A Hybrid Swarm Intelligence Technique for Solving Integer Multi-objective Problems

Ibrahim M. El-henawy
Department of Computer Science, Faculty of Computers and Informatics, Zagazig University, Egypt

Mahmoud M. Ismail
Department of Operations Research, Faculty of Computers and Informatics, Zagazig University, Egypt

ABSTRACT
The multi-objective integer programming problems are considered time consuming. In the past, mathematical structures were used that can get benefits of high processing powers and parallel processing. A general approach to generate all non-dominated solutions of the multi-objective integer programming (MOIP) Problem is developed. In this paper, a hybridization of two different swarm intelligent approaches, stochastic diffusion search, and particle swarm optimization techniques is presented for solving integer multi-objective problems. The hybrid implementation allows us to avoid certain drawbacks and weaknesses of each algorithm, which means that we are able to find an optimal solution in an acceptable computational time. Our hybrid implementation allows the MOIP algorithm to reach the optimal solution in a considerably shorter time than is needed to solve the model using the entire dataset directly within the model. Our hybrid approach outperforms the results obtained by each technique separately. It is able to find the optimal solution in a shorter time than each technique on its own, and the results are highly competitive with the state-of-the-art in large-scale optimization. Furthermore, according to our results, combining the PSO with SDS approach for solving IP problems appears to be an interesting research area in combinatorial optimization.

Keywords
Swarm Intelligence, Integer programming, Multi-objective, Stochastic Diffusion Search and Particle Swarm Optimization

1. INTRODUCTION
Optimization can be viewed as one of the major quantitative tools in network of decision making, in which decisions have to be taken to optimize one or more objectives in some prescribed set of circumstances. In view of the practical utility of optimization problems there is a need for efficient and robust computational algorithms, which can numerically solve on computers the mathematical models of medium as well as large size optimization problem arising in different fields. Heuristics and bioinspired techniques have become efficient and effective alternatives for researchers in solving several complex optimization problems. These techniques are not able to reach the optimal solution for large-scale combinatorial optimization problems in spite of their effectiveness. But these techniques are able to provide satisfactory solutions for most of the applied problems within acceptable computational times. In contrast, mathematical programming techniques, particularly the Integer multi-objective problems, have been studied and developed by scholars over several decades with the main goal of obtaining optimal solutions to difficult problems using as little CPU time as possible. For these reasons, a hybrid swarm intelligence technique has been suggested. In recent years, swarm intelligence, which can be considered as a branch of Artificial Intelligence techniques, has attracted much attention of researchers, and has been applied successfully to solve a variety of problems. A swarm can be viewed as a group of agents cooperating with certain behavioural pattern to achieve some goal [12]. There are a number of different models of swarm intelligence that have been proposed and investigated, and among the most commonly used swarm intelligence models include ant colony optimization [5, 9], particle swarm optimization [7, 14], honey bee swarming [25, 26], stochastic diffusion search, and bacterial foraging [22, 23]. These algorithms have proved their mettle in solving complex and intricate optimization problems arising in various fields. The paper is organized such that the next section provides a brief overview of integer programming and some definitions. Section 3 describes Particle Swarm Optimization technique. Section 4 describes Stochastic Diffusion Search technique. Section 5 describes the method of the proposed hybrid swarm intelligence technique used. Section 6 discusses the computational results. In section 7, a conclusion is introduced.

2. MULTIOBJECTIVE PROBLEMS
[6, 11, 17, 24] Most real world optimization problems involve the optimization of more than one function, which in turn can require a significant computational time to be evaluated. That is, decision makers are not able to base their decisions on a single criterion in order to identify one or more attractive courses of action. Moreover, real decision problems are such that conflicting criteria are often used to evaluate alternative solutions. Multiobjective Programming is the branch of Mathematical Programming that deals with problems for which more than one objective function is required to evaluate the merit of alternative decisions. Formally, the problem is stated as follows:

\[
\begin{align*}
\text{MAX} & \quad (f_1(x), f_2(x), \ldots, f_n(x)) \\
\text{s.t.} & \quad x \in X \\
\text{Where} & \quad X = (x_1, x_2, \ldots, x_n) \\
\end{align*}
\]

\(X\) is the set of feasible solutions

\(f_i\) is the \(i\)th objective function (or decision criterion).

