Pareto Partial Dominance MOEA in Many-Objective Optimization

Hiroyuki Sato\textsuperscript{1}, Hernán E. Aguirre\textsuperscript{2,3} and Kiyoshi Tanaka\textsuperscript{3}

\textsuperscript{1} Faculty of Electro-Communications, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585 JAPAN
\textsuperscript{2} International Young Researcher Empowerment Center, Shinshu University
4-17-1 Wakasato, Nagano, 380-8553 JAPAN
\textsuperscript{3} Faculty of Engineering, Shinshu University
4-17-1 Wakasato, Nagano, 380-8553 JAPAN

Abstract. In this work, we propose a novel multi-objective evolutionary algorithm (MOEA) which improves search performance of MOEA especially in many-objective combinatorial optimization problems (MaOP). Pareto dominance based MOEAs such as NSGA-II and SPEA2 meet difficulty to rank solutions in the population, which noticeably deteriorates search performance as we increase the number of objectives. In the proposed method, we rank solutions by calculating Pareto partial dominance between solutions using \( r \) objective functions selected from \( m \) objective functions to induce appropriate selection pressure in the evolution process of MOEA. Also, we temporally switch \( r \) objective functions among \( mC_r \) combinations in every interval generations \( I_g \) to optimize all of the objective functions throughout the entire evolution process. In this work, we use many-objective 0/1 knapsack problems to verify the search performance of the proposed method and analyze its evolution behavior. Simulation results show that there is an optimum value for the number of objective functions \( r \) to be considered in Pareto partial dominance and the interval (generation numbers) \( I_g \) to maximize the entire search performance. Also, the search performance of the proposed method is superior to recent state-of-the-art MOEAs, i.e., IBEA, CDAS and MSOPS. Furthermore, we analyze the behavior of the proposed algorithm from the viewpoint of front distribution over generations.

1 Introduction
The research interest of the multi-objective evolutionary algorithm (MOEA) \cite{1} community has rapidly shifted to develop effective algorithms for many-objective optimization problems (MaOPs) because more objective functions should be considered and optimized in recent complex applications. However, in general, Pareto dominance based MOEAs such as NSGA-II \cite{2} and SPEA2 \cite{3} noticeably deteriorate their search performance as we increase the number of objectives to more than 4 \cite{4, 5}. This is because these MOEAs meet difficulty to rank solutions in the population, i.e., most of the solutions become non-dominated and the same rank is assigned to them, which seriously spoils proper selection pressure required in the evolution process.
To overcome this problem, recently some states-of-the-art MOEAs such as IBEA [6], CDAS [7] and MSOPS [4] have been proposed and verified the improvement of search performance to solve MaOPs. IBEA [6] introduces fine grained ranking of solutions by calculating fitness value based on some indicators which measure the degree of superiority for each solution in the population. CDAS [7] relaxes the concepts of Pareto dominance by controlling dominance area of solutions to induce appropriate selection pressure for the objective numbers in the problem to be solved [7]. MSOPS aggregates fitness vector with multiple weight vectors, and reflects the ranking of solutions calculated for each weight vector in parent selection [4].

In this work, we propose a novel MOEA to enhance search performance especially for MaOP. While conventional MOEAs select parent solutions by considering all \( m \) objective functions, we propose to rank solutions by calculating Pareto partial dominance which considers only \( r(< m) \) objectives selected from \( m \) objective functions to induce appropriate selection pressure in MaOP. Also, we temporally switch \( r \) objective functions among \( mC_r \) combinations every \( I_g \) generations to optimize all of the objective functions throughout the entire evolution process. In this work, we verify the superiority of the proposed algorithm in many-objective 0/1 knapsack problems with \( m = \{4, 6, 8, 10\} \) objectives by comparing with NSGA-II and recent state-of-the-art MOEAs, i.e., IBEA, CDAS and MSOPS. Furthermore, we analyze the behavior of the proposed algorithm from the viewpoint of front distribution over generations.

