

Mejora en la optimización de la ubicación de trampas para mosquito con Particle Swarm Optimization.

Topiltzin Hernández Mares

Asesor: Dr. Benjamín Valdés

Asesor: Dr. Héctor Sánchez

27 de mayo de 2022

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Estructura del trabajo	1
2. Estado del arte	3
2.1. Conceptos básicos	3
2.1.1. Inteligencia artificial	3
2.1.2. Optimización	4
2.1.3. Heurísticas	4
2.1.4. Metaheurísticas	4
2.1.5. Inteligencia de enjambre	4
2.1.6. Algoritmos evolutivos	5
2.1.7. Particle Swarm Optimization	5
2.2. Trabajos relevantes en el área	6
2.2.1. Optimización de ubicación de una red de sensores inalámbricos en un taller inteligente	6
2.2.2. Asignación óptima de unidades de generación distribuidas en sistemas de energía radiales	7
2.2.3. Optimización de ubicación de trampas para mosquitos . .	9
2.3. Comparación de algoritmos	10
2.4. Conclusión	12
3. Planteamiento del problema	13
3.1. Contexto	13
3.1.1. Marshall Lab	13
3.2. Limitaciones de la herramienta actual	14
3.3. Conclusión	14

4. Solución	15
4.1. Diseño de la solución	15
4.1.1. Selección de algoritmo	15
4.1.2. Stack tecnológico	16
4.2. Desarrollo de la solución	17
4.2.1. Arquitectura del sistema	17
4.2.2. Integración con el paquete MGSurvE	18
5. Metodología de evaluación	20
5.1. Hardware	20
5.2. Selección de parámetros	20
5.3. Comparación de métodos	21
5.3.1. Generación del entorno de evaluación	21
5.3.2. Ejecución de optimizaciones	21
5.3.3. Métricas	22
6. Resultados	24
6.1. Búsqueda de mejores parámetros	24
6.2. Comparación de algoritmos	25
7. Conclusiones y trabajo futuro	33
7.1. Conclusiones	33
7.2. Contribuciones	34
7.3. Trabajo futuro	34

Índice de figuras

2.1.	Inspiración de algoritmos de inteligencia de enjambre.	5
2.2.	Resultados de error de detección de 3D-FOA y 3D-AFOA.	7
2.3.	Resultados de error de detección con diferente cantidad de nodos.	8
2.4.	Comparación de resultados entre el PSO propuesto y el convencional.	9
2.5.	Interacción del Landscape con el algoritmo genético.	10
2.6.	Ventajas y desventajas de PSO y GA basado en [16, 17].	11
4.1.	Diagrama de clases describiendo los módulos del sistema.	19
5.1.	Entorno generado para la evaluación del algoritmo PSO.	22
6.1.	Comportamiento del PSO con diferentes configuraciones. En rojo se muestran las ejecuciones de la configuración 1; en verde la configuración 2; en azul la configuración 3; y en amarillo la 4. . .	26
6.2.	Comportamiento de PSO y GA	27
6.3.	Bandas de percentiles de resultados de PSO y GA	28
6.4.	Mejor resultado PSO	29
6.5.	Mejor resultado GA	30
6.6.	Peor resultado PSO	31
6.7.	Peor resultado GA	32

Índice de cuadros

5.1. Todas las opciones de parámetros	21
6.1. Resultados de comparación de las diferentes configuraciones de parámetros para el PSO (Tabla 5.1).	25

Resumen

A pesar de un incremento en las aplicaciones de algoritmos de optimización en los últimos años, no existen muchas aplicaciones en el ámbito del control y monitoreo de mosquitos. Actualmente solo existe MGSurvE, un paquete diseñado para optimizar la posición de trampas para mosquito, sin embargo, su algoritmo de optimización no es el más adecuado para esta tarea. Se desarrolló un algoritmo PSO que busca mejorar los resultados arrojados por MGSurvE. Se registraron mejoras en la cantidad de generaciones necesarias y en los resultados mínimos de detección de mosquitos.

Capítulo 1

Introducción

La investigación presentada en este trabajo se centra en una problemática existente en el paquete para optimización de la ubicación de trampas para mosquito MGSurvE: no se cuenta con el algoritmo adecuado para generar las posiciones óptimas de las trampas. Se propone un algoritmo como solución a esta problemática y se evalúa su eficacia.

1.1. Motivación

En cualquier proyecto se cuentan con recursos finitos para su desarrollo, ya sea dinero, materiales, o en el caso de lo analizado en este trabajo, trampas para mosquitos. Es de suma importancia lograr los mejores resultados con un número finito de trampas para detectar las migraciones de mosquitos.

1.2. Estructura del trabajo

El presente reporte está compuesto de 7 capítulos que tratan diferentes puntos de la investigación llevada a cabo.

En el Capítulo 2, se presenta una variedad de términos que sirven para dar contexto de este trabajo. En seguida, se presentan y analizan investigaciones relacionadas a la realizada en este reporte.

En el Capítulo 3 se presentan las limitaciones de las herramientas actuales y se detalla la problemática identificada.

A continuación, en el Capítulo 4, se propone una solución a la problemática

identificada en el capítulo anterior. Se detallan todos los componentes de la solución, desde el diseño del algoritmo hasta su arquitectura y el tipo de parámetros necesarios para su correcto funcionamiento.

Una vez propuesta una solución, en el Capítulo 5 se detalla la metodología usada para evaluar la efectividad del algoritmo propuesto. Se describe la búsqueda de la mejor configuración para el algoritmo y cómo será comparado con la implementación actual del paquete MGSurvE.

En seguida, se presentan los resultados de la evaluación en el Capítulo 6. Se presentan una variedad de gráficas y tablas para entender a detalle el comportamiento del algoritmo propuesto en este trabajo.

Finalmente, en el Capítulo 7 se presentan las conclusiones, contribuciones logradas y el posible trabajo futuro para mejorar los resultados obtenidos por la solución propuesta.

Capítulo 2

Estado del arte

Antes de hablar del problema y su solución, es necesario entrar en contexto. Por esto, en la primera sección de este capítulo se definirán conceptos básicos, tales como trampas para mosquitos, las cuales tienen un papel central en el problema. Enseguida, se hablará sobre términos de inteligencia artificial y métodos de optimización, los cuales ayudarán a entender mejor las problemática y su solución.

En la segunda parte de este capítulo se hablará de algunos trabajos recientes para poder entender el estado del arte en problemas de optimización. Sin embargo, por la poca investigación relacionada a la optimización de vigilancia de mosquitos, se analizarán trabajos similares enfocados en redes y monitoreo.

2.1. Conceptos básicos

En este trabajo se tocan temas de optimización, que es una de las ramas de la inteligencia artificial, por lo que es conveniente definir algunos conceptos importantes antes de continuar.

2.1.1. Inteligencia artificial

En el libro Artificial Intelligence: A Modern Approach [1], se define la inteligencia artificial (IA) como máquinas con pensamiento parecido al humano y con acciones racionales.

