Available online at
https://mijuitm.com.my

e-ISSN: 2735-0703

Mathematical Sciences and Informatics Journal 7(1) 2026, 47 – 56.

Mathematical Sciences and Informatics Journal

Strategies Optimization for Rugby Team Using Genetic Algorithms (GA)

Ahmad Harris Hafizin Harde[1], Shaiful Bakhtiar Rodzman[2]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Terengganu, Kampus Kuala Terengganu, Terengganu, Malaysia*
[2]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Pahang, Kampus Raub, Pahang, Malaysia*

ARTICLE INFO

Article history:
Received 8 November 2025
Revised 29 January 2026
Accepted 20 February 2026
Online first
Published 30 April 2026

Keywords:
Genetic Algorithm (GA)
Optimization
Rugby Strategy
Key Performance Indicators (KPIs)
Decision-Making

DOI:
10.24191/mij.v7i1.11571

ABSTRACT

Rugby union is a complex sport where reliance on coaching intuition often leads to predictable gameplay and inconsistent results. To address these inefficiencies, this study develops a strategy optimization prototype utilizing a Genetic Algorithm (GA) to facilitate data-driven tactical decision-making. The system methodology involves pre-processing where irrelevant attributes have been removed, and player datasets are processed to simulate evolutionary phases, including selection, crossover, and mutation. The proposed model categorizes gameplay into Basic, Tactical, and Contingency plays, mapping specific attributes such as player weight and match statistics to optimized roles. Experimental results indicate that the system successfully identifies high-fitness player combinations for specific strategies, such as "Pick and Go," with performance validation achieved through convergence graph analysis. By automating the selection of optimal strategies, this prototype provides coaches with a systematic tool to enhance team adaptability and competitive advantage in high stakes matches.

1. INTRODUCTION

This article presents the development and evaluation of a strategy optimization system for rugby teams utilizing a Genetic Algorithm (GA). Unlike traditional approaches that rely heavily on coaching intuition, this research proposes a computational model to systematically generate and evaluate tactical decisions based on player attributes and game scenarios.

Rugby is an increasingly popular sport in Malaysia, with over 300 clubs and schools actively involved (Morgan et al., 2020). However, strategy planning in rugby still faces several challenges. Most coaches rely

[1*] Corresponding author. *E-mail address*: hafizinharris@gmail.com
https://doi.org/10.24191/mij.v7i1.11571

on intuition and experience, without systematic data support tools for decision-making (Qiu, 2025). This reliance on conventional tactics makes teams predictable to opponents and limits their ability to adapt to different match situations.

Although Genetic Algorithms (GA) have been successfully used in other sports like basketball to optimize training and game tactics (Sadrehaghighi & Ideen, 2025), or in rugby for simpler physical aspects like optimizing kick angles (Ansari & Saubari, 2020), their application for dynamic and complex team strategy optimization remains unexplored. A gap exists in developing a GA model that uses rugby specific indicators like position, player strength, player weaknesses, and player attributes. Therefore, this literature review aims to combine existing research to establish a foundation for a future optimization system by proposing a conceptual framework and validating it through a preliminary prototype. This paper uses real-world indicators identified from the literature to recommend the most effective strategy combinations. This paper systematically reviews existing applications of Genetic Algorithms in sports analytics, with a focus on rugby strategy optimization, and identifies key research gaps that motivate future system development.

This paper is organized into six sections. Section 2 reviews relevant literature and datasets in sports analytics. Section 3 details the proposed methodology, covering the system architecture, dataset selection, and fitness function formulation. Section 4 presents the preliminary results of the prototype. Section 5 discusses the findings, limitations, and comparisons with existing studies. Section 6 concludes the paper and suggests future research directions. While player selection may appear to be a simple decision-making task, optimizing a full 15-player squad involves evaluating millions of possible combinations to achieve maximum team synergy. This research treats squad selection as a multi-objective optimization problem, where Genetic Algorithms (GA) are uniquely suited to navigate the complex search space of 15-player combinations from a 99-player dataset, a task that exceeds the capabilities of static rule-based or decision-tree models.

