Optim Lett (2014) 8:2203–2210
DOI 10.1007/s11590-014-0728-8

ORIGINAL PAPER

Exact model for the cell formation problem

Ilya Bychkov · Mikhail Batsyn · Panos M. Pardalos

Received: 22 March 2013 / Accepted: 28 January 2014 / Published online: 12 February 2014

Abstract The cell formation problem (CFP) consists in an optimal grouping of the given machines and parts into cells, so that machines in every cell process as much as possible parts from this cell (intra-cell operations) and as less as possible parts from other cells (inter-cell operations). The grouping efficacy is the objective function for the CFP which simultaneously maximizes the number of intra-cell operations and minimizes the number of inter-cell operations. Currently there are no exact approaches (known to the authors) suggested for solving the CFP with the grouping efficacy objective. The only exact model which solves the CFP in a restricted formulation is due to Elbenani and Ferland (Cell formation problem solved exactly with the dinkelbach algorithm. Montreal. Quebec. CIRRELT-2012-07, 1–14, 2012). The restriction consists in fixing the number of production cells. The main difficulty of the CFP is the fractional objective function—the grouping efficacy. In this paper we address this issue for the CFP in its common formulation with a variable number of cells. Our computational experiments are made for the most popular set of 35 benchmark instances. For the 14 of these instances using CPLEX software we prove that the best known solutions are exact global optimums.

I. Bychkov · M. Batsyn (✉) · P. M. Pardalos
Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics, 136 Rodionova, Nizhniy Novgorod 603093, Russian Federation
e-mail: mbatsyn@hse.ru

I. Bychkov
e-mail: il.bychkov@gmail.com

P. M. Pardalos
Center of Applied Optimization, University of Florida, 401 Weil Hall, P.O. Box 116595, Gainesville, FL 32611-6595, USA
e-mail: pardalos@ufl.edu

Keywords Cell formation problem · Exact model · Grouping efficacy · Fractional objective function

1 Introduction

The cell formation problem (CFP) consists in an optimal grouping of the given machines and parts into cells so that the grouping efficacy is maximized. The input for this problem is given by m machines, p parts, and a rectangular machine-part matrix $[a_{ij}]$, where $a_{ij} = 1$ if part j is processed by machine i. The grouping efficacy is given by the following formula:

$$\tau = \frac{n_1^{in}}{n_1 + n_0^{in}}.$$

The CFP is NP-hard since it can be reduced from the clustering problem [11]. As a result there appeared many heuristic approaches to the CFP giving solutions of high quality in a reasonable time [12,13,24]. Quality of the CFP solutions is usually determined by using the grouping efficacy measure [17]:

$$\tau = \frac{n_1^{in}}{n_1 + n_0^{in}},$$

where n_1 is the number of ones in the machine-part matrix, n_1^{in} and n_0^{in} are the numbers of ones and zeroes inside the cells. This function is fractional and this is one of the main difficulties of the CFP. To our knowledge currently only one exact model is suggested for the CFP, but in a restricted formulation—Elbenani and Ferland [10]. The restriction of this model consists in fixing the number of production cells. In this paper we suggest the first exact model for the CFP in its common formulation with a variable number of cells. Our computational experiments are made for the most popular set of 35 benchmark instances. For the 14 of these instances using CPLEX software we prove that current best known solutions are exact global optimums.

2 Problem formulation

The CFP is defined by its machine-part matrix. As an example in Table 1 the machine-part matrix of the CFP instance from Waghodekar and Sahu [29] is presented. This CFP instance has five machines and seven parts. Its optimal solution with the grouping efficacy $\tau = 16/(20 + 3) \approx 0.6957$ is presented in Table 2. This solution has two cells, the first cell contains machine 1 together with parts 1, 6, 7 and the second one contains machines 2, 3, 4, 5 together with parts 2, 3, 4, 5.

