

# Raven Roosting Optimization with MPI: A Parallel Metaheuristic Approach

Hector Carlos Flores Reynoso  
Department of Computer Science  
University of Trento  
Student ID 266140  
hector.floresreynoso@unitn.it

**Abstract**—This paper presents the implementation and parallelization of the Raven Roosting Optimization (RRO) algorithm [1], a nature-inspired metaheuristic optimization technique based on the foraging behavior of ravens. We describe both a serial implementation in C and a parallel version using Message Passing Interface (MPI) for distributed memory systems. The algorithm simulates raven behavior including roosting, foraging, and social hierarchy with leader-follower dynamics. Our parallel implementation distributes the population across multiple processes, achieving significant speedup through domain decomposition and collective communication patterns. We present detailed pseudocode for both implementations, discuss key optimizations including early stopping mechanisms and efficient bound checking, and analyze the communication overhead and scalability characteristics of the MPI-based approach.

**Index Terms**—MPI, Parallel Computing, Metaheuristic Optimization, Raven Roosting

## I. INTRODUCTION

### A. Background

Metaheuristic optimization algorithms have emerged as powerful tools for solving complex, NP-hard optimization problems where traditional exact methods become computationally intractable. These nature-inspired algorithms offer several key advantages: they require no gradient information, can escape local optima through stochastic mechanisms, and provide near-optimal solutions with significantly less computational effort than classical approaches. Their versatility and problem-agnostic nature have led to widespread applications across diverse domains including production scheduling, engineering design, and resource optimization [2]. The Raven Roosting Optimization (RRO) algorithm represents a nature-inspired metaheuristic, mimicking the intelligent foraging behavior and social coordination of ravens.

### B. Algorithm Overview

The RRO algorithm is based on three key behavioral patterns observed in ravens [1]:

- **Roosting Behavior:** Ravens return to a common roosting site where they rest and plan their next foraging expedition.
- **Foraging Behavior:** Ravens explore the search space looking for food sources (optimal solutions) through a series of flight steps with exploratory lookouts.

- **Social Hierarchy:** Ravens follow a leader-follower dynamic where the best-performing individual becomes the leader, and a subset of the population acts as followers who explore in the vicinity of the leader.

In real life, ravens have shown a social behaviour mainly through communal roosting that functions as an information-sharing system about food. At night, juvenile, non-breeding ravens gather to successful gather foragers to scout for short-lived carcasses as described by the Information Center Hypothesis. This social system is flexible: different individuals are knowledgeable at different times, and ravens combine socially acquired information with their own private knowledge when deciding whether to follow others or search independently [3].

### C. Problem Statement

The algorithm aims to minimize an objective function  $f(x)$  where  $x \in \mathbb{R}^n$  represents a position in the search space with  $n$  dimensions (features). Each raven in the population represents a candidate solution, and the algorithm iteratively improves these solutions through exploration and exploitation phases.

The algorithm maintains:

- A population of ravens, each with a current position and discovered food source
- Fitness values for each raven based on the objective function
- A global leader (best solution found)
- A roosting site that is kept throughout each iteration.

### D. Motivation for Parallelization

Given the computationally intensive nature of the algorithm, especially with large populations and high-dimensional search spaces, parallelization using MPI becomes essential for:

- Reducing execution time through distributed computation
- Handling larger problem instances
- Scaling across multiple compute nodes in HPC environments

## II. TERMINOLOGY DISCUSSION

Throughout this paper the following notation is used:

- $t$  represents the current iteration number and  $T$  is the maximum number of iterations;
- $P$  is the total population of Ravens and  $F$  the number of features for each Raven;
- $X(t)$  is the position vector of ravens in iteration  $t$ ;
- $RS(t)$  is the roosting site across all  $T$ ;
- $D(t)$  is direction where the raven is moving on  $t + 1$ ;
- $L(t)$  Represent the leader, with the best location of source food (global);
- $FS(t)$  is the best local food source know to a raven. These will be replace by the current positions  $X(t)$  if they yield better fitness values;
- $R(t)$  is the remaining distance to get the final destination represented as a scalar. Which can be  $L(t)$  or  $FS(t)$ , see bellow.
- $U(a, b)$  denotes a random number drawn uniformly in the range  $[a, b]$  based on PCG;
- $UB$  and  $LB$  are the upper and lower bounds of the search space;
- $r_i$  are random numbers drawn from  $U(0, 1)$ ;
- $R$  is the defined looking radius (in a hypersphere) from a specific position  $X(t)$ . Usually is equal to the upper bound  $UB$ ;

### A. PCG Implementation

PCG (Permuted Congruential Generator) is a family of simple, fast, space-efficient, and statistically good random number generators developed by O'Neill [4]. Unlike traditional linear congruential generators (LCGs), PCG applies a permutation function to the output of an LCG to improve statistical quality while maintaining computational efficiency.

