Information entropy based interaction model and optimization method for swarm intelligence

Yunlong Zhu\(^1\), Xiaoxian He\(^1,2\), Kunyuan Hu\(^1\) and Ben Niu\(^1\)

\(^1\)Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China
\(^2\)College of Information Science & Engineering, Central South University, ChangSha, 410083, China

Inspired by social insects, swarm intelligence has been hotly investigated in recent years as an innovative artificial intelligence technique for solving problems. In this paper, we mainly focus on the information interaction of individuals in swarm intelligence. By using information entropy \(H(X)\) and mutual information \(I(X;Y)\) of information theory to evaluate the information quality and interaction efficiency, respectively, the interaction model is proposed. Within this model, individuals’ information is evaluated with uniform standards, so that more excellent individuals can be selected to influence other individuals by interaction. We validated this model with the route-exchange algorithm, which is proposed for combinatorial optimization. Seven benchmarks of the Traveling Salesman Problem are tested in the experiments. The results are compared with other heuristic algorithms.

**Key words:** information entropy; interaction model; route-exchange algorithm; swarm intelligence.

1. **Introduction**

In the natural world, it is believed that all living creatures come from the same origin. Some of them evolved into intelligent individuals with concentrated and developed
neural systems, eg, human beings, while others evolved into swarms of simple and co-operative individuals, eg, social insects. In the latter, because the single individual’s power is very little, they often live in colonies. Swarms of them co-operate by direct or indirect interaction, and get together to achieve their goals. In this way, they have enormous power (Tarasewich and McMullen, 2002). If we regard an individual as a nerve cell, a swarm can be considered a developed neural system, though more distributed and stochastic.

Inspired by the behaviours of social insects, researchers have designed artificial systems to solve problems. This was called swarm intelligence. Since there was research on swarm intelligence beyond social insects later, Bonabeau et al. (1999) extended its definition to include any attempt to design algorithms or distributed problem-solving devices inspired by the collective behaviour of social insect colonies and other animal societies.

In the natural world, information interaction is necessary in social insect colonies and other animal societies for their ecological behaviours. Ants, fish and birds all interact with each other by various ways and means in foraging, nest constructing and colony fighting. In well-known artificial systems such as the ant colony system (ACS; Dorigo and Gambardella, 1997), information interaction also played an important role in improving system’s efficiency. Until now, however, information interaction in swarm intelligence has been ignored to some extent. At least there is not very elaborate analysis about it yet.

On the other hand, environment, individual and behaviour rule are sufficiently considered the three basic elements in swarm systems, but mathematical models still cannot be established based on them. One of the main questions is that there is no uniform standard to evaluate the behaviour of individuals, which leads to difficulties in evaluating and improving individuals’ search behaviours.

In this paper, we introduce Shannon’s information entropy (1948) to depict the uncertainties of individuals’ information, to evaluate and improve their efficiency of interacting behaviours. Based on information entropy and mutual information, an interaction model is proposed, which is also a general model for designing algorithms to solve problems. A new algorithm is designed for combinatorial optimization according to the interaction model, and seven benchmarks of the TSP are tested.

The rest of this paper is organized as follows. Section 2 gives a brief introduction to swarm intelligence. Concepts of information entropy and mutual information are introduced in Section 3. Section 4 describes the application of information entropy and mutual information in swarm systems, and provides the framework of interaction model. The new algorithm for TSP is presented in Section 5. Section 6 tests it on seven benchmarks, and then the results are compared and discussed. Finally, Section 7 outlines the conclusions.
2. Swarm intelligence

As a new way of treating systems, swarm intelligence was noticed decades ago. In 1991, Marco Dorigo (Dorigo and Gambardella, 1997) was inspired by foraging behaviours of ants, and proposed the ant colony optimizer (ACO) for distributed optimization problems. After that, J. Kennedy and R. C. Eberhart (Kennedy and Eberhart, 1995; Kennedy et al., 2001) were inspired by foraging behaviours of birds and fish, and proposed the particle swarm optimizer (PSO) for continuous optimization problems. These methods have been applied to many problems and proved successful. In addition, there are many analogous methods based on swarm intelligence (He et al., 2006; Niu et al., 2006; Tang et al., 2006; Yao, 1999).

As for modelling in swarm intelligence, Han and Cai (2002) developed a general model called the AER model, in which A, E and R denote agent, environment and rule, respectively. In this model, individuals take actions in given environments according to some rules, in which individual, environment and rule are considered the three basic elements of swarm systems. However, that is not enough for theoretical analysis on an individual’s behaviour. For example, the effects of co-operative behaviours cannot be evaluated properly.

