

Ensemble Move Acceptance in Selection Hyper-heuristics

Ahmed Kheiri, Mustafa Mısıır and Ender Özcan

Abstract Selection hyper-heuristics are high level search methodologies which control a set of low level heuristics while solving a given problem. Move acceptance is a crucial component of selection hyper-heuristics, deciding whether to accept or reject a new solution at each step during the search process. This study investigates group decision making strategies as ensemble methods exploiting the strengths of multiple move acceptance methods for improved performance. The empirical results indicate the success of the proposed methods across six combinatorial optimisation problems from a benchmark as well as an examination timetabling problem.

1 Introduction

A selection hyper-heuristic is an iterative improvement oriented search method which embeds two key components; *heuristic selection* and *move acceptance* [1]. The heuristic selection method chooses and applies a heuristic from a set of low level heuristics to the solution in hand, producing a new one. Then the move acceptance method decides whether to accept or reject this solution. The modularity, use of machine learning techniques and utilisation of the *domain barrier* make hyper-heuristics more general search methodologies than the current techniques tailored for a particular domain are. A selection hyper-heuristic or its components can be reused on another problem domain without requiring any change. There is a growing number of studies on selection hyper-heuristics combining a range of simple heuristic selection and move acceptance methods [2, 3]. More on any type of hyper-heuristic, such as their components and application areas can be found in [1].

Hyper-heuristics Flexible Framework (HyFlex) [4] was proposed as a software platform for rapid development and testing of hyper-heuristics. HyFlex is implemented in Java along with six different problem domains: boolean satisfiability, bin-

Ahmed Kheiri
Cardiff University, Operational Research Group, School of Mathematics, Senghennydd Road,
Cardiff, CF24 4AG, UK, e-mail: KheiriA@cardiff.ac.uk

Mustafa Mısıır
University of Freiburg, Department of Computer Science, Georges-Köhler-Allee 052, 79110
Freiburg, Germany, e-mail: misir@informatik.uni-freiburg.de

Ender Özcan
University of Nottingham, ASAP Research Group, School of Computer Science, Jubilee Campus,
Wollaton Road, Nottingham, NG8 1BB, UK, e-mail: ender.ozcan@nottingham.ac.uk

packing, permutation flow-shop, personnel scheduling, travelling salesman problem and vehicle routing problem. HyFlex was used in the first Cross-Domain Heuristic Search Challenge, CHeSC 2011 (<http://www.asap.cs.nott.ac.uk/chesc2011/>) to detect the best selection hyper-heuristic. Following the competition, the results from twenty competing selection hyper-heuristics across thirty problem instances (containing five instances from each HyFlex domain) and the description of their algorithms were provided at the competition web-page.

A recent theoretical study on selection hyper-heuristics in [5] showed that the mixing of simple move acceptance criteria could lead to an improved running-time complexity than using each move acceptance method standalone on some simple benchmark functions. The work in [6] and [3] showed that the performance of the hyper-heuristic on a given instance depends on the selection hyper-heuristic components. This study extends the initial work in Özcan et al. [7] by applying and evaluating four group decision making strategies as ensemble methods using three different move acceptance methods in combination with seven heuristic selection methods on an examination timetabling problem [6]. The same selection hyper-heuristics are then tested on thirty problem instances from six different domains from the HyFlex benchmark.

2 Group Decision Making Selection Hyper-heuristics

An overview of heuristic selection and move acceptance methods as a part of the selection hyper-heuristics as well as the group decision making methods forming an ensemble of move acceptance used in this study is described in this section.

A range of simple heuristic selection methods were studied in [2]. *Simple Random* (SR) selects a heuristic at random at each decision point. *Random Descent* (RD) also selects a heuristic at random, and then applies it to the candidate solution as long as the solution is improved. *Random Permutation* (RP) generates a random permutation of heuristics and applies one heuristic at a time in that order. *Random Permutation Descent* (RPD) is based on the same RP strategy, however similar to RD, applies the same heuristic repeatedly until there is no more improvement. *Greedy* (GR) applies all low level heuristics to the current solution and selects the heuristic which generates the best improvement. *Choice Function* (CF) is an online learning heuristic selection method that scores each low level heuristic based on their utility value and selects the one with the highest score. A *Tabu Search* based hyper-heuristic (TABU) that maintains a tabu list of badly performing low level heuristics to disallow the selection of these heuristics was tested in [8].