The basic PCG algorithm uses a linear congruential generator as its core:

$$s_{t+1} = a \times s_t + c \bmod m \quad (1)$$

where  $s_t$  is the internal state,  $a$  is the multiplier,  $c$  is the increment, and  $m$  is the modulus. The output is then passed through a permutation function that scrambles the bits to eliminate the statistical weaknesses inherent in LCGs. This permutation typically involves bit shifts, rotations, and XOR operations, which are computationally inexpensive but dramatically improve the randomness quality.

PCG generators offer several advantages: they pass stringent statistical test suites (including TestU01's BigCrush), have small state sizes, are extremely fast, and provide strong statistical properties suitable for scientific simulations and optimization algorithms. In this implementation, PCG is used to generate the uniform random numbers  $U(a, b)$  that drive the stochastic behavior of the Raven Roost Algorithm.

### B. Objective Function

The Griewank function is a multimodal optimization benchmark function widely used to test the performance of optimization algorithms. It is particularly challenging due to

its many local minima that can trap optimization algorithms, while having a single global minimum. This function serves as the objective function in the Raven Roost Algorithm (RRA) implementation, providing a rigorous test case for evaluating the algorithm's exploration and exploitation capabilities [1].

The mathematical formulation of the Griewank function is:

$$f(x) = 1 + \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) \quad (2)$$

where  $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$  represents a point in the  $n$ -dimensional search space.

#### a) Function Characteristics:

- **Domain:** The function is typically evaluated on the hypercube  $x_i \in [-600, 600]$  for all dimensions  $i = 1, \dots, n$  [5].
- **Global Minimum:** The unique global minimum occurs at  $x^* = (0, 0, \dots, 0)$  with  $f(x^*) = 0$ .
- **Multimodality:** The function exhibits many widespread local minima that are regularly distributed across the search space. The cosine component creates numerous local optima that can trap gradient-based methods, while the quadratic term creates a parabolic basin structure. This dual nature makes the Griewank function particularly effective for testing an algorithm's balance between exploration (escaping local minima) and exploitation (converging to the global optimum) [5].

### C. Movement of ravens

The movement of ravens in the algorithm is defined by the position update equation  $p_{i,t} = p_{i,t-1} + d_{i,t}$ , where  $d_{i,t}$  represents a random distance vector for raven  $i$  at iteration  $t$ . The distance vector  $d_{i,t}$  is constructed through two key steps in the implementation. A direction, and a random amplifier:

$$d_{i,t} = s_{i,t} \times D(t) \quad (3)$$

To mimic the real behavior,  $D(t)$  is the direction of each bird to find the best food source. At the start of each  $t$ , each raven decides if he is going to pursue his best food source point or follow the leader and scout the area.

$$\vec{d} = \begin{cases} L(t) - X(t) & U(0, 1) \leq 0.2 \\ FS(t) - X(t) & U(0, 1) > 0.2 \end{cases} \quad (4)$$

First we calculate  $\vec{d}$  into a magnitud using norm:

$$\|v\| = \sqrt{\sum_{i=1}^n (x_i)^2} \quad (5)$$

where  $v$  is any given vector ( $\vec{d}$  in this case) that holds the  $n$  number of parameters (dimensions) in the search space. This is later used the computation of a unit vector using:

$$D(t) = \hat{v} = \frac{v}{\|v\|} \quad (6)$$

This vector represents the exploration direction for the current iteration, influenced by the previous movement patterns.

Second, we calculated a random modifier based on the remaining distance to the target location using **Euclidean**

**distance** to get a scalar value between the final destination  $FS(t)$  or  $L(t)$  and the current position  $X(t)$ :

$$d(a, b) = \sqrt{\sum_{i=1}^n (a - b)^2} \quad (7)$$

$a$  and  $b$  represent any vector of  $n$  number of parameters (dimensions) in the search space. We multiply this distance to random step size  $r_i$  to get fraction of this remaining distance using a uniform distribution value, giving the following formula:

$$s_{i,t} = r_i \times R(T) \quad (8)$$

This creates a Lévy flight-like behavior, where ravens take larger steps when they are far from the target and smaller steps as they approach to the target. The actual displacement vector is therefore  $d_{i,t} = s_t \times D(t)$ , resulting in an adaptive movement strategy that balances exploration through randomness with exploitation through proximity-based step sizing.

