Clonal Selection Algorithm for Solving Permutation Optimisation Problems: A Case Study of Travelling Salesman Problem

Wei Pang, Kangping Wang, Yan Wang*, Ge Ou, Hanbing Li, Lan Huang*
College of Computer Science and Technology, Jilin University
Changchun, Jilin, 130012, China
*Corresponding authors: wy6868@hotmail.com; huanglan@jlu.edu.cn

Abstract—As an attempt to solve Permutation Optimisation Problems (POP) by using CLONALG (Clonal Selection Algorithm), a well-established artificial immune system, Traveling Salesman Problem (TSP) is studied as an example in this paper. Operators of CLONALG, especially the hyper-mutation operators are analyzed and modified to make CLONALG adapt to POP. Furthermore, the local search technique is employed to speed up the mature of the repertoire system, and receptor-editing operator is also employed to avoid the premature of the antibody population. Finally, several benchmark problems in TSPLIB are tested to evaluate the best and average performance of the proposed algorithm. Experimental results show the proposed competitive algorithm performs better than both the standard CLONALG and a genetic algorithm.

Keywords- Clonal Selection Algorithm; Artificial Immune System; Travelling Salesman Problem; Optimisation; Local Search

I. INTRODUCTION

As a well-established artificial immune system, CLONALG [1,2] has been widely used to solve a variety of problems. In particular, up to now CLONALG and its variants have been intensively studied for solving Traveling Salesman Problem (TSP) [2,3,4,5]. These algorithms shed light on the application of CLONALG to general permutation optimisation problems. In a permutation problem, all values of the problem have to be assigned to the same number of variables, so that on one hand a certain constraints have to be satisfied, and on the other hand the cost function of the problem achieves the optimal value. TSP is one of the best known and most studied permutation optimisation problems for its simple problem description yet hard-to-solve nature.

Similar to other evolutionary algorithms, one of the advantages of CLONALG is that its underlying search mechanism is problem-independent to some extent: once the operators of CLONALG, including the antibody encoding, hyper-mutation, and affinity evaluation operators, are defined for a particular problem, the search in the problem space will be completely determined by the inherent mechanism of CLONALG, and the algorithm designer does not need to further consider the domain specific details of the problem.

However, it is acknowledged that such general-purpose evolutionary algorithms as CLONALG cannot normally outperform the algorithms specifically designed for TSP (and other permutation optimisation problems), for example, the Lin-Kernighan algorithm [6]. Although special-purpose algorithms are very efficient, the lack of generality limits their use: for different problems researchers may design different algorithms to solve them, which might involve very time-consuming and tedious work. More importantly, the performance of these special-purpose algorithms often very much depends on the unique features of the search space and interactions of variables, which might only exist in the particular problems of interest. This means once the definition of the original problem changes or a new problem is proposed, the special-purpose algorithms like the Lin-Kernighan algorithm will not be suitable because of the changes of the features of the search space as well as variable interactions.

In this paper researchers aim to address the above issue and make use of both the generality of CLONALG and the efficiency of special-purpose algorithm to design a hybrid Artificial Immune System algorithm (AIS) for TSP, termed CLONALG-TSP-2opt. On one hand, researchers hope that CLONALG-TSP-2opt has the potential to be applied to other POP problems because it is based on the CLONALG framework, a general-purpose AIS algorithm. On the other hand, researchers want to make the performance of CLONALG-TSP-2opt outperform the standard CLONALG by incorporating a local search technique, which is problem-specific but easy to implement and generalise under the general POP context.

The rest of this paper is organised as follows: in the remaining part of this section, the original CLONALG is briefly described and related work on TSP is overviewed. In Section 2 the design strategy of the algorithm is presented, and this is followed by the local search technique detailed in Section 3. In Section 4 the algorithm is described in detail. In Section 5 the experimental results are reported. Finally Section 6 concludes the paper and explores future work.

A. Clonal selection algorithm

Inspired by the clonal selection theory [7], de Castro et al. [1,2] proposed the clonal selection algorithm
(CLONALG). Regarding the clonal selection principle, CLONALG is composed of four processes: selection, clonal expansion, hyper-mutation (affinity maturity) and re-selection. The principal idea of CLONALG is that accumulative but blind mutation can direct the search to better solutions. As an important branch in artificial immune system, after the CLONALG was proposed, it has been applied to many fields, such as multi-model function optimisation and combinatorial optimisation problems.

