TACTIQ: Optimal Squad Selection Using Integer Linear Programming

A Multi-Objective Optimization Approach for Football Team Formation

Operations Research Project Report

Version 1.0
October 8, 2025

Preetam Teja B . Vishal AS . Durai Singh . Sarvesh Kannan

Operations Research, Team 10 Batch B

https://tactiq.onrender.com

Abstract

This report presents a comprehensive operations research solution for optimal football squad selection, combining mathematical optimization, constraint satisfaction, and multi-objective decision-making. The TACTIQ (Tactical Intelligent Quantification) system leverages Mixed Integer Linear Programming (MILP) to select the optimal 11-player squad from a larger roster while satisfying formation requirements, position eligibility, budget constraints, and training time limitations. Two distinct optimization approaches are explored: (1) cost as a hard constraint with rating maximization, and (2) multi-objective optimization balancing rating maximization with cost and training time minimization.

Contents

1 Introduction

1.1 Background

Football team selection is a complex combinatorial optimization problem faced by coaches and managers worldwide. The challenge involves selecting a fixed number of players from a larger pool while considering multiple conflicting objectives: maximizing team quality, minimizing financial expenditure, ensuring tactical formation requirements, and managing training logistics.

Traditional approaches to squad selection rely heavily on subjective judgment and heuristic methods. While expertise and intuition remain valuable, they cannot efficiently explore the vast solution space of possible team combinations. For a roster of 50 players selecting 11, there exist approximately $\binom{50}{11} \approx 37$ billion possible combinations, making exhaustive evaluation computationally infeasible without optimization techniques.

1.2 Motivation

The motivation for this work stems from three key challenges in modern football management:

1. **Resource Constraints**: Teams operate under strict budget limitations and salary caps
2. **Tactical Requirements**: Specific formations demand players in precise positions with defined skill sets
3. **Multi-Objective Trade-offs**: Decisions must balance competing priorities (performance vs. cost vs. player development)

Operations research provides a rigorous framework to address these challenges through mathematical modeling and optimization algorithms.

1.3 Objectives

This project aims to:

- Develop a mathematical model for optimal squad selection using integer linear programming
- Implement position eligibility constraints and formation requirements
- Explore two optimization paradigms: single-objective with constraints vs. multi-objective optimization
- Provide decision support through visualization and interactive user interfaces
- Ensure mathematical optimality guarantees through provably optimal solution methods

2 Problem Statement

2.1 Problem Definition

Given a roster of n players, select exactly 11 players to form a starting squad such that:

- Tactical formation requirements are satisfied (e.g., 4-4-2: four defenders, four midfielders, two forwards, one goalkeeper)
- Each player is assigned to an eligible position based on their capabilities
- Budget constraints are respected (if applicable)
- Training time limitations are met (if applicable)
- The overall team quality is maximized while costs and training requirements are minimized

2.2 Decision Variables

Let $x_i \in \{0, 1\}$ be a binary decision variable for each player $i \in \{1, 2, \ldots, n\}$:

$$x_i = \begin{cases} 1 & \text{if player } i \text{ is selected} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Let $y_{ir} \in \{0, 1\}$ be a binary variable indicating role assignment:

$$y_{ir} = \begin{cases} 1 & \text{if player } i \text{ is assigned to role } r \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $r \in \{\text{Goalkeeper, Defender, Mid, Forward}\}$

2.3 Parameters

For each player i:

- R_i: Overall rating (skill level)
- T_i: Training time required (hours)
- C_i: Cost (transfer fee or salary)
- P_i: Position eligibility set (e.g., {Defender, Mid} for versatile players)

System parameters:

- N_{GK}: Number of goalkeepers required (typically 1)
- N_{DEF}: Number of defenders required
- N_{MID}: Number of midfielders required
- N_{FWD}: Number of forwards required
- $B_{\max}$: Maximum budget (optional)
- $T_{\max}$: Maximum total training time (optional)

2.4 Complexity Analysis

The problem belongs to the class of **NP-hard combinatorial optimization problems**, specifically a variant of the multi-dimensional knapsack problem with additional assignment constraints. The computational complexity arises from:

1. Binary decision variables creating a discrete solution space
2. Position eligibility creating assignment sub-problems
3. Formation constraints coupling decisions across positions
4. All-rounder requirements introducing additional logical constraints

3 Mathematical Formulation

3.1 Approach 1: Cost as Constraint

3.1.1 Objective Function

Maximize team quality while minimizing training burden:

$$\text{maximize} \quad Z = \sum_{i=1}^{n} (\alpha R_i - \beta T_i)\, x_i \tag{3}$$

where:

- $\alpha = 1.0$ (weight for rating maximization)
- $\beta = 0.1$ (weight for training time minimization)

3.1.2 Constraints

Formation Requirements:

$$\sum_{i=1}^{n} y_{i,\text{GK}} = N_{\text{GK}} = 1 \tag{4}$$

$$\sum_{i=1}^{n} y_{i,\text{DEF}} = N_{\text{DEF}} \tag{5}$$

$$\sum_{i=1}^{n} y_{i,\text{MID}} = N_{\text{MID}} \tag{6}$$

$$\sum_{i=1}^{n} y_{i,\text{FWD}} = N_{\text{FWD}} \tag{7}$$

Total Squad Size:

$$\sum_{i=1}^{n} x_i = 11 \tag{8}$$

Position Eligibility:

$$y_{ir} \leq x_i \quad \forall i, r \in P_i \tag{9}$$
$$y_{ir} = 0 \quad \forall i, r \notin P_i \tag{10}$$

One Role Per Player:

$$\sum_{r \in R} y_{ir} = x_i \quad \forall i \tag{11}$$

Budget Constraint (Hard Limit):

$$\sum_{i=1}^{n} C_i x_i \leq B_{\max} \tag{12}$$

Training Time Constraint:

$$\sum_{i=1}^{n} T_i x_i \leq T_{\max} \tag{13}$$

All-Rounder Requirements:
Let A_{MD} be the set of mid-defender all-rounders:

$$\sum_{i \in A_{\text{MD}}} x_i \geq \min(|A_{\text{MD}}|, 2) \tag{14}$$

Similarly for defender-mid all-rounders.

3.2 Approach 2: Multi-Objective Optimization

3.2.1 Objective Function

Simultaneously maximize rating while minimizing cost and training time:

$$\text{maximize} \quad Z = \sum_{i=1}^{n} (\alpha R_i - \beta T_i - \gamma C_i)\, x_i \tag{15}$$

where:

- $\alpha = 1.0$ (weight for rating maximization)
- $\beta = 0.1$ (weight for training time minimization)
- $\gamma = 0.001$ (weight for cost minimization)

3.2.2 Weight Selection Rationale

The weights are chosen to normalize different scales:

- **Rating** ($R_i \in [60, 90]$): Primary objective, weight = 1.0
- **Training Time** ($T_i \in [1, 50]$ hours): Secondary consideration, weight = 0.1

- **Cost** ($C_i \in [10^3, 10^6]$): Tertiary consideration, weight = 0.001

The weight γ must be calibrated based on cost magnitude to ensure comparable influence. For costs in thousands, $\gamma = 0.001$ ensures a $10,000 difference has similar impact to a 10-point rating difference.

3.2.3 Constraints

All constraints from Approach 1 remain applicable, except:

- Budget constraint becomes **optional** (cost is optimized, not constrained)
- Training time constraint becomes **optional** (training time is optimized, not constrained)

When hard constraints are removed, the optimization naturally seeks cost-efficient solutions while maintaining quality.

4 Data and Preprocessing

4.1 Dataset Description

The project utilizes the **European Soccer Database**, a comprehensive dataset containing match, player, and team information across multiple seasons. The dataset includes player attributes, match outcomes, and team statistics stored in SQLite format. The preprocessing pipeline transformed and enriched this raw data into a clean, structured format suitable for optimization.