Since the grouping efficacy is a fractional objective function, we look over all the possible values of the denominator (specifically n_0^{in}, because n_1 is constant). We solve the CFP separately for every value of n_0^{in} adding a constraint requiring the number of zeroes inside cells to be equal to the chosen constant n_0^{in}. This way the original

Table 1 Machine-part 5 × 7 matrix for the CFP instance from Waghodekar and Sahu [29]

	1	2	3	4	5	6	7
1	1	0	0	0	1	1	1
2	0	1	1	1	1	0	0
3	0	0	1	1	1	1	0
4	1	1	1	1	0	0	0
5	0	1	0	1	1	1	0

Table 2 Optimal solution for the CFP instance from Waghodekar and Sahu [29]

	1	6	7	2	3	4	5
1	**1**	**1**	**1**	0	0	0	1
2	0	0	0	**1**	**1**	**1**	**1**
3	0	1	0	**0**	**1**	**1**	**1**
4	1	0	0	**1**	**1**	**1**	**0**
5	0	1	0	**1**	**0**	**1**	**1**

Bold values indicates the solution

objective function transforms into the linear function max n_1^{in} and we have to solve several such linear integer problems, one for every fixed value n_0^{in}. Then the optimal solution for the CFP is the optimal solution of the problem which has the greatest grouping efficacy among these linear integer problems.

Our exact model of the CFP can be described using boolean variables x_{ik} and y_{jk}. Variable x_{ik} is equal to 1 if machine i belongs to cell k and is equal to 0 otherwise. Similarly variable y_{jk} is equal to 1 if part j belongs to cell k and is equal to 0 otherwise. Machines index i takes values from 1 to m and parts index j—from 1 to p. Cells index k takes values from 1 to $c = \min(m, p)$ because every cell should contain at least one machine and one part and so the number of cells cannot be greater than m and p. The number of ones inside cells is equal to $\sum_{k=1}^{c}\sum_{i=1}^{m}\sum_{j=1}^{p} a_{ij}x_{ik}y_{jk}$, and the number of zeroes inside cells is equal to $\sum_{k=1}^{c}\sum_{i=1}^{m}\sum_{j=1}^{p}(1-a_{ij})x_{ik}y_{jk}$. We linearize the product $x_{ik}y_{jk}$ in a standard way introducing new boolean variables $z_{ijk} = x_{ik}y_{jk}$ and additional linear constraints (2)–(4). The suggested model is as follows.

$$max \sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{c} a_{ij}z_{ijk} \tag{1}$$

subject to

$$z_{ijk} \le x_{ik} \quad \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p, \ \forall k = 1, \ldots, c \tag{2}$$

$$z_{ijk} \le y_{jk} \quad \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p, \ \forall k = 1, \ldots, c \tag{3}$$

$$z_{ijk} \ge x_{ik} + y_{jk} - 1 \quad \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p, \ \forall k = 1, \ldots, c \tag{4}$$

$$\sum_{k=1}^{c} x_{ik} = 1 \quad \forall i = 1, \ldots, m \tag{5}$$

$$\sum_{k=1}^{c} y_{jk} = 1 \quad \forall j = 1, \ldots, p \tag{6}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{p} z_{ijk} \geq \sum_{i=1}^{m} x_{ik} \quad \forall k = 1, \ldots, c \tag{7}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{p} z_{ijk} \geq \sum_{j=1}^{p} y_{jk} \quad \forall k = 1, \ldots, c \tag{8}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{p}\sum_{k=1}^{c}(1 - a_{ij})z_{ijk} = n_0^{in} \tag{9}$$

$$x_{ik}, y_{jk}, z_{ijk} \in \{0, 1\} \quad \forall i = 1, \ldots, m, \ \forall j = 1, \ldots, p, \ \forall k = 1, \ldots, c \tag{10}$$

Constraints (2)–(4) guarantee that $z_{ijk} = x_{ik}y_{jk}$. Constraints (5), (6) require that every machine and every part is assigned to exactly one cell. Constraints (7), (8) require that there are no cells which have only machines without parts or only parts without machines. Constraint (9) fixes the total number of zeroes inside cells to be equal to the chosen constant n_0^{in}.

We solve model (1)–(10) for all possible values of n_0^{in} using CPLEX 12 solver. To limit the maximum possible number of zeroes inside cells a heuristic solution can be used. Proposition 1 provides an upper bound on the number of zeroes inside cells.