2. RELATED WORKS

The application of Artificial Intelligence in sports has transitioned from simple data tracking to complex strategic forecasting, offering a systematic alternative to traditional intuition-based coaching. This section reviews the evolution of optimization techniques in sports, specifically analysing how Genetic Algorithms (GA) have been utilized in domains such as basketball and cricket to solve team composition problems. By synthesizing these related works, the review identifies a critical research gap in Rugby Union, highlighting the need for a unified framework that integrates individual player attributes into a cohesive, mathematically optimized team strategy.

2.1 Optimization in Sports

The application of Genetic Algorithms (GA) in sports has proven highly effective for automating complex strategic decisions that traditionally rely on coaching intuition. In the context of tactical planning, Qiu (2025) demonstrated that GA could refine basketball game strategies by evolving training variables, resulting in superior team performance compared to manual methods. This success extends to the challenge of team composition, where Babu and Rao (2024) utilized GA for fantasy cricket team selection, efficiently balancing budget and role constraints a methodological approach highly relevant to the problem of optimizing rugby squad lineups.

Within the specific domain of rugby, existing research has largely focused on statistical analysis rather than holistic strategy optimization. For instance, Moolman et al. (2025) established the statistical significance of Key Performance Indicators (KPIs) for match outcomes, while Seo et al. (2006) applied optimization to isolated biomechanical skills like kick mechanics. However, a critical gap remains in

https://doi.org/10.24191/mij.v7i1.11571

synthesizing these elements; while individual player attributes and match KPIs are well-documented, no existing system has yet applied GA to holistically optimize dynamic team strategies in rugby. This research aims to fill that void by adapting the proven team-selection and tactical optimization frameworks from basketball and cricket to the unique physical and tactical demands of rugby.

2.2 Key Performance Indicators (KPIs)

Strategy optimization refers to the process of improving strategy decisions to guarantee the most effective outcomes. The effectiveness of a strategy is evaluated using KPIs. According to a study by Hendricks et al. (2020), several performance indicators (PIs) were identified as strong factors for winning and losing teams in both Rugby Union (RU) and Rugby Sevens (RS).

These indicators, summarized in Table 1 and Table 2, show that winning teams score more tries, win more restarts, and (in RU) commit fewer penalties and miss fewer tackles. Another study (Milat et al., 2022) also identified technical determinants of tackle and ruck performance that relate to possession. These identified indicators determine which aspects of gameplay should be prioritized in the fitness function.

Table 1. Key Performance Indicators (KPIs) Related to Winning in Rugby Sevens

Performance Indicator	Significance (p-value)	Effect size (Cohen's d)	General Benchmark for Winning
Restart received	$p < 0.001$	$d = 1.56$	Winning teams receive more restarts
Restart won	$p = 0.001$	$d = 1.11$	Winning teams win more restarts
Line-outs won	$p = 0.013$	$d = 0.83$	Winning teams gain more lineouts
Tries scored	$p < 0.001$	$d = 2.13$	More tries scored

Table 2. Key Performance Indicators (KPIs) Related to Winning in Rugby Union

Performance Indicator	Significance (p-value)	Effect size (Cohen's d)	General Benchmark for Winning
Restart received	$p < 0.001$	$d = 1.18$	Winning teams receive more restarts
Restart won	$p = 0.001$	$d = 1.21$	Winning teams win more restarts
Line-outs won	$p = 0.048$	$d = 0.65$	Winning teams win more lineouts
Tries scored	$p < 0.001$	$d = 1.43$	More tries scored

3. METHODOLOGY

To address the identified research gap, this study proposes the "RugbyGenius" framework, a computational model designed to automate the complex process of squad selection. The methodology is divided into three strategic phases: the acquisition of granular player data, the formulation of a weighted fitness function based on proven Key Performance Indicators (KPIs), and the configuration of the Genetic Algorithm's evolutionary operators. This section details how each component interacts to transform raw statistical data into optimized tactical solutions, ensuring that the selection process is both objective and adaptable to specific coaching strategies.