## 2 Multi-Objective Evolutionary Algorithms

### 2.1 Multi-Objective Optimization

A multi-objective optimization problem (MOP) including \( m \) kinds of objective functions is defined as follows:

\[
\begin{align*}
\text{Maximize } f(x) &= (f_1(x), f_2(x), \ldots, f_m(x)) \\
\text{subject to } x &\in \mathcal{F}
\end{align*}
\]

(1)

where, \( x \in \mathcal{F} \) is a feasible solution vector in the solution space \( \mathcal{S}(\mathcal{F} \subseteq \mathcal{S}) \), and \( f_i (i = 1, 2, \ldots, m) \) are the \( m \) objectives to be maximized\(^1\). That is, we try to find a feasible solution vector \( x \in \mathcal{F} \) in the solution space maximizing each objective function \( f_i (i = 1, 2, \ldots, m) \) in a vector fitness function \( f \). Here, we define Pareto dominance between solutions \( x, y \in \mathcal{F} \) as follows.

\[
\forall i \in \mathcal{M} : f_i(x) \geq f_i(y) \land \\
\exists i \in \mathcal{M} : f_i(x) > f_i(y).
\]

(2)

where, \( \mathcal{M} = \{1, 2, \ldots, m\} \). If Eq. (2) is satisfied, \( x \) dominates \( y \). In the following, when \( x \) dominates \( y \), we denote \( f(x) \succeq f(y) \). A solution vector \( x \) is said to be Pareto optimal with respect to \( \mathcal{F} \) if it is not dominated by other solution vectors in \( \mathcal{F} \). The presence of multiple objective functions, usually conflicting among

---

\(^1\) Throughout the paper, we consider maximization problems.
them, gives rise to a set of optimal solutions. The set of Pareto optimal solutions (POS) is defined as
\[ \mathcal{POS} = \{ x \in \mathcal{F} \mid \nexists y \in \mathcal{F} : f(y) \succeq f(x) \} , \] (3)
and the Pareto front is defined as
\[ \mathcal{F}_{\text{Front}} = \{ f(x) \mid x \in \mathcal{POS} \} . \] (4)

2.2 Problems in MaOP
NSGA-II [2] is one of the well-known MOEAs that use Pareto dominance, which robust performance has been verified on a wide range of MOP especially for two or three objective optimization problems. In parent selection process, NSGA-II first classifies solutions into several layers (fronts) based on non-dominance level. Then, it selects parent solutions from higher fronts until filling up the half size of entire population. When comparing solutions that belong to a same front, NSGA-II determines superiority of solutions based on crowding distance [2] which considers solution’s distribution in objective space.

When analyzing NSGA-II in MaOP, most of the solutions become non-dominated (front 1 : \( \mathcal{F}_1 \)), and the number of solutions belonging to \( \mathcal{F}_1 \) exceeds the size of parent solutions (half of the entire population) in early stage of the evolution. In such case, parent selection of NSGA-II becomes to strongly rely on crowding distance [2]. Consequently, the obtained POS are well-distributed in objective space, but convergence of POS towards the true POS is substantially deteriorated. To overcome this problem and enhance convergence of the obtained POS, it is necessity to improve selection pressure by discriminating non-dominated solutions using some effective manner.

3 Related Works
3.1 Indicator Based Fitness Assignment
Indicator-based evolutionary algorithm (IBEA) calculates fitness value \( F \) for all solutions by the following equation, and repeats to discard the solution with the lowest \( F \) to select new parent solutions [6].
\[ F(x) = \sum_{y \in \mathcal{P} \setminus \{ x \}} -e^{-I(y,x)/k} \] (5)

