JOURNAL OF APPLIED INTELLIGENT SYSTEMS & INFORMATION SCIENCES

VOL. 5, ISSUE 1, PP. 17-22, FEBRUARY 2024.
Available at: **https://journal.research.fanap.com/**
DOI: **https://doi.org/10.22034/JAISIS.2024.381355.1062**

ENHANCING PROJECT MANAGEMENT THROUGH INTEGRATION OF PERT, MONTE CARLO SIMULATION, NSGA-III, AND DBSCAN

Shahram Hoorshid [1,*], Elaheh Ebrahimi [1]

[1] *Faculty of Management and Accounting, Karaj Branch, Islamic Azad University, Karaj, Iran.*

ABSTRACT

This research introduces an advanced project management framework by seamlessly integrating the Program Evaluation Review Technique (PERT), Monte Carlo simulation, Non-dominated Sorting Genetic Algorithm III (NSGA-III), and the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. The combination of PERT's deterministic scheduling with Monte Carlo's probabilistic modeling establishes an applied foundation for estimating project timelines, particularly in uncertain environments. NSGA-III plays a crucial role in addressing complex trade-offs through multi-objective optimization, while the application of DBSCAN enhances data processing efficiency. Monte Carlo simulation captures the dynamic nature of project completion times, and the collaborative efforts of NSGA-III and DBSCAN result in the generation of a Pareto front, providing diverse solutions for decision-making. This research significantly advances the process of decision-making in project management, offering realistic scheduling and optimized outcomes, thereby embodying a paradigm shift towards adaptive, data-driven methodologies that align with contemporary demands for flexibility and efficiency in project management practices.

Keywords: NSGA III, Program Evaluation Review Technique (Pert), artificial intelligence, Monte Carlo

1. INTRODUCTION

The significance of project scheduling within the realm of project management is of paramount importance. Acknowledged as a temporary endeavor aimed at producing a distinctive product, service, or result, projects necessitate meticulous planning and macro-management. Traditional project management approaches have struggled to keep pace with technological advancements, resulting in a demand for agile and flexible methodologies to adapt to the rapidly evolving landscape. Decision-making, a critical aspect of project management, now relies on algorithms that align with management patterns, particularly in the context of construction and installation projects with uncertain completion times.

The primary challenge addressed in this paper is the inherent uncertainty in project management, particularly in the context of construction and installation projects. Traditional project management approaches, often based on critical path methods, struggle to effectively manage projects marked by time uncertainty. The integration of advanced methodologies, including the Monte Carlo simulation, NSGA-III, and DBSCAN, aims to tackle this

* *Corresponding Author, Email: shahramhourshid@gmail.com*

RECEIVED: 07 FEBRUARY 2023; ACCEPTED: 01 JANUARY 2024; PUBLISHED ONLINE: 28 FEBRUARY 2024

challenge by providing a more dynamic, adaptive, and data-driven framework for project scheduling, optimization, and clustering.

The specific problem addressed includes the limitations of common critical path methods in handling time uncertainty, the need for specialized scheduling techniques, and the demand for agile and flexible methodologies to adapt to the evolving project management landscape. By seamlessly integrating PERT, Monte Carlo simulation, NSGA-III, and DBSCAN, the paper proposes an innovative solution to offer realistic scheduling solutions and optimize project outcomes in the face of uncertainties. The integration of these methods is justified by the aim to address the complexities and trade-offs inherent in project management, especially in projects with uncertain completion times.

The Non-dominated Sorting Genetic Algorithm III (NSGA-III) contributes multi-objective optimization, addressing nuanced trade-offs crucial for decision-making in project management. The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, embedded in NSGA-III, ensures adaptive clustering of data derived from simulations, enhancing accuracy and efficiency. In the meantime, Monte Carlo simulation allows for capturing the dynamic nature of project completion times and providing a more realistic representation of uncertainties. This integration is driven by the aim to holistically address the limitations of individual methods and enhance their collective effectiveness.