Then the problem doesn't have a single solution that could optimize all objectives simultaneously, but it has efficient (Pareto-optimal) solutions that can best attain the objectives as greatly as possible. In this context, deterministic techniques are difficult to apply to obtain the set of Pareto optimal solutions of many multiobjective optimization problems (MOPs), so stochastic methods have been widely used and applied. Among them, the use of evolutionary algorithms for...
solving MOPs has significantly grown in the last years, giving raise to a wide variety of algorithms, such as NSGA-II, SPEA2, PAES, and many others. Most multiobjective programming techniques focus on finding the set of efficient points (E) for a given problem or, in the case of heuristic procedures, an approximation of the efficient set (E*).

To select a suitable compromise solution from all efficient alternatives, a decision process is necessary. Depending on how the computation and the decision processes are combined in search for compromise solution, there exist three broad classes of multiobjective optimization methods:
1. A prior articulation of preferences (a prior methods)
2. A posterior articulation of preferences (posterior method).
3. progressive articulation of preferences (Interactive methods)

The DM provides global preference information during the optimization process.

2.1 Definition 1 (complete optimal solution)
\( \mathbf{X}^* \) is said to be a complete optimal solution, if and only if there exists \( \mathbf{x} \in \mathbf{X} \) such that
\[ z_i(\mathbf{x}^*) \leq z_i(\mathbf{x}), i = 1, \ldots, k, \] for all \( \mathbf{x} \in \mathbf{X} \).

2.2 Definition 2 (Pareto optimal solution)
\( \mathbf{X} \) is said to be a Pareto optimal solution, if and only if there doesn't exist another \( \mathbf{x} \in \mathbf{X} \) such that
\[ z_i(\mathbf{x}) \leq z_i(\mathbf{x}^*), i = 1, \ldots, k, \] for all \( i \) and for all \( \mathbf{x} \)
\[ z_j(\mathbf{x}) \neq z_j(\mathbf{x}^*) \) for at least one \( j \).

2.3 Definition 3 (Weak Pareto optimal solution)
\( \mathbf{X}^* \) is said to be a weak Pareto optimal solution, if and only if there doesn’t exist another \( \mathbf{x} \in \mathbf{X} \) such that
\[ z_i(\mathbf{x}) < z_i(\mathbf{x}^*), i = 1, \ldots, k. \]

2.4 Definition 2.5 (Ideal Vector)
Is the Vector containing the decision variables corresponding to the optima of the objective functions of the problems considering each objective separately.

2.5 Definition 2.5 (Pareto Front)
For a given MOP \( f(\mathbf{x}) \) and Pareto optimal set \( \mathbf{P}^* \), the Pareto front (PF*) is defined as:
\[ \text{PF}^* = \{ \mathbf{u} \rightarrow f = f_1, f_2, \ldots, f_n | \mathbf{x} \in \mathbf{P}^* \}. \]

2.6 Definition 6 Pareto Dominance
A vector \( \mathbf{u} = (u_1, \ldots, u_n) \) is said to dominate
\[ \mathbf{v} = (v_1, \ldots, v_n) \] (denoted by \( \mathbf{u} \leq \mathbf{v} \)) if and only if \( \mathbf{u} \) is partially less than \( \mathbf{v} \).

3. PARTICLE SWARM OPTIMIZATION
The particle swarm optimization (PSO) was inspired by the observations of birds flocking and fish schooling. It differs from other well-known Evolutionary Algorithms (EA) [2, 10, 13]. As in EA a population of potential solutions is used to probe the search space; but, no operators, inspired by evolution procedure, are applied on the population to generate a new promising solution. Instead, in PSO, each individual (named particle) of the population (called swarm), adjusts its trajectory towards its own previous best solution (called pbest) and the previous best solution attained by any member of its topological neighborhood. There are different kinds of sharing information between particles. In the global variant of PSO, the whole swarm is considered as the neighborhood. Thus, global sharing of information takes place and the particles benefit from the discoveries and the previous experiences of all other companions during the search for promising regions of the landscape [16]. Alternatively, there are some local variants of PSO wherein particles only make use of their own information and that of the best of their adjacent neighbors. Each particle in PSO has two main characteristics: its position and its velocity. Assume that the current position and velocity vector of the i-th particle in the d-dimensional search space are denoted as \( \mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{id}) \) and \( \mathbf{v}_i = (v_{i1}, v_{i2}, \ldots, v_{id}) \), respectively. The best earlier position of the i-th particle is represented as \( \text{pbest}_i = (\text{pbest}_{i1}, \text{pbest}_{i2}, \ldots, \text{pbest}_{id}) \). There are different kinds of PSO including global vision of PSO with inertia weight (GWPSO), local vision of PSO with inertia weight (LWPSO), global vision of PSO with constriction factor (GCPSO), and local vision of PSO with constriction factor (LCPSO) [27]. In GWPSO, which is very popular among researchers, there are two methods for updating position and velocity of each particle. The best position of entire group at k-th iteration is used in the first method while in the second method; the best position of entire group up to the current search is employed. In the first method, the position \( \mathbf{x}_{id}^k \) and velocity \( \mathbf{v}_{id}^k \) of particle \( i \) in the k-th iteration are updated as follow:
\[ \mathbf{x}_{id}^{k+1} = \mathbf{x}_{id}^k + \mathbf{v}_{id}^{k+1} \]
\[ \mathbf{v}_{id}^{k+1} = w \mathbf{v}_{id}^k + c_1 r_1 (\text{pbest}_{id}^k - \mathbf{x}_{id}^k) + c_2 r_2 (\text{gbest} - \mathbf{x}_{id}^k) \]

