

# Optimización por manada de Krill: comparación con otros métodos

Bertinetti, Juan; Gómez, Virginia y Estepa, Andrés

*Universidad Nacional del Litoral, Facultad de Ingeniería y Ciencias Hídricas*

**Resumen**—En este trabajo, un nuevo algoritmo bio-inspirado basado en el comportamiento de las Manadas de Krill para encontrar soluciones a problemas de optimización, propuesto por [1], es analizado. En el mismo, el movimiento de los individuos está basado en la posición del alimento y en la alta concentración de individuos en la manada. La posición dependiente del tiempo de cada krill tiene tres componentes principales: (i) El movimiento inducido por la presencia de otros individuos, (ii) el movimiento debido a la búsqueda de comida y (iii) el movimiento difusivo aleatorio. Además, para realizar un modelado más real del comportamiento del krill, el algoritmo incorpora dos operadores genéticos. El algoritmo mencionado es comparado con los métodos PSO (Particle Swarm Optimization) y GA (Genetic Algorithm).

**Palabras clave**—Krill. Problemas de optimización. Algoritmo bio-inspirado.

## I. INTRODUCCIÓN

El desarrollo de métodos metaheurísticos para resolver problemas complejos de optimización ha tenido un gran desarrollo en los últimos años. En general, estos algoritmos suelen ser más poderosos que los métodos convencionales basados en lógica formal o programación matemática [1].

Los algoritmos de optimización metaheurísticos comparten dos características principales, intensificación y diversificación. La primera, también llamada explotación, intenta usar información de las mejores soluciones actuales, buscando en las vecindades de las mismas y seleccionando los mejores candidatos. La segunda, también llamada exploración, garantiza que se explore el área de búsqueda eficientemente, y así se pueda saltar de un óptimo local y generar nuevas soluciones lo más diversas posibles [2].

La principal fuente de inspiración para el desarrollo de estos algoritmos ha sido por excelencia la naturaleza [3], en especial el comportamiento de ciertas especies animales como son las colonias de hormigas, bandadas de pájaros, entre otros. Básicamente, estos algoritmos se basan en la teoría evolutiva, ya que seleccionan los individuos más aptos para dar lugar a una nueva generación de posibles soluciones.

Los algoritmos bio-inspirados pueden ser divididos en tres categorías [4]:

- (i) Algoritmos evolutivos,
- (ii) Algoritmos basados en inteligencia de enjambres, y
- (iii) Algoritmos basados en forraje bacteriano.

El método de manada de krills o KH (por sus siglas en inglés, Krill Herd), es un híbrido entre los paradigmas (i) y (ii). En las siguientes secciones se detallan las ecuaciones que rigen el movimiento del krill, se presenta el esquema de implementación del algoritmo KH y el análisis comparativo con los métodos PSO y GA.

## II. DESCRIPCIÓN DEL ALGORITMO DE MANADA DE KRILLS

Muchos estudios buscan encontrar los mecanismos que gobiernan el comportamiento de especies animales. Los principales mecanismos estudiados tienen que ver con la capacidad de alimentación, la reproducción, la protección contra depredadores, y las condiciones del entorno [5]. El mecanismo estudiado, en el que se inspira el algoritmo KH, es la formación de enjambres de alta densidad después del ataque de los depredadores. Cuando los depredadores del krill, como las ballenas, pingüinos o aves, atacan a los krills, las manadas se dispersan, reduciendo así su densidad. La formación del enjambre después de un ataque es un proceso multiobjetivo con dos objetivos principales: (1) aumentar la densidad de krills, y (2) llegar a la fuente de alimento. Éste es el proceso principal en el que se basa el algoritmo; un krill individual se mueve hacia la mejor solución en su búsqueda de mayor densidad y alimento.

### A. Modelo del movimiento de la manada de krills

El algoritmo de manada de krills simula el movimiento natural considerando tres factores claves del desplazamiento de los individuos en una superficie 2D:

- (i) Movimiento inducido por otros individuos;
- (ii) Movimiento inducido por la búsqueda de alimento;
- (iii) Difusión aleatoria.