Una definición más elaborada es la de [2]: IA es un término utilizado para etiquetar a computadoras que imitan funciones humanas cognitivas, tales como

aprendizaje y resolución de problemas. De igual manera, se llama inteligencia artificial a algoritmos que trabajan con la misma complejidad que expertos humanos.

2.1.2. Optimización

Puede ser definida como el proceso de encontrar la mejor solución posible entre todas las disponibles. Por lo tanto, la tarea de optimización es modelar algún problema en términos de alguna función de evaluación y emplear algún algoritmo de búsqueda para minimizar (o maximizar) esa función objetivo. Sin embargo, no es posible asegurar que el resultado óptimo será encontrado, ya que hay problemas demasiado grandes que es imposible encontrar la solución óptima [3].

2.1.3. Heurísticas

En el área computacional de optimización, se llaman heurísticas a los métodos usados para buscar soluciones de alta calidad, sin asegurar la óptima.

Una definición más elaborada la da [4], pues dice que una técnica heurística es un método que busca soluciones buenas (y casi óptimas) con un costo computacional bajo. Sin embargo, estos métodos no son capaces de indicar qué tan cerca de la optimización perfecta se encuentra la respuesta actual dada por una heurística.

2.1.4. Metaheurísticas

Se refiere a técnicas que coordinan, manipulan y guían el comportamiento y soluciones de métodos heurísticos de optimización. Tales métodos pueden ser desde procedimientos de alto nivel o pueden describir las alteraciones posibles para una heurística [5].

Existe una gran cantidad de metaheurísticas [4, 3], más adelante se hablará de algunos métodos, tal como algoritmos genéticos y optimización de enjambre de partículas.

2.1.5. Inteligencia de enjambre

La inteligencia de enjambre es un campo computacional que diseña y estudia algoritmos para buscar soluciones a problemas. Estos algoritmos están inspira-

dos en el comportamiento complejo y, algunas veces, coordinado de enjambres de cualquier organismo natural [3].

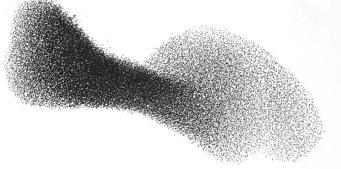


Figura 2.1: Inspiración de algoritmos de inteligencia de enjambre.

2.1.6. Algoritmos evolutivos

Son un tipo de algoritmos que buscan soluciones a problemas, tomando inspiración de la selección natural y genética [6, 7]. Estos algoritmos pueden representar cromosomas como cadenas de texto, los cuales son las soluciones a los problemas; los alfabetos representan genes; y más conceptos tomados de la biología.

2.1.7. Particle Swarm Optimization

Particle Swarm Optimization (PSO) fue inventada por Eberhart y Kennedy en 1995 [8], es un algoritmo estocástico inspirado en el comportamiento de una parvada de aves en busca de alimento. Cada partícula en el algoritmo, que representa una solución al problema, vuela en el espacio de búsqueda, actualizando su velocidad y su posición. La ecuación clásica para actualizar la velocidad de una partícula es 2.1 y para actualizar su posición es 2.2.

$$v_{ij} = v_{ij} + c_1 \text{rand}() (p_{ij} - x_{ij}) + c_2 \text{rand}() (p_{nj} - x_{ij}) \quad (2.1)$$

$$x_{ij} = x_{ij} + v_{ij} \quad (2.2)$$

Actualmente, el algoritmo PSO canónico (CPSO) o PSO con peso inercial (PSO-iw) es considerado como el PSO estándar [9, 10]. En esta forma de PSO, se introduce una nueva variable w , modificando 2.1 a 2.3. Esta nueva variable equilibra la búsqueda local y global de la solución.

$$v_{ij} = wv_{ij} + c_1 \text{rand}() (p_{ij} - x_{ij}) + c_2 \text{rand}() (p_{nj} - x_{ij}) \quad (2.3)$$

2.2. Trabajos relevantes en el área

En esta sección, después de haber definido conceptos importantes, se detallarán y analizarán algunos trabajos relevantes en el área que describen el estado del arte en optimización de posicionamiento de nodos en diferentes tipos de redes.

2.2.1. Optimización de ubicación de una red de sensores inalámbricos en un taller inteligente

En el primer trabajo de esta sección, de 2018, Li et al [11] desarrollaron un nuevo algoritmo con el objetivo de optimizar la ubicación de nodos de una red de sensores inalámbricos (WSN) para minimizar el error de detección de objetos 3D en un taller de manufactura inteligente.

Los autores tomaron inspiración del comportamiento de enjambres de moscas, e implementaron en algoritmo de optimización de moscas (FOA). Sin embargo, este algoritmo presenta un comportamiento 2D, y su problema necesita un análisis en tres dimensiones. Por esto, introdujeron una nueva dimensión en la búsqueda de soluciones, resultando en el algoritmo es 3D-FOA. Como una alternativa, decidieron añadir una variable más a su algoritmo, un peso inercial variable, proveniente de PSO-iw, resultando en un nuevo algoritmo: 3D-AFOA.

Al tener dos algoritmos capaces de optimizar los nodos de la red MSN, realizaron diversos experimentos para comparar dichos métodos. Inicialmente, establecieron la posición fija de 3 nodos de la red y un solo nodo objetivo a detectar en un taller simulado. Establecieron una población de 50 y 500 como número máximo de generaciones en ambos algoritmos. En la figura 2.2 se muestran los resultados obtenidos.

Como se puede ver en la figura 2.2, el algoritmo 3D-AFOA tiene un mejor desempeño, pues converge en una mejor solución en menos generaciones.

Con el objetivo de verificar la efectividad del algoritmo, se probó con 3, 4 y 5 nodos fijos en la red para detectar nuevamente un recurso en el taller simulado. Los resultados se muestran en 2.3.

Analizando los resultados expuestos en 2.3, se puede apreciar que, conforme incrementa el número de nodos en el algoritmo, el tiempo de convergencia ace-