This paper studies ensemble move acceptance methods combining them under a group decision making framework. Considering that a constituent move acceptance method returns either true (1) or false (0) at each decision point, Equation 1 provides a general model for an ensemble of k methods. In this model, each move acceptance carries a certain *strength* (s_i) which adjusts its contribution towards a final acceptance decision.

$$\sum_{i=1}^k s_i \times D(M_i) \geq \alpha \quad (1)$$

where M_i is the i^{th} move acceptance (group member), $D(m)$ returns 1, if a solution is accepted by the move acceptance method m , and 0, if rejected.

In this study, we use group decision making strategies which make an accept/reject decision based on *authority*, *minority* and *majority* rules, namely G-OR (the move acceptance method which accepts the solution has the authority), G-AND (minority decides rejection), G-VOT and G-PVO (considers majority of the votes for the accept/reject decision). G-PVO probabilistically makes the accept/reject decisions. The probability that a new solution is accepted changes dynamically in proportional to the number of members that voted to the acceptance of the new solution. For instance, assuming 6 members in the group out of 10 move acceptance methods accepts a solution at a given step, then G-PVO accepts the solution with a probability of 60%. It is preferable in G-VOT to have an odd number of members for the group decision making move acceptance criteria, where none of the other strategies requires this. More formally, using Equation 1, assuming k move acceptance methods, then for G-AND, G-OR and G-VOT, α is k , 0.5 and $k/2$, respectively, where all s_i values are set to 1. For G-PVO, α equals $k * r$, where r is uniform random number in $[0, 1]$, and s_i values equal $1/k$.

In this study, the heuristic selection methods in {SR, RD, RP, RPD, CF, GR, TABU} are paired with four group decision making move acceptance mechanisms {G-AND, G-OR, G-VOT, G-PVO}, generating twenty eight group decision making selection hyper-heuristics. From this point forward, a selection hyper-heuristic will be denoted as “*heuristic selection method*”_“*move acceptance method*”. For example, SR_G-AND denotes the selection hyper-heuristic using SR as the heuristic selection method and G-AND as the move acceptance method.

Each group decision making move acceptance ensemble tested in this study embeds three move acceptance methods: Improving and Equal (IE), Simulated Annealing (MC) and Great Deluge (GD). IE accepts all non-worsening moves and rejects the rest. Similarly, GD and SA accept all non-worsening moves as well as any worsening move without exceeding a certain threshold and with a certain probability, respectively. These group members are chosen to form the ensemble move acceptance due to their high performance reported in [3].

Simulated Annealing [9] move acceptance criterion, denoted as MC in this paper, accepts all improving moves but the non-improving moves are accepted with a probabilistic formula, p_t , shown in Equation 2.

$$p_t = e^{-\frac{\Delta f}{\Delta F(1-\frac{t}{T})}} \quad (2)$$

where Δf is the fitness change at time or step t , T is the time limit or the maximum number of steps and ΔF is an expected range for the maximum fitness change.

GD acceptance criterion accepts all the improving moves but the non-improving moves are accepted if the objective value of the current solution is not worse than

an expected value, named as level [10]. Equation 3 is used to update the threshold level τ_t at time or step t .

$$\tau_t = F + \Delta f \times \left(1 - \frac{t}{T}\right) \quad (3)$$

where T is the time limit or the maximum number of steps, Δf is an expected range for the maximum fitness change and F is the final objective value.

3 Computational Experiments

Pentium IV 3 GHz LINUX machines having 2.00GB memories are used during the experiments. Following the rules of CHeSC 2011, each trial is run for 10 nominal minutes with respect to the competition machine respecting the challenge rules. The group decision making selection hyper-heuristics are tested on an examination timetabling problem as formulated in [6] and the same termination criterion as in that study is used for the examination timetabling experiments to enable a fair performance comparison of solution methods. The GD and SA move acceptance methods use the same parameter settings as provided in [7].

Two sets of benchmarks are used for examination timetabling: Yeditepe [11, 12] and Toronto benchmarks [13] consisting of eight and fourteen instances, respectively. The mean performance of each group decision making move acceptance method in a selection hyper-heuristic regardless of the heuristic selection method is compared to each other based on their ranks. The group decision making move acceptance methods are ranked from 1 to 4 for each problem instance and heuristic selection method from best to worst based on the mean cost over fifty runs. The approaches are assigned to different ranks if their performances vary in a statistically significant manner for a given instance. Otherwise, their performances are considered to be similar and an average rank is assigned to them all. A similar outcome is observed for the online performances of the group decision making strategies as in the benchmark functions reported in [7]. G-VOT is the best acceptance mechanism based on the average rank over all the problems, while G-PVO, G-AND and G-OR follows it in that order, respectively as illustrated in Figure 1.