Proposition 1 *Let τ be the grouping efficacy value for some feasible CFP solution. Then n_0^{in} in the optimal solution is not greater than $\left\lfloor \frac{1-\tau}{\tau} n_1 \right\rfloor$.*

Proof Since τ is the value of the objective function of a feasible CFP solution, then it is not greater than the optimal value τ^*

$$\tau \leq \tau^* = \frac{n_1^{in}}{n_1 + n_0^{in}}$$

Therefore,

$$n_0^{in} \leq \frac{n_1^{in} - \tau n_1}{\tau}$$

Since $n_1^{in} \leq n_1$ and n_0^{in} is integer, we have the required upper bound

$$n_0^{in} \leq \left\lfloor \frac{1-\tau}{\tau} n_1 \right\rfloor$$

Note that when the considered feasible solution is optimal ($\tau = \tau^*$) and it contains all the ones inside the cells ($n_1^{in} = n_1$) then n_0^{in} is exactly equal to this upper bound. □

The greater the value of τ we have, the tighter is the upper bound for the number of zeroes inside cells. So in this paper, we use the grouping efficacy τ of the best known solution to reduce the number of boolean linear problems (1)–(10) needed to be solved in our approach. Note that in general case, the lower bound on n_0^{in} is trivial: $n_0^{in} \geq 0$.

This bound is reached for example on a solution with no zeroes inside cells and no ones outside cells which has 100 % efficacy.

Another way to improve the model is to add bounds on the number of ones inside cells again using the grouping efficacy of the best known heuristic solution. Such bounds allow CPLEX to prune branches more efficiently. An upper bound on n_1^{in} is trivial: $n_1^{in} \leq n_1$. It is reached on an ideal solution with 100 % efficacy. A lower bound for n_1^{in} is provided in proposition 2.

Proposition 2 *Let τ be the grouping efficacy value for some feasible CFP solution. Then n_1^{in} in the optimal solution is not less than $\lceil \tau(n_1 + n_0^{in}) \rceil$.*

Proof Since τ is the value of the objective function of a feasible CFP solution, then it is not greater than the optimal value τ^*

$$\tau \leq \tau^* = \frac{n_1^{in}}{n_1 + n_0^{in}}$$

Since n_1^{in} is integer, we have the required lower bound

$$n_1^{in} \geq \left\lceil \tau(n_1 + n_0^{in}) \right\rceil$$

In our model, we fix n_0^{in} and thus it is a constant which can be used in the expression for the lower bound on n_1^{in}. □

Using proposition 2, we can add the following inequalities to our model

$$\left\lceil \tau(n_1 + n_0^{in}) \right\rceil \leq \sum_{i=1}^{m} \sum_{j=1}^{p} \sum_{k=1}^{c} a_{ij} z_{ijk} \leq n_1 \tag{11}$$

As an example, let us consider two smallest instances from the literature: 5×7 CFP instance from King and Nakornchai [15] and 5×7 CFP instance from Waghodekar and Sahu [29]. For the first instance, the best known heuristic solution has $\tau = 82.35\,\%$ and the upper bound on n_0^{in} is $[(1 - 0.8235)/0.8235 \cdot 16] = 3$. For the second instance $\tau = 69.57\,\%$ and the upper bound is $[(1 - 0.6957)/0.6957 \cdot 20] = 8$. The efficacy values obtained by solving model (1)–(11) with all possible values of n_0^{in} for these two instances are shown in Table 3. Thus the globally optimal solution for instance 1 has three zeroes inside cells and the efficacy 82.35 %, and for instance 2—also three zeroes inside cells and the efficacy 69.57 %.

Table 3 Solutions for all possible values of n_0^{in}

#	Size	Zeroes inside								
		0	1	2	3	4	5	6	7	8
1	5 × 7	78.57	80.00	81.25	**82.35**					
2	5 × 7	60.00	61.90	63.64	**69.57**	62.50	68.00	63.54	55.56	60.71