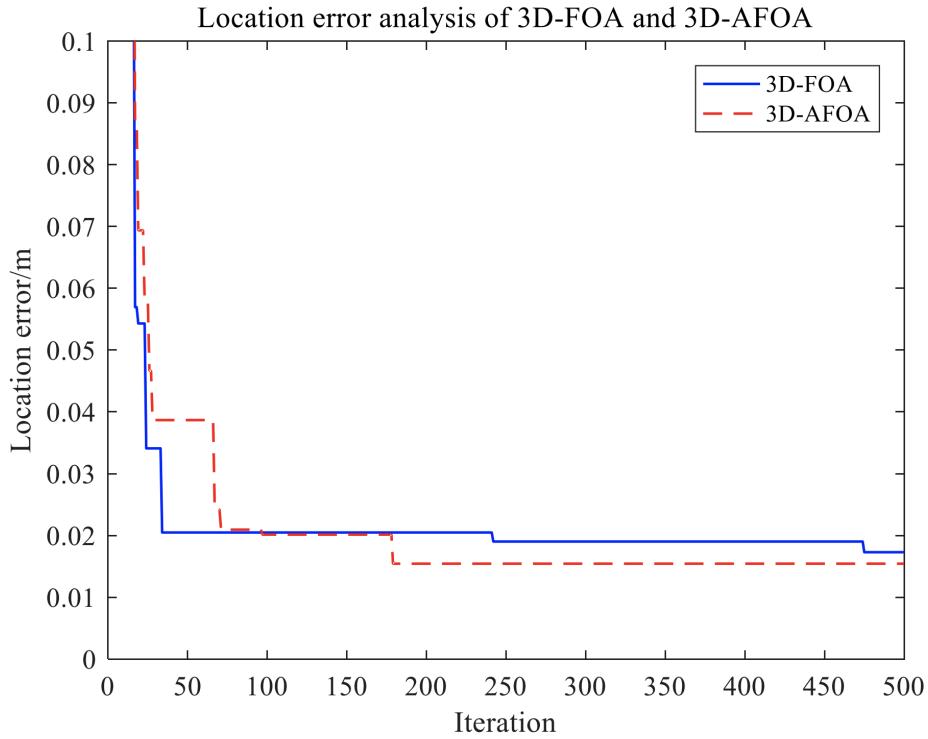


Figura 2.2: Resultados de error de detección de 3D-FOA y 3D-AFOA.

lera. Sin embargo, el error de detección de los objetivos, aunque mejora con más nodos, no existe gran diferencia. Esto es debido a que solamente la ubicación de un solo nodo fue optimizada en los algoritmos.

Para finalizar, los investigadores concluyeron que 3D-AFOA es altamente aplicable al problema de optimización en la ubicación de nodos en una red WSN y no requiere de un alto presupuesto computacional para lograr buenas soluciones.

2.2.2. Asignación óptima de unidades de generación distribuidas en sistemas de energía radiales

Como segundo trabajo relevante presentado en esta sección, Hantash et al, en 2020 [12], publicaron un estudio en el que proponen una modificación al estándar PSO-iw para calcular las posiciones y tamaños de unidades generadoras distribuidas en una red radial de energía eléctrica.

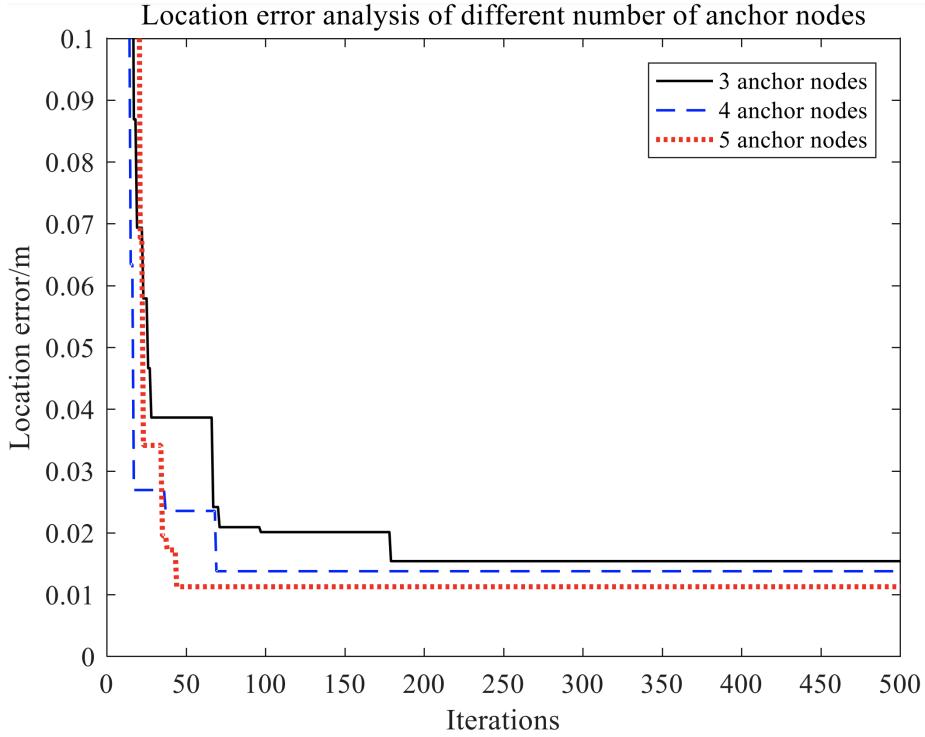


Figura 2.3: Resultados de error de detección con diferente cantidad de nodos.

Los autores partieron del estándar PSO-iw y, después de una investigación, encontraron que valores grandes en el peso inercial (w) facilitan la búsqueda global, mientras que valores pequeños de w favorecen la búsqueda local [9, 13]. Gracias a estos hallazgos, los autores proponen una estrategia no lineal para obtener valores de w variables con el tiempo. Esta estrategia está descrita en la Ecuación 2.4:

$$w = w_{max} e^{(((maxiter - iter) / maxiter) - 1)} - w_{min} \quad (2.4)$$

en donde w_{max} es el peso máximo, w_{min} es el peso mínimo, $iter$ es la iteración actual y $maxiter$ es el número máximo de iteraciones.

Finalmente, para validar sus resultados, el grupo de investigadores decidió comparar su estrategia propuesta contra un PSO convencional. Para comparar los resultados arrojados por ambos algoritmos, se usó la red de energía estándar IEEE 34. Después de la ejecución de los experimentos, los investigado-

res obtuvieron pérdidas de energía 31.6 % menores con su propuesta de PSO a comparación del PSO convencional sin el valor de w variable. Además de esto, la propuesta consumió 62.2325 s para proveer una solución optima, cuando el PSO convencional consumió 62.2325 s. La comparación completa entre el PSO propuesto y el convencional está en la Figura 2.4.

	Proposed PSO	Convectional PSO
DG1_size (MW)	1.6115	1.6722
DG1_Vc (pu)	1.0155	1.0055
DG1_location	24	10
Fitness losses	0.0401	0.0406
Average _elapsed time (s)	62.2325	78.6212

Figura 2.4: Comparación de resultados entre el PSO propuesto y el convencional.

2.2.3. Optimización de ubicación de trampas para mosquitos

Para finalizar esta sección, se describe el trabajo realizado por Sánchez y el Departamento de bioestadística y epidemiología de la universidad de Berkley [14] que aún no ha sido publicado y sigue en progreso, sin embargo, es el estado del arte en el área de este trabajo. Sánchez et al están trabajando en un paquete llamado MGSurvE, el cual está diseñado para optimizar la ubicación de trampas de mosquito en entornos espacialmente heterogéneos.

En el estado actual del paquete cuenta con diversas funcionalidades, tales como: diferentes kernels de movimiento para mosquitos machos y hembras; kernels de atractivo para trampas personalizables; soporte para trampas inamovibles; rutinas de trazado de mapas integradas; integración de rutinas para optimización con GA; y mucho más, con más funciones esperadas en futuras actualizaciones.

Como se mencionó, las rutinas de optimización para la ubicación de las trampas para mosquito se implementa un algoritmo genético (GA), el cual es la variante estándar y es importado directamente del paquete DEAP [15].

