Bilevel Innovization: Knowledge Discovery in Scheduling Systems using Evolutionary Bilevel Optimization and Visual Analytics

Julian Schulte Niclas Feldkamp Sören Bergmann Volker Nissen
Ilmenau University of Technology, Ilmenau, Germany
{julian.schulte, niclas.feldkamp, soeren.bergmann, volker.nissen}@tu-ilmenau.de

ABSTRACT
Determining a scheduling system's framework conditions (e.g. number of vehicles or employees) results in a hierarchical optimization problem, which can be solved through evolutionary bilevel optimization. In this paper, we propose an approach to gain better understanding of a scheduling system's behavior by applying visual analytics on the whole set of evaluated solutions during the bilevel optimization procedure. The results show that bilevel innovization can be used to support the decision making process in a strategic planning context by providing useful information regarding the scheduling system's behavior.

CCS CONCEPTS
• **Applied Computing** → Multi-criterion optimization and decision-making; • **Human-centered computing** → Visual analytics

KEYWORDS
Innovization, evolutionary bilevel optimization, visual analytics, scheduling, staffing

ACM Reference format:
J. Schulte, N. Feldkamp, S. Bergmann, V. Nissen. 2018. Bilevel Innovization: Knowledge Discovery in Scheduling Systems using Evolutionary Bilevel Optimization and Visual Analytics. In *Proceedings of GECCO'18 Companion, Kyoto, Japan, July 2018*, 2 pages.
DOI: 10.1145/3205651.3205726

1 INTRODUCTION
Scheduling systems are subject to a variety of influencing factors, some of which (e.g. number of vehicles or employees) can be determined by the company itself. Since these framework conditions can have a major impact on the scheduling system's performance, their determination is an important management task. The difficulty of this task increases when conflicting objectives have to be considered, such as costs and performance. Even though evolutionary bilevel optimization can be used to solve this kind of strategic multi-objective problems [5], it remains hard to gain deeper insights into the scheduling system's behavior by only analyzing the obtained set of Pareto optimal solutions. To gain a deeper understanding of a considered optimization problem, Deb and Srinivasan [1] introduced the concept of innovization. Following the idea of innovization, in this paper we propose an approach for knowledge discovery in scheduling systems by applying visual analytics on the whole set of evaluated individuals during the evolutionary algorithm. The proposed concept of bilevel innovization is demonstrated by using a nested NSGA-II to solve a strategic personnel planning problem and subsequently applying visual analytics to support decision making regarding the number of employees and implemented shifts.

GECCO '18 Companion, July 15–19, 2018, Kyoto, Japan

ACM ISBN 978-1-4503-5764-7/18/07.
https://doi.org/10.1145/3205651.3205726

2 BILEVEL INNOVIZATION
The here presented concept of bilevel innovization is based on the ideas of simulation-based innovization [2] and knowledge discovery in manufacturing simulations [3]. In general, the bilevel innovization process (see Fig. 1) can be divided into two parts: data generation and data analysis. For data generation, evolutionary bilevel optimization is used. The part of data analysis is based on the visual analytics process.

Starting point is the actual scheduling system to be analyzed and lower-level optimization model, respectively. The system behavior will be represented by the lower-level objective value. The next step is to determine the framework conditions that want to be investigated (e.g. policies or set of available employees). These will serve as decision variables for the upper-level problem. Subsequently, the resulting upper-level problem has to be modeled and a large number of independent runs of the upper-level algorithm should be conducted in order to obtain as many different solutions as possible for the subsequent analysis.

In the context of bilevel innovization, each data record is composed by the objective values of an evaluated individual at the upper-level problem (output data) and the corresponding decision variables (input data). The first step is to visualize the output data and to identify an area of interest for deeper analysis. Now, data mining methods (e.g. clustering) can be applied on the filtered data set both on the output and the input data. Thereafter, the data mining results should be visualized to uncover interesting patterns and to get a better understanding of the analyzed system's behavior. Each of the previously mentioned steps may lead to knowledge, which in turn could be used to start the data generation process at an arbitrary step.