Like genetic algorithms (GAs), CLONALG is known as an evolutionary algorithm and incarnates the Darwin’s evolution theory. However, CLONALG employs clone and hyper-mutation operator instead of selection and crossover operation as in GAs. The optimisation version [1] of CLONALG is briefly described as follows:

Step1: A specific encoding strategy is applied to “encode” an antibody to represent a potential solution to the problem under investigation.

Step2: The antibody repertoire is composed of M such antibodies, and M is a given number determined by the complexity of the problem.

Step3: CSA executes in an iterative manner. For each generation:

i) First, all antibodies are selected to enter the next generation (selection).

ii) Each antibody is cloned for the same number of copies (clonal expansion).

iii) All these cloned copies undergo the hyper-mutation scheme, in which they will be mutated with a high probability.

iv) All the mutated antibodies are evaluated for their affinity values with the antigen (Affinity Evaluation). The evaluation is achieved by setting up a scoring system based on different problem context, which is the same as the fitness evaluation in GAs.

v) M antibodies (the same number as that of antibodies in the initial antibody repertoire) with the highest affinity values will be re-selected to form a new generation of the antibody repertoire (re-selection).

Step4: The algorithm will terminate if the termination conditions are satisfied, such as the optimal solutions have been found or the algorithm has run for a given number of generations.

From the above description, researchers can see that CSA can be considered as a population-based evolutionary algorithm. Each individual (antibody) is essentially a local optimum searcher, and the hyper-mutation can be deemed as a process of blind neighbourhood exploration in the search space in relation to the original antibody. In the re-selection process, the proliferated antibodies are selected according to their affinity values, which is a greedy strategy. After a sufficient number of generations, the antibodies are expected to find the optimal solution(s), or the approximate optimal solution(s).

B. CLONALG for solving TSP

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

As one of the classical and representative permutation optimisation problems, also as a notorious NP-hard problem, TSP has been intensively studied in the last several decades. There have been many effective algorithms specially designed to solve the TSP problems, for instance, the aforementioned Lin-Kernighan algorithm [6].

In addition, TSP has become a benchmark for evaluating the performance of many soft computing algorithms such as Genetic Algorithm (GA) [8,9], Ant Colony Algorithm [10,11] and Particle Swarm Optimisation [12,13]. Although the above soft computing algorithms are not as effective as the specially designed algorithms when dealing with TSP, the research on the design and analysis of these algorithms for solving TSP is still very important. This is because through the design and analysis processes researchers can figure out principles and intrinsic mechanisms that govern these algorithms in the search process and thus find ways to improve the efficiency of these algorithms for solving other permutation optimisation problems in general.

Consequently after the proposal of CLONALG, de Castro and Von Zuben performed some preliminary research on using CLONALG to solve TSP [2] based on the above considerations. In their work, the CLONALG was tested with a randomly generated instance of 30 cities, and the algorithm obtained comparatively satisfying results. Since de Castro and Von Zuben’s first attempt of applying CLONALG to TSP, there have been many variants of CLONALG developed to investigate how CLONALG combined with other mechanisms can better deal with TSP, and among them the following algorithms are noteworthy:

1) inspired by the receptor editing process [14], Gao et al. [4] introduced the receptor-editing operator to CLONALG to obtain better solutions; Zhu et al. [15] employed a CLONALG using receptor editing and an additional self-crossover operator to improve the performance of CLONALG for TSP; recently Liu et al. [5] studied how to apply CLONALG to large-scale TSP problems; in addition, Chen et al. [3] proposed a probabilistic modelling clonal selection algorithm for TSP.