Para el correcto funcionamiento del algoritmo de optimización, 3 grupos de características deben ser definidas: del entorno, de las trampas y del compor-

tamiento de migración de mosquitos. En la figura 2.5 se describe cómo estos tres grupos de características forman el "Landscape", el cual es usado por el algoritmo genético.

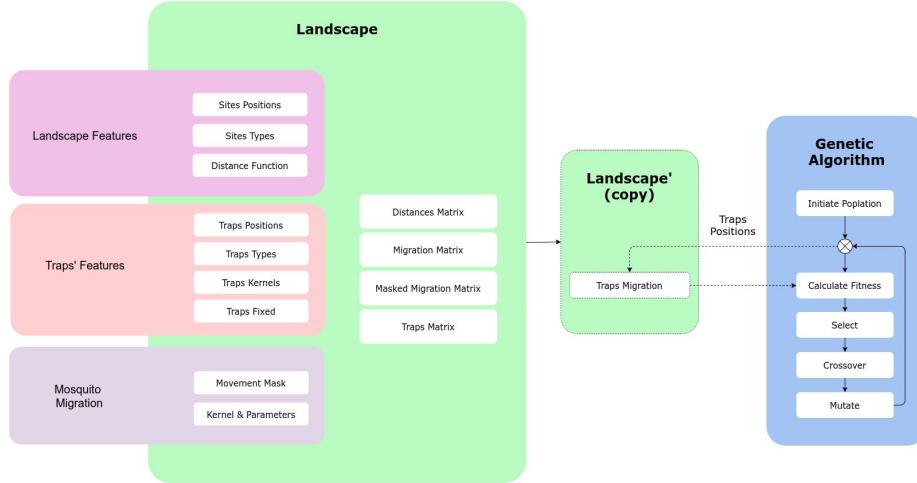


Figura 2.5: Interacción del Landscape con el algoritmo genético.

En la fecha de realización de este capítulo, aunque no hay ningún experimento publicado oficialmente, la documentación del paquete sí cuenta con ejemplos y tutoriales de optimización en la siguiente URL <https://chipdelmal.github.io/MGSurvE/build/html/GA.html>. En los tutoriales, el número de población es de 10 por cada trampa a optimizar. Si se desean optimizar 4 trampas, se utiliza una población de entre 40 o 50. Además de definir la población, también se limita a 500 generaciones para la optimización.

Para finalizar, el paquete permite obtener el mejor cromosoma y trazar en un diagrama el mapa de las trampas y la migración de los mosquitos.

2.3. Comparación de algoritmos

Los trabajos mencionados en la sección anterior proponen diferentes algoritmos para solucionar problemas muy similares: encontrar la mejor ubicación de nodos en una red de nodos para optimizar algún resultado de una medición. Un aspecto en común de las soluciones propuestas por los diferentes investigadores es que todos hacen uso de algoritmos metaheurísticos iterativos para

optimización. Otra similitud es que sus soluciones están basadas en inteligencia de enjambre y más específicamente en PSO.

Para resaltar las ventajas y desventajas de los diferentes tipos de algoritmos, en la tabla 2.6 se muestra una comparación detallada entre PSO y GA.

Algoritmo	Ventajas	Desventajas
Algoritmos Genéticos	<ul style="list-style-type: none"> Resultados sobresalientes con presupuesto computacional alto. Procesos más exploratorios. 	<ul style="list-style-type: none"> Requiere poblaciones grandes (al menos 50 individuos). Malos resultados con presupuesto computacional bajo. Lenta convergencia.
Particle Swarm Optimization	<ul style="list-style-type: none"> Requiere poblaciones pequeñas (20-40 partículas). Resultados sobresalientes con presupuesto computacional bajo. Sencillo de implementar. 	<ul style="list-style-type: none"> Malos resultados con presupuesto computacional alto. Riesgo de convergencia prematura.

Figura 2.6: Ventajas y desventajas de PSO y GA basado en [16, 17].

Además de los algoritmos propuestos por las investigaciones mencionadas en la sección anterior, se pueden encontrar una gran variedad usados en la actualidad para resolver problemas optimización. En un estudio realizado en 2021 por Piotrowski et al [16], encontraron que las principales familias de algoritmos usados para resolver problemas de optimización relacionados a la pandemia de COVID-19 fueron Swarm Intelligence (SI) y Evolutionary Algorithms (EA).

En 2017, un estudio similar realizado por Piotrowski et al [17], sometieron a 33 metaheurísticas propuestas entre 1960 y 2016 a 22 experimentos, cada una con diferentes presupuestos computacionales. Al finalizar los experimentos, encontraron que algoritmos basados en inteligencia de enjambre tienen un mejor desempeño cuando el presupuesto computacional es bajo. A diferencia de los algoritmos basados en inteligencia de enjambre, los basados en procesos evolutivos tienen un mejor desempeño cuando el presupuesto computacional es alto.

Aunque lo anterior es cierto, en un estudio de 2020 realizado por Piotrowski et al [18], se encontró que variantes modernas de PSO y SI sin peso inercial logran mejores resultados con una población de entre 70 y 500 partículas, contradiciendo las recomendaciones para lograr mejores resultados con variantes clásicas. Debido a esto, es de gran importancia conocer las condiciones adecuadas para cada variante de algoritmos basados en PSO. Si la cantidad adecuada de partículas para un problema en cuestión es desconocido, la recomendación

es usar entre 70 y 100 partículas.

2.4. Conclusión

En la actualidad, existe una gran variedad de algoritmos de optimización los cuales tienen comportamientos diferentes para cada tipo de problema [17]. Por esto, es de suma importancia escoger correctamente el algoritmo que será empleado en alguna solución, para así obtener resultados realmente óptimos.

Como se ve en la tabla 2.6 y se discutió en la sección anterior, los algoritmos genéticos tienen un mejor comportamiento y desempeño cuando el presupuesto computacional y la población es grande (más de 50), por lo que el trabajo realizado en MGSurvE [14] no está optimizando correctamente las ubicaciones de las trampas para mosquitos, pues usan poblaciones de entre 40 y 50, además de un presupuesto computacional bajo. Esta situación da paso a una oportunidad de mejora, para implementar un algoritmo que se ajuste en mejor manera a la situación de MGSurvE, con la posibilidad de mejorar la optimización de la ubicación de las trampas para mosquitos.

Capítulo 3

Planteamiento del problema

Después de analizar diferentes métodos de optimización aplicados en el área de este trabajo, el objetivo de este capítulo es describir la problemática en la optimización de la ubicación de trampas para mosquito. Para lograrlo, en este capítulo primeramente se dará un poco de contexto sobre la herramienta actual, enseguida se hablará sobre problemas y limitaciones con la herramienta actual y finalmente se concluirá con el planteamiento del problema.

3.1. Contexto

3.1.1. Marshall Lab

Marshall Lab [19] es un laboratorio perteneciente al departamento de epidemiología y bioestadística de la universidad de Berkeley en California. Su principal contribución es el desarrollo y simulación de estrategias basadas en genética para controlar enfermedades transmitidas por mosquitos. Para controlar estas enfermedades, el laboratorio simula la liberación de mosquitos modificados genéticamente en un entorno controlado, para así estudiar y desarrollar estrategias que permitan el comportamiento deseado de estos mosquitos en el ambiente. El control de los mosquitos liberados se hace con un correcto entendimiento de los patrones de movimiento de estos mosquitos. Para detectar el movimiento de mosquitos, se utilizan trampas, las cuales detectan la migración de los mosquitos.