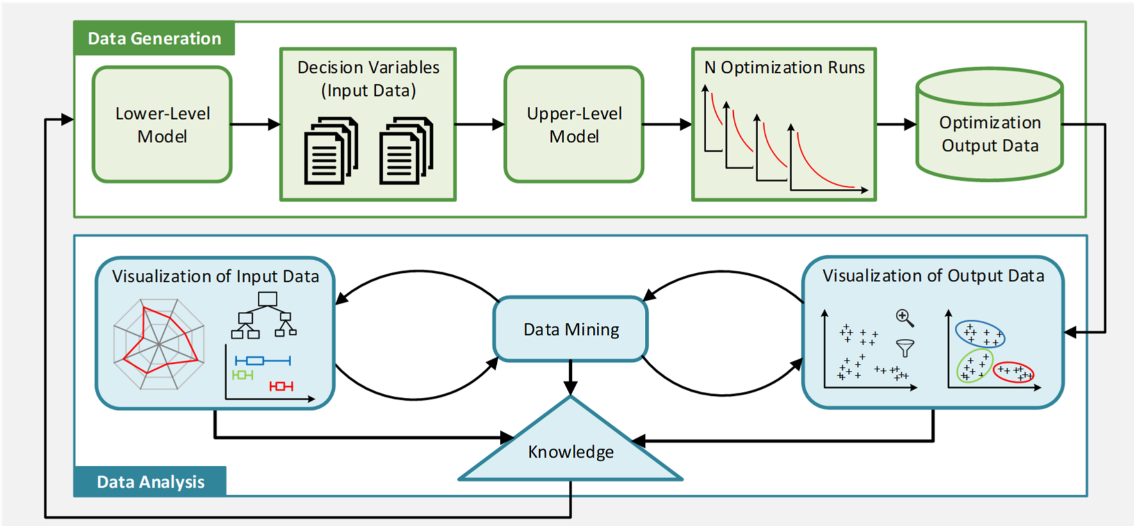

Figure 1: Bilevel innovization process.

3 EXPERIMENTAL STUDY

The strategic workforce planning problem considered in this study is based on [4]. The objective is to determine the number of employees in different categories (6 discrete decision variables) as well as the shift patterns to be implemented (25 binary decision variables). For the upper-level NSGA-II, a population size of 30, a generation number of 100, and n=30 restarts were chosen, yielding 88,762 unique explored solutions and data records, respectively. Subsequently an area of interest with a maximum scheduling penalty of 10,000 was chosen, limiting the remaining solutions to 57,531.

Prior to an examination of the input data, target areas within the selected area of interest were identified by clustering the output data regarding overall penalty and staffing costs. The best structuring was found with five clusters, the k-means clustering algorithm and a cosine-based distance measure (see Fig. 2). In the further process, we focus on the three clusters along the Pareto front: blue (16,563 solutions), green (10,521) and violet (13,378).

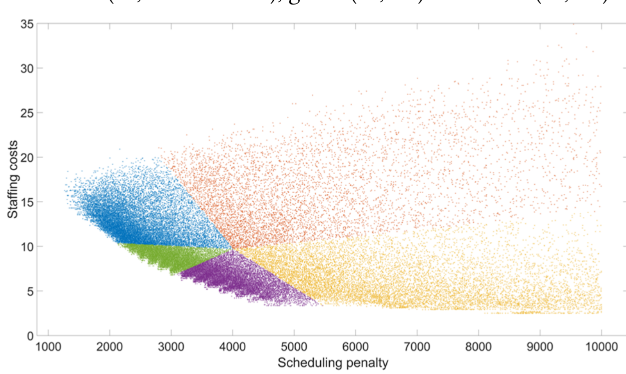

Figure 2: Clustered area of interest.

We now analyze the clusters regarding the six discrete input variables affecting workforce size and structure by using radar charts (see Fig. 3). It becomes apparent that solutions in the blue cluster have significantly more staff involved, especially flexible and part-time workers. The green and violet clusters form a similar shape, differing mainly in the number of employees.

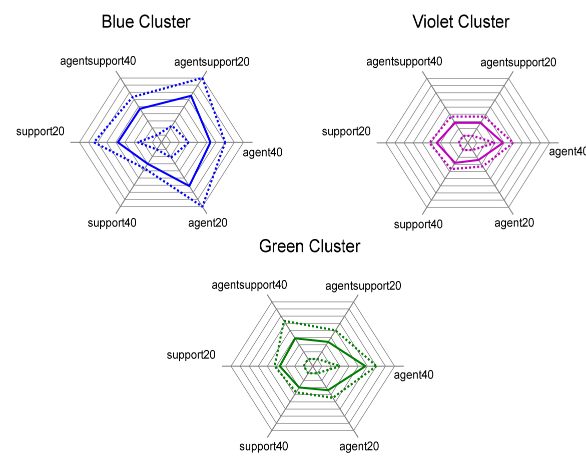

Figure 3: Employee distribution within the target clusters.

REFERENCES

[1] Kalyanmoy Deb and Aravind Srinivasan. 2006. Innovization. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation - GECCO '06*. ACM Press, New York, New York, USA, 1629.

[2] Catarina Dudas, Marcus Frantzén, and Amos H.C. Ng. 2011. A synergy of multi-objective optimization and data mining for the analysis of a flexible flow shop. *Robotics and Computer-Integrated Manufacturing* 27, 4, 687–695.

[3] Niclas Feldkamp, Soeren Bergmann, and Steffen Strassburger. 2015. Knowledge Discovery in Manufacturing Simulations. In *Proceedings of the 3rd ACM Conference on SIGSIM-Principles of Advanced Discrete Simulation - SIGSIM-PADS '15*. ACM Press, New York, New York, USA, 3–12.

[4] Julian Schulte, Maik Günther, and Volker Nissen. 2017. Evolutionary Bilevel Approach for Integrated Long-Term Staffing and Scheduling. In *Proceedings of the 8th Multidisciplinary International Conference on Scheduling: Theory and Applications (MISTA 2017)*, Kuala Lumpur, Malaysia, 144–157.

[5] Ankur Sinha, Pekka Malo, and Kalyanmoy Deb. 2017. A Review on Bilevel Optimization. From Classical to Evolutionary Approaches and Applications. *IEEE Trans. Evol. Computat.*, 1.