A Relationship between Network Topology and Search Performance of PSO

Takahiro Tsujimoto, Takuya Shindo, Takayuki Kimura, and Kenya Jin’no
Dept. Electrical and Electronics Eng. ,
Nippon Institute of Technology
Email: jinno@nit.ac.jp

Abstract—Particle swarm optimization (abbr. PSO) is one of the most effective optimization algorithms. The PSO contains many control parameters, therefore, the performance of the searching ability of the PSO is significantly alternated. In order to analyze the dynamics of such PSO system rigorously, we have analyzed a deterministic PSO (abbr. D-PSO) systems which does not contain any stochastic factors, and its coordinate of the phase space is normalized. The found global best information influences the dynamics. This situation can be regarded as the full-connection state. On the other hand, there is the case where the best information in a limited population. Such information is called as ibest. How to get the ibest information from any population is equivalent to a network structure. Such network structure influences the performance of searching ability. In order to clarify a relationship between network structures of the PSO and its performance, we pay attention to the degree and the average distance used in graph theory. We consider the two cases where the D-PSO has an extended cycle structure and a Small World network structure. Our numerical simulation results indicates the searching performance of the D-PSO is depended on the average distance of the node. Especially, the long average distance exerts the search performance on the D-PSO. We confirm that the search performance properties of the D-PSO and the conventional stochastic PSO are completely different to the average distance. The search performance of the D-PSO is improved according to the average distance. On the other hand, the search performance of the conventional stochastic PSO is deteriorated according to the average distance. We consider that the slow transmission of the beneficial information leads to the diversification of the particles of the D-PSO. Also, we clarify the small perturbation of the random range of the stochastic PSO is important.

Index Terms—PSO, deterministic, complex network, small world, graph theory, average distance, stochastic factor, perturbation, diversity

I. INTRODUCTION

Searching for an optimal value of a given evaluation function of various problems is very important problem in engineering fields. In order to solve such optimization problems speedily, various heuristic optimization algorithms have been proposed. Particle swarm optimization (PSO), which was originally proposed by J. Kennedy and R. Eberhart [1],[2], is one of such heuristic algorithms. The PSO algorithm is a useful tool for optimization problems[3]-[8].

The original PSO is described as
\[ v_{j}^{t+1} = w v_{j}^{t} + c_{1} r_{1} (\text{pbest}_{j}^{t} - x_{j}^{t}) + c_{2} r_{2} (\text{gbest}^{t} - x_{j}^{t}) \] (1)
\[ x_{j}^{t+1} = x_{j}^{t} + v_{j}^{t+1} \] (2)
where \( w \geq 0 \) is an inertia weight coefficient, \( c_{1} \geq 0 \), and \( c_{2} \geq 0 \) are acceleration coefficients, and \( r_{1} \) and \( r_{2} \) are two diagonal matrices which diagonal components are separately generated uniformly distributed random numbers in the range \([0, 1]\). \( x_{j}^{t} \in \mathbb{R}^{N} \) denotes the location of the \( j \)-th iteration on the \( t \)-th particle in the \( N \)-dimensional space, and \( v_{j}^{t} \in \mathbb{R}^{N} \) denotes a velocity vector of the \( j \)-th particle on the \( t \)-th iteration.

\( \text{pbest}_{j}^{t} \in \mathbb{R}^{N} \) means the location that gives the best value of the evaluation function of the \( j \)-th particle on the \( t \)-th iteration. In this article, \( \text{pbest}_{j}^{t} \) is called as a personal best location information of \( j \)-th particle on \( t \)-th iteration. \( \text{pbest}_{j}^{t} \) can be given as
\[ \text{pbest}_{j}^{t} = \min_{\tau} f(x_{j}^{\tau}) \quad \tau \leq t \] (3)
\( \text{gbest}^{t} \in \mathbb{R}^{N} \) means the location which gives the best value of the evaluation function on the \( t \)-th iteration in the swarm. \( \text{gbest}^{t} \) is called as a global best location information \( t \)-th iteration. \( \text{gbest}^{t} \in \mathbb{R}^{N} \) can be given as
\[ \text{gbest}^{t} = \min_{j} \text{pbest}_{j}^{t} = \min_{j, \tau} f(x_{j}^{\tau}) \quad \tau \leq t \] (4)