Para la resolución de problemas de dimensión arbitraria, el siguiente modelo Lagrangiano es generalizado para un espacio de búsqueda  $n$  dimensional:

$$\frac{dX_i}{dt} = N_i + F_i + D_i \quad (1)$$

donde  $N_i$  es el movimiento inducido por otros individuos,  $F_i$  es el movimiento inducido por la posición del alimento, y  $D_i$  es el movimiento difusivo aleatorio.

**A.1. Movimiento inducido por otros individuos:** De acuerdo a argumentos teóricos, los krills intentan mantener una alta concentración y se mueven debido a la influencia entre ellos mismos. La dirección del movimiento  $\alpha_i$  se estima a partir de la densidad del enjambre local (efecto local), la densidad del enjambre objetivo (efecto global) y una densidad repulsiva (efecto repulsivo). Para un individuo este movimiento es definido como:

$$N_i^{nuevo} = N^{max} \alpha_i + \omega_n N_i^{viejo} \quad (2)$$

donde

$$\alpha_i = \alpha_i^{local} + \alpha_i^{obj} \quad (3)$$

y  $N^{max}$  es la máxima velocidad inducida,  $\omega_n$  es la inercia del movimiento inducido y toma valores en el rango [0,1],

$N_i^{viejo}$  es el último movimiento inducido,  $\alpha_i^{local}$  es el efecto local debido a los vecinos y  $\alpha_i^{obj}$  es la dirección objetivo debido al mejor individuo en la manada. La máxima velocidad inducida es  $0.01(ms^{-1})$ . El efecto de los vecinos sobre el  $i$ -ésimo krill se modela como:

$$\alpha_i^{local} = \sum_{j=1}^{NV} \hat{K}_{ij} \hat{X}_{ij} \quad (4)$$

$$\hat{X}_{ij} = \frac{X_j - X_i}{\|X_j - X_i\| + \epsilon} \quad (5)$$

$$\hat{K}_{ij} = \frac{K_i - K_j}{K^{peor} - K^{mejor}} \quad (6)$$

donde  $K^{mejor}$  y  $K^{peor}$  son el mejor y el peor fitness respectivamente.  $K_i$  representa el valor de fitness del  $i$ -ésimo krill; de igual manera  $K_j$  representa el fitness del  $j$ -ésimo vecino ( $j = 1, 2, \dots, NV$ );  $X$  es el vector posición; y  $NV$  es el número de vecinos. La ecuación (5) representa el vector dirección (normalizado) entre el  $i$ -ésimo krill y el  $j$ -ésimo krill vecino, éste vector puede tener un efecto atractivo o repulsivo dependiente del signo que tome. Por su parte la ecuación (6) es el valor del fitness normalizado y actúa como peso del vector dirección de cada vecino. La ecuación (4) muestra la influencia de los vecinos sobre cada krill.

Para la elección de los vecinos, se utiliza una distancia de sensado ( $d_s$ ) que determina si un krill es vecino o no. La distancia de sensado para cada iteración se define como:

$$d_{s,i} = \frac{1}{5N} \sum_{j=1}^N \|X_i - X_j\| \quad (7)$$

donde  $d_{s,i}$  es la distancia de sensado del  $i$ -ésimo krill y  $N$  es el número total de individuos. Así, si la distancia entre dos individuos es menor que la distancia de sensado, ellos son vecinos.

El vector objetivo de cada krill está dado por el individuo que posee el menor (mejor) valor de fitness en la manada, provocando que cada krill sea atraído hacia él, lo que lleva al óptimo global. Se modela este comportamiento como:

$$\alpha_i^{obj} = C^{mejor} \hat{K}_{i,mejor} \hat{X}_{i,mejor} \quad (8)$$

donde,  $C^{mejor}$  es un coeficiente que se define como:

$$C^{mejor} = 2(rand + \frac{I}{I_{max}}) \quad (9)$$

y  $rand$  es un número aleatorio entre 0 y 1 utilizado para mejorar la exploración,  $I$  es la iteración actual y  $I_{max}$  es el máximo número de iteraciones. La ecuación (9) provoca que, con el avance de iteraciones, los individuos sean más atraídos por el mejor krill de la manada, lo que hace que en los primeros pasos prevalezca la exploración.