\( F(x) \) is calculated by considering all of the relations between a solution \( x \) and other solutions in the population \( \mathcal{P} \) based on indicator function \( I \). \( I \) determines a relative superiority (or inferiority) between two solutions. The additive binary indicator \( I_{+} \) [6] gives the minimum distance by which a Pareto set approximation needs to or can be translated in each dimension in objective space such that another approximation is weakly dominated.
\[ I_{+}(y,x) = \min \{ f_i(y) + \epsilon \geq f_i(x) : i \in 1, \cdots, m \} \] (6)
A solution \( x \) having a higher value of \( I_{+}(y,x) \) is considered as a better solution. In this way, IBEA achieves fine grained solution ranking by discriminating non-dominated solutions using the indicator function.
3.2 Controlling Dominance Area of Solutions

To induce appropriate selection pressure in MOEA, controlling dominance area of solutions (CDAS) [7] contracts or expands the dominance area of solutions. CDAS modifies the fitness value for each objective function by changing a parameter $S_i$ in the following equation

$$f'_i(x) = r \cdot \frac{\sin(\omega_i + S_i \cdot \pi)}{\sin(S_i \cdot \pi)} \quad (i = 1, 2, \cdots, m),$$

(7)

where $\varphi_i = S_i \cdot \pi$, $r$ is the norm of $f(x)$, $f_i(x)$ is the fitness value in the $i$-th objective, and $\omega_i$ is the declination angle between $f(x)$ and $f_i(x)$. When $S_i < 0.5$, the $i$-th fitness value $f_i(x)$ is increased to $f'_i(x) > f_i(x)$. On the other hand, when $S_i > 0.5$, $f_i(x)$ is decreased to $f'_i(x) < f_i(x)$. When $S_i = 0.5$, $f'_i(x) = f_i(x)$ which is equivalent to conventional dominance. CDAS shows better search performance in MaOP than conventional NSGA-II due to convergence by the effects of fine grained ranking of solutions using $S_i < 0.5$ [7].

3.3 Multiple Single Objective Pareto Sampling

Multiple single objective Pareto sampling (MSOPS) utilizes multiple scalarization function for parent selection [4]. This method creates $N \times W$ score matrix where $N$ is the population size and $W$ is the number of weight vectors. First, MSOPS fills up all the elements of the matrix by calculating

$$s = \min_{i=1}^{m}(w_i f_i(x)),$$

(8)

where $w_i$ $(i = 1, 2, \cdots, m)$ are weights in a weight vector and $x$ is a solution. Then, for each columns, MSOPS gives rank $1 \sim N$ to all solutions in descending order based on score values, i.e., a solution with the highest score is assigned rank 1, and a solution with the lowest score is assigned rank $N$ for each columns. Then, for each row, the indices of column are sorted in ascending order based on rank value. MSOPS selects solutions which have higher ranks for weight vectors as parent solutions. Compared with conventional NSGA-II, MSOPS shows better search performance on continuous MaOPs [4].

In this work, we compare the search performance of the proposed method in MaOP with conventional NSGA-II, IBEA$_{ϵ+}$, CDAS and MSOPS.

4 Proposed method

4.1 Concept of Pareto Partial Dominance

In this work, we rank solutions based on Pareto partial dominance which considers $r$ objectives selected from $m$ objective functions in the calculation of dominance. Pareto partial dominance is defined as follows:

$$\forall i \in R \subset M : f_i(x) \geq f_i(y) \land$$

$$\exists i \in R \subset M : f_i(x) > f_i(y),$$

(9)

where $R$ contains $r$ kinds of indices selected from $M = \{1, 2, \ldots, m\}$. There are $mC_r$ combinations when we select $r$ objectives from $m$ objective functions. Pareto
Procedure of Proposed Algorithm

initialize and evaluate($R_0$)