#### D. Lookout function

As mentioned earlier, raven's move a fraction of the total remaining distance governed by  $d_{i,t} = s_t \times D(t)$ . This so that each point they stop they can scout the area with a Hyper-sphere vision across  $n$  dimensions. The looking radii ( $r$ ) is governed by:

$$r = \frac{R}{3.6 \times P^{\frac{1}{F}}} \quad (9)$$

When a raven finds a promising food source, he updates it and decides to continue or report it back to the rooster immediately.

$$p_{i,t+1} = \begin{cases} p_{i,t} + d_{i,t} & U(0, 1) \leq 0.1 \\ RS(t) & U(0, 1) > 0.1 \end{cases} \quad (10)$$

### III. SERIAL IMPLEMENTATION

#### A. Algorithm Structure

The serial implementation follows a straightforward execution model where a single process handles the entire population of ravens.

##### a) Main Parameters:

The algorithm accepts the following parameters:

- **Dataset Path:** CSV file containing initial population positions (NxM matrix)
- **Lower/Upper Bounds:** Search space boundaries for each dimension
- **Iterations:** Number of complete optimization cycles
- **Flight Steps:** Number of steps each raven takes toward its destination
- **Lookout Steps:** Number of exploratory searches per flight step
- **Radius:** Looking radius for exploratory searches ( $r_{\{popt\}}$ )

##### b) Core Algorithm Pseudocode:

```

1 Algorithm: Raven Roosting Optimization (Serial)
2 Initialize population with clamping
3 Evaluate best food source with Objective Function
4 Set roosting site and current leader
5 Define followers (excluding leader)
6 for t= 1 to iterations do
7   for n = 1 to pop_size do
8     Set it's current position to the roosting site
9     Determine it's destination based on is he a follower
10    for step = 1 to flight steps do
11      Move to  $p_{i,t+1}$  with clamping
12      for lookout = 1 to lookout steps do
13        Pick a random spot with clamping  $\leq$  radii
14        if finds a better food source then
15          Set it as personal best
16          if p < 0.1 then
17            Stop early and go back to the roosting site
18          end if
19        end if
20      end for
21    end for
22  end for
23  Set current new leader
24  Define followers (excluding leader)
25 end for
26 return current leader

```

Fig. 1. Pseudocode for serial RRO implementation

##### c) Key Implementation Details:

###### Leader Selection:

The leader is determined as the raven with the minimum fitness value, representing the best solution found so far:

$$\text{leader} = \min(FS(t)) \quad (11)$$

###### d) Time Complexity:

For a iteration of size  $I$ , population of size  $P$  with  $F$  features. Each on taking  $F$  flight steps and  $L$  lookout steps:

- **Overall:**  $O(I \cdot P \cdot F \cdot F \cdot L)$
- **Objective function evaluations:**  $O(I \cdot P \cdot F \cdot L)$

Note: The early stopping mechanism ( $p < 0.1$ ) can reduce the number of objective function evaluations by terminating lookout steps prematurely when a significantly better solution is found, potentially reducing evaluations by up to 10% in practice.

## IV. PARALLEL IMPLEMENTATION WITH MPI

### c) *Parallel Algorithm Pseudocode:*

#### A. *Parallelization Strategy*

The parallel implementation in the `parallel/` directory uses MPI to distribute the raven population across multiple processes. The key parallelization approach is **domain decomposition** [6], where each MPI process manages a subset of the total population.

##### a) *MPI Context and Configuration:*

The implementation initializes an MPI context structure that tracks process rank and world size, communicator, and global configuration parameters broadcast to all processes. The context is initialized at program startup to establish the distributed computing environment.

##### b) *Data Distribution:*

Each process handles a local subset of ravens:

- Process 0 reads parameters and broadcasts to all processes
- Population is divided among processes
- Each process maintains local copies of: food sources, positions, fitness values
- Global leader information is shared across all processes

```

1 Algorithm: Raven Roosting Optimization (Parallel MPI)
2 Input: Same as serial + MPI context
3 Output: Best solution found
4 MPI_Init()
5 rank ← MPI_Comm_rank()
6 size ← MPI_Comm_size()
7 if rank == 0 then
8   | Parse command-line arguments
9 end if
10 MPI_Bcast(parameters, root=0)
11 local_rows ← pop_size / size
12 Initialize local population segment
13 Evaluate local fitness values
14 Set roosting_site (shared across all processes)
15 MPI_Barrier()
16 local_best ← argmin(local_fitness)
17 global_best ← MPI_Allreduce(local_best, op=MINLOC)
18 MPI_Bcast(leader_position, root=global_best.rank)
19 num_local_followers ← 0.2 × local_rows
20 local_followers ← random_select(num_local_followers)
21 for iter = 1 to iterations do
22   for i = 1 to local_rows do
23     | current_position[i] ← roosting_site
24     if is_follower[i] then
25       | target ← random_point_near(leader, r_pcpt)
26     else
27       | target ← food_source[i]
28     end if
29     for step = 1 to flight_steps do
30       | direction ← normalize(target - current_position[i])
31       remaining ← distance(current_position[i], target)
32       step_size ← uniform(0, remaining)
33       current_position[i] ← current_position[i] + step_size × direction
34       enforce_bounds(current_position[i])
35       for lookout = 1 to lookout_steps do
36         | candidate ← random_point_near(current_position[i], r_pcpt)
37         enforce_bounds(candidate)
38         if f(candidate) < fitness[i] then
39           | food_source[i] ← candidate
40           | fitness[i] ← f(candidate)
41         end if
42       end for
43     end for
44   end for
45   random_select(num_local_followers)

```

Fig. 2. Pseudocode for parallel MPI-BRO optimization (MPI-AMIX, root=0)

d) *Communication Patterns:*

*Leader Selection:*

Finding the global leader requires collective communication. Each process first finds its local minimum:

$$\text{local\_best}_p = \arg \min_{i \in \text{local\_ravens}_p} f(x_i) \quad (12)$$

Then a global reduction determines the overall minimum across all processes, followed by broadcasting the leader position from the owning process to all others. This requires  $O(\log k)$  communication time for  $k$  processes.

*Parameter Broadcasting:*

Configuration is broadcast from rank 0 to all processes. The global parameter structure is transmitted as a byte buffer, distributing algorithm configuration uniformly across the distributed system in  $O(\log k)$  time.

*Timing Reduction:*

Maximum execution time across all processes is gathered using a reduction operation:

$$T_{\{\max\}} = \max_{p \in \{0, \dots, k-1\}} T_p \quad (13)$$

This determines the wall-clock time as the slowest process duration.

e) *Synchronization Points:*

The parallel implementation uses several synchronization barriers:

- Before starting computation
- After initialization
- Before timing measurements

f) *Load Balancing:*

Population distribution aims for balanced workload. Ravens are divided as evenly as possible with each process handling

$$P_{\{\text{local}\}} = \left\lfloor \frac{P}{k} \right\rfloor \quad (14)$$

ravens (where  $P$  is total population and  $k$  is number of processes). Each process handles the same number of iterations, and computational work scales linearly with local population size.

g) *Scalability Considerations:*

**Strong Scaling:** For fixed problem size, speedup is limited by:

- Communication overhead (leader updates, broadcasts)
- Load imbalance if population doesn't divide evenly
- Synchronization barriers

**Weak Scaling:** Increasing problem size proportionally with processes maintains efficiency better, as communication-to-computation ratio remains favorable.

h) *Random Number Generation:*

Each process uses an independent random number generator seeded with rank-dependent values. The seed for process  $p$  is computed as

$$s_p = s_{\{\text{base}\}} + p \quad (15)$$

where  $s_{\{\text{base}\}}$  is the initial seed. This ensures reproducibility while maintaining statistical independence across processes.

## V. OPTIMIZATIONS

## VI. CONCLUSION

## REFERENCES

- [1] A. Brabazon, W. Cui, and M. O'Neill, "The Raven Roost Algorithm: A social foraging-inspired algorithm for optimisation," *Journal of Metaheuristic Algorithms*, vol. 1, pp. 1–20, 2020.
- [2] X.-S. Yang, *Nature-Inspired Metaheuristic Algorithms*, 2nd ed. Luniver Press, 2010.
- [3] J. Wright and others, "Communal Roosts as Structured Information Centres in Ravens," *Animal Behaviour*, 2003.
- [4] M. E. O'Neill, "PCG: A Family of Simple Fast Space-Efficient Statistically Good Algorithms for Random Number Generation," technical report HMC-CS-2014-0905, Sept. 2014. [Online]. Available: <https://www.pcg-random.org/pdf/toms-oneill-pcg-family-v1.02.pdf>
- [5] M. Jamil and X.-S. Yang, "A Literature Survey of Benchmark Functions For Global Optimization Problems," *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, no. 2, pp. 150–194, 2013.
- [6] E. Alba, "Parallel Metaheuristics: Recent Advances and New Trends," *International Transactions in Operational Research*, vol. 20, no. 1, pp. 1–48, 2013.