https://upright.pub/index.php/ijrstp/

Original Contribution

Data-driven Approach to Enhance Roster of Operations: A Review

Md Saikat Hosen[1], Harshini Priya Adusumalli[2], Mahesh Babu Pasupuleti[3]

Keywords: Data-driven Approach, Data Science, Operations Roster, Management Decision Tool

International Journal of Reciprocal Symmetry and Theoretical Physics

Vol. 7, Issue 1, 2020 [Pages 9-13]

When it comes to today's business environment, especially in the operations-centered industry, the most important goal is to improve the roster of activities and serve those operations in the most efficient manner possible. Powered by data Modern optimization algorithms in combination with models based on production factors have been proved to be useful in the manufacturing business for optimizing production schedules and increasing profitability. We have constructed time and cost models based on real-world data that we have collected. Make use of the information provided to determine the most effective solution to the Job-Shop Scheduling Problem utilizing three algorithms: particle filter, particle swarm optimization, and the genetic algorithm (if applicable). When we want to create operational rosters that are based on a combination of time and cost optimization, the method comes in handy.

INTRODUCTION

Modern business uses data-driven operational segmentation. The data-driven approach has the advantage of not relying on the business users' intuition and expertise. It provides a level of classification that surpasses human thought and conventional wisdom. It recorded more detailed and broader applicant features. Data-driven market segmentation is now widely used in strategic marketing. This method is part of data mining or machine learning. The data-driven approach uses many techniques such as soft computing, clustering, data visualization, and K-Means. They do indeed produce hundreds of variables to describe the emerging market segments. Adusumalli (2018) investigated a scoring system for bank activities. The model assigns a score to each operation based on their deposit and transaction activity. Author (Adusumalli, 2019) used RFM (Recency Frequency, Monetary Value) model to implement K-means clustering. The model identified the most profitable and risky businesses for an auto insurer. Data could also be used to reduce operational attrition. First, they gathered demographic data (gender, age, etc.) and correlated them. Then they used several selection approaches to include the most critical factors. In another study, Fadziso et al. (2018) compare decision trees with logistic regression to see which is better at forecasting operations' behavior. As a result, the decision tree outperforms the logistical model by 70%. Many businesses predict improved results from further research on data-driven operational segmentation (Pasupuleti, 2017). This paper summarizes recent studies on data-driven market segmentation.

The ideal manufacturing schedule depends not only on cost and duration choices, but also on estimation of these factors based on the processing furnace. This work aims to propose an algorithmic approach to resource allocation and scheduling. Notably, the cost vs. time optimization approach for various furnaces is data driven, based on current furnaces and tasks handled so far. The number of furnaces determines the necessity to update the model. The model may need to be updated more frequently if the number of furnaces or parts

[1]College of Management, Capital Normal University, Haidian District, Beijing, CHINA
[2]Software Developer, iMinds Technology systems, Inc., 1145 Bower Hill Rd, Pittsburgh, PA, USA
[3]Data Analyst, Department of IT, Iminds Technology Systems INC, 1145 Bower Hill Rd, Pittsburgh, PA 15243, USA
maheshbp.gs@gmail.com

treated changes frequently. This is also true if the model was made using a different sort of part. This can be daily or monthly. The data driven model can be applied to any facility and once calculated, can only be used for that facility (Pasupuleti & Adusumalli, 2018).

Some methods combine data mining algorithms. To solve the customer segmentation problem, the authors presented a hybrid soft computing methodology based on classical clustering, rule mining, and decision tree analysis. The researcher's model has two modules. After data purification, we may use k-means clustering to group consumers of a firm based on their buying behavior. The Davies-Bouldin index should be used to choose the optimal number of clusters. A subset selection approach is used to identify customer features that are statistically significant to predict the outcome. Following the selection, each client trait would be scored, showing its value in customer segmentation. The model uses IF criteria to divide customers into three categories: High Value, Potential Value, and Low Value. The researchers believe that machine learning could increase the model's consumer segmentation accuracy.

REVIEW OF RELATED LITERATURE

Pasupuleti (2018) goes into great detail about the heat treatment process as well as the procedures that take place prior to the preparation of molten metal and its casting. Using heuristic methods, they were able to optimize the operation of two furnaces. They make the following assumptions about the world:

- Due to technological limitations, it is not possible to handle works from various families in the same batch at this time. These job-families will be referred to as incompatible. Furthermore, these operations will have to be performed without interruption on parallel and non-identical BPs (BPs with varied capacities), which will be available continuously with the goal of maximizing the use of the BPs, which would be available continuously.
- There is a one-week planning period for scheduling.
- There is continuous availability of all batch processors, and all jobs must pass through the operation(s) that will be carried out at the batch processors.