To do so, in Section 2 the previous related studies are reviewed. Section 3 discusses the methodology whereas Section 4 highlights the results. Section 5 also concludes the paper and devises optional streams of research for prospective scientific works.

2. Literature Review

This literature review critically examines relevant studies that contribute to the foundational concepts and methodologies explored in the present paper. The integration of the Non-dominated Sorting Genetic Algorithm III (NSGA-III), Monte Carlo simulation, and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is at the core of this research. Various studies highlight the unique capabilities of each method and their collective application in addressing challenges within project management and uncertainty modeling.

In the realm of project scheduling and uncertainty management, Yu et al. (2022) underscore the significance of the Monte Carlo method for obtaining precise results amidst parameter uncertainty, particularly in project phases marked by heightened uncertainty. The work of Loizou and French (2012) explores the integration of Monte Carlo simulation in project timelines, emphasizing its role in capturing the dynamic nature of uncertainties. Their findings support the idea that Monte Carlo simulation serves as a pivotal tool for probabilistic modeling, enabling a more realistic representation of project completion times. In a similar vein, the study conducted by Avlijaš (2019) investigates the practical implementation of Monte Carlo simulation in project scheduling. The researchers highlight its importance in handling uncertainties associated with activity durations, resource availability, and external factors. By utilizing Monte Carlo simulation, they demonstrate how project managers can obtain a distribution of potential project completion times, thereby enhancing decision-making in the face of uncertainty. Additionally, Kwak and Ingall (2007) provide a comprehensive review of Monte Carlo simulation applications in project management, offering its advantages and disadvantages for the industrial as well as academic practitioners.

The integration of NSGA-III into scheduling applications has been also a subject of interest. Peng et al. (2022) show NSGA-III's ability to handle a complex trade-off and provides a significant basis for its potential to solve integrated hierarchical scheduling processes in project management. Liu et al. (2022) utilized NSGA-III to address the flexible earthwork scheduling problem in construction sites, demonstrating superior performance compared to NSGA-II and the strength Pareto evolutionary algorithm. NSGA-III produces a set of Pareto-optimal results for each case study, automating the scheduling process and aiding project managers and dispatchers in the graphical analysis of solutions. There are also typical studies that suggest spanning the usage of NSGA-III with clustering methods (Liu et al. 2019, Cai et al. 2015).

DBSCAN, as a clustering algorithm, has been widely studied for its effectiveness in handling complex data structures and identifying patterns in noisy datasets. In this respect, one of the scopes of the work of Orabi (2023) is related to DBSCAN to classify a business application. Focusing on task scheduling, Mustapha, and

Gupta (2024) use DBSCAN in the cloud environment. To enhance the efficiency of users' tasks, their proposed algorithm triggers DBSCAN when it comes to computing the fitness function.

While individual studies contribute to the understanding of these methodologies, the novelty of this paper lies in the simultaneous and cohesive integration of NSGA-III, Monte Carlo simulation, and DBSCAN in the given context of project scheduling. By combining these methods, this research offers a comprehensive and adaptive framework for project scheduling, optimization, and clustering, addressing challenges posed by uncertainties in project management.

3. Methodology

In this Section, the corresponding clustering method and adopted Metaheuristic algorithm are described.

3.1 Clustering Method - DBSCAN Algorithm

The Davies-Bouldin Index (DBI), as introduced by Das et al. (2007), plays a crucial role in the clustering methodology. To activate the cluster centers, a minimum threshold is determined, taking the example of 0.6 as the limit for cluster centers. This initiates the process of locating data within the clusters. Following the establishment of cluster centers, the algorithm focuses on minimizing the distance between data points and the center of the cluster through multi-objective functions. Simultaneously, the algorithm maximizes the distance between clusters to ensure clear borders. The iterative process culminates in the presentation of results as a Pareto front, a pivotal aspect in solving multi-objective problems.