In the second equation w is the inertia weight. \( c_1 \) and \( c_2 \) are positive constants called cognitive and social parameters, respectively; and \( r_1 \) and \( r_2 \) are random numbers selected in the interval [0 1]. The constants \( c_1 \) and \( c_2 \) represent the weighting of the stochastic acceleration terms that pull each particle towards pbest and gbest positions and usually are set \( c_1 = c_2 = 2 \). In the second method gbest is replaced by gbestk. As will be shown later, in the numerical examples of mixed-variables or in the problems that only have discrete variables, usage of gbestk is more suitable compared to the use of gbest. In other words, the success rate of gbestk is higher than that of gbest. The reason is firstly due to the fast convergence of gbest and secondly, the inability of particles to escape from local minima in gbest method. In other words, the success rate of gbestk is higher than that of gbest. The reason is firstly due to the fast convergence of gbest and secondly, the inability of particles to escape from local minima in gbest method. In other words, the success rate of gbestk is higher than that of gbest. The reason is firstly due to the fast convergence of gbest and secondly, the inability of particles to escape from local minima in gbest method. In other words, the success rate of gbestk is higher than that of gbest. The reason is firstly due to the fast convergence of gbest and secondly, the inability of particles to escape from local minima in gbest method.
Where $w_{\text{max}}$ and $w_{\text{min}}$ are the maximum and minimum values of the inertia term, respectively, and $k_{\text{max}}$ is the maximum number of iterations. These parameters are assumed to be:

$$w_{\text{max}} = 1, \quad w_{\text{min}} = 0$$

Sometimes as particle oscillations become wider, the system will decrease an average number of PS and PSO is the partial function evaluation - convergence criteria [18] and linear time complexity [21]. The convergence to global optimum [20], robustness and minimal overheads entailed when deploying a PSO onto a problem domain, have some disadvantages of PSO Algorithm as follow:

- The method easily suffers from the partial optimism, which causes the less exact at the regulation of its speed and the direction.
- The method cannot work out the problems of scattering and optimization.
- The method cannot work out the problems of non-coordinate system, such as the solution to the energy field and the moving rules of the particles in the energy field.

4. STOCHASTIC DIFFUSION SEARCH

This section introduces Stochastic Diffusion Search (SDS) [3], a multi-agent global search and optimization algorithm, which is based on simple interaction of agents. A high-level description of SDS is presented in the form of a social metaphor demonstrating the procedures through which SDS allocates resources. SDS introduced a new probabilistic approach for solving best-fit pattern recognition and matching problems. SDS, as a multi-agent population-based global search and optimization algorithm, is a distributed mode of computation utilizing interaction between simple agents [8]. Unlike many nature inspired search algorithms, SDS has a strong mathematical framework, which describes the behavior of the algorithm by investigating its resource allocation [19], convergence to global optimum [20], robustness and minimal convergence criteria [18] and linear time complexity [21]. The SDS algorithm commences a search or optimization by initializing its population. In any SDS search, each agent maintains a hypothesis, $h$, defining a possible problem solution. After initialization two phases are followed, test phase, and diffusion phase. In the test phase, SDS checks whether the agent hypothesis is successful or not by performing a partial hypothesis evaluation which returns a boolean value. Later in the iteration, contingent on the precise recruitment strategy employed, successful hypotheses diffuse across the population and in this way information on potentially good solutions spreads throughout the entire population of agents. In the Test phase, each agent performs partial function evaluation, pFE, which is some function of the agent’s hypothesis; $pFE = f(h)$. In the diffusion phase, each agent recruits another agent for interaction and potential communication of hypothesis.