The particles in the swarm fly through the \( N \)-dimensional space according with Eqs. (1) and (2). Each particle memorizes its record of the best evaluation value and the corresponding best location. Also, each particle shares such best evaluation information among the swarm. On the basis of such information, the moving direction and the velocity are calculated by Eq. (1). Namely, all particles will move toward a coordinate that gives the current best value of the evaluation function.

In such PSO system, the parameters play very important role to the searching ability. Therefore, many researchers study about adequate parameters selecting[5]. The searching ability of such PSO depends on the inertia weight coefficient and the acceleration coefficients. Since the acceleration coefficients are multiplied by a random number, the system can be regarded as a stochastic system. The rigorous theoretical analysis of the dynamics of such stochastic system is difficult. In order to analyze the dynamics of such PSO, M. Clerc, and J. Kennedy proposed a simple deterministic PSO system, and analyzed its dynamics theoretically[4]. The simple deterministic PSO system does not contain stochastic factors, namely, the random coefficients have been omitted from the original PSO system. The analysis of such a deterministic PSO is very important for determining the effective parameters of the standard PSO[4]-[6]. The dynamics of the deterministic PSO is depended on...
II. DETERMINISTIC PSO

In this section, we introduce the deterministic PSO (abbr. D-PSO). The D-PSO can be derived from the conventional stochastic PSO to omit stochastic factors. For simplicity, \( \text{pbest}_j \) and \( \text{gbest}_t \) are combined as the following

\[
\begin{align*}
\text{p}_j^t &= \frac{c_1 \text{pbest}_j^t + c_2 \text{gbest}_t}{\psi} \\
\psi &= c_1 + c_2
\end{align*}
\]

The parameter \( \psi \) means an integrated acceleration parameter.

Since each dimension variable of the particle is independent, we can consider one dimensional case without loss of generality. Therefore, we consider one dimensional system, hereafter. For one dimensional deterministic PSO, the dynamics is described by a canonical form. The coordinate of the phase space of the CD-PSO becomes stable if the eigenvalues exist within the unit circle on the complex plane. In this situation, each particle must converge to a fixed point, which is represented by Eq. (5) with rotational motion in the phase space.

Such deterministic system is suitable for an analog electronic circuit implementation. Now, we try to implement the D-PSO by analog electronic circuit[16]. This is one reason that we pay attention to the deterministic system.

III. NETWORK STRUCTURE

For the conventional stochastic PSO system, the location information which gives the best value of the evaluation function is shared with the other particles in the swarm. On the other hand, there is the case where the best information in a limited population. Such information is called as a local best location information, \( \text{lbest} \). How to get the local best information from any population is equivalent to a network structure. Namely, the population corresponds to the connection relations which can be expressed as a network structure in graph theory. Such network structure influences the performance of searching ability. Therefore, we analyze a relationship between network structures of PSO and its performance by using graph theory of features. In this article, we pay attention to the degree and the average distance used in graph theory. The degree of a vertex of a graph is the number of connected particles. The distance between two vertices in a graph is the number of edges in the shortest path connecting them. If the degree is increased, the average distance is decreased. The average distance can be calculated as

\[
\lambda = \frac{(w - \psi + 1) \pm \sqrt{\psi^2 - 2(w + 1)\psi + (w - 1)^2}}{2}
\]
average distance $L_{\text{average}}$ is described as

$$L_{\text{average}} = \frac{1}{P(P-1)} \sum_m \sum_{m \neq n} d_{mn}$$

where, $P$ denotes the number of particles, and $d_{mn}$ represents the distance between $m$-th particle and $n$-th particle.

