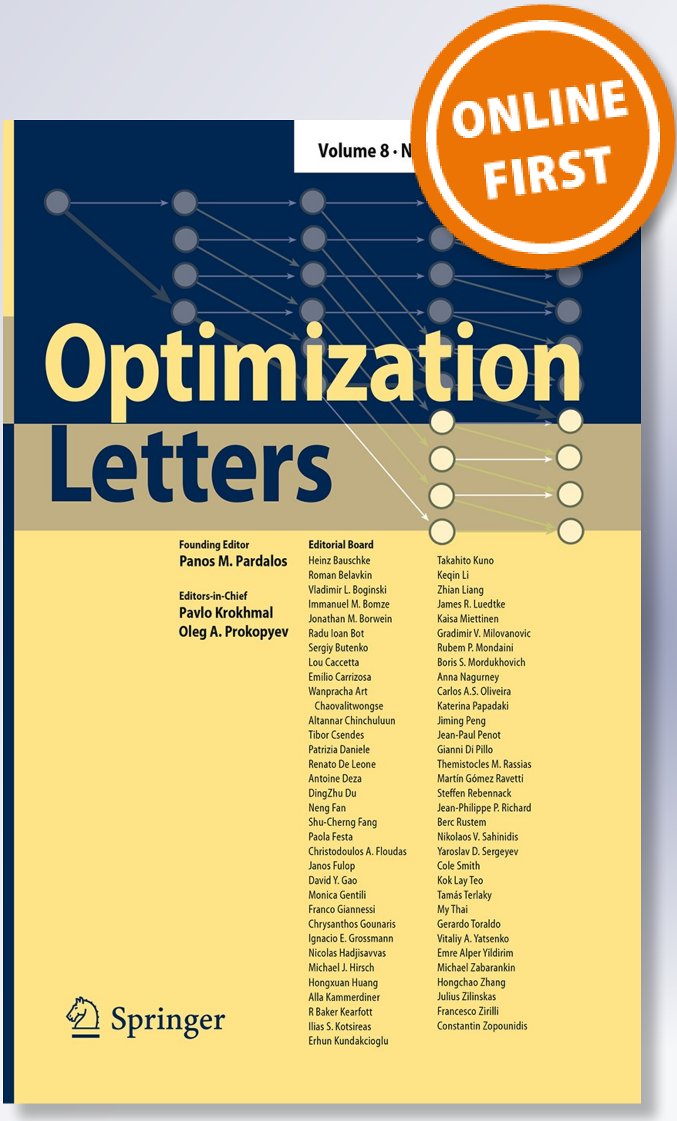
Book review "Cell formation in industrial engineering: theory, algorithms and experiments" by Boris Goldengorin, Dmitry Krushinsky, Panos M. Pardalos

Boris Mirkin

Optimization Letters

ISSN 1862-4472

Optim Lett
DOI 10.1007/s11590-014-0842-7



Springer

Optim Lett
DOI 10.1007/s11590-014-0842-7

BOOK REVIEW

Book review "Cell formation in industrial engineering: theory, algorithms and experiments" by Boris Goldengorin, Dmitry Krushinsky, Panos M. Pardalos

Springer, 2013

Boris Mirkin

Received: 28 October 2014 / Accepted: 5 December 2014

When I saw the title of the book, I thought of that as another sortie at bioinformatics by the operations research community. Indeed by googling "Cell formation" one gets within a lot of stuff on the biological cell processes. What a group of combinatorial optimization experts can do to advance a murky business of life developments? Yet by looking at the intro, I see the book is on "industrial cell formation", which is a quite different business. It appears the cell formation problem (CFP) here belongs in a growing field of research related to the issues of partitioning a set of technology units in clusters that concentrate as many operations as possible within themselves so that those remaining in-between are minimized.

