldots,n} \{Net_{x,i}\}$. Therefore, it is verified that

$$\begin{aligned} d_{12} + d_{13} &< K + d_{23}, \\ d_{12} + d_{13} &< K + d_{2n} + d_{3n}. \end{aligned}$$

To satisfy both conditions, it is sufficient for $K$ to satisfy

$$K > d_{12} + d_{13} - d_{23}, \quad (11)$$

since $d_{2n} + d_{3n} > d_{23}$ straightforwardly.

Let $d_{max} = \max\{d_{xy}\}$, $d_{min} = \min\{d_{xy}\}$. Firstly, assume the 'worst' scenario for deriving (11), i.e., $d_{12} = d_{13} = d_{max}$, $d_{23} = d_{min}$, it holds

$$K > 2d_{max} - d_{min}. \quad (12)$$

Secondly, assume the 'best' scenario, i.e., $d_{12} = d_{13} = d_{min}$, $d_{23} = d_{max}$, the following is obtained

$$K > 2d_{min} - d_{max}. \quad (13)$$

Obviously, condition (12) is the sufficient condition for guaranteed convergence to fully valid solutions, while (13) is the necessary condition for convergence to some valid solutions. Although a specific case has been assumed, the results obtained are regardless of the specific case.

**(ii)** To investigate the dynamical stability of the CCM, a $n \times n$ network (for $n$ cities) is considered. After the $n$-th WTA updating, the network would have reached the state with only one '1' per column, but may have more than one '1' per row.

Suppose $\mathbf{v}^t$ and $\mathbf{v}^{t+1}$ are two states before and after WTA updating respectively. Consider $p$-th column, and let neuron $(a, p)$ be the only active neuron before updating, i.e., $v^t_{a,p} = 1$ and $v^t_{i,p} = 0, \forall i \neq a$. After updating, let neuron $(b, p)$ be the winning neuron, i.e., $v^{t+1}_{b,p} = 1$, $v^{t+1}_{i,p} = 0, \forall i \neq b$.

The energy function (6) can be further broken into two terms $E_p$ and $E_o$, i.e., $E = E_p + E_o$, where $E_p$ stands for the energy of the consecutive columns $p-1, p$ and $p+1$ and of the rows $a$ and $b$. $E_o$ stands for the energy of the rest columns and rows.

$E_p$ is calculated by

$$E_p = \frac{K}{2}\left(\sum_i v_{a,i}\sum_{j\neq i} v_{a,j} + \sum_i v_{b,i}\sum_{j\neq i} v_{b,j}\right) + \sum_x\sum_y d_{xy}v_{x,p}(v_{y,p+1} + v_{y,p-1}). \quad (14)$$

Accordingly, $E_o$ is computed by

$$\begin{aligned} E_o \quad = \quad & \frac{K}{2}\sum_{x\neq a,b}\sum_i (v_{x,i}\sum_{j\neq i} v_{x,j}) \\ & + \frac{1}{2}\sum_x\sum_y\sum_{i\neq p-1,p,p+1} d_{xy}v_{x,i}(v_{y,i+1} + v_{y,i-1}) \\ & + \frac{1}{2}\sum_x\sum_y d_{xy}v_{x,p-1}v_{x,p-2} \\ & + \frac{1}{2}\sum_x\sum_y d_{xy}v_{x,p+1}v_{x,p+2}. \end{aligned} \quad (15)$$

Thus, it can be seen that only $E_p$ will be affected by the states of the column $p$. It is assumed that the neuron $(a, p)$ is the only active neuron in the $p$-th column before updating, i.e., $v^t_{a,p} = 1$ and $v^t_{i,p} = 0$ for all $i \neq a$, and the neuron $(b, p)$ wins the competition after updating, i.e., $v^{t+1}_{b,p} = 1$ and $v^{t+1}_{i,p} = 0$ for all $i \neq b$. Let $E^t$ and $E^{t+1}$ be the energy before updating and after updating, respectively. To investigate how $E$ changes under the WTA learning rule, the following two cases are considered.

Case 1: Row $a$ contains $m$ '1' ($m > 1$), row $b$ is an all-zero row.

