

# Genetic Algorithm: Bin Packing Problem

Matthew D. Branson

*Department of Computer Science*

*Missouri State University*

Springfield, MO

branson773@live.missouristate.edu

## Abstract

This paper presents the implementation and analysis of a...

## Index Terms

Genetic Algorithms, Bin Packing Problem, Optimization, Heuristic Search

## I. Q1: ENCODING AND INITIALIZATION

An integer-based representation is used to encode bin assignments. For an instance with  $N$  items, each individual is a vector of length  $N$ , where the value at index  $i$  denotes the bin to which item  $i$  is assigned.

For example, the encoding  $[0, 1, 0, 2, 1, 1, 2, 1, 3, 0]$  represents a solution for 10 items in which items 0, 2, and 9 are assigned to bin 0; items 1, 4, 5, and 7 to bin 1; items 3 and 6 to bin 2; and item 8 to bin 3, resulting in the use of 4 unique bins.

Initial populations are generated by assigning each item independently to a random bin in the range  $[0, N - 1]$ . This produces diverse initial solutions while satisfying the constraint that the maximum number of bins does not exceed the number of items.

## II. Q2: FUNCTION EVALUATION

### A. Objective Function

The objective function  $f$  measures the number of unique bins used in a given individual. For instance, the encoding  $[0, 1, 0, 2, 1, 1, 2, 1, 3, 0]$  uses bins  $\{0, 1, 2, 3\}$ , yielding  $f = 4$ . The

algorithm minimizes  $f$  to reduce the total number of shipping boxes required.

### *B. Constraint Handling*

The constraint violation function  $g$  counts the number of bins whose total weight exceeds the 10 kg capacity. For each bin, the weights of assigned items are summed. A bin contributes 1 to  $g$  if its cumulative weight exceeds the limit. A solution is considered feasible when  $g = 0$ .

### *C. Evaluation Function*

Each individual is evaluated as a tuple  $(f, g)$ , where  $f$  represents the number of bins used and  $g$  the number of constraint violations. This formulation enables the algorithm to balance minimization of bin usage with satisfaction of feasibility constraints.

## III. Q3: GA OPERATIONS

### *A. Tournament Selection*

Tournament selection is used to construct a mating pool of  $N$  individuals. For each selection, two candidates are drawn at random, and the winner is determined by a constraint-aware comparison rule. This process is repeated  $N$  times. Mating pairs are then selected at random from the resulting pool.

### *B. Winner Selection Rules*

Selection is guided by the following criteria:

- If both individuals are feasible ( $g_1 = 0, g_2 = 0$ ), the one with lower  $f$  is preferred.
- If only one is feasible, the feasible individual is selected.
- If both are infeasible, the one with fewer violations ( $g$ ) is selected.

These rules guide selection toward feasible individuals with fewer bins and penalize constraint violations when feasibility is not yet achieved.

### *C. Crossover*

Two-point crossover is applied with a probability of 0.9. Two cut points are chosen to divide each parent chromosome into three segments. Offspring are generated by exchanging the middle segments, preserving the integer-based encoding and promoting genetic diversity.

#### *D. Mutation*

Each gene (bin assignment) is mutated independently with probability  $p_m = 1/N$ , where  $N$  is the number of items. When selected, a gene is reassigned to a random integer in  $[0, N - 1]$ . This maintains valid encodings while introducing variation for local exploration.

### IV. Q4: GA EXECUTION

#### *A. Baseline Configuration*

The genetic algorithm was initially executed using the specified baseline parameters: population size of 20 and 50 generations. This configuration was tested on four problem instances containing 10, 25, 50, and 100 orders, with order weights uniformly distributed between 0 and 2 kg and a bin capacity of 10 kg.

#### *B. Parameter Variation Study*

To assess the impact of population size and generation count on performance, additional configurations were tested. Population sizes of 10 and 40 individuals were compared against the baseline of 20, while generation counts of 25 and 100 were compared against the baseline of 50. Each configuration was evaluated across all four problem sizes using three random seeds (42, 773, 2025).

#### *C. Convergence Analysis*

Plot the best and average fitness (i.e., number of bins used) per generation.

#### *D. Best Solutions*

Report the best solution found and its bin-wise packing configuration.

### V. Q5: ANALYSIS AND COMPARISON

#### *A. Convergence Analysis*

Compare the convergence behavior of different runs (e.g., different random seeds or order sizes).

#### *B. Performance Comparison*

Discuss whether the algorithm tends to find feasible solutions early or late in the process.

## VI. CONCLUSION

This study successfully implemented...