3.1 Overview of Research Framework

This study proposes a "RugbyGenius" prototype system. The methodology involves the conceptual design of the optimization engine, the formulation of a fitness function, and the simulation of evolutionary phases to generate optimal strategies. The system architecture is illustrated in Fig. 1.

https://doi.org/10.24191/mij.v7i1.11571

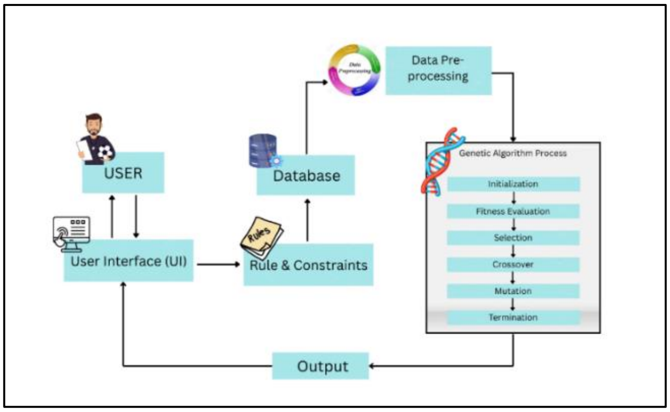

Fig. 1. System Architecture of the proposed optimization model

3.2 Data Collection and Pre-processing

To ensure the replicability of the study, the dataset was sourced from a public repository to simulate real-world constraints.

- Dataset Source: The data were obtained from Kaggle, titled "Statistics on the Best Rugby Players, 2023-2024)".
- Dataset Description: The dataset has 99 player instances and 38 attributes, including Name, Position, Weight, Height, Club_Starter, Club_Tries, Red_Card, Yellow_Card and National_match.
- Pre-processing: Data cleaning involved removing irrelevant columns (e.g., salary, nationality) and handling missing values to ensure the GA receives consistent numerical inputs.

3.3 Genetic Algorithm Implementation

The proposed system utilizes a standard GA flow: initialisation, evaluation, selection, crossover, mutation, and termination.

- Chromosome Encoding: To address the complexity of rugby strategy, this study employs a specific encoding scheme. A single chromosome represents a full team lineup (e.g., 15 players). Each gene within the chromosome represents a specific player assigned to a specific position (e.g., Gene 1 = Prop, Gene 2 = Hooker). Unlike simple binary encoding (0s and 1s), this system uses integer encoding, where each gene is an index pointing to a player in the dataset. This ensures that tactical decisions are linked directly to specific player identities.
- Chromosome Definition: This chromosome is explicitly defined as a vector where each index is a fixed position (e.g., Index 0 = Prop, Index 1 = Hooker).
- Crossover: This study applies the uniform crossover method. This justifies why GA is used: it allows the system to swap "tactical units" between two "parent" teams to see if they perform better together.
- Mutation: This study applies swap mutation. Explain that this introduces diversity by randomly trying a new player in a position to see if it improves the total fitness.

https://doi.org/10.24191/mij.v7i1.11571

3.4 Fitness Function Formulation

The fitness score $\boldsymbol{f(c)}$ is a weighted mathematical model as shown in Equation (1), not a simple rule set:

$$f(c) = \sum_{k=1}^{n} \left(\frac{a_k}{b_k} \times w_k \right) + S(c) \tag{1}$$

where:

- a_k represents the candidate player's actual attribute value.
- b_k denotes the ideal benchmark value required for the specific strategy.
- w_k is the assigned weight, reflecting the attribute's priority.
- n is the total number of attributes considered.
- $S(c)$ represents the consistency score, derived from the player's scoring history (Total Tries / Total Matches × 10).

This formula ensures that players who closely match the physical and technical demands of a strategy (such as weight for a scrum or speed for a winger) contribute higher scores to the overall team fitness. The fitness function serves as the objective evaluator within the evolutionary loop. While the benchmarks (b_k) provide the 'rules' for success, the GA optimization process is what allows the system to efficiently explore the global search space of 15 player combinations from 99 candidates, a task that exceeds the capabilities of static rule-based or decision-tree models.

4. RESULTS AND FINDINGS

The findings of this study highlight the essential components required to develop the proposed optimization system, namely the selection of appropriate performance benchmarks for the fitness function and the application of Genetic Algorithms validated in similar contexts.

4.1 Calculation of Fitness Score

The fitness score is used to validate the logic of the proposed algorithm, through a manual calculation conducted on a sample dataset for the strategy "Pick and Go". This strategy prioritizes physical dominance, therefore, attributes such as weight, height, and experience, which are determined as club starts inside the dataset, are heavily weighted. Table 3 demonstrates the calculation for a single player ("Prop") against the ideal benchmarks (Weight > 120kg, Height > 1.95m).