Jeffrey E. Boyd et al. (2004) have simulated insects and animals’ co-operative behaviours by computer. They called it ‘interactive art’ because complex co-operative behaviours in swarms can be realized by interaction between individuals. Taking social insects for example, individuals can obtain local information and interact with their geographical neighbours. They can also change the local environments or mark in the local environments to interact with the remote individuals indirectly. This interaction approach is called stigmergy, which was simulated in the ACS by laying pheromone (Dorigo and Gambardella, 1997).

When solving problems, there may be a lot of information for each individual. Some may be useful to its neighbours, while the rest is not. Some given information may be useful to neighbour (a), but may be useless to neighbour (b). As a result, the quality of information should be evaluated before interaction. In this case, information entropy may be a good choice to do that.

3. Information entropy and information interaction

Originally, the concept of entropy in thermodynamics referred to the amount of energy that is inaccessible for work. Shannon (1948) introduced entropy into information theory and proposed information entropy to describe uncertainty of information. High information entropy is equivalent to high uncertainty, disorganization or variability in a system. Information entropy has been broadly used in communication engineering, statistics and artificial intelligence systems.
Let $S = \{a_1, a_2, \ldots, a_M\}$ be a finite set and $p$ be a probability density function on $S$. The amount of information needed to fully characterize all of the elements of this set consisting of $M$ discrete elements is defined by (Golan, 2002)

$$I(S_M) = \log_2 M$$

(1)

and is known as Hartley’s formula. Based on Hartley’s formula, Shannon developed its information criterion within the context of a communication process. Its criterion, called information entropy (Thomas and Joy, 1991), is

$$H(X) = - \sum_{x \in S} p(x) \log p(x)$$

(2)

with $x \log(x)$ tending to zero as $x$ tends to zero. This information entropy measures the uncertainty or information content that is implied by $p$. $H(p)$ reaches its maximum when $p_1 = p_2 = \cdots = p_M = 1/M$ (and is equal to Hartley’s formula) and minimum when $p(x_i) = 1$, $\exists x_i \in S$.

Considering two probability variants $X$ and $Y$, if their joint probability density is $p(x,y)$, marginal probability is $p(x)$ and $p(y)$, respectively, and the conditional probability is $p(y|x)$, then the conditional entropy is defined as:

$$H(Y|X) = - \sum_{x,y} p(x,y) \log p(y|x)$$

(3)

Their mutual information $I(X;Y)$ is defined as (Thomas and Joy, 1991):

$$I(X; Y) = \sum_{x \in S} \sum_{y \in S} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

(4)

Because the information entropy and the mutual information can be used to evaluate the uncertainty of an individual’s information and the relationship of two individuals’ information, respectively, they are applied in the interaction model as follows.

### 4. Interaction model

Assuming there is an ideal environment. In this environment, firstly, there is only finite information. Secondly, heuristic information always helps to solve problems. Thirdly, the solution of problems is implied in the information of environment. A swarm of individuals searches for the solution in the environment independently.
and simultaneously according to the rules and heuristic information. The searching process of an individual is described by transition of states within $S$ over time, and each state can be visited only once by the same individual, where $S = \{s_1, s_2, \ldots, s_N\}$ is a finite state space. Each individual also has a finite set of actions $A = \{a_1, a_2, \ldots, a_N\}$ available to it. Formally, an action $a_k$ maps a state $s_i$ into another state $s_j$. Whether to take the action is determined by behaviour rules, the past states and heuristic information. Therefore, an action can be represented by a transition matrix with elements $p_{ij}^{kt}$, $i, j = 1, 2, \ldots, N$. This is the probability that the $k$th individual enters state $s_j$ at discrete time step $t + 1$ when at time $t$ it was in state $s_i$, in which $p_{ii} = 0$, $i = 1, 2, \ldots, N$. Mathematically speaking,

$$p_{ij}^{kt} = \Pr (S^{t+1} = s_j | S^t = s_i, S^{t-1} = s_{i-1}, \ldots, S^0 = s_0)$$

(5)

Because the next state selected depends not only on the current state, but also past states the individual has experienced, the system is not necessarily Markovian. Moreover, in the searching process, an individual has different state sequences at different time steps.