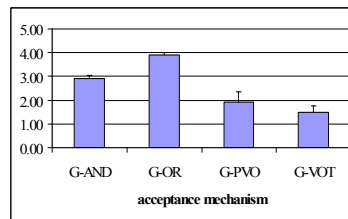


Fig. 1 Mean rank (and the standard deviation) of each group decision making move acceptance mechanism considering their average performance over all runs

Similarly, all twenty eight hyper-heuristics are ranked from 1 to 28 (best to worst) based on the best objective values obtained over fifty runs for each instance. The ranks are averaged/shared in case of a tie. Figure 2 illustrates the performance of six group decision making selection hyper-heuristics with a better mean performance that are significantly better as compared to the rest, from the best to the worst; GR_G-VOT, TABU_G-VOT, RP_G-VOT, GR_G-PVO, SR_G-VOT and CF_G-VOT.

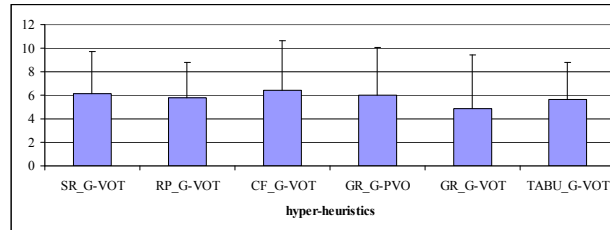


Fig. 2 Mean rank (and standard deviation) of the group decision making hyper-heuristics that generate statistically significant performance variance from the rest over all examination timetabling problems.

Table 1 compares the average performances of the best six group decision making hyper-heuristics (see Figure 2) to the best hyper-heuristic for each problem instance reported in [6]. Hyper-heuristics with multiple move acceptance methods under decision making models generated superior performance compared to the hyper-heuristics where each utilises a single move acceptance method. This performance variation is statistically significant within a confidence interval of 95% based on the Wilcoxon signed-rank test. In eighteen out of the twenty one problems, hyper-heuristics with the majority rule voting as their acceptance criterion, namely G-VOT and G-PVO deliver the best performances. There is a tie between the simulated annealing based hyper-heuristics and group decision making hyper-heuristics for sta83 I and yue20013. It is also known that there is an optimal solution for yue20023 [11]. GR_G-PVO improves the average performance of CF_MC for yue20023, still, all the hyper-heuristics seem to get stuck at local optima while solving sta83 I, yue20013 and yue20023. Excluding yue20032, the group decision making hyper-heuristics improve the average performance of previous best hyper-heuristics by 30.7% over all problem instances. RP_G-PVO delivers a similar average performance to CF_MC for yue20032, yet CF_MC is slightly better. Large improvements are observed for large problem instances, such as car91 I and car92 I. Overall, the experimental results confirm that group decision making hyper-heuristics have great potential.

The twenty eight hyper-heuristics are implemented as an extension to HyFlex to check their level of generality across the CHeSC 2011 problem domains. Each experiment is repeated thirty one times following the competition rules. All hyper-heuristics are ranked using the Formula 1 scoring system. The best hyper-heuristic obtaining the best median objective value over all runs for each instance gets 10 points, the second one gets 8, and then 6, 5, 4, 3, 2, 1 and the rest gets zero point.

Table 1 *%imp.* denotes the percentage improvement over the average best cost across fifty runs that the ‘current’ best hyper-heuristic(s) (investigated in this work) produces over the ‘previous’ best hyper-heuristic (reported in [6]) for each problem instance. If a hyper-heuristic delivers a statistically significant performance, it appears in the ‘current’ column. Bold entries highlight the best performing method. The hyper-heuristics that have a similar performance to the bold entry are displayed in parentheses. “+” indicates that all hyper-heuristics in {GR_G-VOT, TABU_G-VOT, RP_G-VOT, GR_G-PVO, SR_G-VOT, CF_G-VOT} has similar performance. “/” excludes the hyper-heuristic from this set that is displayed afterwards.