![Initialising agents](image1)

---

**Fig 1: The pseudocode of PSO algorithm**

Although PSO has been used mainly to solve unconstrained, single-objective optimization problems, PSO algorithms have been developed to solve constrained problems, multi-objective optimization problems, problems with dynamically changing landscapes, and to find multiple solutions. There are some disadvantages of PSO Algorithm as follow:

- The method easily suffers from the partial optimism, which causes the less exact at the regulation of its speed and the direction.
- The method cannot work out the problems of scattering and optimization.
- The method cannot work out the problems of non-coordinate system, such as the solution to the energy field and the moving rules of the particles in the energy field.

**Fig 2: The pseudocode of SDS algorithm**

Passive recruitment mode is employed in SDS algorithm. In this mode, if the agent is inactive, a second agent is randomly selected for diffusion; if the second agent is active, its hypothesis is communicated to the inactive one. Otherwise a completely new hypothesis is generated for the first inactive agent at random. The main disadvantage of the SDS is in the case of search spaces distorted heavily by noise, diffusion of activity due to disturbances will decrease an average number of inactive agents taking part in random search and in effect will increase the time needed to reach the steady state.

**Fig 3: Passive Recruitment Mode**

5. PROPOSED SDS-PSO TECHNIQUE

Because of the drawbacks of PSO algorithm, a hybrid swarm intelligence technique called SDS-PSO technique has been proposed for solving integer programming problems to produce better solution by improving the effectiveness and reducing the limitations. The motivating thesis justifying the merging SDS and PSO is the partial function evaluation deployed in SDS, which may mitigate the high computational overheads entailed when deploying a PSO onto a problem with a costly fitness function. In the hybrid algorithm, each PSO particle has a position, and a velocity; each SDS agent, on the other hand, has hypothesis and status. Every PSO particle is an SDS agent too together termed psAgents. In the psAgent, SDS hypotheses are defined by the PSO particle positions and a status which determines whether the psAgent is active or inactive (see Figure 4).
Figure 5 shows the pseudocode of the proposed SDS-PSO technique. In the test-phase of a stochastic diffusion search, each agent has to partially evaluate its hypothesis. The fitness of each psAgent’s particle’s personal best is compared against that of a random psAgent; if the selecting psAgent has a better fitness value, it will become active, otherwise it is flagged inactive. On average, this mechanism will ensure 50% of psAgents remain active from one iteration to another. In the Diffusion Phase, each inactive psAgent picks another psAgent randomly, if the selected psAgent is active, the selected psAgent communicates its hypothesis to the inactive one; if the selected psAgent is inactive too, the selecting psAgent generates a new hypothesis at random from the search space. In the proposed technique, after each n number of PSO function evaluations, one full SDS cycle is executed.

5.1 Different Performance Measures
The use of performance measures (or metrics) allows a researcher or computational scientist to assess in a quantitative way the performance of their algorithms. The SDS-PSO field has no different. SDS-PSO performance measures tend to focus on the phenotype or objective domain as to the accuracy of the results. This is different to what most operations researchers do. They tend to use metrics in the genotype domain. But since there is an explicit mapping between the two, it doesn’t really matter in which domain that define that metrics. There are many techniques for measuring the performance of Multiobjective problems [15] such as Hyperarea and Ratio, and Generational Distance.

Figure 5. Pseudocode of the proposed technique
5.1.1 Hyperarea and Ratio (H, HR)
These metrics define the area of coverage PF\textsubscript{KNOWN} has with respect to the objective space. This would equate to the summation of all the areas of rectangles, bounded by the origin and \((f_1(x), f_2(x))\), for a two-objective SDS-PSO.

\[ H = \left( \bigcup_{i} a_i \mid v_i \in PF\textsubscript{KNOWN} \right) \]

Where \(v_i\) is a nondominated vector in \(PF\textsubscript{KNOWN}\) and \(a_i\) is the hyperarea calculated between the origin and vector \(v_i\). But if \(PF\textsubscript{KNOWN}\) is not convex, the results can be misleading. It is also assumed in this model that the origin is (0, 0) [4].