**Definition 1.** If the $k$th individual is in state $s_i$ at time step $t$, then its information entropy is:

$$H_i^{kt} = - \sum_{j \in \text{allowed}_{i}^{kt}} p_{ij}^{kt} \log p_{ij}^{kt}$$

(6)

where $\text{allowed}_{i}^{kt}$ denotes the set of candidate states when the $k$th individual at the state $s_i$ at time step $t$.

**Deduction 1.** Considering $H_1^{kt}, H_2^{kt} \in H^k$, if $t_2 > t_1$, $H_2^{kt} \leq H_1^{kt}$.

**Proof.** If the $k$th individual has not obtained any heuristic information from time step $t_1$ to $t_2$, then its probability distribution for selecting the next state is even distribution, namely

$$p_1 = p_2 = \cdots = p_M = 1/M$$

where $M$ is the number of states that the $k$th individual has not visited. Because $S$ is a finite set, and each state can be visited only once by the same individual, so $M_{t_1} \leq M_{t_2}$ Therefore,

$$H_1^{kt} = - \sum_{j} p_j^{kt} \log p_j^{kt} = \log M_{t_1} \geq \log M_{t_2} = H_2^{kt}$$
If the $k$th individual has obtained any heuristic information from time step $t_1$ to $t_2$, then according to the property of heuristic information, the same results can be obtained.

Deduction 1 indicates that in an episode, an individual’s information entropy will decrease over time.

As assumed above, because the heuristic information always helps to solve the problems, according to the definition of $H$, the more heuristic information an individual obtains, the lower the value of $H$ is, and the more effective the behaviour of the individual is.

According to deduction 1, if $H_{kt}^{li} > H_{lt}^{li}$, then the $l$th individual should have visited more states, or have obtained more heuristic information than the $k$th individual.

**Definition 2.** Considering two state sequences $A$ and $B$, if there exists the same state sequence as $A$ in $B$, then we call $A \subset B$.

**Definition 3.** Considering the $k$th individual and $l$th individual, the state sequences they have experienced are $S_k$ and $S_l$, respectively. If their initial state is the same, and $S_k \subset S_l$, then the information of the $k$th individual is redundant to the $l$th individual.

If the information of the $k$th individual is redundant to the $l$th individual, then the information of the $k$th individual is useless to the $l$th individual.

**Deduction 2.** If the information of the $k$th individual is redundant to the $l$th individual, and vice versa, then $S_k = S_l$.

**Definition 4.** Let $p(y^l_t)$ be distribution function of the $l$th individual at time $t$; if its information has influence on the $k$th individual’s action, then $p(x^k_t | y^l_t)$ denotes the distribution of the $k$th individual after information of the $l$th is given, or

$$p(x^k_t | y^l_t) = p(x^k_t)$$  \hspace{1cm} (7)

**Definition 5.**

$$p(x^k_t, y^l_t) = p(y^l_t)p(x^k_t | y^l_t)$$  \hspace{1cm} (8)

It should be noticed that

$$p(x^k_t, y^l_t) \neq p(y^l_t, x^k_t)$$  \hspace{1cm} (9)

Now we can give the definition of mutual information of the $k$th individual and $l$th individual.

**Definition 6.**

$$I(X^k_t; Y^l_t) = \sum_{x^k_t, y^l_t} p(x^k_t, y^l_t) \log \frac{p(x^k_t, y^l_t)}{p(x^k_t)p(y^l_t)}$$  \hspace{1cm} (10)
Deduction 3. If information of the lth individual is given, there exists the following equation:

\[
I(X^{kt}, Y^{lt}) = H(X^{kt}) - H(X^{kt}|Y^{lt})
\]  

(11)

Proof.

\[
I(X^{kt}, Y^{lt}) = \sum_{x^{kt}, y^{lt}} p(x^{kt}, y^{lt}) \log \frac{p(x^{kt}, y^{lt})}{p(x^{kt})p(y^{lt})}
\]

\[
= \sum_{x^{kt}, y^{lt}} p(x^{kt}, y^{lt}) \log \frac{p(x^{kt}|y^{lt})}{p(x^{kt})}
\]

\[
= - \sum_{x^{kt}, y^{lt}} p(x^{kt}, y^{lt}) \log p(x^{kt}) + \sum_{x^{kt}, y^{lt}} p(x^{kt}, y^{lt}) \log p(x^{kt}|y^{lt})
\]

Because the information of the lth individual is given, we have

\[
p(x^{kt}, y^{lt}) = p(y^{lt})p(x^{kt}|y^{lt})
\]

\[
= p(x^{kt})
\]

so,

\[
I(X^{kt}, Y^{lt}) = -\sum_{x^{kt}} p(x^{kt}) \log p(x^{kt}) - \left( -\sum_{x^{kt}, y^{lt}} p(x^{kt}, y^{lt}) \log p(x^{kt}|y^{lt}) \right)
\]

\[
= H(X^{kt}) - H(X^{kt}|Y^{lt})
\]

Formula (11) denotes that \(I(X^{kt}; Y^{lt})\) measures the decrease of uncertainty for the kth individual at time step t after information on the lth individual is given.