The way forward would be to be able to schedule jobs at more regular intervals that they are using due to the fact that jobs arrive on the shop floor at shorter intervals, they recommended after successfully using data from an existing foundry (Adusumalli & Pasupuleti, 2017).

For a periodic job shop scheduling problem, Rahman and colleagues (2019) proposed a hybrid solution that used particle swarm optimization and simulation-annealing techniques, with the latter being the most efficient of the two. They assert that "for a variety of reasons, including convergence speed, evolutionary algorithms are finding greater application than classical algorithms." Their comparison of algorithms demonstrates the benefit of PSO's memory, which allows all particles to retain the characteristics of the best solutions, regardless of whether or not their populations have changed during the course of the experiment.

Pasupuleti and Amin (2018) recently developed heuristics to deal with problems concerning lot sizing and scheduling, which they published in Nature Communications. This distinction between the operational and tactical aspects of task scheduling is important because it is necessary to fulfill both cost and time objectives at the same time. The two cannot be pursued independently. Our objectives include closing the gap and incorporating preferences for a particular work into the tactical production plan. We also make decisions about facility optimization methods based on the tactical production strategy.

Adusumalli and Pasupuleti (2017) conducted an in-depth review of studies on the application of advanced optimization methods to address scheduling challenges in the manufacturing industry, which resulted in a comprehensive assessment. Both Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) are widely used optimization techniques. Particle Swarm Optimization (PSO) is one of the most widely used optimization techniques. We chose Particle Filter, Particle Swarm Optimization, and Genetic Algorithm for our research because of the following reasons:

It is proposed in this study to use three commonly used approaches to solve the optimization problem, which is a novel approach to problem solution. The Particle Filter, Particle Swarm Optimization, and Genetic Method algorithms, which were previously described, are examples of this type of algorithm. Pasupuleti et al. (2019) proposed a random-key representation of the priority of the jobs, which we combine with a scheduling approach to produce a robust representation that is independent of the optimization algorithm used. Another key advantage of this strategy is that, in contrast to other heuristics, the criteria that influence the optimal decision are very clearly stated. This is one of the most significant advantages of this approach.

MOCKUPS FOR PERIOD AND BUDGETS

It is necessary to clean the data before a full analysis can be performed. We begin by creating the mockup from the data collected by the automation system placed in the factory, which is then tested. Data from an existing facility has been utilized to develop a model that predicts the time of the heat treatment process. To prepare for the current situation, the following steps were taken for the following reasons:

- Unaccounted-for fixture weights have been included to guarantee that fittings which are also heated in the furnace and hence require more time and energy are taken into consideration.
- It has been decided to eliminate the charges related with rework from the bill. In order to maintain data purity, this procedure must be followed because repeating it could result in an unsatisfactory final product.
- Corrections have been made to part weights to remedy errors and missing values, ensuring that the data is accurate and full as possible. There is information on the part weight stored in the plant database.
- Charges that had been processed in an insufficiently thorough manner were removed.
- The gross weight based on fixed values has been changed to account for inflation.
- The length of time it would take to complete each process step was determined.
- Furthermore, only charges that were handled completely in automatic mode were taken into account. Charges with more than one element were removed from the system.

The quality of the data stored is critical in order to maximize the amount of data available for usage in the model. The optimization of the acquired data should be considered while putting our methodology into action. It is planned to include data from a greater number of furnaces over a greater length of time and with a greater number of parts in the next phase of the program. I think it's important to underline that both quality and quantity are important, and that recognizing this would help to increase the usefulness of this work even further. We will give a brief overview of the data in this section. The steps of the progression are depicted in Figure 1.