According to equation (9), the input to neuron $(a, p)$ and $(b, p)$ is computed by (16) and (17), respectively.

$$Net_{a,p} = -K\sum_{j\neq p} v_{a,j} - \sum_y d_{ay}(v_{y,p-1} + v_{y,p+1}) \quad (16)$$

$$Net_{b,p} = -K\sum_{j\neq p} v_{b,j} - \sum_y d_{by}(v_{y,p-1} + v_{y,p+1}) \quad (17)$$

Considering the WTA learning rule (10), it leads to $Net_{b,p} = \max_i\{Net_{i,p}\}$. It implies that

$$K\sum_{j\neq p} v_{b,j} + \sum_y d_{by}(v_{y,p-1} + v_{y,p+1}) < K\sum_{j\neq p} v_{a,j} + \sum_y d_{ay}(v_{y,p-1} + v_{y,p+1}) \quad (18)$$

Obviously, $\sum_{j\neq p} v_{b,j} = 0$, $\sum_{j\neq p} v_{a,j} = m - 1$. Therefore, it is derived from (18) that

$$\sum_y d_{by}(v_{y,p-1} + v_{y,p+1}) - \sum_y d_{ay}(v_{y,p-1} + v_{y,p+1}) < K(m-1) \quad (19)$$

Now considering the energy function (14), $E^t_p$ and $E^{t+1}_p$ is computed by (20) and (20), respectively.

$$E^t_p = \frac{K}{2}m(m-1) + \sum_y d_{ay}v_{a,p}(v_{y,p+1} + v_{y,p-1}).$$

$$E^{t+1}_p = \frac{K}{2}(m-1)(m-2) + \sum_y d_{by}v_{b,p}(v_{y,p+1} + v_{y,p-1})$$

Thus, it is clear that

$$\begin{aligned} E^{t+1}_p - E^t_p \quad = \quad & -K(m-1) + \\ & \sum_y d_{by}v_{b,p}(v_{y,p+1} + v_{y,p-1}) \\ & - \sum_y d_{ay}v_{a,p}(v_{y,p+1} + v_{y,p-1}) \end{aligned} \quad (20)$$

Recall equation (19), it is obtained $E^{t+1}_p - E^t_p < 0$. It implies that

$$E^{t+1} - E^t < 0 \quad (21)$$

Case 2: Row $a$ contains only one '1', row $b$ is an all-zero row.

According to the WTA updating rule, again it holds that $Net_{b,p} = \max_i \{Net_{i,p}\}$. Similar to case 1, it is obtained

$$K \sum_{j \neq p} v_{b,j} + \sum_y d_{by}(v_{y,p-1} + v_{y,p+1}) < \\ K \sum_{j \neq p} v_{a,j} + \sum_y d_{ay}(v_{y,p-1} + v_{y,p+1}) \quad (22)$$

Obviously, $\sum_{j \neq p} v_{b,j} = 0$, $\sum_{j \neq p} v_{a,j} = 0$. Therefore, it is obtained that

$$\sum_y d_{by}(v_{y,p-1} + v_{y,p+1}) - \\ \sum_y d_{ay}(v_{y,p-1} + v_{y,p+1}) < 0. \quad (23)$$

On the other hand, $E_p^t$ and $E_p^{t+1}$ are now computed as follows,

$$E_p^t = \sum_y d_{ay} v_{a,p}(v_{y,p+1} + v_{y,p-1}) \quad (24)$$

$$E_p^{t+1} = \sum_y d_{by} v_{b,p}(v_{y,p+1} + v_{y,p-1}) \quad (25)$$

Again, it is obtained

$$E^{t+1} - E^t < 0. \quad (26)$$

Based on the above two cases, the CCM is always convergent under the WTA updating rule. This completes the proof. ■

It is noted that the WTA tends to drive the rows with more than one active neurons to reduce the number of active neurons, and drive one of the neurons in all-zero rows to become active. Once there is no row that contains more than one active neurons (equivalently there is no all-zero row), then a valid state is reached and the state of the network stops the transition (in this case $E^{t+1} - E^t = 0$).

## IV. Simulated Annealing Applied to Competitive Model

Simulated Annealing (SA), a stochastic process, is known to improve the quality of solutions when solving combinatorial optimization problems [7], [8]. It derives its name from an analogy of thermodynamics and metallurgy (the interested reader is directed to [8]) – th metal is first heated to a high temperature $T_0$, causing the metal atoms to vibrate vigorously, hence resulting in a highly disordered structure. The metal is then cooled slowly, allowing the atoms to rearrange themselves in an orderly fashion, which in turn corresponds to a structure of minimum energy.

Optimum or near optimum results can usually be obtained, but at the expense of long computational time, due to the slow convergence of the SA algorithm. WTA, on the other hand, offers fast convergence, but produces solutions that are of lower quality. Wth this inmind, it is beneficial to combine these 2 techniques, such that a fully valid solution set, preferably one with mostly optimum or near optimum tour distances, can be obtained within a reasonably short time.