4.1.1 Dataset Sources and Structure

- **Original Format:** SQLite database
- **Extracted Tables:** `Player`, `Player_Attributes`, `Match`, `Team`
- **Final Format:** Excel and CSV files used for model input
- **Total Features:** 47 attributes per player
- **Key Columns:** `player_id`, `player_name`, `team_id`, `overall_rating`, `potential`, `training_time`, `cost`, `final_position`

The attributes can be broadly categorized into:

1. **Identification** – Player ID, Name, Team, and Team ID
2. **Performance Metrics** – Overall rating, potential, and skill attributes
3. **Financial Metrics** – Cost derived from player rating and training needs
4. **Physical Attributes** – Acceleration, sprint speed, stamina, strength, jumping
5. **Technical Skills** – Dribbling, ball control, passing accuracy, finishing
6. **Mental Attributes** – Positioning, vision, aggression, reactions
7. **Positional Data** – Preferred position, role flexibility, goalkeeper-specific skills

4.2 Data Extraction and Integration Process

The preprocessing pipeline involved extensive data extraction, transformation, and feature engineering from the European Soccer Database. The following steps were executed to create the final optimization-ready dataset:

4.2.1 Step 1: Database Extraction

The initial data source was the European Soccer Database stored in SQLite format. Raw player and match data were extracted from the database using SQL queries and converted to Excel format for subsequent processing. This extraction provided the foundation for player attribute analysis and team roster construction.

4.2.2 Step 2: Player Attribute Integration

Player attributes (technical skills, physical characteristics, mental attributes) were mapped to the player identification table using unique player IDs as the primary key. The `Player` and `Player_Attributes` tables were merged on `player_id` to associate each player's attributes with their identity. This integration created a comprehensive player profile dataset containing player identification information and 47 distinct attributes covering technical, physical, and mental capabilities.

4.2.3 Step 3: Team Roster Assignment

Team affiliations were determined by analyzing the match table's home team player listings:

1. Extracted home team player IDs from match records (fields `home_player_1` through `home_player_11`)
2. Mapped extracted player IDs to the integrated player attributes dataset
3. Assigned team IDs based on match participation patterns
4. Grouped players by their respective team IDs to form preliminary team rosters

This process ensured accurate team-player associations based on actual match participation data.

4.2.4 Step 4: Temporal Deduplication

The dataset contained multiple time-series instances for individual players due to attribute updates across different match dates. To retain the most accurate and current player profile:

1. Identified all duplicate player entries with different attribute timestamps
2. Retained only the **latest attribute instance** for each unique player
3. Removed all earlier temporal records to eliminate redundancy
4. Verified data integrity post-deduplication

This step ensured that each player appeared exactly once in the final dataset with their most recent attribute values.

4.2.5 Step 5: Positional Clustering

Using attribute-based analysis, players were clustered into position categories. Based on player attributes (particularly goalkeeper-specific skills, defensive stats, passing abilities, and attacking metrics), players were classified into:

- **Goalkeeper** – Players with high goalkeeper-specific attributes
- **Defender** – Players with strong defensive and tackling capabilities
- **Midfielder** – Players with balanced passing, vision, and stamina
- **Forward** – Players with high finishing, pace, and attacking attributes

Multi-positional players (all-rounders) were assigned flexible roles based on dominant attribute patterns, creating the final `final_position` column that maps to the position eligibility sets used in optimization.

4.3 Feature Engineering

After obtaining the clean, merged player dataset, several new features were engineered to support optimization modeling.

4.3.1 Training Time Calculation

The training time metric was designed to reflect player development effort and potential growth:

$$\text{Training Time} = (\text{Potential} - \text{Overall Rating}) + 5 \tag{16}$$

This formula ensures:

- A minimum training requirement of 5 hours for all players
- Scaling proportionally with the gap between current ability and maximum potential
- Higher training times for players with growth potential
- Lower training times for players at or near their peak performance

4.3.2 Cost Estimation

The player cost metric combines performance quality and development investment:

$$\text{Cost (\$)} = (\text{Overall Rating} \times 1000) + (\text{Training Time} \times 500) \tag{17}$$

This formulation:

- Assigns higher costs to high-performing players (rating component)
- Accounts for development investment required (training component)
- Produces realistic cost values ranging from $60,000 to $100,000+
- Creates trade-offs between immediate quality and development potential