**A.2. Movimiento inducido por el alimento:** Este movimiento es formulado a partir de dos parámetros principales. El primero es la localización del alimento y el segundo es la experiencia previa acerca de la misma. Este movimiento es formulado como:

$$F_i = V_f \beta_i + \omega_f F_i^{viejo} \quad (10)$$

donde

$$\beta_i = \beta_i^{alim} + \beta_i^{mejor} \quad (11)$$

y  $V_f$  es la velocidad de forraje,  $\omega_f$  es la inercia del movimiento y toma valores en el rango  $[0,1]$ ,  $F_i^{viejo}$  es el último movimiento inducido por el alimento,  $\beta_i^{alim}$  es la atracción del alimento, y  $\beta_i^{mejor}$  es el efecto del mejor fitness obtenido por el  $i$ -ésimo krill hasta el momento. El valor de  $V_f$  es  $0.02 (ms^{-1})$ . La posición del alimento no se puede determinar, pero se puede estimar. El centro virtual de la concentración de alimento puede ser estimado como un centro de masa de acuerdo a la distribución de fitness de los individuos, y está dado por:

$$X^{alim} = \frac{\sum_{i=1}^N \frac{1}{K_i} X_i}{\sum_{i=1}^N \frac{1}{K_i}} \quad (12)$$

La atracción que produce el alimento sobre el  $i$ -ésimo krill se modela como:

$$\beta_i^{alim} = C^{alim} \hat{K}_{i,alim} \hat{X}_{i,alim} \quad (13)$$

donde

$$C^{alim} = 2(1 - \frac{I}{I_{max}}) \quad (14)$$

Esto hace que el efecto atractivo del alimento disminuya con el tiempo.

El efecto del mejor fitness histórico del  $i$ -ésimo krill se modela como sigue:

$$\beta_i^{mejor} = \hat{K}_{i,mejor} \hat{X}_{i,mejor} \quad (15)$$

donde  $K_{i,mejor}$  hace referencia al fitness de la mejor posición visitada hasta el momento ( $X_{i,mejor}$ ).

**A.3. Movimiento difusivo aleatorio:** La difusión física de un individuo es considerado un proceso aleatorio. Este movimiento puede ser expresado en términos de una máxima velocidad de difusión y un vector dirección aleatorio. Se puede formular este movimiento como:

$$D_i = D^{max} \delta \quad (16)$$

donde  $D^{max}$  es la máxima velocidad de difusión, y  $\delta$  es el vector dirección aleatorio. Las componentes de este vector son números aleatorios en el rango  $[-1,1]$ . El rango para  $D^{max}$  es  $[0.002, 0.010](ms^{-1})$ . Es deseable que la difusión aleatoria disminuya con el paso del tiempo, por lo que se reformula la ecuación (16) como:

$$D_i = D^{max} (1 - \frac{I}{I_{max}}) \delta \quad (17)$$

Esto hace que la velocidad difusiva decrezca linealmente con el paso de las iteraciones.

## B. Proceso de Movimiento

A partir de los diferentes parámetros de movimientos, se modela el movimiento de un individuo con el paso del tiempo de la siguiente manera:

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \quad (18)$$

donde  $t + \Delta t$  es la iteración siguiente. El valor de  $\Delta t$  se define como:

$$\Delta t = C_t \sum_{j=1}^N (UB_j - LB_j) \quad (19)$$

donde  $N$  es el número total de variables, y  $LB_j$  y  $UB_j$  son los límites inferior y superior de la  $j$ -ésima variable ( $j = 1, 2, \dots, N$ ). El valor de  $C_t$  pertenece al rango  $[0,2]$ .

### C. Operadores genéticos

Para mejorar el rendimiento del algoritmo, mecanismos de reproducción genética son incorporados en el mismo: la cruza y la mutación.