When $K$ is at least $2d_{max} - d_{min}$, all states are valid, but the solutions are of average quality (since there are fewer states with an optimum tour distance). As $K$ decreases, it becomes more probable for the network to converge to states of optimum tour distance, but there is also a corresponding increase in the number of spurious states. When $K$ is less than $2d_{min} - d_{max}$, the network never converges to valid states, regardless of the initial network states. Therefore, by having a small $K$ and slowly increasing it, the quality of solutions can be increased, while preserving the validity of all states at the same time. The algorithm in conjunction with the simulated annealing schedule for the parameter $K$ can be implemented in the following way:

**Competitive Model with SA**

1) Initialized the network, with $v_{x,i}$ having a small initial value, added with small random noise. Initialize $K = d_{max}$ and $\epsilon > 0$ that determines how fast $K$ reaches $2d_{max} - d_{min}$.
2) Do $N$ times: Select a column $i$ to compute the input $Net_{x,i}$ for each neuron in that column, then apply WTA and update the output of the neurons. This constitutes a whole epoch.
3) Increase $K$ by

$$K = d_{max} + (0.5 \cdot \tanh(t - \epsilon) + 0.5)(d_{max} - d_{min}),$$

where $t$ is the current epoch number. Go to step 2 until convergence.

## V. Simulation Results

It is known that the solutions of double-circle problem are very hard to obtain by the original Hopfield network [5]. In this section, the 24-city example that was analyzed in [5] is employed to verify the theoretical results. There are 12 cities uniformly distributed in the outer circle, while the rest are allocated in the inner circle, with the radius equals to 2 and 1.9, respectively. In this work, 500 simulations have been performed for each case of Hopfield network and CCM. The simulation results are given in Tables I and II. The item good indicates the number of tours with distance within 150% of the optimum distance. Figures 1 and 2 depict the optimum tour and the near-optimum tour generated by the CCM.

TABLE I
The performance of original Hopfield model for the 24-city example

| $C$ | Valid | Invalid | Good | Minimum length | Average length |
|---|---|---|---|---|---|
| 10 | 302 | 198 | 111 | 35.2769 | 50.9725 |
| 1 | 285 | 215 | 108 | 26.2441 | 37.3273 |
| 0.1 | 366 | 134 | 168 | 25.2030 | 36.3042 |
| 0.01 | 360 | 140 | 159 | 27.5906 | 36.9118 |
| 0.001 | 372 | 128 | 153 | 21.6254 | 32.5432 |

TABLE II

THE PERFORMANCE OF CCM FOR THE 24-CITY EXAMPLE

| $K$ | Valid | Invalid | Good | Minimum length | Average length |
|---|---|---|---|---|---|
| $d_{max}$ | 438 | 62 | 248 | 13.3220 | 18.1172 |
| $d_{max}$ $+d_{min}$ | 453 | 47 | 256 | 13.3220 | 17.9778 |
| $2d_{max}$ $+d_{min}$ | 500 | 0 | 18 | 15.1856 | 24.1177 |