4.4 Final Dataset Summary

Attribute	Description
`player_id`	Unique identifier for each player
`player_name`	Full name of the player
`team_id`	Identifier of the team the player belongs to
`team_name`	Name of the team
`overall_rating`	Overall performance score of the player (60-90 range)
`potential`	Maximum achievable rating of the player
`training_time`	Derived metric representing training effort required (hours)
`cost`	Player cost in dollars derived from rating and training parameters
`final_position`	Final assigned positional cluster (GK, Defender, Mid, Forward)

Table 1: Summary of the final processed dataset

4.5 Position Mapping Logic

Players are categorized into position eligibility sets that enable tactical flexibility:

Player Classification	Eligible Roles
Forward	{Forward}
Midfielder	{Mid}
Defender	{Defender}
Goalkeeper	{Goalkeeper}
Forward-Mid All-rounder	{Forward, Mid}
Mid-Back All-rounder	{Mid, Defender}
All All-rounder	{Forward, Mid, Defender}

Table 2: Position eligibility mapping for tactical flexibility

4.6 Preprocessing Outcomes

The complete preprocessing pipeline produced:

- Clean, deduplicated, and enriched dataset containing each player's latest attributes
- Structured positional mappings enabling formation-based optimization
- Realistic financial (`cost`) and training (`training_time`) metrics suitable for Integer Linear Programming modeling
- Team rosters with verified formation feasibility (at least one goalkeeper per team)
- Comprehensive player profiles with 47+ attributes for potential future enhancements

5 Solution Methodology

5.1 Optimization Algorithm

The problem is solved using the **Branch and Bound algorithm** implemented via the PuLP library with the CBC (COIN-OR Branch and Cut) solver.

5.1.1 Algorithm Overview

Step 1: Linear Relaxation

Relax integer constraints: $x_i \in [0, 1]$ instead of $x_i \in \{0, 1\}$

Solve the relaxed Linear Program (LP) using Simplex or Interior Point methods to obtain upper bound Z^*_{LP}

Step 2: Branching

If the relaxed solution has fractional values:

- Select a fractional variable $x_k = 0.7$
- Create two sub-problems (branches):
  - Branch A: $x_k = 0$
  - Branch B: $x_k = 1$

Step 3: Bounding

For each branch, compute the LP relaxation upper bound

Maintain the best integer solution found (incumbent) $Z^*_{\text{incumbent}}$

Step 4: Pruning

Prune branch if:

- Upper bound $\leq$ incumbent (no better solution possible)
- Sub-problem is infeasible
- Integer solution found (update incumbent if better)

Step 5: Termination

Algorithm terminates when all branches are explored or pruned

The incumbent solution is provably optimal

5.1.2 Optimality Certification

The Branch and Bound algorithm guarantees optimality through:

$$\text{Gap} = \frac{Z^*_{\text{upper}} - Z^*_{\text{incumbent}}}{Z^*_{\text{incumbent}}} \times 100\% \tag{18}$$

When the Gap is close to 0%, the solution is certified optimal.

5.2 Computational Complexity

- **Worst-case complexity**: $O(2^n)$ (exponential in number of players)
- **Practical performance**: Branch and bound with intelligent pruning typically explores $O(n^3)$ to $O(n^4)$ nodes for problems of this structure
- **Problem size**: For $n = 50$ players, CBC typically solves to optimality in milliseconds to seconds

5.3 Choice of Platform: Justification for Using Python over Excel

A crucial aspect of implementing the TACTIQ optimization system involved selecting the appropriate computational platform. While Microsoft Excel offers a user-friendly interface and a built-in Solver capable of handling linear and integer optimization, it is inherently limited when applied to large-scale, multi-objective, and constraint-heavy problems such as football roster optimization. The following points summarize the comparative rationale for selecting Python with the PuLP library and CBC solver as the chosen framework.

5.3.1 Scalability and Computational Performance

The TACTIQ dataset comprises over 10,500 players across 400 teams, resulting in thousands of binary decision variables and constraints. Excel Solver, though equipped with a basic Branch and Bound algorithm, lacks advanced heuristics, pruning, and parallel computation support. In practice, Excel begins to slow significantly beyond 50–100 binary variables, often taking several hours or failing to converge. In contrast, Python with the CBC solver efficiently handles thousands of variables within seconds to a few minutes due to its optimized memory management and advanced search strategies.