For simplicity, we assume that the network structure of PSO is an extended cycle as shown in Fig. 1. The degree of all nodes of the extended cycle structure is uniform. Therefore, the average distance is depended on the degree and the number of particles. Each figure in Fig. 1 corresponds to each degree. In Fig. 1, $\circ$ represents the particle, and it share the best location information with the connected particles. Table I shows the average distance in the case where the number of particles is 50.

Figure 1(e) illustrates the case where the information is shared with all other particles. In this case, they share the global best location information $g_{\text{best}}$. Each particle uses this global best location information and its personal best location information $p_{\text{best}}$ in Eqs. (1) and (2).

On the other hand, particles in a limited population are shared the best information as shown in Figs. 1(a), (b), (c), and (d). In these cases, the local best position of each particle is given as

$$l_{\text{best}}^t_{j} = \min_{\tau, j \in N_{R_j}} f(x^\tau_{j}) , \tau \leq t$$

where $N_{R_j}$ means the neighborhood of $j$-th particle which connects each other, and $l_{\text{best}}^t_{j}$ represents the local best location information of the $j$-th particle on $t$-th iteration within the neighborhood $N_{R_j}$.

We use the local best information instead of the global best information in Eq. (5). Namely, the desired fixed point is changed as

$$p^t_{j} = c_1 p_{\text{best}}^t_{j} + c_2 l_{\text{best}}^t_{j}$$

In our numerical simulations, the best value is selected in the local best values.

IV. RELATIONSHIP BETWEEN NETWORK STRUCTURE AND THE PERFORMANCE

In order to clarify a relationship between the network structures and its searching performance, we carry out some numerical simulations. First, we apply a simple network structure, that is extended cyclic graph, to the connection structure of the PSO. The extended cyclic graph is shown in Fig. 1, has a strong symmetric property, namely, all nodes have the same average distance. In order to confirm whether the same characteristics can be obtained in other structures, we apply a small world structure. In this article, we apply the Watts and Strogatz model[17]. The the Watts and Strogatz model can be derived from the extended cyclic structure to exchange some edges. For example, the extended cyclic as shown in Fig. 2(a) derives the Watts and Strogatz model as shown in Fig. 2(b).

We apply such Watts and Strogatz model to the connection of the D-PSO. The average distance of the Watts and Strogatz model becomes short comparing with the extended cyclic graph. Depend on the random selection, the average distance of the Watts and Strogatz model is changed. Thus, we carry out numerous numerical simulations, and the results are sorted by each average distance. In the simulations, we apply 50-dimensional benchmark functions.

The parameters set as shown in Table II. The number of particles of our simulations is 50. Thus, if a particle is regarded as a node, its maximum degree is 49. We carry out some numerical experiments that changed the average distance of the node by changing the network structure. Figure 3 shows the results. The horizontal axis denotes the average distance, and the vertical axis denotes the error between the optimal solution and the given solution. Blue curve denotes the results of the network structure is the extended cyclic graph. Red
Fig. 4. Comparison of the D-PSO and the conventional stochastic PSO. (a) and (b) are the results of the D-PSO. (c) and (d) are the results of the conventional stochastic PSO. Rastrigin function is applied as a benchmark function in (a) and (c). Rosenbrock function is applied in (b) and (d).

In previous section, we found the long average distance leads to the good search performance on the D-PSO. We have to confirm whether this is a specific property of the D-PSO. Thus, we carry out some numerical simulations for the conventional stochastic PSO which is described as Eqs. 1 and 2. The same parameters are applied to the conventional stochastic PSO. Figure 4 shows the results of the numerical simulations. In each graph, the horizontal axis denotes the average distance, and the vertical axis denotes the error between the optimum solution and the finding solution. Figures 4(a) and (b) show the results of the D-PSO. Figures 4(c) and (d) show the results of the conventional stochastic PSO. Rastrigin function is applied to (a) and (c), and Rosenbrock function is applied to (b) and (d).