I cannot help but bring in my memories related to this "discovery". That was exactly the problem posed before me by Arkady Gelman who unexpectedly entered my office in the Summer 1967 at which I, a fresh PhD in Computational Mathematics, was trying to penetrate the mathematics of competitive equilibrium and fixed point theorems. This was the starting point of my lifelong addiction to cluster analysis research. Back then I formalized the Gelman's problem as of finding a partition with the maximum within cluster summary similarities or, equivalently, the minimum between-cluster summary similarities. Jointly with my only collaborator of that time, late Volodia Kupershtokh (1945–2001), we came to the conclusion that the problem as is, at non-negative similarities, has trivial solutions only. One way to go was to restrict ourselves with pre-specified cluster cardinalities at which we came up with what later was called Kernighan-Lin algorithm, and estimates that hold "almost always" ([1], Our first results: [2]). Another way to move on was found when we noticed that the problem

B. Mirkin (✉)
Birkbeck University of London, London, UK
e-mail: bmirkin@hse.ru

B. Mirkin
Higher School of Economics Moscow, Moscow, Russian Federation

Published online: 30 December 2014

relates to the min-cut concept emerged in the Ford and Fulkerson theory for maximal flows in networks. Unfortunately, our efforts in the min-k-cut problem led to nowhere except for a beautiful theorem on multipolar flows by Kupershtokh (1971), which was the first in a series of "multiflow" results by Russian mathematicians and of which probably only A.V. Karzanov remembers by now ([3], Citing [4]). Then we went a step deeper in two directions: (a) real-world applications to organization structure design involving an original model and criterion of structured partition for network aggregation (see [5], pp. 268–277) leading to a systematic methodology [6]; (b) data clustering involving either similarity data or quantitative features or nominal features or mixed-scale features–or consensus clusters—using the data recovery approximation criteria, see [7]. Roughly speaking, this all is an extensive development; almost all the criteria involved lead to non-polynomial problems, thus, to using heuristics, however effective, for computation.

This explains my great interest to this book. The book contains 8 chapters including a detailed summary/conclusion part. Among many mathematical formulations for the cell formation problem (CFP) in group technology the authors choose the three main formulations: (a) the p-Median Problem (PMP), a version of optimal location problem (Chapters 2 and 3), (b) multi-min-cut, along with appropriate constraints (Chapter 4), and (c) what they refer to as bi-criterion cell formation (Chapter 7). Unlike themes (a) and (b), theme (c) makes a proper use of the twofold nature of the cell formation problem in which entities of two types are considered, machines and parts, so that rows and columns of the data matrix can be dealt with independently. It should be pointed out that this is probably the very first monograph at which the presence of the two types is taken rather seriously so that two chapters (Chapter 5 and Chapter 6) contain a less formal discussion of possible ways for developing more realistic models and criteria for the cell formation problem.

Chapter 1 provides the readers with an introduction to the CFP and its variations as well as a historical overview of the CFP's numerous motivations. Chapters 2 and 3 focus on the p-median problem (PMP) and its application to the CFP. The authors coined a term "optimal modeling" of the PMP within the class of Mixed Integer Linear Programming (MILP) models based on the following criteria: (i) the objective function of PMP has the minimum number of non-zero entries; (ii) the number of linear constraints is minimal. Since these numbers are equal to each other, the optimality is preserved for the dual problem to the authors' relaxation of the primal PMP formulation. This elegant result makes the problem much easier to solve. Applications to other PMP models, including the Simple Plant Location Problem, are discussed in Sect. 3.2.3. The optimal modeling makes PMP based models flexible towards introduction of additional constraints such as availability of workforce, production capacity, workload balancing, etc.

Two other themes, (b) the Minimum Multi-Cut Problem (chapter 4) and (c) two models for Bi-Criterion (chapters 5 and 7) contain formulations addressing the practical needs to solve many versions of CFP. These needs are illustrated by mentioning a real world industrial case based on the data from a small company that manufactures high-precision tools (p. 112). This case study is likely among the largest published industrial datasets to date. The authors provide a critical analysis with respect to reconfiguration of cells, impact of identical machines, balancing the cells by means of their loading, and working hours, demonstrating the efficiency of cell formation in an industrial environment. I think that this case study, if referred to properly, might be used as a template for solving versions of the CFP in various factory settings. Potentially this can be used for reducing production costs and increasing profits by using an optimal design for production cells within their manufacturing systems.