5.3.2 Automation and Experimentation

Excel requires manual setup and execution for each optimization scenario, making it unsuitable for experimentation involving multiple objectives or constraint variations. Python enables complete automation—allowing batch execution across multiple teams, constraint settings, and weight combinations—with results logged, visualized, and compared programmatically. This capability reduces repetitive effort and accelerates analysis by over 90% compared to manual Excel workflows.

5.3.3 Multi-Objective and Analytical Flexibility

Excel Solver supports only a single objective function per run, limiting its ability to model multi-objective trade-offs between performance metrics such as player rating, training time, and cost. Python allows defining composite objectives or running multi-objective sweeps automatically, generating Pareto-efficient solutions without repeated manual intervention. This enables richer analytical insight and better decision-making flexibility.

5.3.4 Integration and Extensibility

Python offers seamless integration with data analytics, visualization, and AI libraries. This allows TACTIQ to:

- Incorporate AI-driven inference and natural language-based interpretation for insights.
- Generate automated dashboards for performance comparison across teams.
- Deploy the optimization pipeline as a web-based system accessible from any device.

5.3.5 Summary

While Excel Solver can serve as a demonstrative tool for small-scale models (e.g., 5–10 teams), it is computationally infeasible for large-scale, production-level problems. Python with PuLP and the CBC solver offers superior scalability, automation, and extensibility, reducing execution time from hours to seconds and enabling AI integration and deployment capabilities that Excel cannot support.

5.4 Solver Configuration

The CBC solver is invoked with:

- **Presolve**: Enabled (reduces problem size)
- **Cut generation**: Enabled (strengthens LP relaxations)
- **Heuristics**: Enabled (finds good feasible solutions quickly)
- **Optimality tolerance**: 0% gap (exact optimality required)

6 Results and Analysis

6.1 Solution Quality Metrics

Optimality Status Categories:

- **Optimal**: Mathematical guarantee that no better solution exists
- **Feasible**: Valid solution found but optimality not proven
- **Infeasible**: No solution satisfies all constraints
- **Unbounded**: Objective can improve indefinitely (rare)

6.2 Approach Comparison

Both optimization approaches were tested on the AS Nancy-Lorraine roster with a 4-4-2 formation configuration. The following sections present detailed results from actual system runs.

6.2.1 Approach 1 Results (Multi-Objective: Cost in Objective Function)

Figure 1: Figure 1: TACTIQ dashboard showing optimal squad with cost minimization in objective function

Configuration:

- **Formation**: 4-4-2 (4 Defenders, 4 Midfielders, 2 Forwards, 1 Goalkeeper)
- **Team**: AS Nancy-Lorraine
- **Max Training Time Constraint**: 100 hours
- **Max Cost Constraint**: $800,000

Optimization Results:

- **Total Rating**: 737 points
- **Total Cost**: $764,500 (4.4% below budget limit)
- **Total Training Time**: 55.0 hours (45% below time limit)
- **Number of Players**: 11
- **Optimization Status**: Optimal
- **Objective Value**: -33.00

Selected Squad Composition:

Player	Position	Role	Rating	Training (hrs)	Cost
Landry N'Guemo	Defender	Defender	72	5	$74,500
Andre Luiz	Defender	Defender	70	5	$72,500
Michael Basser Chretien	Defender	Defender	69	5	$71,500
Bocundji Ca	Defender	Defender	66	5	$68,500
Romain Grange	Midfielder	Mid	66	5	$68,500
Pascal Berenguer	Midfielder	Mid	66	5	$68,500
Benjamin Gavanon	Midfielder	Mid	65	5	$67,500
Moncef Zerka	Midfielder	Mid	63	5	$65,500
Jo-Gook Jung	Forward	Forward	67	5	$69,500
Chris Malonga	Forward	Forward	67	5	$69,500
Damien Gregorini	Goalkeeper	Goalkeeper	66	5	$68,500

Table 3: Approach 1: Selected squad composition

Key Observations:

1. **Cost Efficiency**: The optimizer selected a squad costing $764,500, utilizing 95.6% of the available $800,000 budget while achieving strong overall rating
2. **Balanced Distribution**: Player ratings are relatively balanced (63-72 range), avoiding extremely expensive star players in favor of cost-effective team composition
3. **Training Time Optimization**: Total training time of 55 hours is well within the 100-hour constraint, demonstrating efficient selection
4. **Formation Compliance**: Exact formation requirements satisfied (4-4-2) with appropriate position assignments
5. **Multi-Objective Balance**: The negative objective value (-33.00) reflects the weighted sum where cost minimization (gamma weight) pulls the value down, indicating successful cost-consciousness

6.2.2 Approach 2 Results (Cost as Hard Constraint)

Figure 2: Figure 2: TACTIQ dashboard showing optimal squad with cost as budget constraint

Configuration:

- **Formation**: 4-4-2 (4 Defenders, 4 Midfielders, 2 Forwards, 1 Goalkeeper)
- **Team**: AS Nancy-Lorraine
- **Max Training Time Constraint**: 100 hours
- **Max Cost Constraint**: $800,000

Optimization Results:

- **Total Rating**: 770 points
- **Total Cost**: $798,500 (0.2% below budget limit)
- **Total Training Time**: 57.0 hours (43% below time limit)
- **Number of Players**: 11
- **Optimization Status**: Optimal
- **Objective Value**: 764.30

Selected Squad Composition:

Player	Position	Role	Rating	Training (hrs)	Cost
Benjamin Moukandjo	Forward	Forward	77	6	$80,000
Thomas Mangani	Midfielder	Mid	73	5	$75,500
Marama Vahirua	Forward	Forward	73	5	$75,500
Landry N'Guemo	Defender	Defender	72	5	$74,500
Samba Diakite	Defender	Defender	70	6	$73,000
Andre Luiz	Defender	Defender	70	5	$72,500
Michael Basser Chretien	Defender	Defender	69	5	$71,500
Alexandre Cuvillier	Midfielder	Mid	68	5	$70,500
Romain Grange	Midfielder	Mid	66	5	$68,500
Pascal Berenguer	Midfielder	Mid	66	5	$68,500
Damien Gregorini	Goalkeeper	Goalkeeper	66	5	$68,500

Table 4: Approach 2: Selected squad composition

Key Observations:

1. **Rating Maximization**: Total rating of 770 points represents 4.5% improvement over Approach 1, demonstrating the impact of prioritizing rating over cost minimization
2. **Budget Utilization**: Nearly complete budget utilization ($798,500 of $800,000), maximizing investment in player quality
3. **Star Player Inclusion**: Benjamin Moukandjo (rating 77, cost $80,000) is the highest-rated player selected, exemplifying willingness to invest in top talent
4. **Training Time Impact**: Slightly higher training time (57.0 vs 55.0 hours) due to inclusion of higher-rated players requiring more preparation
5. **Positive Objective Value**: Objective of 764.30 reflects pure rating maximization minus training time penalty (no cost penalty in objective)

6.3 Comparative Analysis

Metric	Approach 1	Approach 2	Difference
Total Rating	737	770	+33 (+4.5%)
Total Cost	$764,500	$798,500	+$34,000 (+4.4%)
Training Time	55.0 hrs	57.0 hrs	+2.0 hrs (+3.6%)
Budget Utilization	95.6%	99.8%	+4.2%
Objective Value	-33.00	764.30	Different scales
Cost per Rating Point	$1,037	$1,036	-$1 (-0.1%)
Highest Player Rating	72	77	+5 points
Rating Range	63-72	66-77	Wider spread

Table 5: Detailed comparison of optimization approaches

6.4 Detailed Performance Analysis

6.4.1 Rating Distribution

Approach 1 (Cost in Objective):

- Average player rating: 67.0
- Standard deviation: 2.7
- Most expensive player: Landry N'Guemo ($74,500, Rating 72)
- Best value player: Moncef Zerka ($65,500, Rating 63)

Approach 2 (Cost as Constraint):

- Average player rating: 70.0
- Standard deviation: 3.4
- Most expensive player: Benjamin Moukandjo ($80,000, Rating 77)
- Best value player: Damien Gregorini ($68,500, Rating 66)

6.4.2 Cost-Quality Trade-off

The marginal cost of gaining 33 additional rating points (Approach 2 vs Approach 1) is $34,000, equivalent to:

- **$1,030 per rating point improvement**
- **4.4% cost increase for 4.5% rating increase**
- **Nearly proportional trade-off** (cost efficiency maintained)

This indicates both approaches operate near the efficiency frontier, with Approach 2 willing to pay proportionally for quality gains.