Para la simulación del correcto posicionamiento de trampas, el laboratorio ha creado el software MGSurvE [14], el cual ya fue descrito en el capítulo anterior

2.2.3.

Es de suma importancia el trabajo realizado por este laboratorio, pues el correcto posicionamiento de trampas para mosquitos significaría un ahorro económico en la batalla para el control de enfermedades transmitidas por mosquito.

3.2. Limitaciones de la herramienta actual

En la sección de comparación de algoritmos del capítulo del estado del arte 2.3 se analizaron diferentes tipos de algoritmos de optimización y en la tabla 2.6 se compararon los algoritmos de la familia de PSO contra algoritmos de la familia de GA. Se concluyó en esa sección que los algoritmos PSO generalmente tienen un mejor desempeño cuando el presupuesto computacional es bajo, lo contrario de algoritmos de tipo GA. Como se habló en la sección 2.2.3, el paquete MGSurvE implementa un algoritmo genético (GA) para sus tareas de optimización de tiempos de detección de mosquitos. Sin embargo, se hace uso de una población pequeña, así como un límite de 500 generaciones, por lo que el presupuesto computacional es bajo y el rendimiento del algoritmo y sus resultados obtenidos posiblemente no sean los óptimos.

3.3. Conclusión

Después de establecer un estado del arte en el capítulo anterior y detallar las limitaciones de la herramienta actual, se puede concluir la problemática como: El software MGSurvE no genera las ubicaciones óptimas para trampas de mosquitos.

Capítulo 4

Solución

En este capítulo se detalla la solución propuesta para la problemática expuesta en el capítulo anterior. Se explica el diseño y desarrollo de la solución, así como el algoritmo implementado y configuración de hyperparámetros.

4.1. Diseño de la solución

4.1.1. Selección de algoritmo

Como se habló en el capítulo pasado, la problemática a resolver es que el paquete MGSurvE no genera las ubicaciones óptimas para trampas de mosquitos, siendo la causa principal del problema que el algoritmo de optimización usado actualmente en MGSurvE (un algoritmo genético, o GA) no tiene un correcto comportamiento con el bajo presupuesto computacional [17]. Por esto, un algoritmo perteneciente de tipo PSO es adecuado para resolver este problema [17, 18]. Los algoritmos de tipo PSO tienden a comportarse de manera ideal con una población pequeña (de 20 a 50 partículas), por lo que necesitan un presupuesto computacional bajo, adecuado para resolver el problema presentado anteriormente.

Se implementó un algoritmo de PSO como solución a la problemática. Este algoritmo es el propuesto en 2020 por Hantash et al [12]. En este algoritmo introducen una variación al peso inercial original propuesto por Shi et al [9] en 1998 (descrito anteriormente en la sub sección 2.1.7). Con esta modificación, la convergencia del algoritmo es controlada por el peso. Al empezar con valores mayores y terminar con valores menores, se prioriza la búsqueda global inicial-

mente y, en las últimas etapas de la búsqueda, se mejora la búsqueda local en cada partícula [9].

Este algoritmo se divide en dos partes principales: la primera, actualización de la velocidad y posición de una partícula; y segunda, la evaluación de la función de aptitud y actualización de valores óptimos. Las partes del algoritmo están detalladas en las ecuaciones 4.2, 4.3 y en el algoritmo 1.

$$w = w_{max} e^{(((maxiter - iter) / maxiter) - 1)} - w_{min} \quad (4.1)$$

$$v_{ij} = w v_{ij} + c_1 rand() (p_{ij} - x_{ij}) + c_2 rand() (p_{nj} - x_{ij}) \quad (4.2)$$

$$x_{ij} = x_{ij} + v_{ij} \quad (4.3)$$

En la ecuación 4.1 se muestra el cálculo y actualización del peso inercial. Al ser una función no lineal, se aseguran valores de w más grandes en etapas tempranas del algoritmo, así como una rápida transición a valores más pequeños conforme progresá la ejecución. Por lo tanto, se asegura una búsqueda global de mínimos con pocas posibilidades de quedar atrapado en un mínimo local [13].

Las ecuaciones 4.2 y 4.3 corresponden a la primer parte del PSO. En la primer ecuación, se realiza la actualización de la velocidad de la partícula i en la o j con respecto a la mejor partícula n . La segunda ecuación describe la actualización de la posición de la partícula con la nueva velocidad calculada.

4.1.2. Stack tecnológico

A continuación, se describen todas las dependencias usadas para la implementación del algoritmo.

Como herramienta principal para la manipulación de datos está la biblioteca Pandas [21]. Esta herramienta de código libre permite el análisis y manipulación de datos en la solución, así como la creación de gráficas rápida y sencillamente.

Además de análisis y manipulación de datos, la solución depende de una herramienta que permite realizar operaciones con vectores y matrices. El paquete por default en el ecosistema de Python es NumPy, el cual ofrece funciones matemáticas y de vectorización. Escrito en código optimizado de C y con una interfaz para Python, ofrece gran eficiencia y oportunidades de paralelización en sus operaciones [22].

Algorithm 1 Evaluación de función de aptitud y actualización de mejor partícula

```
Initialize mejorParticula ← None
for each generacion ∈ generaciones do
    for each particula ∈ particulas do
        particula.aptitud.valores ← funcionAptitud(particula)
        if not particula.mejor or particula.mejor.aptitud < particula.aptitud then
            particula.mejor ← nueva particula basada en particula
            particula.mejor.aptitud.valores ← particula.aptitud.valores
        end if
        if not mejorParticula or particula.aptitud < particula.aptitud then
            mejorParticula ← nueva particula basada en particula
            mejorParticula.aptitud.valores ← particula.aptitud.valores
        end if
    end for
    for each particula ∈ particulas do
        actualizaParticula(particula, mejorParticula, generacion)
    end for
end for
```

Por último, para la implementación del algoritmo se basa en el paquete DEAP [23], el cual provee un marco de trabajo para la creación e implementación de algoritmos evolutivos en Python. La implementación del algoritmo genético con el que funciona actualmente MGSurvE está construida con DEAP, por lo que la reutilización de este paquete ayuda el desarrollo de la solución propuesta en este trabajo.

4.2. Desarrollo de la solución

4.2.1. Arquitectura del sistema

El sistema está dividido en diferentes módulos, los cuales, junto con sus interacciones y relaciones, están descritos en la figura 4.1. Como se puede apreciar en el diagrama, el módulo principal es el llamado *Optimization*. Este cuenta con tres propiedades principales: *toolbox*, el cual almacena todas las herramientas y funciones del algoritmo; *population*, que guarda la lista de todas las partículas involucradas en el algoritmo de optimización; y *best*, representando la mejor partícula de la búsqueda. De igual manera, el módulo *Optimization* cuenta con diferentes funciones y métodos: *optimize_traps_pso*, el cual funciona como

wrapper para configurar e inicializar el algoritmo de optimización con los datos de las trampas y un *Landscape*; *pso_variable_weight*, el algoritmo con peso inercial variable; *generate_particle*, crea una partícula nueva para ser usada en el algoritmo; *update_particle_variable_weight*, que implementa la función 4.1 para actualizar los pesos de velocidad de una partícula; y finalmente *create_toolbox*, el cual crea la caja de herramientas y registra las funciones mencionadas anteriormente para ser usadas por el algoritmo.