In the searching process, an individual should evaluate its information itself first. Considering the lth individual, the smaller the value of \(H(Y^{lt})\) is, the better the quality of its information is, and the more reliable the information is for other individuals that the lth individual would interact with. Then, the value of \(I(X^{kt}; Y^{lt})\) can help to determine whether the kth individual interacts with the lth, because the larger the value of \(I(X^{kt}; Y^{lt})\) is, the more advantages the kth individual obtains according to deduction 3. In this way, the individuals can be more and more intelligent to find the solution.

The framework of interaction model is illustrated in Figure 1. According to this framework, an individual should calculate its information entropy before interacting.
with others, so that individuals can interact selectively. As for two individuals who
have met interaction conditions, it is better to evaluate their mutual information before
they interact, because it can be used to assess effectiveness of interaction.

5. The route-exchange algorithm

5.1 The basic route-exchange algorithm (BREA)

According to the interaction model, the route-exchange algorithm is designed for
solving the Traveling Salesman Problem (TSP) (Grefenstette et al., 1985). The problem
is to find the shortest route for a travelling salesman to visit all the cities once and only
once, and return to the starting city, which is also known as a Hamiltonian cycle.
Because the exchange of route information is the most important idea in the algorithm,
it is called the route-exchange algorithm. The problem can be described as follows.

Given a set of cities $C = \{c_1, c_2, \ldots, c_n\}$, for each pair $(c_i, c_j)$, $i \neq j$, let $d(c_i, c_j)$ be the
distance between city $c_i$ and $c_j$. Solving the TSP entails finding a permutation $\pi^*$ of the
cities $(c_{\pi(i)}, \ldots, c_{\pi(n)})$, such that

$$
\sum_{i=1}^{n} d(c_{\pi(i)}, c_{\pi(i+1)}) \leq \sum_{i=1}^{n} d(c_{\pi(i)}, c_{\pi(i+1)}), \quad \forall \pi \neq \pi^*, n + 1 \equiv 1
$$

(12)

In a symmetrical TSP, $d(c_i, c_j) = d(c_j, c_i), \forall i, j$, while in an asymmetrical TSP this
condition is not satisfied. In this work, we consider the symmetrical TSP.

The route-exchange algorithm is a heuristic approach inspired by the social
behaviours of people, in which individuals can interact with others and learn from
others to improve their own performance. The algorithm is designed according to the
concept of an interaction model. Individuals evaluate their own information before

![Figure 1 The framework of interaction model](image_url)
they interact with others. Then, the algorithm interacts selectively with more excellent individuals so that it can do better by learning from interaction.

Given a symmetrical TSP of \( n \) nodes, \( m \) individuals are required to find a closed tour independently and simultaneously from different starting nodes with the same velocity \( v \), where \( m \leq n \). They could encounter others somewhere in the process. Each of them has two storages: one is for the sequence of cities \( (c_{\pi(1)}, \ldots, c_{\pi(n)}) \) visited by itself in the current repetition, another is for the best sequence \( (c_{\pi^*(1)}, \ldots, c_{\pi^*(q)}) \) visited by itself or others ever encountered by now in the current repetition, where \( q \leq n \). For the convenience of depiction, we denote the former route-storage-A, RS_A for short, and the latter RS_B. When two individuals meet, they will compare the quality of two routes visited in the current repetition with each other, put the better one into RS_B, and continue to tour. In the next repetition, the individual would prefer to select nodes of RS_B of previous repetition and follow it first, than previous RS_A.