6.5 Strategic Implications

When to Use Approach 1 (Multi-Objective):

- Flexible budgets where cost savings are valued
- Development-focused teams prioritizing financial sustainability
- Scenarios requiring balanced squad composition without star players
- Long-term team building with emphasis on cost efficiency

When to Use Approach 2 (Cost as Constraint):

- Fixed budget scenarios with "use it or lose it" dynamics
- Win-now situations prioritizing immediate performance
- Competitive environments requiring maximum quality within budget
- Transfer windows with specific financial allocations

6.6 Sensitivity Analysis

Budget Constraint	Estimated Rating	Utilization	Marginal Impact
$700,000	~740-750	~100%	High binding
$750,000	~755-765	~100%	Moderate binding
$800,000	770	99.8%	Near-optimal
$850,000	~770-775	~94%	Diminishing returns
$900,000	~770-778	~89%	Minimal improvement

Table 6: Budget sensitivity analysis

6.7 Constraint Violation Analysis

When budget is set below the minimum cost of 11 players meeting formation requirements, the system correctly identifies infeasibility:

> *"Insufficient budget! Minimum cost for 11 players is $652,500, but your budget is $600,000"*

This proactive error handling prevents wasted computation and guides users toward feasible parameter configurations.

7 System Architecture and User Interface

7.1 Interactive Dashboard

The TACTIQ system features a professional web-based interface with three main components:

1. **Configuration Panel**: User inputs for formation, constraints, and preferences
2. **Visualization Module**: Interactive pitch diagram showing optimal formation
3. **Analytics Dashboard**: Real-time metrics and optimization statistics

7.2 AI-Powered Coaching Assistant with Internet Access

The TACTIQ system integrates an advanced AI coaching assistant powered by Google Gemini 2.0 Flash with comprehensive capabilities:

7.2.1 Real-Time Internet Search Integration

The assistant has **live access to Google Search**, enabling it to:

- Retrieve up-to-date player performance statistics from recent matches
- Access current transfer market valuations and player availability
- Fetch injury reports and player fitness updates

- Search for tactical analyses from professional football databases
- Pull historical match data and head-to-head statistics
- Find expert commentary on formations and tactical strategies

7.2.2 Complete Team Roster Intelligence

The assistant has full access to:

- **All 30+ players** in the team roster (not just the selected 11)
- Detailed player attributes including ratings, positions, costs, and training requirements
- Position versatility and all-rounder capabilities
- Both the optimized starting 11 and the complete bench roster

7.2.3 Tactical and Strategic Analysis

The AI assistant provides expert-level football insights:

- **Player Synergy Analysis**: Evaluates how selected players complement each other
- **Formation Evaluation**: Analyzes strengths and weaknesses of tactical setups
- **Match Strategy Development**: Creates game plans tailored to squad capabilities
- **Training Drill Recommendations**: Suggests exercises to improve team weaknesses
- **Tactical Adjustments**: Proposes in-game modifications
- **Alternative Lineup Suggestions**: Recommends roster changes for different scenarios

7.2.4 Multilingual Support

The assistant supports **over 30 languages**, including:

- European languages: English, Spanish, French, German, Italian, Portuguese, Dutch, Polish, Russian
- Asian languages: Chinese, Japanese, Korean, Hindi, Arabic, Vietnamese, Thai, Indonesian
- Other languages: Turkish, Swedish, Norwegian, Danish, Finnish, Greek, Hebrew

8 Validation and Testing

8.1 Correctness Validation

Mathematical Verification:

1. All selected players satisfy position eligibility constraints
2. Formation requirements exactly met (no over/under selection)
3. Budget and training constraints never violated when specified
4. Objective values computed correctly according to weight formulations

CBC solver provides certificate of optimality through dual variables and gap analysis. All test cases achieved 0.0% optimality gap.