4.2.2. Integración con el paquete MGSurvE

Para terminar esta sección de descripción de la solución, a continuación se describe la integración del algoritmo propuesto al paquete MGSurvE. El paquete, al ya contar con una función de optimización entre sus herramientas, integrar un nuevo algoritmo con la misma funcionalidad no fue complicado ni laborioso. Se añadió una clase llamada *PSO*, la cual tiene una interfaz similar a la del algoritmo de GA actual de MGSurvE. Los argumentos que recibe el constructor de la clase son: *landscape*, el objeto del paquete MGSurvE que contiene información del entorno, trampas y migración de mosquitos; *pop_size*, el tamaño de la población; *generations*, en donde se especifican el número de generaciones en las que el algoritmo correrá; *speed_min*, la velocidad mínima de las partículas; *speed_max*, la velocidad máxima de las partículas; *phi1*, límite superior para generar números aleatorios para influenciar la búsqueda local; *phi2*, límite superior para generar números aleatorios para influenciar la búsqueda global; *w_max*, peso inercial máximo; y finalmente, *w_min*, peso inercial mínimo.

El PSO, además de ser compatible con la configuración del Landscape, también genera un logbook con los mismos datos generados por el GA, para así permitir la comparación de ambos métodos de optimización.

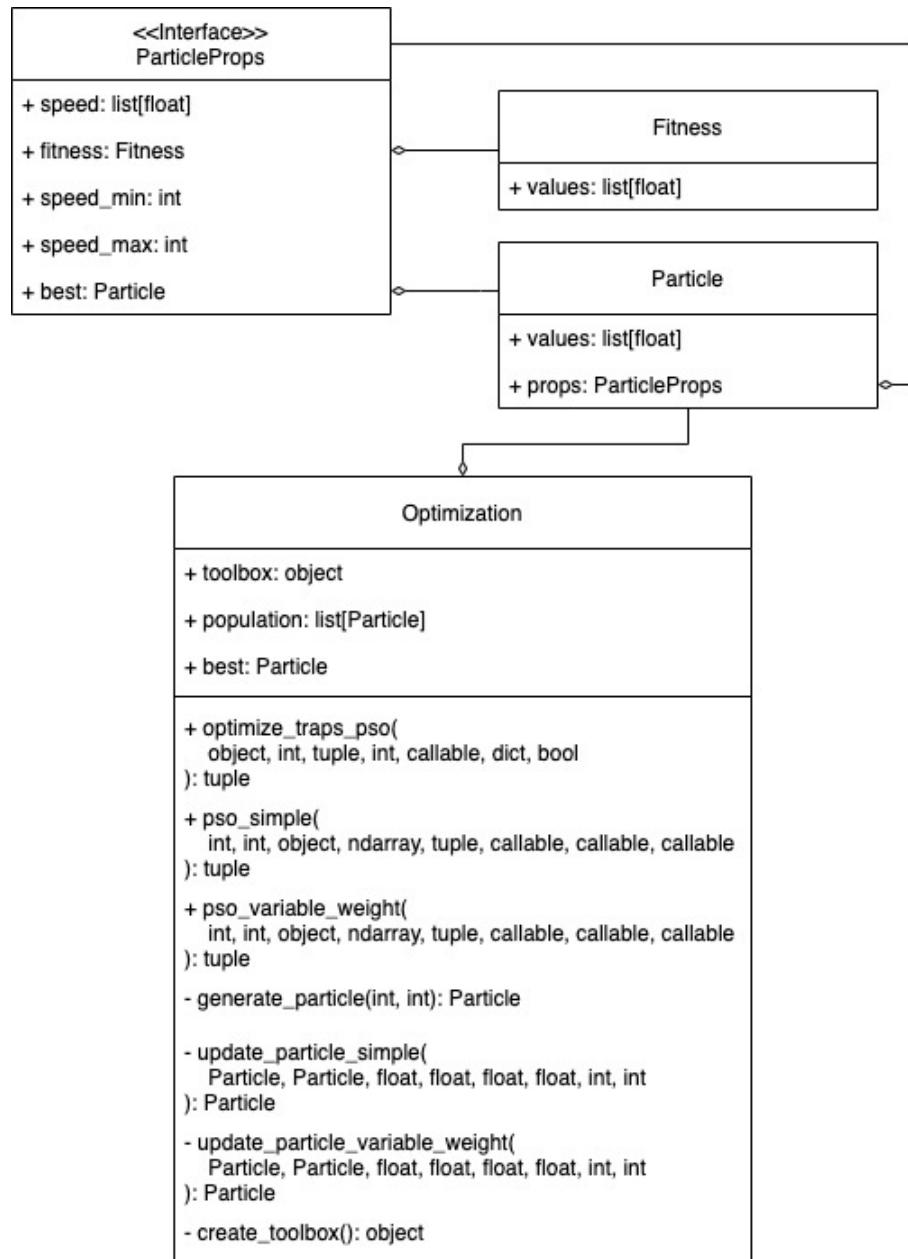


Figura 4.1: Diagrama de clases describiendo los módulos del sistema.

Capítulo 5

Metodología de evaluación

En este capítulo se describe la metodología de evaluación seleccionada para validar que se ha resuelto la problemática identificada en este trabajo. A continuación, se describe la técnica usada para evaluar el desempeño del algoritmo implementado como solución, así como las métricas que se obtendrán para confirmar la efectividad de la solución.

5.1. Hardware

La evaluación del algoritmo propuesto en este trabajo se realizó en una MacBook Pro 2020, con un procesador Intel Core i7 de 2.6 GHz con 6 núcleos y memoria de 16 GB 2667 MHz DDR4.

5.2. Selección de parámetros

Los algoritmos de tipo PSO necesitan una configuración de parámetros, como el tamaño de la población con la que se ejecutará la búsqueda y las velocidades mínima y máxima.

Se realizó una búsqueda de parámetros comparando los resultados de ejecutar la optimización con un mismo entorno entorno y con distintas configuraciones de parámetros. La Tabla 5.1 contiene las diferentes configuraciones con que el algoritmo de optimización fue ejecutado. Se ejecutó 5 veces un PSO con cada configuración, para así obtener promedios de los resultados arrojados por cada una. En la sección 6.1 se explicarán a detalle los resultados de esta búsqueda.