$O_{table} \leftarrow$ objects combination table($m$, $r$)

for $t \leftarrow 1 \ldots T$

\begin{itemize}
  \item $m \choose r$ non-dominated sorting ($R_t$, $O_{table} (k \mod m \choose r)$)
  \item crowding distance ($R_t$)
  \item $P_t \leftarrow$ truncation ($R_t$)
  \item $Q_t \leftarrow$ mating and reproduction ($P_t$)
  \item evaluate($Q_t$)
  \item $R_{t+1} \leftarrow Q_t \cup P_t$
\end{itemize}

if $t \mod I_g = 0$ or $t = T$ then

\begin{itemize}
  \item $m \choose r$ non-dominated sorting ($R_{t+1} \cup A$
  \item crowding distance ($R_{t+1} \cup A$
  \item $R_{t+1} \leftarrow$ truncation ($R_{t+1} \cup A$
  \item $A \leftarrow$ copy($R_{t+1}$)
  \item $k \leftarrow k + 1$
\end{itemize}

end

end

Fig. 1. Pseudo-code of the proposed algorithm

Partial dominance makes discrimination of solutions easier than conventional Pareto dominance using all $m$ objectives defined in Eq. (2) even in MaOPs, since dominance is partially considered for only $r$ objectives between solutions. In this way, we can expect to change selection pressure in the evolution process by fine-grained raking of solutions.

4.2 Algorithm

We show the pseudo-code of the proposed algorithm using Pareto partial dominance in Fig. 1. In this paper, we implement Pareto partial dominance based on the framework of NSGA-II. In our algorithm, the number of objectives $r$ considered in partial dominance and the interval generation $I_g$ to switch combination of $r$ objective functions are parameters. Before starting the evolution, we generate $O_{table}$ that lists the selection order of all the possible $m \choose r$ combinations of $r$ objectives from $m$ objective functions. We temporally switch combination in every $I_g$ generation by following the order in $O_{table}$. We perform non-dominated sorting by using the selected $r$ objectives, and reflect the result in parents selection. Furthermore, the proposed algorithm updates the archive population $A$ to maintain POS when we switch combination. Here, we select $R_{t+1}$ from $P_t \cup Q_t \cup A$ by using conventional non-dominated sorting which considers $m$ objective functions. Then $R_{t+1}$ is copied to $A$. Although the conventional NSGA-II has mechanism to keep non-dominated solution in the population as parent population $P_t$, the solutions kept in $P_t$ in the proposed algorithm are partially non-dominated solutions. Thus, to keep conventional non-dominated solutions that obtained by partial dominance during $I_g$ generations, the proposed algorithm updates archive population $A$.

4.3 Expected Effects of Pareto Partial Dominance

Here, we show the expected effects of Pareto partial dominance by a simple experiment. We generate a hundred of individuals having $m = 8$ random fitness
values in the range \([0, 1]\), and then classify all the solutions into fronts by the way of NSGA-II \([2]\). **Fig. 2** shows the number of solutions belonging to each front obtained by partial dominance as we vary the number of objectives \(r\) considered in partial dominance. All the plots are average of \(10^4\) times runs.

For example, for \(r = 8\), 82\% of solutions are belonging to \(F_1\). This result reveals the difficulty to discriminate solutions by conventional dominance in MaOP. On the other hand, as we decrease the number of objectives \(r\) to be considered in Pareto partial dominance, the number of solutions belonging to \(F_1\) decreases and the number of fronts increases. Thus, we can expect that Pareto partial dominance makes the solutions ranking more fine grained, which induces appropriate selection pressure in dominance based MOEAs.

Next, we observe that relation between front distributions obtained by conventional dominance and partial dominance. **Fig. 3** shows the front distribution obtained by partial dominance for the solutions classified into front 1 ∼ 3, \(F_1 \sim F_3\), by conventional dominance. The solutions belonging to conventional \(F_1\) are distributed from front 1 ∼ 24, \(F_1^{(p)} \sim F_24^{(p)}\) by partial dominance. On the other hand, some solutions belonging to conventional \(F_2\) and \(F_3\) become higher fronts by partial dominance than some solutions belonging to conventional \(F_1\). That is, most of solutions belonging to conventional higher fronts are classified into higher fronts by partial dominance with high probability, but sometime superiority of solutions replaced by partial dominance. That is, when optimizing only \(r\) objectives among \(m\) objective functions, some solutions belonging to lower fronts by conventional dominance survive in parent solutions.