Table 3. Example of Fitness Calculation for Player P1 (Strategy: Pick and Go)

Attribute	Player Data	Ideal Value	Weight	Calculation
Position	prop	Yes	30	30
Weight	110 kg	120 kg	25	(110/120)25=22.9
Height	192 m	1.95 m	25	(192/1.95)25=24.6
Club starter	14	15	10	(14/15)10=9.3
Consistency = tries per match	4/8 = 0.5	× 10	10	0.5 × 10=5

Based on the calculation in Table 3, player P1 achieved a high fitness score of 91.8. This result confirms that the player's attributes are highly suitable for the "Pick and Go" tactical play. It validates that the underlying mathematical logic of the fitness function is sound. The genetic algorithm then leverages this scoring mechanism to select and recombine chromosomes (strategies) containing similar high-fitness players.

https://doi.org/10.24191/mij.v7i1.11571

Following this individual validation, the system scales the process to optimize a complete squad. In the proposed framework, a single chromosome represents a full team lineup (e.g., 15 players). The system calculates the fitness score for every player in the chromosome using the validated formula and sums the scores to derive the Total Team Fitness. To identify the global optimum, the Genetic Algorithm executes an iterative evolutionary cycle: selection retains the teams with the highest cumulative scores, crossover exchanges tactical combinations between parent teams, and mutation introduces random player swaps to maintain diversity. This cycle repeats across multiple generations, ensuring that the final output is not merely a collection of good individuals, but a cohesive unit optimized for the specific strategy.

4.2 Prototype Implementation

“RugbyGenius” is the name of the prototype that successfully integrates the Genetic Algorithm module with a user-friendly dashboard. As shown in Fig. 2, the interface allows users (coaches or analysts) to select specific strategy combinations such as attacking play alone or in combination with phase play or counterattack and display the results of optimization. It also has three types of rugby formats which are 7’s, 10’s, or 15’s that the user can choose from.

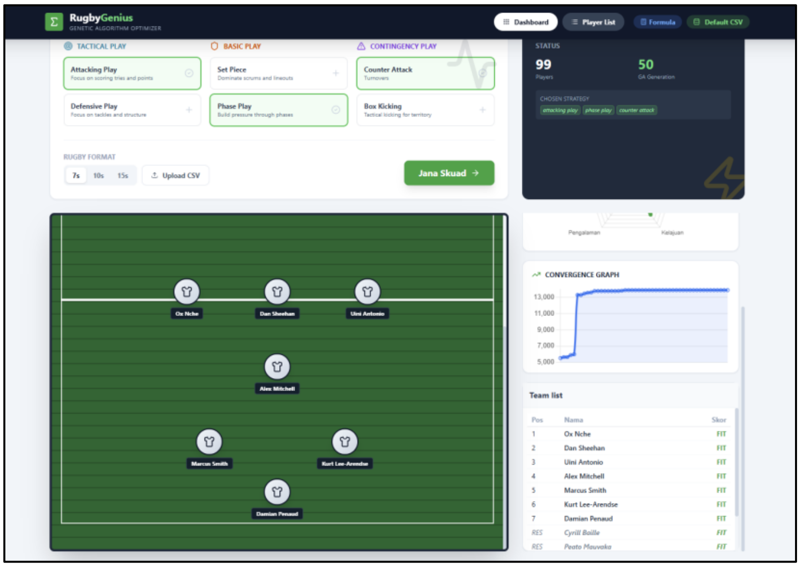

Fig. 2. System Interface Design

4.3 Convergence Analysis

The efficiency of the GA was tested over 50 generations. As shown in the system output (Fig. 3), the system demonstrates rapid convergence.

https://doi.org/10.24191/mij.v7i1.11571

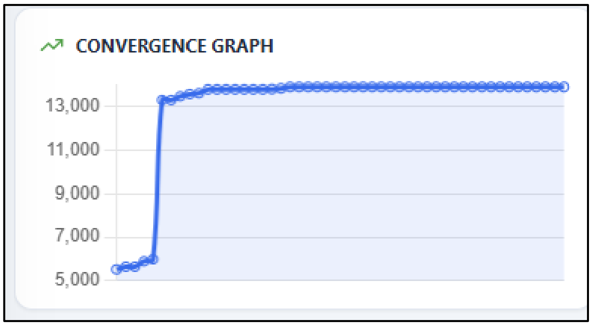

Fig. 3. Convergence Analysis

To validate the efficiency of the Genetic Algorithm, the system tracks the fitness value of the best strategy over 50 generations. As shown in the convergence graph in Fig. 3, the algorithm demonstrates a rapid optimization capability. The graph shows a steep incline in fitness scores during the initial phase (Generations 1–10) before stabilizing around 13,800. This "convergence" confirms that the system has successfully found an optimal or near-optimal solution set, indicating high reliability.