Different from the above-mentioned existing algorithms, in this paper researchers focus on investigating how to incorporate problem-dependent local search techniques into CLONALG. In particular, researchers are interested in how to design the local search algorithm so that the balance between efficiency and the quality of solutions can be achieved. In addition, researchers aim to investigate the interactions between the local search technique and other immune operators, especially the hyper-mutation operator when performing the search. The research on the above mentioned topics would result in the proposal of CLONALG-TSP-2opt, a novel hybrid CLONALG approach to TSP. To verify the effectiveness of this novel algorithm, researchers would like to make comparisons between the proposed algorithm and other two algorithms, namely GA and the original CLONALG. This is achieved by testing these three algorithms with the benchmark TSPs taken from the well-established TSPLIB [16].
II. ALGORITHM DESIGN

In this section researchers present a solution of adapting CLONALG to solve TSP by designing the antibody coding strategy as well as the selection, clonal expansion, and hyper-mutation operators. In addition, researchers also propose a receptor editing approach suitable for POP.

A. Antibody encoding strategy

Similar to GA for solving TSP, the intuitive integer coding strategy is adopted. For a TSP with N cities, an antibody is represented by a vector $A_b = \{s_1, s_2, \ldots, s_N\}$, where each element of the vector $s_i (i \in [1, N])$ is an integer in the range of 1 to N and denotes a city in TSP.

B. Selection operator

The selection operator follows the original de Castro’s CLONALG for TSP. Due to the complexity of TSP problem space [17], the corresponding shape space [18] of the immune system is extraordinary complicated. To acquire better search ability, the exploration of the shape space should be strengthened. Consequently, all the antibodies in the population will participate in the processes of clonal expansion and hyper-mutation, that is, all the antibodies will be selected to explore the shape space.

C. Clonal expansion

An excellent style manual for science writers is [7]. The clone size is an important factor that largely influences the algorithm speed as well as the quality of solutions. With the increase of the clone size, the speed of the algorithm will slow down, while the solution quality may be improved. Consequently an appropriate clone size should be chosen to make a trade off between the convergence speed and the solution quality.

Based on the above consideration, the dynamic self-adaptive clone size is proposed: in each generation, each antibody is cloned with the same size. However, with the increase of the iterations, the clone size may increase by a given step until it attains the limitation of the maximum clone size.

D. Hyper-mutation

Considering the context of POP, two hyper-mutation operators are investigated: the point mutation and inverse mutation. Although the point mutation is commonly used in the original CLONALG, in this research the inverse mutation is employed for solving TSP because it is more intensive than the former and more suitable for exploration in complicated shape space such as the one in TSP. Later experiments will show that the inverse mutation performs better than the point mutation.

E. Receptor editing

Since all the antibodies are involved in the searching process, after the hyper-mutation the receptor-editing operator is employed to increase the probability of jumping out from the local optima and thus to avoid the premature of the antibody population. Researchers design the receptor-editing operator as follows: first choose a random continuous subsequence of length $d$ $(d \in [1, N),$ and $N$ is the number of the cities in TSP) from the antibody, and then the subsequence of $d$ elements is rearranged in a range into one of $d!$ possible arrangements selected at random.

III. LOCAL SEARCH

With the increase of the clone size, the algorithm will be more computationally expensive. It is necessary to speed up the convergence of the algorithm by introducing additional operations. Local search is one good option, and as mentioned before, the local search can make better use of the domain specific knowledge. For solving the TSP, the classical 2-opt approach [19] can effectively eliminate the cross edges and obtain a better solution, but it is very complicated and time-consuming. Apparently this kind of computationally expensive operation is not suitable for population-based iterative search, as during the search process the local search function will be executed many times, and an expensive local search function will result in a significant increase in the algorithm running time. Based on this consideration, researchers propose a compromise approach termed quasi-2opt, which partially improves the quality of solutions and thus is not as expensive as the original 2-opt approach. The pseudo code of the quasi-2opt is given as follows:

```
For an antibody S of length N, $S = \{s_1, s_2, s_3, \ldots, s_N\}$
BEGIN: For i=1 to N-1, i++
    For j=i+1 to N, j++
    {
        Swap and , get the new sequence SS;
        Calculate the cost of S;
        If (Cost (SS) < Cost (S)) {
            S = SS;
            Goto BEGIN, Restart the process;
        }
    }
```

In the above, $S$ is an antibody as described in Section 2.1, and SS is the newly obtained antibody after a swap operation. Cost ($S$) and Cost (SS) are the cost functions (the total length of the tour) of $S$ and SS, respectively. It should be pointed out that when researchers calculate the cost of SS generated by the swap operator, there is no need to recalculate all the edges, because the swap operator only influences the edges that are related to the two swapped cities. Furthermore, from the above pseudo code researchers can draw the conclusion that after the quasi-2opt process, swaps of any two cities in the final resulting antibody $S$ cannot further improve the solution.