Cluster analysis, an unsupervised learning method, categorizes data points based on similar characteristics. The DBSCAN algorithm, among various cluster analysis methods, employs different distance measurements, such as K-Means, propensity distribution, mean shift, and Gaussian mixture. The Davies-Bouldin Index (DBI) serves as a benchmark for evaluating clustering algorithms, providing an internal validation scheme based on inherent dataset properties. Combining meta-innovative algorithms with studied indices aids in reducing dimensions and distinguishing data early in the processing stage (Heidari et al., 2019). The advantages of DBSCAN include its ability to find arbitrary clusters without the need for predetermined cluster numbers, making it robust and adaptable to varying data densities. However, it is not entirely deterministic, and choosing a meaningful distance threshold can be challenging in certain situations.

In terms of the implementation, the DBSCAN algorithm necessitates determining two parameters, the radius, and the minimum points in a cluster. The algorithm identifies clusters based on the Epsilon radius, forming clusters when a sufficient number of neighboring points are found. The calculation of the DBSCAN Index involves determining the distance within clusters and between clusters. This process, transferred to the NSGA-III algorithm as a multi-objective function, aims to minimize the distance within clusters while maximizing the distance between clusters (Gueorguieva et al., 2017).

3.2 NSGA-III Algorithm

The NSGA-III algorithm is a powerful tool for multi-objective optimization, seeking non-dominated solutions representing various trade-offs between multiple objectives (Mkaouer et al., 2015). It involves steps such as generating a normal distribution function of random numbers, creating an initial population, defining genes, and chromosomes, and performing intersection and mutation operations. Linear intersection and simple mutation are employed in this study. The algorithm's effectiveness is realized through iterations, leading to optimal conditions, and the number of iterations depends on achieving optimal conditions, as determined through 200 repetitions in this investigation. The NSGA-III algorithm's entry into data mining has significantly reduced costs and processing time, particularly in the era of big data.

Indeed, the NSGA-III follows a systematic process for multi-objective optimization. It begins with the initialization of a population of solutions, each represented by a set of decision variables. The algorithm then performs non-dominated sorting, categorizing solutions into different fronts based on their dominance relationships. Crowding distance is calculated to measure the density of solutions around each candidate. Environmental selection is employed to prioritize higher-ranked solutions with greater crowding distances for the next generation. A mating pool is created from the selected solutions, and genetic operators like crossover and mutation are applied to produce offspring, introducing diversity. Parent and offspring populations are combined, and the process of non-dominated sorting and crowding distance assignment is repeated. The algorithm iteratively refines the population through environmental selection until termination criteria, such as a maximum number of generations or achievement of desired solutions, are met. NSGA-III's objective is to find a set of non-dominated solutions that represent a trade-off between conflicting objectives in the optimization problem, ensuring diversity and a balanced representation of the Pareto front.

4. RESULTS

In the conducted applied study of a high-rise tower construction project, the initial schedule was established, and the project's failure structure was scrutinized through Monte Carlo simulation. The simulation process involved setting activities, determining their relationships, and initializing durations. Random numbers between zero and one were generated and assigned to activities, and optimistic, probable, and pessimistic times were created with variations of +20 and -20 days compared to the probable time. Cumulative distributions for each state were formed, and the final project completion time was derived. The corresponding scheduling is given in Table 1.

Table 1. Typical Scheduling

Task	Start	Duration	Finish	Random No	Optimistic time	Probable time	Pessimistic time	Cumulative % of optimistic time	Cumulative % of probable time	Cumulative % of pessimistic time
Foundation	2023-01-17	136	2023-06-02	0.49543675	78	107	136	0.102	0.14	0.177
Structure	2023-06-12	140	2023-10-30	0.419106356	78	135	140	0.102	0.176	0.183
Ceiling	2023-11-04	105	2024-02-17	0.16791145	67	99	105	0.087	0.129	0.137
Heading	2024-02-17	102	2024-05-29	0.918884413	62	86	102	0.81	0.112	0.133
Mechanical phase 1	2024-05-29	107	2024-09-13	0.025864669	76	98	107	0.099	0.128	0.14
Electrical Phase 1	2024-09-13	107	2024-12-29	0.17252228	62	84	107	0.081	0.11	0.14
Joinery	2025-02-07	136	2025-06-23	0.132218447	70	73	136	0.091	0.095	0.177
Mechanical phase 2	2025-06-23	136	2025-11-06	0.54648772	60	133	136	0.075	0.173	0.177
Electrical phase 2	2025-11-06	131	2026-03-17	0.3466772	73	89	131	0.095	0.116	0.171
Exterior of the building	2026-03-17	121	2026-17-16	0.284464844	62	113	121	0.081	0.147	0.158
Area	2026-07-16	128	2026-11-21	0.3326385	69	117	128	0.09	0.153	0.167
Finishing	2026-11-21	49	2027-01-09	0.220181909	10	47	49	0.013	0.061	0.064
Total duration	**1453**									