Table 4 Computational results

#	Source	Size	Efficacy (%)	Time (s)	Zeroes in
1	King and Nakornchai [15]	5 × 7	82.35	0.63	3
2	Waghodekar and Sahu [29]	5 × 7	69.57	2.29	3
3	Seifoddini [25]	5 × 18	79.59	5.69	3
4	Kusiak [19]	6 × 8	76.92	1.86	4
5	Kusiak and Chow [20]	7 × 11	60.87	9.14	0
6	Boctor [2]	7 × 11	70.83	5.15	3
7	Seifoddini and Wolfe [26]	8 × 12	69.44	13.37	1
8	Chandrasekharan and Rajagopalan [6]	8 × 20	85.25	18.33	0
9	Chandrasekharan and Rajagopalan [7]	8 × 20	58.72	208.36	18
10	Mosier and Taube [22]	10 × 10	75.00	6.25	4
11	Chan and Milner [5]	10 × 15	92.00	2.93	4
12	Askin and Subramanian [1]	14 × 23	72.06	259.19	10
13	Stanfel [28]	14 × 24	71.83	179.21	10
14	McCormick et al. [21]	16 × 24	51.61[a]	20,829.38[a]	8
15	Srinivasan et al. [27]	16 × 30	69.00[a]	13,719.99[a]	13
16	King [14]	16 × 43	57.53[a]	24,930.93[a]	20
17	Carrie [4]	18 × 24	57.73[a]	13,250.01[a]	8
18	Mosier and Taube [23]	20 × 20	38.71[a]	43,531.77[a]	44
19	Kumar et al. [16]	20 × 23	46.72[a]	33,020.13[a]	9
20	Carrie [4]	20 × 35	77.85[a]	11,626.98[a]	22
21	Boe and Cheng [3]	20 × 35	46.75[a]	33,322.08[a]	1
22	Chandrasekharan and Rajagopalan [9]	24 × 40	100.00	1.64	0
23	Chandrasekharan and Rajagopalan [9]	24 × 40	85.11[a]	6,916.24[a]	11
24	Chandrasekharan and Rajagopalan [9]	24 × 40	56.49[a]	14,408.88[a]	0
25	Chandrasekharan and Rajagopalan [9]	24 × 40	46.56[a]	34,524.47[a]	0
26	Chandrasekharan and Rajagopalan [9]	24 × 40	43.51[a]	41,140.94[a]	0
27	Chandrasekharan and Rajagopalan [9]	24 × 40	41.22[a]	44,126.76[a]	0
28	McCormick et al. [21]	27 × 27	54.02[a]	22,627.28[a]	31
29	Carrie [4]	28 × 46	24.65[a]	71,671.08[a]	4
30	Kumar and Vannelli [18]	30 × 41	48.44[a]	22,594.20[a]	0
31	Stanfel [28]	30 × 50	50.65[a]	31,080.82[a]	0
32	Stanfel [28]	30 × 50	38.32[a]	48,977.01[a]	0
33	King and Nakornchai [15]	30 × 90	39.41[a]	99,435.64[a]	29
34	McCormick et al. [21]	37 × 53	59.60[a]	47,744.04[a]	17
35	Chandrasekharan and Rajagopalan [8]	40 × 100	84.03[a]	24,167.76[a]	37

[a] Some of the subproblems have not been solved to optimality within time limit of 300 s
Bold values indicates the optimal values

3 Computational results

For the computational experiments, we use the set of the most popular 35 CFP instances from the literature [13,24]. The first 13 problems and the 22nd problem could be solved exactly in quite a little computational time (Table 4). For all of these instances, the solutions found by our model are equal to the best known solutions. Thus, for the 14 instances, the global optimality of the best known solutions is proved. Solving the remaining instances requires too much memory and computational time. For such instances when we add inequalities (11) CPLEX solver cannot find any solution within 10 h. So for these instances we use our model without inequalities (11), run CPLEX with 300 s time limit for every subproblem, and report the best found value of grouping efficacy. All the computations are performed on an Intel Core i7 processor running at 2.2 GHz with 8 GB RAM.

4 Conclusion

There are a lot of papers devoted to the CFP with the grouping efficacy as an objective function, but almost all of them present different heuristic algorithms. This is probably connected with the complexity of this problem caused by its fractional objective and large amount of feasible solutions even for small numbers of machines and parts. In this paper we have suggested an exact approach which allows us to replace the original fractional programming problem with several integer programming problems. We have also provided two propositions which help to reduce the number of the IP problems to be solved and to improve the IP model itself. As a result we are able to solve 14 of the 35 most popular benchmark instances with sizes from 5×7 to 24×40.

Acknowledgments The authors are partially supported by LATNA Laboratory, NRU HSE, RF government grant, ag. 11.G34.31.0057.