5. DISCUSSION

The development of the "RugbyGenius" prototype provides a preliminary validation of the theoretical framework, confirming the feasibility of using evolutionary algorithms for rugby strategy optimization. This section interprets the experimental findings, specifically analysing the system's ability to converge on stable solutions and the logical validity of the generated squad line-ups against tactical benchmarks. Furthermore, the discussion addresses the current scope of the study, acknowledging the limitations of using historical data and outlining the necessary steps to transition the model from a conceptual prototype to a field-ready application supported by expert qualitative feedback.

5.1 Interpretation of Findings

The development of the "RugbyGenius" prototype serves as a proof of concept to validate the theoretical framework proposed in this review. The primary finding is the demonstrated feasibility of automating complex tactical decisions using Genetic Algorithms. The system successfully translated abstract coaching concepts (e.g., "attacking play") into quantifiable fitness scores, as evidenced by the high selection priority given to players with appropriate attributes (e.g., high speed for Wingers) during the manual validation process. Alternative techniques such as Decision Trees or A* algorithms were considered; however, these are typically limited to sequential decision-making or pathfinding. Because rugby strategy requires the simultaneous optimization of 15 interdependent roles, the GA's ability to evolve entire populations of teams makes it the superior choice for finding a global optimum in a high-dimensional data environment. Furthermore, the convergence behaviour observed in the prototype indicates that the algorithm can efficiently navigate the search space of potential team combinations to find a stable solution, confirming that the proposed fitness function formulation is mathematically sound and computationally viable for future full-scale implementation.

5.2 Comparison with Previous Studies

Consistent with Qiu (2025) in basketball, this study confirms that evolutionary algorithms can handle the multi-objective constraints of team sports. However, unlike Qiu's study, which focused on training loads, this research has successfully applied GA to tactical role assignment, bridging the gap identified in the literature review.

https://doi.org/10.24191/mij.v7i1.11571

5.3 Limitations

Despite the promising results, this model has specific limitations:

1. Static Data: The current system relies on historical CSV. It does not account for real-time player fatigue or injuries during a match.
2. Lack of Opponent Modelling: The current fitness function optimizes the team in isolation. It does not currently account for the specific weaknesses of an opponent (e.g., optimizing for speed if the opponent has a heavy, slow pack).
3. Lack of Field Validation: The current evaluation is based on theoretical fitness scores. The prototype has not been tested in real-world scenarios, meaning the impact of external variables (such as weather or player psychology) on the match result cannot yet be measured.
4. Absence of Expert Verification: The prototype relies solely on quantitative data. It lacks qualitative validation from experienced coaches, which is necessary to confirm that the generated strategies are practically viable and not just mathematically optimal.

6. CONCLUSION

This paper has presented a comprehensive review of strategy optimization in rugby, identifying a critical gap in current coaching methods, which rely heavily on intuition rather than systematic data analysis. By synthesizing existing literature on Genetic Algorithms, this study proposed a novel conceptual framework that categorizes gameplay into Basic, Tactical, and Contingency plays, establishing a theoretical basis for automating squad selection. The review validates that mapping player attributes to specific tactical roles via a weighted fitness function is a scientifically sound approach to enhance decision-making efficiency without relying on subjective bias. Consequently, this framework serves as a foundational model to transition Malaysian rugby from manual to computational strategy planning. Future work will focus on the full implementation of this design by integrating real-time match data and opponent modelling to develop a dynamic, field-ready tactical assistant. To address current evaluation limits, future research will also involve expert reviews by professional coaches. This will be followed by testing the system in real match scenarios to clearly evaluate its impact on actual game performance.

7. ACKNOWLEDGEMENTS/FUNDING

The authors would like to acknowledge the support of the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Terengganu and Pahang Branches, for providing the opportunity and support to complete this study.