The research findings from ten thousand simulations in the Monte Carlo environment provided the project duration metrics. Table 2 provides the simulation results for the project duration, offering key metrics to assess the variability and characteristics of the construction project. The "Average Project Completion" time, calculated at 1342 days, represents the central tendency, or mean completion time across the ten thousand simulations. This value serves as an indicator of the expected or typical duration of the project. The "Standard Deviation" of 63 days measures the degree of variability or dispersion in the completion times obtained from the simulations. A higher standard deviation suggests greater variability in the data points, indicating a wider range of possible completion times. In this case, the standard deviation of 63 days provides insight into the spread of completion times around the average. The "Most Recent Completion Time" is recorded at 1539 days, representing the duration observed in the most recent simulation. This metric reflects the latest result among the ten thousand repetitions, offering a snapshot of the project completion time under current conditions. Finally, the "Minimum Completion Time" is reported as 1057 days. This represents the shortest completion time observed across all simulations. While the average provides an estimate of the typical completion time, the minimum completion time highlights the best-case scenario, indicating the shortest duration the project could potentially achieve based on the simulation parameters. Thus, it provides a comprehensive overview of the project duration simulation results, incorporating measures of central tendency, variability, and extreme values to assist in understanding the distribution and potential outcomes of the construction project.

Table 2. Project duration simulation results

Simulation metric	Value
Average Project Completion	1342 days
Standard Deviation	63 days
Most Recent Completion Time	1539 days
Minimum Completion Time	1057 days

Upon examining the Monte Carlo simulation results, it was evident that the NSGA-III artificial intelligence algorithm needed to be employed. Table 3 shows its parameters setting.

Table 3. Parameter setting of NSGA-III

Algorithm Parameter	Value
Primary Population	10,000
Number of Repetitions	200
Intersection Ratio	0.08
Mutation Ratio	0.05
Mutation Probability	0.02

After executing the NSGA-III algorithm with 200 repetitions, the outputs of Table 4 were obtained. The algorithm generated optimal random numbers (X) that represent the decision variables for the scheduling problem. These numbers contribute to finding solutions on the Pareto front. The objective function values (F) associated with these random numbers constitute the Pareto front, indicating trade-offs between conflicting objectives. In the meantime, the role of DBSCAN was instrumental in the clustering of data points obtained from the NSGA-III algorithm. This clustering process, which was crucial for the automatic categorization of simulation results conducted in ten thousand repetitions, involved classifying the data into three distinct states: optimistic, probable, and pessimistic. DBSCAN played a pivotal role in processing and optimizing the data derived from Monte Carlo simulations. The primary function of DBSCAN was to dynamically categorize the data into clusters, contributing to the adaptive and efficient classification of simulation outcomes. This dynamic clustering process ensured that the dimensions of the categories changed based on the evolving characteristics and borders of the data, allowing for a responsive and effective categorization approach. Integrated into the NSGA-III algorithm, DBSCAN significantly contributed to refining the data clustering model by minimizing the least squares error function. This optimization process enhanced the accuracy of the clustering model, underscoring the adaptability and efficiency of project management methodologies, particularly in the face of uncertainty.