instance	current	previous	<i>%imp.</i>
yue20011	GR_G-VOT+	SR_GD	20.84
yue20012	RP_G-VOT+	SR_GD	24.93
yue20013	+	SR_MC	0
yue20021	TABU_G-VOT+	SR_GD	17.97
yue20022	GR_G-PVO	CF_MC	3.97
yue20023	GR_G-PVO	CF_MC	1.97
yue20031	GR_G-PVO (GR_G-VOT, SR_G-VOT)	CF_MC	4.4
yue20032	n/a	CF_MC	n/a
car91 I	GR_G-VOT+	TABU_IE	81.37
car92 I	GR_G-VOT+/GR_G-PVO	TABU_IE	196.89
ear83 I	GR_G-PVO (GR_G-VOT)	CF_MC	1.1
hecs92 I	GR_G-PVO (GR_G-VOT, SR_G-VOT, TABU_G-VOT)	CF_MC	21.46
kfu93	GR_G-VOT+	SR_GD	30.88
lse91	GR_G-PVO+	CF_MC	13.38
pur93 I	GR_G-PVO (SR_G-VOT)	SR_IE	15.6
rye92	TABU_G-VOT+	CF_MC	41.67
sta83 I	+	SR_MC	0
tre92	GR_G-VOT+	SR_GD	92.93
uta92 I	GR_G-VOT+/GR_G-PVO	TABU_IE	36.36
ute92	GR_G-PVO	CF_MC	0
yor83 I	GR_G-PVO+	CF_MC	9.01

These points are accumulated over all instances across all domains forming the final score for each hyper-heuristic.

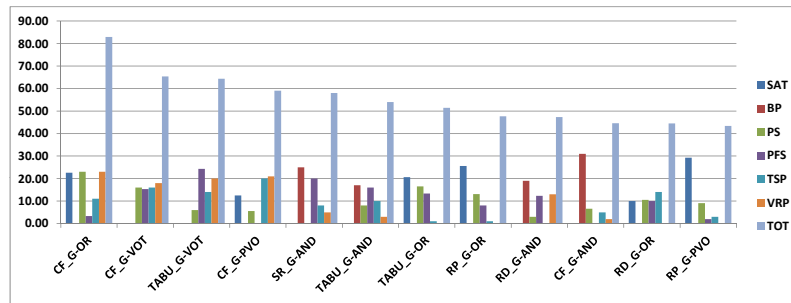


Fig. 3 Median performance comparisons between different group decision making hyper-heuristics based on their Formula 1 scores.

Firstly, performance of all group decision making hyper-heuristics are compared to each other. Figure 3 summarises the results including top twelve out of twenty eight approaches. In the overall, CF_G-OR, CF_G-VOT and TABU_G-VOT are the top three group decision making methods, while GR_G-AND and GR_G-OR are the worst. RP_G-PVO, CF_G-AND, CF_G-OR, TABU_G-VOT, CF_G-PVO and CF_G-OR perform the best on boolean satisfiability (SAT), bin-packing (BP), personnel scheduling (PS), permutation flow-shop (PFS), travelling salesman (TSP) and vehicle routing problems (VRP), respectively. Table 2 summarises the ranking of those six group decision making hyper-heuristics and all competing hyper-heuristics at CHeSC 2011, including the top ranking method, denoted as AdapHH. The top ten ranking hyper-heuristics from the competition remains in their positions and group decision making methods perform relatively poor. CF_G-AND is the third best approach for BP. TABU_G-VOT comes sixth for PS. TABU_G-VOT, CF_G-AND and CF_G-VOT score better than the CHeSC 2011 winner for the same problem. CF_G-OR is the best among the group decision making methods for SAT, ranking the eighth. The best group decision making hyper-heuristic for TSP, i.e. CF_G-OR, takes the ninth place. For VRP, CF_G-VOT as the best hyper-heuristic with group decision making is the sixth best approach among the CHeSC 2011 competitors. However, its performance on VRP is still better than the winning approach. The performance of all group decision making methods is poor on the PFS problem. CF_G-AND is the group decision making hyper-heuristic winner and it ranks the eleventh when compared to the CHeSC 2011 hyper-heuristics with a total score of 39.00.

Table 2 Ranking of selected group decision making hyper-heuristics to the CHeSC 2011 competitors based on Formula 1.