5 Experimental Results and Discussion

To verify the search performance of the proposed algorithm, in this paper we use many-objective 0/1 knapsack problems \([9]\) as benchmark problem. We generate problems with \(m = \{4, 6, 8, 10\}\) objectives, \(n = 100\) items, and feasibility ratio \(\phi = 0.5\). For all algorithms to be compared, we adopt two-point crossover with a crossover rate \(P_c = 1.0\), and apply bit-flipping mutation with a mutation rate \(P_m = 1/n\). In the following experiments, we show the average performance with 30 runs, each of which spent \(T = 2,000\) generations. Population size is set to
Hypervolume
Interval generation ($I_g$)
0 100 200 300 400 500
1.8
1.9
2
2.1
($\times 10^{21}$)

(a) 6 objectives

(b) 8 objectives

Fig. 4. Hypervolume as we increase $r$ and $I_g$ (KP100-m)

$N = 200$ ($|P_t| = |Q_t| = 100$). In IBEA+$\epsilon$, scaling parameter $k$ is set to $0.05$ similar to [6]. In CDAS, we adopt an optimal parameter $S^*$, which maximizes $HV$ in the experiments. Also, in MSOPS, we use $W = 100$ uniformly distributed weight vectors [8], which maximizes $HV$ in the experiments. In the proposed algorithm, the size of archive population is set to $N_A = 200$.

In this work, to evaluate search performance of MOEA we use Hypervolume ($HV$) [10], which measures the $m$-dimensional volume of the region enclosed by the obtained non-dominated solutions and a dominated reference point in objective space. Here we use $r = (0,0,\ldots,0)$ as the reference point. Obtained POS showing a higher value of hypervolume can be considered as a better set of solutions from both convergence and diversity viewpoints. To provide additional information separately on convergence and diversity of the obtained POS, in this work we also use Norm [11] and Maximum Spread (MS) [10], respectively. Higher value of Norm generally means higher convergence to true POS. On the other hand, Higher MS indicates better diversity in POS which can be approximated widely spread Pareto front.

5.1 Performance Varying Parameters $r$ and $I_g$

First, we observe the effects of varying parameters $r$ that is the number of objectives to be considered in partial dominance, and $I_g$ that is the interval generation to switch combination of $r$ objectives in the proposed algorithm. Fig. 4 shows $HV$ achieved for various $r$ and $I_g$. $HV$ obtained by conventional NSGA-II, IBEA+$\epsilon$, CDAS, and MSOPS are also plotted different as horizontal lines.

From the results, we can see that there is an optimal parameter $r^*$ to maximize $HV$, i.e., $r^* = 2$ for $m = \{6,8\}$, respectively. Also, there is an optimal interval generation $I_g$ to maximize $HV$, i.e., $I^*_g = 50$ for $m = \{6,8\}$, respectively. Except the case of $r = 1$, we can increase $HV$ as we increase the interval generation $I_g$. However, in the case of $r = 1$ which uses single objective function for parent selection, $HV$ seriously deteriorates as we increase $I_g$. Also, we can see that peak regions of $HV$ become narrow as we increase the number of objectives. Although MSOPS shows the highest $HV$ among conventional MOEAs, the proposed algorithm with $r = 2$, $I_g = 50$ shows significantly better $HV$ than other algorithms compared in Fig. 4.
5.2 Performance Varying the Number of Objectives

Second, we observe the search performance of each algorithm as we vary the number of objectives. Fig. 5 (a) - (c) show results on HV as a comprehensive metric, Norm as a measure of convergence, and MS as a measure of diversity. In these figures, all the plots are normalized by the results of NSGA-II.