8. CONFLICT OF INTEREST STATEMENT

The authors declare there is no conflict of interest in the subject matter or materials discussed in this manuscript.

9. AUTHORS' CONTRIBUTIONS

Ahmad Harris Hafizin Harde carried out the research, designed and developed the Genetic Algorithm system prototype, performed data analysis, and wrote the manuscript. Shaiful Bakhtiar Rodzman supervised the research progress, reviewed the conceptual framework, and approved the article submission.

https://doi.org/10.24191/mij.v7i1.11571

REFERENCES

Ansari, R., & Saubari, N. (2020). Application of genetic algorithm concept on course scheduling. *IOP Conference Series: Materials Science and Engineering, 821*(1). https://doi.org/10.1088/1757-899X/821/1/012043

Ardana, N. B., Hastomo, W., & Arman, S. A. (2024). Development of adaptive lecture scheduling system using genetic algorithm: A case study of Ahmad Dahlan Institute of Technology and Business. *Journal of Computer Science Advancements, 2*(4), 200–212. https://doi.org/10.70177/jsca.v2i4.1310

Babu, P. V., & Rao, D. M. C. S. (2024). Evolving fantasy cricket teams: Applying genetic algorithms for optimal player selection. *Design, Construction, Maintenance, 4*, 270–278. https://doi.org/10.37394/232022.2024.4.29

Edison, F. N. E., Mosquera, E. M., Arguero, T. B., & Zambrano, J. A. (2024). Experimental study of convergence and stability of a genetic algorithm using different selection methods. *In 2024 IEEE Eighth Ecuador Technical Chapters Meeting (ETCM)* (pp. 1–6). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/ETCM63562.2024.10746169

Guerrero, C., Lera, I., & Juiz, C. (2017). Genetic algorithm for multi-objective optimization of container allocation in cloud architecture. *Journal of Grid Computing, 16*(1), 113-135. https://doi.org/10.1007/s10723-017-9419-x

Hendricks, S., Van Niekerk, T., Sin, D. W., Lambert, M., Den Hollander, S., Brown, J. C., Maree, W., Treu, P., Till, K., & Jones, B. (2020). Technical determinants of tackle and ruck performance in international rugby sevens. *European Journal of Sport Science, 20*(7), 868–879. https://doi.org/10.1080/17461391.2019.1675764

Kanyilmaz, A., Tichell, P. R. N., & Loiacono, D. (2022). A genetic algorithm tool for conceptual structural design with cost and embodied carbon optimization. *Engineering Applications of Artificial Intelligence*, 112, 104711. https://doi.org/10.1016/j.engappai.2022.104711

Milat, M., Knezić, S., & Sedlar, J. (2022). Application of a genetic algorithm for proactive resilient scheduling in construction projects. *Designs, 6*(1). https://doi.org/10.3390/designs6010016

Moolman, B., Broodryk, R., & Van Den Berg, P. H. (2025). Shared benefits for Rugby Union and Rugby Union Sevens players: A comparative analysis of performance indicators. *International Journal of Performance Analysis in Sport.* https://doi.org/10.1080/24748668.2025.2491961

Morgan, K., Mouchet, A., & Thomas, G. (2020). Coaches' perceptions of decision making in rugby union. *Physical Education and Sport Pedagogy, 25*(4), 394-409. https://doi.org/10.1080/17408989.2020.1725458

Qiu, D. (2025). Using Genetic Algorithm to optimize the training plan and game strategies of basketball players. *Scalable Computing: Practice and Experience, 26*(1), 441-449. https://doi.org/10.12694/scpe.v26i1.3845

Sadrehaghighi, & Idden. (2025). *Optimization problem (CFD Open Series, Patch 2.60).* CFD Open Series. https://doi.org/10.13140/RG.2.2.10973.69605/1

Seo, K., Kobayashi, O., & Murakami, M. (2006). Multi-optimization of three kicks in rugby. *The Engineering of Sport 6*, *1*, 223–228. https://doi.org/10.1007/978-0-387-46050-5_40

https://doi.org/10.24191/mij.v7i1.11571

Suanpang, P., & Jamjuntr, P. (2024). Optimizing service scheduling by genetic algorithm support decision-making in smart tourism destinations. *Decision Making: Applications in Management and Engineering, 7*(1), 624-650. https://doi.org/10.31181/dmame7120241273