Rank	HH	Total	SAT	BP	PS	PFS	TSP	VRP
1	AdapHH	170.00	33.75	43.00	6.00	37.00	40.25	10.00
7	HAHA	65.75	31.75	0.00	19.50	3.50	0.00	11.00
11	CF_G-AND	39.00	0.00	25.00	10.00	0.00	0.00	4.00
14	CF_G-OR	27.50	9.50	0.00	2.00	0.00	8.00	8.00
15	CF_G-VOT	23.50	0.00	0.00	8.50	0.00	4.00	11.00
20	CF_G-PVO	16.14	0.14	0.00	1.00	0.00	6.00	9.00
22	TABU_G-VOT	11.50	0.00	0.00	11.50	0.00	0.00	0.00
23	RP_G-PVO	7.00	6.50	0.00	0.50	0.00	0.00	0.00

4 Conclusion

The experimental results show that the ensemble move acceptance methods based on group decision making models can exploit the strength of constituent move acceptance methods yielding an improved performance. In general, learning heuristic selection performs well within group decision making hyper-heuristics. Considering their performance over the examination timetabling benchmark problems, Greedy performs the best as a heuristic selection method. Combining multiple move acceptance methods using a majority rule improves the performance of Greedy as com-

pared to using a single move acceptance method. On the other side, CF outperforms other standard heuristic selection schemes on the CHeSC 2011 benchmark, performing reasonably well in combination with AND-operator group decision making move acceptance. The proposed ensemble move acceptance methods enable the use of the existing move acceptance methods and do not introduce any extra parameters other than the constituent methods have. Discovering the best choice of move acceptance methods in the ensemble as well as their weights is left as a future work. More interestingly, new adaptive ensemble move acceptance methods, which are capable of adjusting the weight/strength of each constituent move acceptance during the search process, can be designed for improved cross domain performance.

References

1. Burke, E.K., Gendreau, M., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., Qu, R.: Hyper-heuristics: A survey of the state of the art. *Journal of the Operational Research Society* **64** (2013) 1695–1724
2. Cowling, P., Kendall, G., Soubeiga, E.: A hyperheuristic approach to scheduling a sales summit. In: Selected papers from the Third International Conference on Practice and Theory of Automated Timetabling, London, UK, Springer-Verlag (2001) 176–190
3. Özcan, E., Bilgin, B., Korkmaz, E.E.: A comprehensive analysis of hyper-heuristics. *Intelligent Data Analysis* **12** (2008) 3–23
4. Ochoa, G., Hyde, M., Curtois, T., Vazquez-Rodriguez, J., Walker, J., Gendreau, M., Kendall, G., McCollum, B., Parkes, A., Petrovic, S., Burke, E.: Hyflex: A benchmark framework for cross-domain heuristic search. In Hao, J.K., Middendorf, M., eds.: *European Conference on Evolutionary Computation in Combinatorial Optimisation*. Volume 7245 of LNCS. (2012) 136–147
5. Lehre, P.K., Özcan, E.: A runtime analysis of simple hyper-heuristics: To mix or not to mix operators. In: *Proceedings of the Twelfth Workshop on Foundations of Genetic Algorithms XII. FOGA XII '13*, New York, NY, USA, ACM (2013) 97–104
6. Bilgin, B., Özcan, E., Korkmaz, E.: An experimental study on hyper-heuristics and exam timetabling. In Burke, E., Rudová, H., eds.: *Practice and Theory of Automated Timetabling VI*. Volume 3867 of LNCS. (2007) 394–412
7. Özcan, E., Misir, M., Kheiri, A.: Group decision making hyper-heuristics for function optimisation. In: *Computational Intelligence (UKCI), 2013 13th UK Workshop on*. (2013) 327–333
8. Cowling, P., Chakhlevitch, K.: Hyperheuristics for managing a large collection of low level heuristics to schedule personnel. In: *Proceedings of the 2003 Congress on Evolutionary Computation*. (2003) 1214–1221
9. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. *Science* **220** (1983) 671–680
10. Kendall, G., Mohamad, M.: Channel assignment optimisation using a hyper-heuristic. In: *Proceedings of the 2004 IEEE Conference on Cybernetic and Intelligent Systems (CIS2004)*, Singapore (2004) 790–795
11. Parkes, A.J., Özcan, E.: Properties of yeditepe examination timetabling benchmark instances. In: *Proceedings of the 8th International Conference on the Practice and Theory of Automated Timetabling*. (2010) 531–534
12. Özcan, E., Ersoy, E.: Final exam scheduler - fes. In Corne, D., Michalewicz, Z., McKay, B., Eiben, G., Fogel, D., Fonseca, C., Greenwood, G., Raidl, G., Tan, K.C., Zalzal, A., eds.: *Proceedings of 2005 IEEE Congress on Evolutionary Computation*. (2005) 1356–1363
13. Carter, M.W., Laporte, G., Lee, S.Y.: Examination timetabling: Algorithmic strategies and applications. *Journal of The Operational Research Society* **47** (1996) 373–383