As for the \( k \)th individual, if it is at node \( i \) sometime, then the probability that it selects the node \( j \) as its next node is defined as

\[
p_{ij}^k = \begin{cases} 
(\eta_{ij})^\alpha \cdot \rho_B, & l_{ij} \in \text{RS}_B_k \\
\sum_{r \in \text{allowed}_k} (\eta_{ir})^\alpha, & (l_{ij} \in \text{RS}_A_k) \land (l_{ij} \notin \text{RS}_B_k) \\
(\eta_{ij})^\alpha \cdot \rho_A, & (l_{ij} \in \text{RS}_A_k) \land (l_{ij} \notin \text{RS}_B_k) \\
\sum_{r \in \text{allowed}_k} (\eta_{ir})^\alpha, & \text{else}
\end{cases}
\]

where \( \eta_{ij} = 1/d_{ij} \), \( l_{ij} \) denotes the edge \( (i, j) \) and \( \text{allowed}_k \) denotes the set of candidate nodes when the \( k \)th individual is at node \( i \). \( \rho_A, \rho_B \) and \( \rho_C \) are selection parameters, \( 0 < \rho_C < \rho_A < \rho_B \). \( \alpha(>0) \) is a parameter to control the power of greedy selection.

Given that the \( k \)th individual reaches the node \( i \) at time \( T_k \), and the \( l \)th individual reaches the same node at time \( T_l \). If the following inequality is satisfied, then the two individuals are judged to have met.

\[
|T_k - T_l| \leq \frac{\bar{d}_{ij}}{\beta_0}, \quad j \in \text{Neighbor}_i
\]

where \( \text{Neighbor}_i \) is the set of nodes connected directly with node \( i \), \( \bar{d}_{ij} \) is the average length of all edges that connected with node \( i \), and \( \beta(>0) \) is a control parameter.

In order to evaluate the quality of the route \( (c_{\pi(i)}, \ldots, c_{\pi(q)}) \), we define a quality factor \( G \), thus

\[
G = \left( \frac{1}{q} \right)^\delta \cdot \frac{1}{q} \sum_{i=1}^{q-1} d_{i,i+1}
\]

(15)
where $\sum_{i=1}^{q-1} \frac{d_{i,i+1}}{q}$ is average distance between two nodes, $(1/q)^{\delta}$ is used to emphasize that the more nodes the route includes, the more important the information is, and $\delta$ is a control parameter, $0<\delta$. The smaller the value of quality factor $G$ is, the better the quality of the route is.

A positive feedback mechanism is designed to avoid vibrations. After each repetition, the individual having the best results in the current repetition preserves its experience in the next repetition. In this way, the best individual can keep its capability of influence on others, and accelerate others to improve their fitness. The route-exchange algorithm is illustrated in Figure 2.

### 5.2 The route-exchange algorithm with information entropy (IEREA)

In this problem, considering the complexity of computation, we still take the time that individuals reach the same node as interaction criteria instead of $I(X; Y)$. The value of $G$ is calculated as

$$G = H^{kt} \cdot \frac{\sum_{i=1}^{q-1} d_{i,i+1}}{q} \quad (16)$$

where $H^{kt}$ is the information entropy of the $k$th individual at time $t$. 

---

**Figure 2** The route-exchange algorithm
6. Experiments and results

The algorithm was tested on seven TSP instances defined in the TSPLIB (http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95). The experiments were run on a Windows XP Pentium IV 1.60GHz PC with 256 Mbytes of memory. The results of computation are summarized in Table 1, where $\rho_A = 2$, $\rho_B = 3$ and $\rho_C = 1$. The first column stands for the name of problems; the second for exact optimal tour length for each problem given in the TSPLIB. The third column stands for the average value of three run times with BREA, the fifth with IEREA, the fourth and the sixth for the relative error (Err), respectively, where the relative error is calculated as

$$Err = \frac{Ave - Opt}{Opt} \times 100\%$$ (17)

The seventh column denotes the repetition times required, and the last denotes the number of individuals employed in the program.

