

Applied optimization and data mining

W. Art Chaovalitwongse¹ · Chun-An Chou² ·
Zhe Liang³ · Shouyi Wang⁴

Published online: 17 January 2017
© Springer Science+Business Media New York 2017

This special volume is dedicated to Professor Panos Pardalos in honor of his 60th birthday (in 2014) and of his fundamental contributions to the research areas of applied optimization and data mining. Pardalos's distinguished pioneer work also established an important interdisciplinary field of applied optimization and data mining, which has been attracting broad attention from both academia and industry.

In this era of big data, it has become increasingly important to develop large-scale optimization and data mining models to solve challenging data-driven problems in engineering and sciences. Successful, practical applications have been increasingly visible in various domains such as transportation, logistics, healthcare, energy, etc.

To honor Professor Pardalos, we invited a selected number of his distinguished colleagues as well as leading researchers in applied optimization and data mining for their contributions. The twenty papers in this special volume can be categorized into several major areas including applied mathematical optimization in data mining, transportation, scheduling, energy, healthcare, and finance. There are five papers mainly focusing on modeling and algorithms of different optimization problems, such as multi-objective optimization, network optimization, and risk averse optimization; four papers are linked to transportation problems; three papers are devoted to scheduling problems; four papers investigate supervising learning (or classification) problems particularly; three papers are related to energy; one paper is moti-

Dedicated to Dr. Panos Pardalos on the occasion of his 60th birthday.

✉ W. Art Chaovalitwongse
artchao@uark.edu

¹ Institute for Advanced Data Analytics, Department of Industrial Engineering, University of Arkansas, Fayetteville, AR, USA

² Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA, USA

³ Department of Management Science and Engineering, Tongji University, Shanghai, China

⁴ Department of Industrial, Manufacturing and Systems Engineering, University of Texas, Arlington, TX, USA

vated by optimal investment decision-making strategy problems; and one paper studies a healthcare problem.

Multi-objective optimization has attracted increasing attention in recent years, and most researches are interested in transforming a multi-objective optimization problem into a single objective optimization problem. Zarepisheh and Pardalos propose a new transformation technique using a new definition of proper efficiency that could improve the computational efficiency. Motivated by various real-life applications, network optimization is always a focused research area in operations research society. Balasubramaniam and Butenko study two clique relaxation models, k -blocks and k -robust 2-clubs, used to describe structurally cohesive clusters with good robustness and reachability properties. Stozhkov et al. investigate a network interdiction problem, a bi-level extension of the classical linear assignment problem, in which the objective is to minimize the total cost of the interdiction actions and the cost of the assignments made by the follower. Uncertainty is involved in most realistic decision-making problems. Pajouh et al. study network problems with uncertainty that are associated with the existence of an edge. The problem is modeled as a CVaR-constrained single-stage stochastic program. The new Benders decomposition algorithm shows outperformance on a test bed of randomly generated instances, and real-life biological and social networks, when compared to an existing decomposition approach. Another paper, presented by Vinel and Krokmal, proposes a new representation to introduce a class of coherent or convex measures of risk that can directly incorporate rational risk preferences by the corresponding utility function, through the concept of certainty equivalence. Two practical cases are demonstrated—flood insurance claims and financial portfolio optimization.

In the twenty-first century, transportation systems have become more complex than ever. The interdependencies among various stockholders and infrastructures and the gigantic size of the transportation system have created great challenges in analyzing transportation problems. In response, we include four related papers in this volume. Wang and Zhang address various transportation network design strategies among multiple regions by using multiple-level games to describe the behaviors of different regions. Stefanello et al. study the problem of minimizing traffic congestion by assigning tolls to streets and roads. They use piecewise linear functions to approximate the congestion cost, and a random-key genetic algorithm to solve the problem. Huang et al. study the traveling salesman and the shortest path problem with uncertain topologies modeled by arc failures. Mixed integer programming models are proposed and three cutting plane algorithms are employed to speed up the solution procedure. Burdakov et al. consider the problem of scheduling replacement of unmanned aerial vehicles, and the optimal replacement strategies are justified.