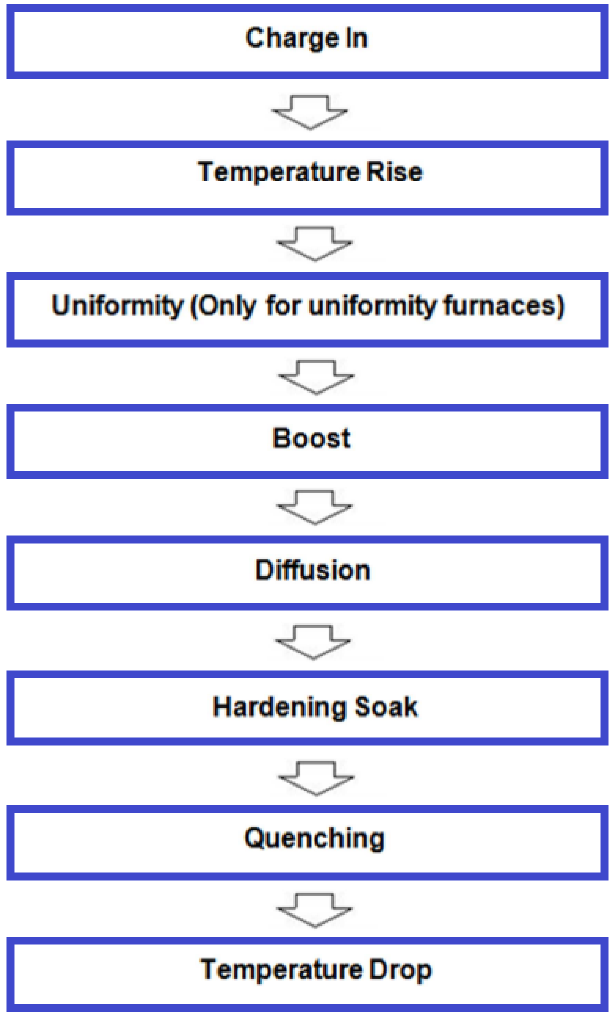

Figure 1: Development Steps

During the course of processing a work, it is not necessary to complete all of the steps outlined in the procedure. Please have a look at the following graphs to get a sense of the complexity of the information: For each of the phases listed above, time graphs of stage completion have been created. We should also underline that we are willing to make the raw data available to anyone who is interested in doing so because it would allow for a better understanding of the Modeling process, and that we are always open to hearing different points of view on the subject. The Models that we produced are still in use and are updated on a regular basis; however, they are currently limited to the furnaces that have been covered previously. An automated method that will allow the Models provided here to be automatically updated based on the data collected is something we are working to put in place. A bigger amount of data collected under a broader range of conditions would aid in the reduction of these uncertainties and the enhancement of the value of this Model in the long run.

OPTIMAL OPERATIONS ROSTER

On the basis of a Cost and Duration Model, we demonstrate how to construct an optimal production schedule in this section. We rely on the models that have been developed here, but they can be substituted by any other model that is suitable.

A single decision-making step

To begin, we will explore a single selection for scheduling a single job, and then we will proceed to the overall optimization of the scheduling problem. To this point, we've spoken about models for duration and cost for a single charge. We can make decisions based on these assumptions, presuming that all or a specific collection of furnaces is available. There is a possibility that the furnaces will be unavailable for a period of time due to the demands of other jobs. In order to take this into consideration, one can include a schedule that takes past decisions into account. As a result, we regard previously planned jobs to be limits on our ability to solve the optimization problem.

Penalty Function with Multiple Parameters

In this phase, we broaden the scope of the optimization problem to include costs and duration at the same time as the constraints, as well as the constraints themselves. We apply the penalty technique to loosen the restrictions and incorporate them into the goal function in the case of the constraints. We develop a new objective function that includes the duration and cost of each job, with the duration and cost weighted according to the preferences for the job. Due to the fact that each task has its own set of preferences, we take into consideration the fact that for certain occupations, duration is vital, while for others, duration is not. In addition to reflecting real-world economic realities, eliminating the premise that preferences are the same for every job has the advantage of being more accurate. The resulting schedule is optimized based on the priority of each job and attempts to satisfy as many restrictions as possible.

Scheduling many jobs at the same time

We have now scheduled a list of jobs in the most efficient manner. It is not possible to assume in our model that each task would have the same parameters on each furnace or that the decision criteria will be the same on each job. The total penalty is the sum of the penalties for each work; hence, by employing the scheduling rule, the problem may be reduced to finding the best sequence in which to schedule the jobs. The following rule is proposed: if we have the jobs in an ordered list, we always choose the single best decision for the penalty function for each task, regardless of the order in which they are listed. For a global optimal, it can be argued that a non-optimal decision for a single job can only be accepted if the advantages of the optimal option for another job are greater than the benefits of the non-optimal decision for the single job. If this is the case, the job that would benefit the most from being prioritized, i.e. scheduled first, should be done first. As a scheduling rule, we propose that if we have the tasks in an ordered list, we always make the single best judgment possible for the penalty function for each work in the order list. Because of the random-key representation of the ordered list, we have a very strong representation of the ordered list that leads to a possible solution for every set of random keys. When it comes to finding the optimal collection of random keys, which corresponds to the optimal schedule, we can apply any optimization technique available to us.