5.1.2 Generational Distance (GD)
It reports how far, on average, \(PF\textsubscript{KNOWN}\) is from \(PF\textsubscript{TRUE}\). This metric requires that the researcher knows \(PF\textsubscript{TRUE}\). It is mathematically defined as in the following equation

\[ GD = \frac{1}{n} \left( \sum_{i=1}^{n} d_i^p \right)^{1/p} \]
Where \( n \) is the number of vectors in \( PF_{\text{KNOWN}} \), \( p=2 \) and \( d \) is the Euclidean distance between each member and the closest member of \( PF_{\text{TRUE}} \) in the phenotype space. When \( GD = 0 \), \( PF_{\text{KNOWN}} = PF_{\text{TRUE}} \).

6. COMPUTATIONAL RESULTS

This section presents computational results of the proposed SDS-PSO technique to solve integer multi-objective problems. The computational results show performance comparison between the Non-Dominated Sorting Genetic Algorithm (NSGAII) and our proposed SDS-PSO technique. The proposed SDS-PSO technique runs on computer (Core 2 Due CPU, 2.2GHz, 2048MB RAM.). The technique has been compared to NSGAII with different measures. The first measure used is: Hyperarea and Ratio, the results are shown in table 1.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Hyperarea</th>
<th>Problem size</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSGAII</td>
<td>149674</td>
<td>10</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>149294</td>
<td>10</td>
</tr>
<tr>
<td>NSGAII</td>
<td>309855</td>
<td>50</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>335271</td>
<td>50</td>
</tr>
<tr>
<td>NSGAII</td>
<td>263880</td>
<td>100</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>263475</td>
<td>100</td>
</tr>
<tr>
<td>NSGAII</td>
<td>844427</td>
<td>500</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>852897</td>
<td>500</td>
</tr>
<tr>
<td>NSGAII</td>
<td>15647268</td>
<td>1000</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>17156790</td>
<td>1000</td>
</tr>
</tbody>
</table>

It’s appears from the previous table that the hyperarea covered by SDS-PSO is larger than the hyperarea covered by NSGAII. Another comparative measure have been used is the average of the nearest neighbor, the results of this measure are shown in table 2.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Average Distance for Nearest Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>NSGAII</td>
<td>40.6</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>82.6</td>
</tr>
</tbody>
</table>

For more analysis, in 1000 variables problem size we made parametric analysis to see the appropriate technique in defined range of weights.

Table 3. Parametric analysis for 1000 variables problem

<table>
<thead>
<tr>
<th>Technique</th>
<th>0 ≤ w₁ ≤ 0.3</th>
<th>0.3 &lt; w₁ ≤ 0.7</th>
<th>0.7 &lt; w₁ ≤ 1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSGAII</td>
<td>28.076</td>
<td>44.237447</td>
<td>43.9664</td>
</tr>
<tr>
<td>SDS-PSO</td>
<td>36.03377</td>
<td>41.8739</td>
<td>31.6227</td>
</tr>
</tbody>
</table>

The result was: for \( 0 \leq w_1 \leq 0.3 \) and \( 1.0 \geq w_2 \geq 0.7 \) NSGAII produces the smallest average distance to the nearest neighbor. For \( 0.7 < w_1 \leq 1.0 \) and \( 0.3 > w_2 \geq 1.0 \), SDS-PSO produces the smallest average distance to the nearest neighbor. For \( 0.3 < w_1 \leq 0.7 \) and \( 0.7 > w_2 \geq 0.3 \), the SDS-PSO produces the smallest average distance to the nearest neighbor. The results of parametric analysis for 1000 variables problem are shown in table 3.

7. CONCLUSION

This paper proposed a hybrid swarm intelligence technique called SDS-PSO to solve integer multi-objective problems, and compares its performance with the Non-Dominated Sorting Genetic Algorithm (NSGAII). The proposed technique effectively overcomes the drawbacks of PSO technique, such as the partial optimism, which causes the less exact at the regulation of its speed and the direction. It also, overcomes the drawbacks of SDS technique. SDS-PSO technique also, increases the efficiency of the solution process, improves the performance scalability, and increases the diversification of solutions at the same time, reducing the execution time in comparison with NSGAII technique. The proposed technique has been tested by solving a set of different knapsack problems. It is capable to provide a considerable reduction of time compared with other techniques most obviously at lower scale problems. In general, the proposed SDS-PSO technique seems an efficient alternative for solving Integer multi-objective problems, when deterministic approaches fail, or it could be considered as an algorithm for providing good initial points to deterministic methods, as the Branch and Bound technique, and thus, help them converge to the global minimizer of the integer problem.

8. REFERENCES