For scheduling decision-making problems, three interesting papers are included. Pei et al. study a serial-batching machine with multiple job types and sequence-dependent set-up times. They propose optimization algorithms to minimize the makespan and the number of tardy jobs. Pan and Liang take a dual viewpoint of the time-indexed integer liner programming formulation for the min-sum scheduling problem. They develop a procedure to compute optimal dual solutions using the solution information from the Dantzig-Wolfe decomposition and column generation methods. Rowe et al. investigate a procurement decision-making problem for a polystyrene packaging manufacturer to determine optimal purchase quantities between two suppliers of virgin material and recycled material to maximize the manufacturer's expected profit. The proposed optimization approach identifies optimal break points of purchase prices to select a sole-supplier or a dual-supplier strategy, and to determine the optimal purchasing quantities from different suppliers.

In data mining, supervising learning (or classification) is a typical optimization problem with general objectives to maximize the classification accuracy and minimize the complexity

of learned models. The challenges include computational complexity and inaccurate prediction due to data volume, variety, veracity, and velocity. Şeref et al. found that classification results were largely biased by outliers in imbalanced data, and thus propose a new variant of support vector machines (SVM), called weighted relaxed support vector machines (WRSVM), integrating a cost-sensitive strategy. They demonstrate more accurate and robust experimental results than SVM and Weighted SVM. Le Thi and Nyuyen study a regularization approach in multi-class SVM (MSVM). They propose two models of using L_0 and L_2-L_0 regularizations solved by nonconvex approaches based on Difference of Convex functions (DC) programming and DC Algorithms (DCA). Gotoh and Uryasev study unified formulations of SVM. In their SVM optimization model, two objectives are mainly taken into account: convex empirical risk and convex regularizer. The two perspectives on robust optimization modeling are (1) that the primal formulation can be viewed as a robust empirical risk minimization; and (2) that the dual formulation is compatible with the distributionally robust modeling. Chou et al. formulate and solve the classification problem as a combinatorial optimization model of categorical data. They propose a column generation framework to decompose the classification problem, where the classification accuracy can be optimized in the master problem while decision rules can be generated in a sub-problem.

Scheduling and delivery of renewable energy must be planned efficiently and smartly to avoid waste. Roni et al. study a supply chain for biofuels that considers costs and environmental and social impacts of delivering biofuels. In their work, they formulate the supply chain system as a multi-objective optimization problem that can be solved by an augmented constraint method. The solution could help in the design of policies that encourage and support renewable energy production. Chen et al. investigate a Contingency-Constrained Unit Commitment problem (CCUC) in electric power operation systems. They propose a mixed-integer programming model for CCUC that takes account of both transmission and generation element failures, and solves the optimization problem by a cutting plane algorithm.

Problems regarding cooperative investment among multiple agents have both practical and theoretical interest. It is required to take into account risk preferences and sharing gain/losses of all participating agents while planning an optimal portfolio in investment strategies. Grechuk and Zabaranin investigate this as a theoretical problem, characterizing a set of all Pareto-optimal allocations, selecting a fair Pareto-optimal allocation with the suggested equilibrium-based method, and demonstrate cooperative investment with agents using drawdown constraints in multi-period investment strategies.

Finally, Chaovalitwongse et al. develop a decision support system in healthcare using a mathematical programming technique. The objective is to optimize the staff allocation across many hospitals in Thailand with major considerations of transportation costs and massive heterogeneous data distributed in hospitals. The developed system enables decision makers to allocate medical resources (e.g., staffs) more effectively and efficiently.

We are truly delighted to bring together recent developments of applied optimization and data mining from a group of prominent researchers who had scientific and personal relationships with Professor Pardalos. This dedicated special volume is our best wish and thanks for his tremendous contributions in the field. We are grateful to all the authors and reviewers. Especially, we express our sincere thanks to the Editor-in-Chief, Professor Endre Boros, for accepting this special volume in the *Annals of Operations Research*. We are also thankful to Katie D'Agosta for her administrative efforts and kind assistance during the preparation and publication of the volume.