NSGA-II achieves the highest MS for \( m = \{6, 8, 10\} \) objectives, but the minimum Norm instead. That is, NSGA-II can obtain well distributed POS, but its convergence is insufficient in MaOP. Consequently, the values of HV are also the lowest for \( m = \{6, 8, 10\} \) objectives. IBEA achieves extremely high convergence, but scarce diversity. MSOPS achieves balanced search on both convergence and diversity, and consequently achieves higher HV among conventional MOEAs compared in Fig. 5. On the other hand, the proposed algorithm achieves highest diversity close to NSGA-II but also enhances convergence considerably from conventional NSGA-II. Consequently, the proposed algorithm achieves the highest values of HV compared to all algorithms. In other words, the proposed algorithm maintains well distributed POS similar to conventional NSGA-II, but its convergence is considerably improved. These results suggest that proposed partial dominance is an effective method to enhance the search performance of MOEAs in MaOP by achieving well-balance search with proper selection pressure. The effects become significant especially for the large number of objectives.

5.3 Behavior Analysis on the Proposed Algorithm

Fig. 6 (a) - (c) show the front distribution over generation by the conventional NSGA-II and by the proposed algorithm with \( r = \{2, 3\} \) and \( I_g = 50 \) in \( m = 8 \) objectives problem, respectively.

Fig. 6 (a) reveals that the difficulty to discriminate solutions by using conventional dominance in MaOP, since the number of solutions belonging to \( F_1 \) exceeds the parent population size \( |P| = 100 \) in early generations. On the other hand, the proposed algorithm can discriminate solutions into multiple fronts (layers) by using partial dominance with a small number of objectives \( r \). About the periodical dynamics in every \( I_g \) generations, once the combination of \( r \) objectives changes, the front distribution become fine grained based on new criterion. However, the number of solutions belonging to \( F_1^{(p)} \) rapidly increases and the
number of fronts decreases. This gradually deteriorates the selection pressure, and the combination of objectives should be changed to refresh the front distribution. Summarizing, the proposed algorithm can induce fine-grained ranking of solutions by Pareto partial dominance with a small number of objectives \( r (< m) \), and maintain proper (high) selection pressure throughout all generations of the solution search by switching the criterion to select parent solutions in every \( I_g \) generations.

Next, Fig. 7 show that among solutions belonging to partial front \( \mathcal{F}_1^{(p)} \) the fraction of solutions belonging to conventional front \( \mathcal{F}_1 \sim \mathcal{F}_3 \) over generation. The fraction of solutions belonging to conventional front \( \mathcal{F}_1 \) increases in the solutions belonging to partial front \( \mathcal{F}_1^{(p)} \) over generations. That is, the number of solutions belonging both to partial front \( \mathcal{F}_1^{(p)} \) and conventional front \( \mathcal{F}_1 \) increases along with the evolution. In the case of \( r = 3 \), after about 500 generations, we can see that most solutions in partial front \( \mathcal{F}_1^{(p)} \) are selected from conventional front \( \mathcal{F}_1 \).

### 6 Conclusions

In this paper, we have proposed a novel MOEA using Pareto partial dominance that considers \( r (< m) \) objectives selected from \( m \) objective functions to improve the search performance of MOEA especially for MaOP. We verified the search performance of the proposed algorithm on many-objective 0/1 knapsack
problems with $4 \sim 10$ objectives. Simulation results showed that significant improvement on $HV$ was achieved by the proposed algorithm with the optimum number of objectives $r^* = 2$ to be considered in partial dominance and the optimum interval generation $I_g^* = 50$ to switch the combination of objectives. Also, POS obtained by the proposed algorithm achieved higher diversity compared with conventional IBEA, CDAS and MSOPS. Furthermore the convergence was improved compared with conventional NSGA-II. That is, the proposed algorithm achieved well-balanced search between convergence toward true POS and diversity in objective space. Also, we revealed that the effectiveness of the proposed algorithm become significant as we increased the number of objectives.

As future works, we are planning to apply Pareto partial dominance to other Pareto based MOEAs. Also, we want to design the algorithm that can control parameters $r$ and $I_g$ adaptively.

References