Cuadro 5.1: Todas las opciones de parámetros

Configuración	Tamaño de población	Velocidad Mínima	Velocidad Máxima
1	12	-5	5
2	12	-20	20
3	40	-5	5
4	40	-20	20

5.3. Comparación de métodos

Se determinó que la mejor manera de evaluar el desempeño del algoritmo propuesto en este trabajo es una comparación directa entre el algoritmo PSO y el algoritmo existente en MGSurvE. Para esto, primero se realizó una búsqueda de mejores parámetros, para así tener una referencia de con qué tipo de parámetros el algoritmo PSO funciona mejor. Después, con los mejores parámetros encontrados, se comparará contra el GA en un entorno generado por el mismo MGSurvE. A continuación se detallará todo el procedimiento.

5.3.1. Generación del entorno de evaluación

Para generar el entorno de evaluación se utilizó la función *ptsDonut* de MGSurvE, la cual genera una distribución de puntos en forma de dona. Como parámetros, se le indicó que generara 200 puntos en un radio mínimo de la dona de 75 unidades y un máximo de 100 unidades. El resultado de ejecutar esta función es el mostrado en la Figura 5.1. Este entorno cuenta con 4 trampas, cuyas ubicaciones serán optimizadas. Para facilitar el uso de este entorno, se guardaron los datos generados por MGSurvE en un archivo pickle, para después ser cargado por el programa de comparación de métodos.

5.3.2. Ejecución de optimizaciones

Para facilitar la ejecución de ambos algoritmos con un mismo entorno, se escribió un programa que primero lee la información del entorno generado previamente, para después correr los dos métodos de optimización (PSO y GA). Este programa ejecuta 20 veces el PSO y 20 veces el GA, para así obtener mejores estadísticas de los datos generados (como los promedios, y percentiles). Los datos generados por el programa se almacenaron en estructuras de datos para,

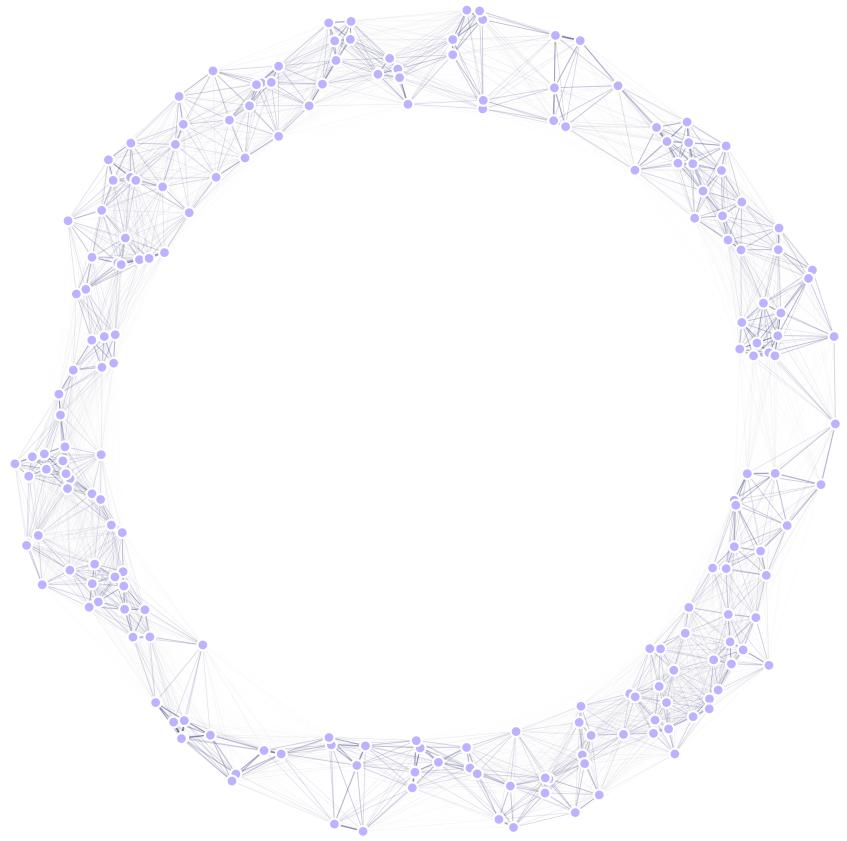


Figura 5.1: Entorno generado para la evaluación del algoritmo PSO.

posteriormente, generar visualizaciones de estos.

5.3.3. Métricas

Una vez exportados los datos de las optimizaciones, se obtuvieron tres métricas que ayudan a validar el desempeño del algoritmo propuesto en este trabajo: tiempo de ejecución de la optimización, esto ayuda a verificar si el algoritmo PSO es más rápido que la implementación actual en MGSurvE; generación en la que se encuentra el valor óptimo, así se sabe qué algoritmo encuentra el valor mínimo más rápido; y valor mínimo de detección de mosquitos, esta métrica es

la principal para saber si el PSO es mejor optimizando las ubicaciones de las trampas a comparación del GA y determinará si se cumplió el objetivo.

En el siguiente capítulo se presentan los resultados obtenidos de esta evaluación.

Capítulo 6

Resultados

En este capítulo se describirán a detalle los resultados obtenidos de la evaluación del algoritmo PSO presentado en este trabajo.

6.1. Búsqueda de mejores parámetros

Los resultados de la búsqueda de mejor configuración para el PSO se muestran en la tabla 6.1. En color verde se resalta la configuración con el mejor promedio de resultados. Como se puede ver en la tabla, la configuración 4 de parámetros fue la que arrojó un mejor valor mínimo de detección de mosquitos, seguida de cerca por la configuración 2, enseguida la 3 y muy por detrás la configuración 1. La principal causa de que la configuración 1 y 2 obtuvieran resultados muy similares son las velocidades mínima y máxima, pues ocasionan búsquedas amplias al inicio del algoritmo, encontrando así los mínimos globales rápidamente. Sin embargo, hay una gran diferencia de más de 45 segundos en el tiempo de ejecución promedio, esto debido a la diferencia entre el tamaño de población.

Para comprender y analizar mejor el comportamiento de cada configuración, se presenta la Figura 6.1, en la cual se grafica el comportamiento del PSO con cada configuración. En esta gráfica se aprecian 5 líneas por cada configuración, para así entender su comportamiento general. En la gráfica que se ve la similitud entre las configuraciones 2 y 4, sin embargo, la configuración 4 comienza a encontrar mejores resultados que la 2 en menos generaciones. Así mismo, se aprecia en la gráfica que las configuraciones 2 y 4 se atoran en menos mínimos

locales, y la configuración 4 lo hace menos que la 2.

Por lo tanto, se seleccionaron 20 y -20 como velocidad máxima y mínima. Igualmente, como población para la etapa de comparación contra el algoritmo genético se optó por 40 partículas, pues se obtuvieron mejores resultados en menor cantidad de generaciones con este tamaño de población. Además, como parámetros de peso (w) máximo y mínimo se seleccionaron 1.5 y 0.9, para lograr aún mejores búsquedas globales al principio del algoritmo y mejores búsquedas locales en las etapas finales. Para las generaciones, se puede apreciar en la Figura 6.1 que entre más generaciones, mejores son los resultados, por esto, se usarán 500 generaciones para la comparación de algoritmos.

Cuadro 6.1: Resultados de comparación de las diferentes configuraciones de parámetros para el PSO (Tabla 5.1).