Table 4. NSGA-III Outputs

NSGA-III Output	Value
Number of Repetitions	200
Optimal Random Numbers (X)	[-0.0336, -0.0766, 0.0832, ..., -0.0665, 0.0534]
Objective Function Value (F)	0.3339

5. Conclusion

This paper presented an innovative and integrated framework for project management, addressing the challenges posed by uncertainties in construction projects. The combination of PERT, Monte Carlo simulation, NSGA-III, and DBSCAN provided an applied approach for scheduling, optimization, and dynamic clustering. The contributions of this research lie in advancing decision support in project management, offering realistic scheduling solutions and optimized outcomes.

For future research, several avenues can be explored to enhance and extend the proposed framework. Firstly, the integration of real-time data and adaptive algorithms could further improve the responsiveness of the framework to changing project conditions. Additionally, investigating the applicability of the framework to different types of construction projects and industries could broaden its scope. Further exploration of the parameter settings for NSGA-III and DBSCAN, as well as their sensitivity analysis, could provide valuable insights into optimizing the performance of these algorithms in project management contexts. Overall, the proposed framework establishes a foundation for future research endeavors aimed at advancing adaptive methodologies in project management and optimizing resource allocation in the face of uncertainty.

References

Avlijaš, G. (2019). Examining the value of Monte Carlo simulation for project time management. Management: Journal of Sustainable Business and Management Solutions in Emerging Economies, 24(1), 11-23.

Cai, L., Qu, S., Yuan, Y., & Yao, X. (2015). A clustering-ranking method for many-objective optimization. Applied Soft Computing, 35, 681-694.

Das, S., Abraham, A., & Konar, A. (2007). Automatic clustering using an improved differential evolution algorithm. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 38(1), 218-237.

Gueorguieva, N., Valova, I., & Georgiev, G. (2017). M&MFCM: fuzzy c-means clustering with Mahalanobis and Minkowski distance metrics. Procedia computer science, 114, 224-233.

Heidari, S., Alborzi, M., Radfar, R., Afsharkazemi, M. A., & Rajabzadeh Ghatari, A. (2019). Big data clustering with varied density based on MapReduce. Journal of Big Data, 6, 1-16.

Kwak, Y. H., & Ingall, L. (2007). Exploring Monte Carlo simulation applications for project management. Risk management, 9, 44-57.

Liu, Q., Liu, X., Wu, J., & Li, Y. (2019). An improved NSGA-III algorithm using genetic K-means clustering algorithm. Ieee Access, 7, 185239-185249.

Liu, Y., You, K., Jiang, Y., Wu, Z., Liu, Z., Peng, G., & Zhou, C. (2022). Multi-objective optimal scheduling of automated construction equipment using a non-dominated sorting genetic algorithm (NSGA-III). Automation in Construction, 143, 104587.

Loizou, P., & French, N. (2012). Risk and uncertainty in development: A critical evaluation of using the Monte Carlo simulation method as a decision tool in real estate development projects. Journal of Property Investment & Finance, 30(2), 198-210.

Mkaouer, W., Kessentini, M., Shaout, A., Koligheu, P., Bechikh, S., Deb, K., & Ouni, A. (2015). Many-objective software modularization using NSGA-III. ACM Transactions on Software Engineering and Methodology (TOSEM), 24(3), 1-45.

Mustapha, S. D. S., & Gupta, P. (2024). DBSCAN inspired task scheduling algorithm for cloud infrastructure. Internet of Things and Cyber-Physical Systems, 4, 32-39.

Orabi Alkhen, W. (2023). Exploring Advanced Clustering Techniques for Business Descriptions: A Comparative Study and Analysis of DBSCAN, K-Means, and Hierarchical Clustering.

Peng, W., Lin, J., Zhang, J., & Chen, L. (2022). A bi-objective hierarchical program scheduling problem and its solution based on NSGA-III. Annals of Operations Research, 308(1-2), 389-414.

Yu, H., Lin, M., Peng, W., & He, C. (2022). Seasonal changes of heavy metals and health risk assessment based on Monte Carlo simulation in alternate water sources of the Xinbian River in Suzhou City, Huaibei Plain, China. Ecotoxicology and Environmental Safety, 236, 113445.