References

1. Askin, R.G., Subramanian, S.P.: A cost-based heuristic for group technology configuration. Int. J. Prod. Res. **25**(1), 101–113 (1987)
2. Boctor, F.F.: A linear formulation of the machine-part cell formation problem. Int. J. Prod. Res. **29**(2), 343–356 (1991)
3. Boe, W., Cheng, C.H.: A close neighbor algorithm for designing cellular manufacturing systems. Int. J. Prod. Res. **29**(10), 2097–2116 (1991)
4. Carrie, S.: Numerical taxonomy applied to group technology and plant layout. Int. J. Prod. Res. **11**, 399–416 (1973)
5. Chan, H.M., Milner, D.A.: Direct clustering algorithm for group formation in cellular manufacture. J. Manuf. Syst. **1**(1), 64–76 (1982)
6. Chandrasekharan, M.P., Rajagopalan, R.: MODROC: an extension of rank order clustering for group technology. Int. J. Prod. Res. **24**(5), 1221–1233 (1986a)
7. Chandrasekharan, M.P., Rajagopalan, R.: An ideal seed non-hierarchical clustering algorithm for cellular manufacturing. Int. J. Prod. Res. **24**(2), 451–464 (1986b)
8. Chandrasekharan, M.P., Rajagopalan, R.: ZODIAC: an algorithm for concurrent formation of part families and machine cells. Int. J. Prod. Res. **25**(6), 835–850 (1987)
9. Chandrasekharan, M.P., Rajagopalan, R.: Groupability: analysis of the properties of binary data matrices for group technology. Int. J. Prod. Res. **27**(6), 1035–1052 (1989)

10. Elbenani, B., Ferland, J.A.: Cell Formation Problem Solved Exactly with the Dinkelbach Algorithm. Montreal. Quebec. CIRRELT-2012-07, pp. 1–14 (2012)
11. Ghosh, S., Mahanti, A., Nagi, R., Nau, D.S.: Manufacturing cell formation by state-space search. Ann. Oper. Res. **65**(1), 35–54 (1996)
12. Goncalves, J.F., Resende, M.G.C.: An evolutionary algorithm for manufacturing cell formation. Comput. Ind. Eng. **47**, 247–273 (2004)
13. James, T.L., Brown, E.C., Keeling, K.B.: A hybrid grouping genetic algorithm for the cell formation problem. Comput. Oper. Res. **34**(7), 2059–2079 (2007)
14. King, J.R.: Machine-component grouping in production flow analysis: an approach using a rank order clustering algorithm. Int. J. Prod. Res. **18**(2), 213–232 (1980)
15. King, J.R., Nakornchai, V.: Machine-component group formation in group technology: review and extension. Int. J. Prod. Res. **20**(2), 117–133 (1982)
16. Kumar, K.R., Kusiak, A., Vannelli, A.: Grouping of parts and components in flexible manufacturing systems. Eur. J. Oper. Res. **24**, 387–397 (1986)
17. Kumar, K.R., Chandrasekharan, M.P.: Grouping efficacy: a quantitative criterion for goodness of block diagonal forms of binary matrices in group technology. Int. J. Prod. Res. **28**(2), 233–243 (1990)
18. Kumar, K.R., Vannelli, A.: Strategic subcontracting for efficient disaggregated manufacturing. Int. J. Prod. Res. **25**(12), 1715–1728 (1987)
19. Kusiak, A.: The generalized group technology concept. Int. J. Prod. Res. **25**(4), 561–569 (1987)
20. Kusiak, A., Chow, W.S.: Efficient solving of the group technology problem. J. Manuf. Syst. **6**(2), 117–124 (1987)
21. McCormick, W.T., Schweitzer, P.J., White, T.W.: Problem decomposition and data reorganization by a clustering technique. Oper. Res. **20**(5), 993–1009 (1972)
22. Mosier, C.T., Taube, L.: The facets of group technology and their impact on implementation. OMEGA **13**(6), 381–391 (1985a)
23. Mosier, C.T., Taube, L.: Weighted similarity measure heuristics for the group technology machine clustering problem. OMEGA **13**(6), 577–583 (1985b)
24. Paydar, M.M., Saidi-Mehrabad, M.: A hybrid genetic-variable neighborhood search algorithm for the cell formation problem based on grouping efficacy. Comput. Oper. Res. **40**(4), 980–990 (2013)
25. Seifoddini, H.: A note on the similarity coefficient method and the problem of improper machine assignment in group technology applications. Int. J. Prod. Res. **27**(7), 1161–1165 (1989)
26. Seifoddini, H., Wolfe, P.M.: Application of the similarity coefficient method in group technology. IIE Trans. **18**(3), 271–277 (1986)
27. Srinivasan, G., Narendran, T.T., Mahadevan, B.: An assignment model for the part-families problem in group technology. Int. J. Prod. Res. **28**(1), 145–152 (1990)
28. Stanfel, L.: Machine clustering for economic production. Eng. Costs Prod. Econ. **9**, 73–81 (1985)
29. Waghodekar, P.H., Sahu, S.: Machine-component cell formation in group technology MACE. Int. J. Prod. Res. **22**, 937–948 (1984)