It should be also noticed that the hyper-mutation and selection operations imply the hill-climbing process, and this process is problem-independent. The local search method is also a hill-climbing strategy, but it is problem-dependent. Take the TSP as an example: the quasi-2opt operation considers the features of TSP, that is, the connections between the node cities and the network structure of the problem. This makes the quasi-2opt only explore feasible solutions of TSP without violating problem constraints.
Having described all the design details of CLONALG-TSP-2opt, in this section researchers give a formal description of the algorithm:

**A. Population initialisation**

Set the population size $M$, the initial clone size $\text{initial\_clone\_size}$, current clone size $\text{clone\_size}$, maximum generation number $\text{gen\_num}$, maximum clone size $\text{max\_clone\_size}$, the increase step of the clone size $\text{increase\_step}$, and the probability of local search and receptor editing. Let $\Pi_i$ denote the antibody population in the $i$-th generation. First all the antibodies are generated randomly: $P_0 = \{A_{b_1}, A_{b_2}, A_{b_3}, \ldots, A_{b_M}\}$, where $P_0$ is the initial population, and for each antibody $A_{b_i} \in P_0$, $A_{b_i} = \{s_1, s_2, \ldots, s_N\}$, where $i \in [1,M], s_j \in [1,N], j \in [1,N]$, and $N$ is the number of the cities in TSP.

**B. Selection**

All the antibodies in $\Pi_i$ will be selected.

**C. Clonal expansion**

Adjustment of the clone size: after a given number of generations ($\text{increase\_generation}$) is exceeded, if there is still no improvement of the solution, the clone size will increase by $\text{increase\_step}$, but the maximum clone size $\text{max\_clone\_size}$ cannot be exceeded.

Each antibody in the population will be cloned $\text{clone\_size}$ copies. After the clonal expansion the cloned population is represented by $P_i = \{Q_1, Q_2, Q_3, \ldots, Q_M\}$, where $Q_i (i \in [1,M])$ is a $\text{clone\_size}$-dimensional vector denoted by $Q_i = (A_{b_1}, A_{b_2}, \ldots, A_{b_N})'$.

**D. Hyper-mutation**

All the antibodies will suffer a hyper-mutation process: the inverse sequence operator described in Section 2.4 is taken in this algorithm. Let the symbol $\otimes$ denote the hyper-mutation operator, and the population after hyper-mutation is represented by $R_i = \otimes Q_j = (\otimes A_{b_1}, \otimes A_{b_2}, \ldots, \otimes A_{b_N})'$, $i \in [1,M]$.

**E. Local search**

The quasi-2opt operation is applied to $P_i$ with a given probability, if the symbol $\chi$ denotes the local search operator, the population after the local search is as follows: $P_R = \chi P_i$.

**F. Receptor editing**

The receptor editing is applied to $P_L$ with a very small given probability, after this operation, if the symbol $\theta$ denotes the receptor-editing operator, the population after receptor editing is as follows: $P_R = \theta P_L$.

**G. Re-selection**

Select the $M$ best antibodies from $P_R$, and finally the new generation $P'$ is obtained.

Step 2~Step 7 are executed repeatedly until the maximum number of generations is reached or the quality of solutions is acceptable.

**V. EXPERIMENTAL RESULTS**

In this section the best and average performance of the algorithm will be evaluated by benchmark problems. All the algorithms were implemented in C++ with STL (Standard Template Library), and all experiments were performed on a Macbook Pro with an Intel Core i7 3.0GHz CPU and 8GB RAM.

**A. Best Performance Comparison**

Several Benchmark problems from TSPLIB are selected to evaluate the best performance of the algorithms. All the algorithms were run ten trials and the best results were recorded.