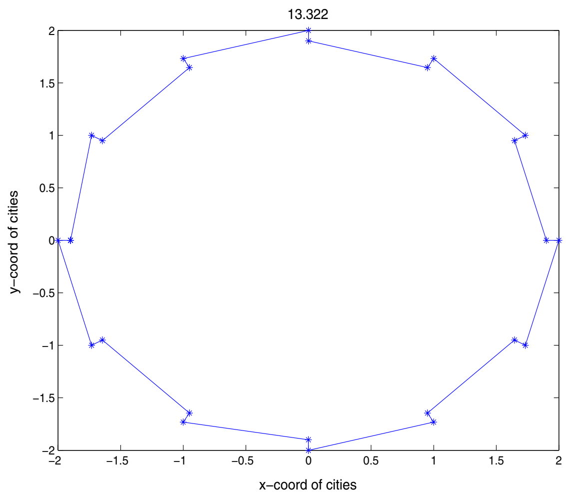



Fig. 1. Optimum tour generated by CCM

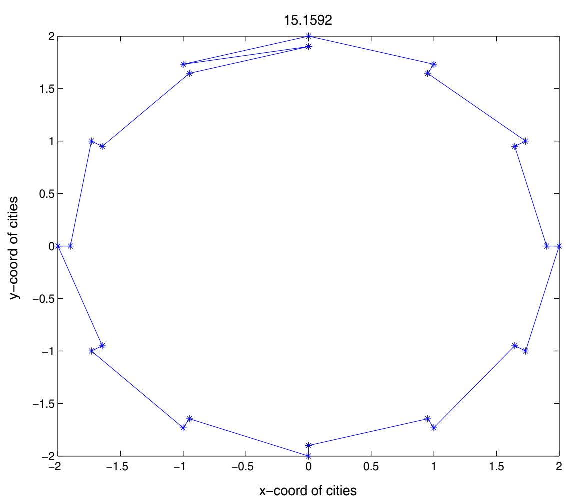



Fig. 2. Near-optimum tour generated by CCM

As can be seen from the simulation results, the original Hopfield network generated a large number of invalid solutions with respect to various weighting factors $C$. Its minimum and average distances are also far away from the optimum solution of $13.322$. Comparatively, the new competitive model generated less spurious states and produced better quality solutions in terms of the minimum and average tour lengths. In addition, the CCM has a rapid convergence rate of less than $5$ epochs on the average. It can be seen that when $K$ was set to a small value (e.g., $d_{max}$), with a small portion of invalid solutions, the CCM can easily reach the optimum solution. When $K$ was increased to $2d_{max} - d_{min}$, all the spurious states were eliminated, which was achieved at minor expense of the solution quality. Obviously, these findings are in agreement with the theoretical results obtained.

With SA applied to the new WTA model, its performance increases significantly III. All the states generated are valid, and has approximately $65\%$ good solutions (optimum or near optimum states) for all cases. The average tour distance is about $18$, which is much better than the Hopfield model and that of the WTA model (without SA). Its convergence rate is slightly slower than the WTA model (without SA) and requires $9$ epoches on the average. Nevertheless, it is much faster that of the Hopfield model.

TABLE III

THE PERFORMANCE OF CCM WITH SA FOR THE 24-CITY EXAMPLE

| $\epsilon$ | Valid | Invalid | Good | Minimum length | Average length |
|---|---|---|---|---|---|
| 0 | 500 | 0 | 18 | 15.1856 | 24.1177 |
| 1 | 500 | 0 | 203 | 13.3404 | 22.2688 |
| 2 | 500 | 0 | 218 | 13.3220 | 19.9771 |
| 3 | 500 | 0 | 259 | 13.3220 | 18.7153 |
| 5 | 500 | 0 | 256 | 13.3220 | 18.6785 |
| 10 | 500 | 0 | 301 | 13.3220 | 18.4080 |
| 15 | 500 | 0 | 307 | 13.3220 | 18.3352 |
| 20 | 500 | 0 | 289 | 13.3220 | 18.4643 |

The theoretical results are also validated upon the 48-city example which has a similar coordinates configuration as the 24-city one. 100 experiments are performed for each case and the results are given in Table IV. It can be seen that when $K$ was increased from $d_{max}$ to $2d_{max} - d_{min}$, all invalid solutions were suppressed with the minimum and average lengths increased, which is consistent with the findings observed in Table IV.

TABLE IV

THE PERFORMANCE OF CCM FOR THE 48-CITY EXAMPLE

| $K$ | Valid | Invalid | Good | Minimum length | Average length |
|---|---|---|---|---|---|
| $d_{max}$ | 90 | 10 | 0 | 24.6860 | 32.7950 |
| $d_{max}$ $+d_{min}$ | 92 | 8 | 0 | 24.9202 | 33.8892 |
| $2d_{max}$ $-d_{min}$ | 100 | 0 | 0 | 35.9438 | 42.4797 |

*Remark 1:* Since for non-proper values of $K$ (e.g., $d_{max}$ and $d_{max} + d_{min}$), the convergence of the CCM is not guaranteed, thus for the sake of comparing the performance of the CCM with different parameter settings, only the results of the convergent cases in the experiments recorded are recorded. It is also noted that when the coordinates configuration becomes more complex, good solutions become more difficult to be achieved as well.

The WTA with SA model is now applied to TSP of various city sizes, with their city topology similar to that of 24 cities, i.e., their coordinates evenly distributed around 2 circles of radius 1.9 and 2.0. The $\epsilon$ is set to 5 for the simulations. The results are shown in Table V and VI. The performance ratio is defined to be the ratio of the minimum length to the average length of the tours found, i.e. (Perf. ratio=Minimum length/Average length).