CONCLUSION

The study findings show that practitioners need more than just clustering to determine client segmentation. The models with more layers and processes, such the hybrid model with clustering, rule mining, and decision trees, did better. The framework has multiple uses. First, get an optimal production schedule based on real-time production data. Adding data or parameters to the models for time and expenses could enhance them. However, the Model is generic enough to be employed at different institutions. Efforts are underway to establish and obtain equivalent data from another site. The method and models are generic. They work for any institution. Because the data will be different, the model parameters will be different, and hence the schedules. We will also note that the models must be updated. The frequency depends on the facility's adjustments. This technology benefits insurers, healthcare providers, and legislators alike. Segmentation could provide insights about their patients' and constituents' health-related activities and status. Using the data, users could predict future customer/constituent wants and actions, promoting general social welfare. If the modifications are minor, the outcomes should be similar. However, frequent additions/changes necessitate updating the model. The facility manager makes the call. The technique is coded in an Excel sheet, so updating the model is as simple as adding/changing data and making a request. The data can be immediately linked to the automation system, as we have done, making model updates easier. We examined several optimization methods and discovered that they may be used to tackle real-world optimization problems. Other uses for our framework are left for future research. Our framework has only a few applications displayed. It could be used to answer various questions about resource allocation in industry. More detailed models for process duration and costs, containing extensive data of furnaces, parts or production process could be developed. The system might also be readily tested in a real-world setting with more tasks and preferences.

REFERENCES

Adusumalli, H. P. (2018). Digitization in Agriculture: A Timely Challenge for Ecological Perspectives. *Asia Pacific Journal of Energy and Environment*, *5*(2), 97-102. https://doi.org/10.18034/apjee.v5i2.619

Adusumalli, H. P. (2019). Expansion of Machine Learning Employment in Engineering Learning: A Review of Selected Literature. *International Journal of Reciprocal Symmetry and Physical Sciences*, *6*, 15–19. Retrieved from https://upright.pub/index.php/ijrsps/article/view/65

Adusumalli, H. P., & Pasupuleti, M. B. (2017). Applications and Practices of Big Data for Development. *Asian Business Review*, *7*(3), 111-116. https://doi.org/10.18034/abr.v7i3.597

Fadziso, T., Adusumalli, H. P., & Pasupuleti, M. B. (2018). Cloud of Things and Interworking IoT Platform: Strategy and Execution Overviews. *Asian Journal of Applied Science and Engineering*, *7*, 85–92. Retrieved from https://upright.pub/index.php/ajase/article/view/63

Pasupuleti, M. B. (2017). AMI Data for Decision Makers and the Use of Data Analytics Approach. *Asia Pacific Journal of Energy and Environment*, *4*(2), 65-70. https://doi.org/10.18034/apjee.v4i2.623

Pasupuleti, M. B. (2018). The Application of Machine Learning Techniques in Software Project Management- An Examination. *ABC Journal of Advanced Research*, *7*(2), 113-122. https://doi.org/10.18034/abcjar.v7i2.626

Pasupuleti, M. B., & Adusumalli, H. P. (2018). Digital Transformation of the High-Technology Manufacturing: An Overview of Main Blockades. *American Journal of Trade and Policy*, *5*(3), 139-142. https://doi.org/10.18034/ajtp.v5i3.599

Pasupuleti, M. B., & Amin, R. (2018). Word Embedding with ConvNet-Bi Directional LSTM Techniques: A Review of Related Literature. *International Journal of Reciprocal Symmetry and Physical Sciences*, *5*, 9–13. Retrieved from https://upright.pub/index.php/ijrsps/article/view/64

Pasupuleti, M. B., Miah, M. S., & Adusumalli, H. P. (2019). IoT for Future Technology Augmentation: A Radical Approach. *Engineering International*, *7*(2), 105-116. https://doi.org/10.18034/ei.v7i2.601

Rahman, M. M., Pasupuleti, M. B., & Adusumalli, H. P. (2019). Advanced Metering Infrastructure Data: Overviews for the Big Data Framework. *ABC Research Alert*, *7*(3), 159-168. https://doi.org/10.18034/abcra.v7i3.602

--0--