**C.1. Cruza:** La operación de cruza está gobernada por una probabilidad de cruza  $Cr$ . El esquema de cruza binomial utilizado puede realizar una cruza sobre cada componente del vector posición, dependiendo de la probabilidad. La cruza se realiza de la siguiente manera:

$$x_{i,m} = \begin{cases} x_{r,m} & \text{si } rand_{i,m} < Cr \\ x_{i,m} & \text{caso contrario} \end{cases} \quad (20)$$

donde  $x_{i,m}$  es la  $m$ -ésima componente del vector posición  $X_i$ ,  $r \in \{1, 2, \dots, i-1, i+1, \dots, N\}$ . La probabilidad de cruza se define como:

$$Cr = 0,2\hat{K}_{i,mejor} \quad (21)$$

Esto hace que el mejor individuo tenga probabilidad de cruza igual a cero, por lo que se podría decir que el algoritmo usa elitismo en el sentido de que el mejor individuo no es modificado genéticamente. Por su parte la probabilidad de cruza de los peores individuos es más elevada.

**C.2. Mutación:** El operador de mutación es controlado por una probabilidad de mutación  $Mu$ . El esquema de mutación utilizado es:

$$x_{i,m} = \begin{cases} x_{gbes,m} + \mu(x_{p,m} - x_{q,m}) & \text{si } rand_{i,m} < Mu \\ x_{i,m} & \text{caso contrario} \end{cases} \quad (22)$$

$$Mu = 0,05/\hat{K}_{i,mejor} \quad (23)$$

donde  $p, q \in \{1, 2, \dots, i-1, i+1, \dots, K\}$  y  $X_{gbes}$  es la posición del mejor individuo global. Notar que la probabilidad de mutación del mejor es nula, y la probabilidad aumenta con el decrecimiento del fitness.

### III. METODOLOGÍA DEL ALGORITMO

El algoritmo KH puede implementarse mediante los siguientes pasos:

1. Estructura de datos: Definir el rango del problema, determinar los parámetros del algoritmo, etc.
2. Inicializar: Crear aleatoriamente las posiciones iniciales de la población en el espacio de búsqueda.
3. Evaluar fitness: Evaluar la función objetivo para cada individuo.
4. Cálculo del movimiento:
  - Movimiento inducido por otros individuos
  - Movimiento inducido por el alimento
  - Movimiento difusivo aleatorio
5. Aplicar operadores genéticos.

TABLA I  
COMPARACIÓN PARA LA FUNCIÓN ACKLEY

Método	Mejor	Iteración	Promedio
KH-cg	$8,2901 \times 10^{-04}$	1000	0,1290
KH-sg	$2,1149 \times 10^{-04}$	996	0,4872
GA	$4,4409 \times 10^{-16}$	140	$4,4409 \times 10^{-16}$
PSO	$5,5502 \times 10^{-04}$	460	9,1998

6. Actualizar la posición de los individuos.

7. Repetir: Ir al paso 3 mientras no se satisfaga la condición de corte.

### IV. ANÁLISIS COMPARATIVO

Para analizar el algoritmo KH, se comparó con dos algoritmos clásicos de optimización bio-inspirados, como son GA y PSO. Los métodos propuestos por [1] para la prueba del algoritmo son conjuntos de funciones de prueba bien conocidas, por este motivo, en este trabajo investigamos las prestaciones de los diferentes algoritmos en base a problemas de la misma complejidad. Así, analizamos por cada método un problema de cada uno de los siguientes grupos:

- (i) Funciones de prueba en altas dimensiones.
- (ii) Entrenamiento de una red neuronal para clasificación de patrones.
- (iii) Problemas ingenieriles con restricciones.