8.2 Robustness Testing

Edge Cases Tested:

- Empty player pool → Correct infeasibility detection
- Insufficient budget → Clear error messaging
- No eligible goalkeeper → Graceful failure with explanation
- All-rounder scarcity → Appropriate constraint relaxation

Stress Testing:

- Rosters up to 200 players solved in under 2 seconds
- Simultaneous constraint tightening handled effectively
- Numerical stability maintained across diverse cost scales

9 Conclusions

9.1 Key Contributions

This work demonstrates the successful application of integer linear programming to football squad optimization, achieving:

1. **Provably Optimal Solutions**: Mathematical guarantee of best possible team selection
2. **Multi-Objective Balance**: Effective trade-offs between quality, cost, and logistics
3. **Practical Implementation**: User-friendly system with sub-second solve times
4. **AI-Enhanced Decision Support**: Internet-connected coaching assistant with multilingual capabilities

9.2 Practical Impact

The TACTIQ system provides quantitative decision support for:

- **Budget Planning**: Understanding cost-quality trade-offs
- **Tactical Flexibility**: Exploring formation alternatives
- **Resource Allocation**: Optimizing training time distribution
- **Transfer Strategy**: Identifying high-value acquisition targets

9.3 Methodological Insights

Approach 1 vs. Approach 2:

For the AS Nancy-Lorraine roster analyzed, **Approach 2** delivers superior results, providing 4.5% higher rating for only 4.4% additional cost. However, **Approach 1** remains valuable when budget flexibility allows trading performance for cost savings.

Optimization vs. Heuristics:

Compared to greedy selection or manual choice, mathematical optimization provides:

- Guaranteed global optimality (not just good solutions)
- Efficient exploration of combinatorial space
- Consistency and reproducibility
- Quantified trade-offs

9.4 Final Remarks

Operations research transforms football squad selection from subjective art to data-driven science while preserving coaching expertise. The integration of mathematical optimization with domain knowledge, real-time internet data access, and AI-powered multilingual assistance creates a powerful decision support framework applicable beyond football to any resource allocation problem with multiple objectives and complex constraints.

References

[1] Nemhauser, G. L., & Wolsey, L. A. (1988). *Integer and Combinatorial Optimization.* Wiley-Interscience.

[2] Williams, H. P. (2013). *Model Building in Mathematical Programming* (5th ed.). Wiley.

[3] Trick, M. A., Yildiz, H., & Yunes, T. (2012). Scheduling Major League Baseball umpires and the traveling umpire problem. *Interfaces*, 42(3), 232-244.

[4] Boon, B. H., & Sierksma, G. (2003). Team formation: Matching quality supply and quality demand. *European Journal of Operational Research*, 148(2), 277-292.

[5] Tavana, M., Azizi, F., Azizi, F., & Behzadian, M. (2013). A fuzzy inference system approach for rule-based supplier selection. *Expert Systems with Applications*, 40(18), 7178-7191.

A Mathematical Notation Summary

Symbol	Description
n	Total number of players in roster
x_i	Binary selection variable for player i
y_{ir}	Binary role assignment variable
R_i	Overall rating of player i
C_i	Cost of player i
T_i	Training time required for player i
α, β, γ	Objective function weights
$N_{\text{GK}}, N_{\text{DEF}}, N_{\text{MID}}, N_{\text{FWD}}$	Formation requirements
$B_{\max}$	Maximum budget constraint
$T_{\max}$	Maximum training time constraint
P_i	Position eligibility set for player i

Table 7: Complete mathematical notation reference

B Constraint Summary

Hard Constraints (Always Enforced)

1. Formation requirements (exact player counts per position)
2. Position eligibility (players only in valid roles)
3. One role per player
4. Exactly 11 players selected

Optional Constraints (User-Configurable)

1. Budget limitation: $\sum C_i x_i \leq B_{\max}$
2. Training time limitation: $\sum T_i x_i \leq T_{\max}$
3. All-rounder requirements (minimum versatile players)

Report compiled: October 8, 2025
TACTIQ Version 1.0
Deployment Link: https://tactiq.onrender.com