TABLE V

THE PERFORMANCE OF CCM WITH SA FOR VARIOUS CITY SIZES

| City size | Simulations | Valid | Invalid |
|---|---|---|---|
| 30 | 250 | 250 | 0 |
| 100 | 250 | 250 | 0 |
| 500 | 100 | 100 | 0 |
| 1000 | 50 | 50 | 0 |
| 3000 | 5 | 5 | 0 |

TABLE VI

THE PERFORMANCE OF CCM WITH SA FOR VARIOUS CITY SIZES (CONTINUATION FROM TABLE ABOVE)

| City size | Minimum length | Average length | Perf. ratio |
|---|---|---|---|
| 30 | 13.6922 | 22.3159 | 0.6136 |
| 100 | 28.2552 | 35.6920 | 0.7916 |
| 500 | 257.1380 | 283.2810 | 0.9060 |
| 1000 | 454.3827 | 490.6963 | 0.9260 |
| 3000 | 1395.5 | 1454.2 | 0.9455 |

The proposed approach of using a WTA with SA model is seen to perform satisfactory, with an average performance ratio of 0.85. Less computing resources are used, thereby allowing city sizes of up to 3000 to be successfully solved using the proposed approach. This is quite remarkable, as none of the existing algorithms proposed in the literature has been able to solve such a large city size in practice using a computational intelligence framework, primarily due to the extensive amount of time and resources required.

## VI. CONCLUSION

In this article, a new columnar competitive model (CCM) incorporating the WTA learning rule has been proposed for solving the combinatorial optimization problems, which has guaranteed convergence to valid states and total suppression of spurious states in the network. Consistently good results, better than that of the Hopfield network in terms of both the number as well as quality of valid solutions were obtained using computer simulations. It also has the additional advantage of faster convergence, while utilizing comparatively lower computing resources, thereby allowing it to solve large-scale combinatorial problems. The dynamics analysis of the CCM implied that the CCM is incapable of hill-climbing transitions and thus being trapped at local minima, which is still a problem and a possible direction for future work. With the full validity of solutions ensured by competitive learning, probabilistic optimization techniques such as LME [12] can be incorporated into the competitive model to further improve the solution quality. Simulated annealing [7], [8], when incorporated into the proposed approach showed even better results. In addition, with some modification to the energy function, the competitive model can also be extended to solve other polytopes of combinatorial optimization problems, such as the 4-color map.

## REFERENCES


[1] S.V.B. Aiyer, M. Niranjan and F. Fallside, A theoretical investigation into the performance of the Hopfield network *IEEE Trans. Neural Networks*, vol. 1(2), pp. 204-215, 1990.

[2] K.A. Smith, Neural networks for combinatorial optimization: A review of more than a decade of research, *INFORMS Journal on Computing*, vol. 11(1), pp. 15-34, 1999.

[3] P. M. Talavn and J. Yez, *Parameter setting of the Hopfield network applied to TSP*, Neural Netw., vol. 15, pp. 363373, 2002.

[4] D.E. Rumelhart and D. Zipser, Feature discovery by competitive learning, *Cognitive Science*, vol. 9, pp. 75-112, 1985.

[5] Y. Tachibana et al., A simple steepest descent method for minimizing Hopfield energy to obtain optimal solution of the TSP with reasonable certainty, *in Proc. IEEE Int. Conf. Neural Networks*, 1995, pp. 18711875.

[6] Z. Yi, P. A. Heng, and P. F. Fung, Winner-take-all discrete recurrent neural networks, *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 15841589, Dec. 2000.

[7] G. Papageorgiou, A. Likas, and A. Stafylopatis, Improved exploration in Hopfield network state-space through parameter perturbation driven by simulated annealing, *Eur. J. Oper. Res.*, vol. 108, pp. 283292, 1998.

[8] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, Optimization by Simulated Annealing, *Science*, 220, 671-680, 1983.

[9] L. Wang, On competitive learning, *IEEE Trans. Neural Networks*, 8(5), 1214–1216, 1997.

[10] J.P.F. Sum and C.S. Leung and P.K.S. Tam and G.H. Young and W.K. Kan and L.W. Chan, Analysis for a class of winner-take-all model, *IEEE Trans. Neural Networks*, 10(1), 64–70, 1999.

[11] J. J. Hopfield and D.W. Tank, 'Neural' Computation of Decisions in Optimization Problems, *Biological Cybernetics*, vol. 52, pp. 141-152, 1985.

[12] M. Peng, K. Narendra and A. Gupta, An investigation into the improvement of local minimum of the Hopfield network *Neural Networks*, vol. 9, pp. 1241-1253, 1993.