Chapter 6 discusses more realistic models and algorithms for the CFP. The authors show its flexibility to include many decision-making problems including such conventional combinatorial optimization models as linear ordering (triangulation), specific structures in networks (such as weighted trees), special scheduling, etc. They formulate a very broad concept of pattern–anything to be found by optimizing a combinatorial criterion over a pre-specified class of admissible structures. I think, to be a working tool in further research, they need to somewhat narrow down the concept of "pattern". On the other hand, it is useful to show practitioners and engineers the unity among the many problems emerging in areas such as organization design, data analysis, computer science, and bioinformatics.

Overall, I would say that the material of the book is relevant in many other engineering problems such as VLSI or other applications related to the need in partitioning a large system into smaller blocks to organize an effective activity around these systems. This of course includes such subjects in industrial engineering as Operations Management, Logistics, Supply Chain Management, etc. This is especially important with respect to current tendencies in manufacturing–an increasing fraction of high low-volume orders, short delivery times, increased complexity and precision, etc. I cannot help but mention that the models and methods presented have much similarity to problems of clustering and biclustering of data of various phenomena. Yet there is an important difference as well, first of all, in the level of "depth" in solving a combinatorial problem. Main goals of data analysis are related to enhancing knowledge of the phenomena and processes to which the data relate to. Currently, the relation between the combinatorial criteria and knowledge is much unclear, which in my view is behind the currently widely accepted low standards in solving the problems: why one should try to approximate an exact solution if the validity of the criterion is not clear and may be subject to debate? Say, our recent experiments give enough ground to doubt validity of even such a well-appreciated criterion as the squared error in k-means clustering [8]. Yet in the engineering applications the criteria have clear operational meaning, so that their minimization should be pursued rigorously. In this aspect, the book under review is a very good source (including a set of solved examples, pp.179–194) continuing the long term passion of Panos Pardalos and his collaborators in exact solution of difficult problems (see [9] as an intermediate milestone). It also adds a very valuable contribution to the body of knowledge of the cell formation problem that can be used in both teaching and research including in bioinformatics.

References

1. Kernighan, B.W., Lin, S.: An efficient heuristic procedure for partitioning graphs. Bell Syst Tech J **49**(2), 291–307 (1970)
2. Kupershtokh, V.L., Mirkin, B.G.: On an automatic classification problem. In: Bagrinovsky C.,(ed.) Mathematical methods in economics, Institute of Economics, Novosibirsk, pp. 39–49 (1968) (in Russian)

3. Karzanov, A.V.: Èkonomnye metody resheniya dvykh izvestnykh zadach o mul'tipotokakh v neorientirovannykh setyakh. In: Kombinatornye Methody v Potokovykh Zadachakh, Inst. for System Studies (VNIISI) Press, Moscow, Issue 3, pp. 96–103 (1979) in Russian. (Title transl.: Fast methods for solving two known problems on multiflows in undirected networks, In: Combinatorial Methods for Flow Problems)
4. Kuperschtokh, V.L.: An extension of Ford and Fulkerson theorem to multipolar networks. Cybernetics. 3 (1971)
5. Mirkin, B.: Mathematical classification and clustering (nonconvex optimization and its applications). Kluwer, p. 448 (1996)
6. Aganbegian, A. (ed.): Analysis and design of organization structure in manufacturing enterprise: methodology and method. Nauka, p. 182 (1983) (in Russian)
7. Mirkin, B.G.: Clustering: a data recovery approach. CRC Press (2012)
8. Mirkin, B.G.; Shestakov, A.: Least squares consensus clustering: criteria, methods, experiments. In: Advances in information retrieval. vol. 7814, pp. 764–768. Springer, Berlin-Heidelberg, (2013)
9. Pardalos, P.M. (ed.): Approximation and complexity in numerical optimization: continuous and discrete problems, vol. 42, Springer (2000)