Eilin problems, including Eil51, Eil75, and Eil101, are three problems from TSPLIB and have been used as benchmark problems by many researchers such as Fogel [20]. First the mutation strategy in hyper-mutation is investigated. The point mutation and inverse sequence mutation operators are tested with the Eil51 problem, of which the best-known solution is 427.855. In the experiment the parameters are taken as follows: the population size $M=100$, $\text{initial\_clone\_size}=10$, $\text{gen\_num}=1000$, $\text{max\_clone\_size}=20$, $\text{increase\_step}=1$, $\text{increase\_generation}=100$. In addition, the probability of receptor editing is 0.1%. The experimental results are shown in Fig. 1.

![Figure 1. The Effect of Different Hyper-mutation Strategies on the Search.](image-url)
short period. Finally the CLONALG+2opt can find better solution than CLONALG without 2opt.

Lastly the three algorithms: standard CLONALG, GA and CLONALG+2opt are compared together and the results are shown in Fig. 4.

Figure 2. The Effect of Quasi-2opt On the Search.

Third, the algorithm is compared with a general genetic algorithm designed for TSP. In this GA the population size is 100; the most commonly used Partial Match Crossover (PMX) [8] is employed and the cross probability is 0.8; the inverse sequence mutation strategy is selected and the mutation probability is 0.5. The results were shown in Fig. 3, from which researchers can see that the best solution obtained by CLONALG+2opt is also better than that from GA.

B. Average Performance Comparison

The proposed algorithm is then compared with standard CLONALG and GA in terms of average performance using four benchmark problems from TSPLIB: the three Eilon problems and Berlin52 problem. The Berlin52 problem is the most difficult problem among the four problems. The parameter settings of all algorithms are the same as those described in Section 5.1. All the algorithms were run in ten times and the average values were recorded. The results are shown in Table 1. From this table researchers can see that the average performance of CLONALG+2opt is always better than GA and CLONALG, which demonstrates the effectiveness of CLONALG+2opt. In addition, from all experiments researchers can see that the standard deviation of CLONALG+2opt tends to be smaller than the other two algorithms, which shows its stability.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Generations</th>
<th>GA</th>
<th>CLONALG</th>
<th>CLONALG+2opt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cost</td>
<td>Deviation</td>
<td>Cost</td>
</tr>
<tr>
<td>Eil51</td>
<td>1000</td>
<td>443.15</td>
<td>3.574%</td>
<td>441.36</td>
</tr>
<tr>
<td>Eil75</td>
<td>1000</td>
<td>566.32</td>
<td>3.027%</td>
<td>565.23</td>
</tr>
<tr>
<td>Eil101</td>
<td>1000</td>
<td>777.31</td>
<td>23.58%</td>
<td>704.32</td>
</tr>
<tr>
<td>Berlin52</td>
<td>1000</td>
<td>7704.67</td>
<td>2.136%</td>
<td>7752.47</td>
</tr>
</tbody>
</table>

Figure 3. Comparison between GA and CLONALG+2opt.

Figure 4. Comparison of the Three Algorithms.
VI. CONCLUSIONS

In this paper, researchers proposed a hybrid AIS system that incorporates the local search technique into CLONALG for solving TSP. The immune operators of CLONALG were also modified to make CLONALG adapt to TSP. Experimental results show that the proposed approach, CLONALG with quasi-2opt and sequence mutation strategy, is a promising algorithm for solving TSP and has the potential to be applied to other permutation optimisation problems.

The work presented in this paper is a preliminary research on solving general permutation optimisation problems by immune-inspired algorithms combined with problem-dependent local search. The design of the quasi-2opt operator is an example showing that by making use of the information about problem features the performance of immune-inspired algorithms could be significantly improved. The fact that there are already existing local search techniques specially designed for other POP problems encourages us to further investigate the potential of the hybrid approach to POP, which combines CLONALG with the local search approach. More work in this field could be done in the near future. One promising future work is to apply CLONALG to the Vehicle Routing Problem with Time Windows (VRPTW) [21], a more complicated and realistic POP than TSP. The local search techniques for VRPTW, such as PFH (Push Forward Insertion Heuristic) [22] and 2-interchange [23] techniques, and their variants could be integrated into CLOANLG, and the performance of the resulting algorithms can be compared with the previous work on using GA to solve the same problem [24].

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China (No. 61472159), Jilin Innovation Team Project (No. 20122805) and Development Project of Jilin Province of China (No. 20130522111JH, 20140101180JC)

REFERENCES