Table 1 indicates that the proposed route-exchange algorithm can be used to solve the symmetrical TSP effectively. Through the comparison of the two algorithms, it can be seen that the results of BREA are a little better than IEREA when treating small-size problems. This is because, in BREA, the emphasis of control is to evaluate the quality of the route one has visited. It is beneficial when the search space is not very large, but individuals only can make local optimal decisions. In IEREA, the control is rough and global, but individual can obtain more comprehensive information by computing its information entropy. This is necessary in searching large-scale space. So, it is reasonable to believe that IEREA can perform better in solving large-size problems. Moreover, in some swarm systems, the number of individuals is huge. In these systems, information entropy is a very good approach for evaluating individuals' behaviours.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Opt</th>
<th>BREA</th>
<th>Err (%)</th>
<th>IEREA</th>
<th>Err (%)</th>
<th>NC&lt;sub&gt;max&lt;/sub&gt;</th>
<th>Ind.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burma14</td>
<td>33.23</td>
<td>32.77</td>
<td>−1.38</td>
<td>31.80</td>
<td>−4.30</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Ulysses22</td>
<td>75.67</td>
<td>76.12</td>
<td>0.59</td>
<td>76.12</td>
<td>0.59</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Bayg29</td>
<td>9074</td>
<td>9389</td>
<td>3.47</td>
<td>9450</td>
<td>4.14</td>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>Bays29</td>
<td>2020</td>
<td>2067</td>
<td>2.33</td>
<td>2113</td>
<td>4.60</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td>Swiss42</td>
<td>1273</td>
<td>1319</td>
<td>3.61</td>
<td>1352</td>
<td>6.21</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Att48</td>
<td>33524</td>
<td>35413</td>
<td>5.63</td>
<td>36240</td>
<td>8.10</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Eil51</td>
<td>426</td>
<td>455</td>
<td>6.81</td>
<td>461</td>
<td>8.22</td>
<td>30</td>
<td>40</td>
</tr>
<tr>
<td>Berlin52</td>
<td>7542</td>
<td>7900</td>
<td>4.75</td>
<td>8087</td>
<td>7.23</td>
<td>30</td>
<td>40</td>
</tr>
</tbody>
</table>
Most of heuristic algorithms are likely to stagnate in local optima, i.e., they cannot converge to the global optima necessarily in each repetition. IEREA, though it often performs well, sometimes still stagnates in local optima. Figure 3 indicates the evolution of best tour length of Burma14 and Swiss42 with IEREA. It is obvious that in Figure 3(a), the algorithm converges quickly to the optimum. In Figure 3(b), however, it vibrates in several local optimal results.

Compared with other heuristic approaches, IEREA can converge to a good result with fewer repetitions. Table 2 indicates the comparison results with other heuristic algorithms on the number of tours before the algorithms get best results, where ACS denotes the ant colony system, GA denotes the genetic algorithm, EP denotes evolutionary programming and SA denotes simulated annealing. The data of problems are from http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95, and results for other heuristic algorithms are from Dorigo and Gambardella (1997).

From Table 2 we can see that IEREA has searched the least tours while it achieves results as good as other approaches. As a result, IEREA occupies fewer computing resources, and converges to the optima more quickly. On the other hand, according to No-free-lunch theory (Wolpert and Macready, 1997), the algorithm will have weak diversity, so it also needs to be intensified in the future work.

**Table 2** Comparison with other heuristic algorithms

<table>
<thead>
<tr>
<th>Problem</th>
<th>Opt</th>
<th>IEREA</th>
<th>ACS</th>
<th>GA</th>
<th>EP</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eil50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best result</td>
<td>425</td>
<td>428</td>
<td>425</td>
<td>428</td>
<td>426</td>
<td>443</td>
</tr>
<tr>
<td>Number of tours</td>
<td>1185</td>
<td>1830</td>
<td>25,000</td>
<td>100,000</td>
<td>68,512</td>
<td></td>
</tr>
<tr>
<td>Eil75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best result</td>
<td>535</td>
<td>542</td>
<td>535</td>
<td>545</td>
<td>542</td>
<td>580</td>
</tr>
<tr>
<td>Number of tours</td>
<td>2024</td>
<td>3480</td>
<td>80,000</td>
<td>325,000</td>
<td>173,250</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3** Evolution of best tour length
7. Conclusions

Interaction behaviours exist widely in most biological species from insects to mammals in the natural world, which is necessary for them to fulfil tasks co-operatively. Investigation in information flow and information processing in artificial swarm systems is significant. In this paper, we introduced information entropy and mutual information into artificial intelligence systems, which can be applied to analyse interaction behaviour mathematically. After that, the interaction model was proposed as a general model for swarm intelligence, then a new algorithm was proposed. As can be seen in the experiments and results, this algorithm can converge quickly to a good solution with quite low computing complexity. By applying the information entropy to the algorithm, the results are still very good. This shows that our approach is feasible. Because the information entropy is global and easy to calculate, the interaction model has great potential in treating large swarm systems.

Acknowledgement

This work is supported by the National Natural Science Foundation, China (No. 70431003) and the National Basic Research Program, China (2002CB312204).

References


http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95