In these figures, the vertical axis denotes the error between the optimal solution and the given solution. The horizontal axis denotes the average distance of each node. The number of particles is 50, thus, the average distance is 1 means full

### TABLE II

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>the number of particles $P$</td>
<td>50</td>
</tr>
<tr>
<td>the dimension</td>
<td>50</td>
</tr>
<tr>
<td>$w$</td>
<td>0.729</td>
</tr>
<tr>
<td>$\psi$</td>
<td>1.39445</td>
</tr>
<tr>
<td>iteration number</td>
<td>3000</td>
</tr>
<tr>
<td>trial number</td>
<td>30</td>
</tr>
</tbody>
</table>
connection. Namely, the full connection corresponds to the conventional PSO which uses the \textit{gbest}.

Based on these results, we confirmed that the search performance properties of the D-PSO and the conventional stochastic PSO are completely different to the average distance. The search performance of the D-PSO is improved according to the average distance. On the other hand, the search performance of the conventional stochastic PSO is deteriorated according to the average distance. The difference point between these two systems is the presence of the stochastic factors \(r_1\) and \(r_2\) in Eq. 1. The stochastic factor is a uniform random number whose range is \([0, 1]\). The D-PSO treats this stochastic factor as 1 as a constant. Therefore, we consider a stochastic system that links these two systems. We regard the lower limit of the random range as a parameter. Namely, the range is set as \([r, 1]\) where \(r\) denotes the lower limit.

Figures 5 and 6 show the results of such systems. In Figs. 5 and 6, the horizontal axis denotes the average distance, and the vertical axis denotes the error between the optimal value and the given evaluation value. Each line represents the each case of \(r\). (a) shows the cases from \(r = 0.1\) to \(r = 0.5\), and (b) shows the cases from \(r = 0.6\) to \(r = 0.9\). Note that the case of \(r = 0.0\) corresponds to the conventional stochastic PSO, and the case of \(r = 1.0\) corresponds to the D-PSO.

The results of Figs. 5 and 6 indicate that the search performance is improved according to an increase of \(r\). Especially, the search performance exhibits superior results when the average distance becomes long. The stochastic factor increases the diversification of each particle, thus, the stochastic system exhibits good search performance. On the other hand, the particles of the D-PSO do not have the diversity, because the particle of the D-PSO does not have stochastic factor. However, if the average distance is long, the transmission speed of the beneficial information of each particle becomes
Fig. 7. The relationship between the lower limit of the random number range and the search performance. Each range of the random number range is $[r, 1]$.

slow. We consider that the slow transmission of the beneficial information leads to the diversification of the particles of the D-PSO. Figure 7 shows the relationship between the lower limit of the random number range and the search performance. In these numerical simulations, the lower limit of the random range $r$ is changed from 0.05 to 0.95 with 0.05 step. 50-dimensional Rastrigin function is applied to the case of Fig. 7(a), and 50-dimensional Rosenbrock function is applied to the case of Fig. 7(b). These simulation results indicate the large lower limit exerts the superior search performance on each particle. The large lower limit denotes that the random range becomes narrow band. Namely, we consider the small perturbation is effective for search. However, the theoretical analysis of this point is not sufficient, thus, the rigorous theoretical analysis is a future problem.

VI. CONCLUSIONS

In this article, we paid attention to the network structure of the PSO. The extended cyclic graph and the small-world structure are applied to the D-PSO and the conventional stochastic PSO. Our numerical simulation results indicates the searching performance is depended on the average distance of the node. In the case of the D-PSO, the long average distance leads to superior search result. In the case of the conventional stochastic PSO, however, the log average distance is deteriorated according to the average distance. The more theoretical analysis is one of our future problems.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (C) 22560389.

REFERENCES