#### A. Funciones de alta dimensión

El motivo de la elección de una función de alta dimensión se basa en el hecho de que los tres métodos se comportan bien para funciones de baja dimensión, encontrando el valor óptimo con gran exactitud. La función analizada es llamada *Ackley Function*, de 20 dimensiones, que tiene el mínimo global  $f^*(X) = 0$ . Sus características pueden encontrarse en [6]. Para la prueba de dicha función, se utilizaron en los tres métodos, 15 individuos y 1000 iteraciones como máximo, además el algoritmo KH se probó con operadores genéticos (*KH-cg*) y sin ellos (*KH-sg*). En la tabla I se muestran, para 100 corridas, la mejor solución lograda, la iteración en la cual se logró esa mejor solución, y el promedio de solución para cada uno de los métodos. Puede observarse que el algoritmo genético es mucho más preciso y rápido en converger en comparación con los demás. Por su parte, el algoritmo KH sin operadores genéticos en su mejor corrida obtuvo un mejor resultado que con

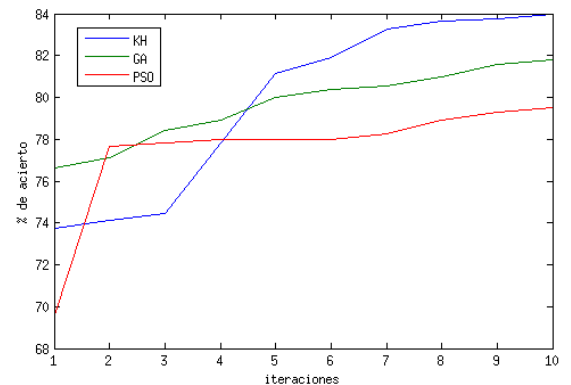


Fig. 1. Red neuronal entrenada por los diferentes métodos

TABLA II  
COMPARACIÓN PROBLEMA RECIPIENTE DE PRESIÓN

Método	Mejor Solución	Promedio	Tiempo
KH	11164,0	17283,0	3m 10.576s
GA	6556,7	7104,1	0m 30.986s
PSO	7787,4	13263,0	0m 0.280s

operadores genéticos, sin embargo en promedio la solución con operadores genéticos es mejor.

### B. Entrenamiento de red neuronal

Para este problema, se entrenó una red neuronal consistente en dos capas ocultas de 10 neuronas, y una neurona en la capa de salida. El archivo *clouds.csv* [7] fue utilizado para la clasificación. El objetivo de este estudio es acotar la cantidad de iteraciones y observar con cuál de todos los métodos la red entrenada tiene un mayor porcentaje de aciertos. El número de individuos usado para cada método es de 25 y la cantidad de iteraciones máximas 10. Se realizaron 10 corridas distintas para cada uno, y las mejores corridas de cada uno pueden visualizarse en la figura 1. Puede apreciarse que el algoritmo KH asciende más rápidamente que los demás métodos. En promedio, los porcentajes de acierto alcanzados son de 79.41 % para el KH, 77.75 % para el GA y 78.37 % para el PSO. Esto no significa que el algoritmo KH sea el más efectivo para el entrenamiento de la red, sino que se logra en pocas iteraciones resultados aceptables. Además, los tiempos promedios de los métodos para ejecutar las 10 iteraciones son de 1m 25.333s para el KH, 31.670s para el GA y 28.798s para el PSO. Como puede observarse, el KH es mucho más lento que los demás, siendo previsible este resultado teniendo en cuenta la cantidad de cálculos que posee el algoritmo según su definición.

### C. Problemas ingenieriles con restricciones

En este grupo el problema elegido es el de optimizar (minimizar) el costo de diseño de un recipiente de presión (ver Fig. 2) sujeto a una serie de restricciones. Las variables involucradas son el espesor ( $T_s$ ), el espesor de la cabeza ( $T_h$ ), el radio interior ( $R$ ) y la longitud de la sección cilíndrica del recipiente ( $L$ ). La solución óptima del problema es  $f^*(X) = 6059,714$ . Detalles del problema pueden encontrarse en [2].

En este análisis, se usaron 250 individuos por método, y se limitó la cantidad máxima de iteraciones a 1000 para observar cuál alcanzó una solución más aceptable. Los resultados pueden visualizarse en la tabla II. Se muestra la mejor solución que logró cada método en 10 corridas distintas y el promedio de soluciones, además del tiempo promedio en ejecutar las 1000 iteraciones. Como se observa, el algoritmo genético logró una mejor solución que los demás, y en promedio también. Las soluciones encontradas por el algoritmo KH no son muy aceptables. Como en el caso anterior, el algoritmo KH es el más lento de los tres, y el PSO el más rápido.

## V. RESULTADOS Y DISCUSIÓN

Comparando el algoritmo de manada de krills o KH con GA (algoritmo genético) y PSO (inteligencia de enjambre) en los distintos problemas presentados, podemos decir lo siguiente sobre la performance del mismo. En el caso

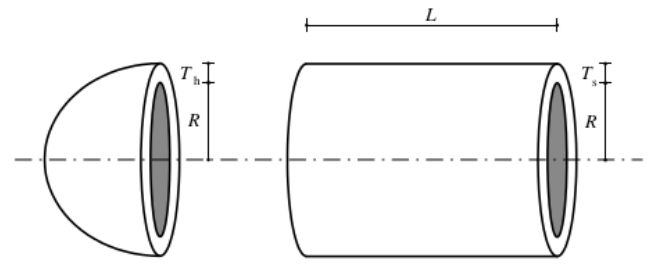


Fig. 2. Recipiente de presión

de optimizar una función de prueba de altas dimensiones (función *Ackley* en este caso), la mejor performance la tuvo el algoritmo genético, llegando en pocas iteraciones en comparación con los demás al óptimo global (el orden de  $10^{-16}$  podría considerarse como 0). El algoritmo KH aunque se acercó al óptimo, no tuvo tanta precisión ni rapidez de convergencia. Cabe destacar que en corridas posteriores que se hicieron y aumentando el número de iteraciones, el KH no fue capaz de superar el orden de  $10^{-4}$ .

En el caso del entrenamiento de una red neuronal (perceptrón multicapa) para clasificación, el KH fue el método que dio mejores porcentajes de acierto durante el entrenamiento, sin embargo como se mencionó anteriormente, la cantidad de cálculos que posee lo hace muy lento. De igual manera hay que tener en cuenta que el algoritmo y el código de implementación se podrían tratar de optimizar para reducir el costo de cálculos.

En el caso del problema con restricciones, los resultados brindados por el KH no son aceptables, estando demasiado lejos de la solución óptima. Se podría decir que para problemas con restricciones el algoritmo KH no es adecuado; los autores del mismo no presentan ninguna prueba con algún problema con restricción, solamente funciones sin restricciones.

## VI. CONCLUSIONES

Aunque los resultados obtenidos en el presente trabajo indican que usar el algoritmo KH no presenta ninguna ventaja comparativa con respecto a otros algoritmos de optimización como GA y PSO, el algoritmo KH es muy reciente y es posible de ser objeto de muchas investigaciones futuras que pueden mejorarlo sustancialmente.

## REFERENCIAS

- [1] Amir Hossein Gandomi, Amir Hossein Alavi, "Krill herd: A new bio-inspired optimization algorithm", *Communications in Nonlinear Science and Numerical Simulation*, Volume 17, Issue 12, December 2012, Pages 4831-4845, ISSN 1007-5704, 10.1016/j.cnsns.2012.05.010.
- [2] Gandomi A.H., Yang X., Alavi A.H., Talatahari S. "Bat algorithm for constrained optimization tasks". *Neural Computing Applications*. (20 July 2012), pp. 1-17.
- [3] Zang H, Zhang S, Hapeshi K. "A review of nature-inspired algorithms". *J Bionic Eng* 2010;7:S232-7.
- [4] Tang WJ, Wu QH. "Biologically inspired optimization: a review". *Trans Inst Meas Control* 2009; 31:495-515.
- [5] Hofmann EE, Haskell AGE, Klinck JM, Lascara CM. "Lagrangian modelling studies of Antarctic krill (*Euphasia superba*) swarm formation". *ICES J Mar Sci* 2004; 61:617-31.
- [6] Andries P. Engelbrecht. "Computational Intelligence: An Introduction". Second Edition, 2007.
- [7] Nayer Wanas. "Feature Based Architecture for Decision Fusion". Tesis, University of Waterloo, 2003.