Configuración	Resultado promedio (días)	Tiempo de ejecución promedio (segundos)
1	1.654883190977614	15.095740842819215
2	1.5887967821719775	15.379290246963501
3	1.5999659131506383	50.010193014144896
4	1.5770302194541654	60.24051342010498

6.2. Comparación de algoritmos

En las Figuras 6.2 y 6.3 se presentan los comportamientos generales de cada algoritmo con el entorno generado previamente. En la primer imagen se muestran los resultados óptimos encontrados por generación durante las 20 ejecuciones de cada algoritmo. Se puede apreciar como, el PSO, es capaz de encontrar muy buenos resultados tempranamente, pues entre las generaciones 50 y 100 está alcanzando los mínimos. Sin embargo, un 20 % de las ocasiones, el algoritmo propuesto se encuentra atorado en mínimos locales desde etapas tempranas de la búsqueda y no es capaz de salir. A diferencia del PSO, el algoritmo genético es capaz de salir de los mínimos locales y encontrar un valor cercano al mínimo global en las etapas tardías de la optimización. A pesar de esto, el PSO logra encontrar mejores resultados la mayoría de las veces.

En la figura 6.3 se puede apreciar mejor la calidad de los resultados encontrados por cada algoritmo. En la gráfica de bandas de percentiles se muestra una línea sólida con el promedio de los resultados por generación y en colores

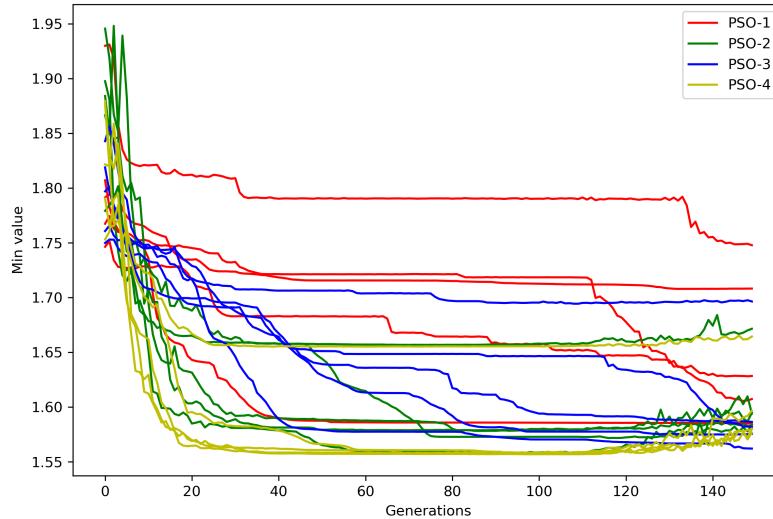


Figura 6.1: Comportamiento del PSO con diferentes configuraciones. En rojo se muestran las ejecuciones de la configuración 1; en verde la configuración 2; en azul la configuración 3; y en amarillo la 4.

más claros se presentan los percentiles 75 y 25. Se puede apreciar que la mayoría de las veces el PSO logra optimizar de mejor manera la posición de las trampas en el entorno generado. La línea de promedio del PSO se ve separada de los percentiles debido a las ejecuciones en donde este algoritmo se vió atorado en un mínimo local.

El tiempo de ejecución promedio del GA fue de 97.807 segundos con el equipo de cómputo presentado anteriormente (5.1) y el PSO presenta un tiempo de ejecución de 200.9426 segundos. El tiempo de ejecución del PSO se ve impactado en su mayoría por el número de evaluaciones por generación (40), a diferencia del GA (de 30 a 50). Además, la implementación del PSO realiza operaciones a nivel de Python, cuando el GA realiza la mayoría de sus operaciones de una manera más eficiente (con NumPy principalmente).

Adicionalmente, en las Figuras 6.4 y 6.5 se presentan las mejores posiciones encontradas para las trampas por el PSO y el GA (respectivamente). Así mismo, las Figuras 6.6 y 6.7 muestran los peores resultados de estos algoritmos. En la Figura 6.6 se puede apreciar la consecuencia de quedar en un mínimo local, pues

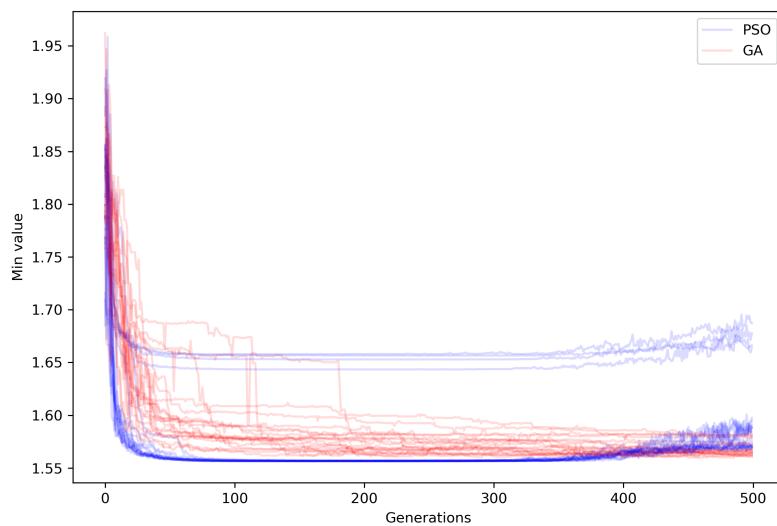


Figura 6.2: Comportamiento de PSO y GA

el resultado es significativamente peor que el del GA.

Gracias a estos resultados, se puede afirmar que el algoritmo propuesto en este trabajo logra encontrar mejores posiciones para las trampas, a comparación del algoritmo genético implementado en MGSurvE.

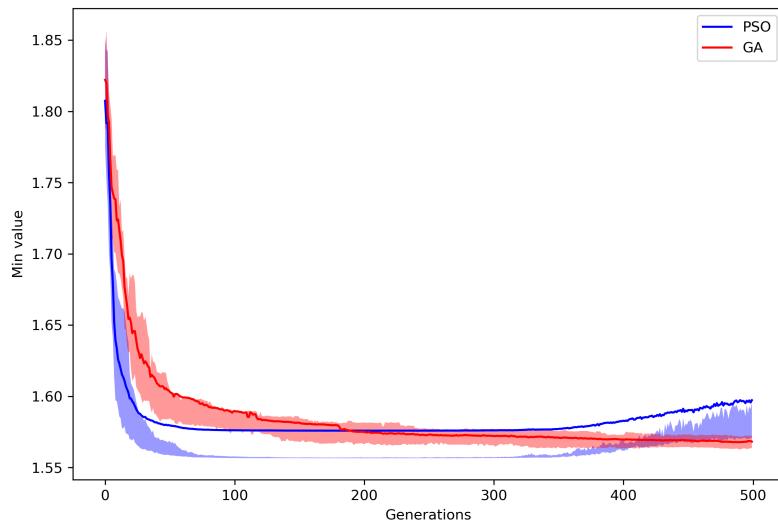


Figura 6.3: Bandas de percentiles de resultados de PSO y GA

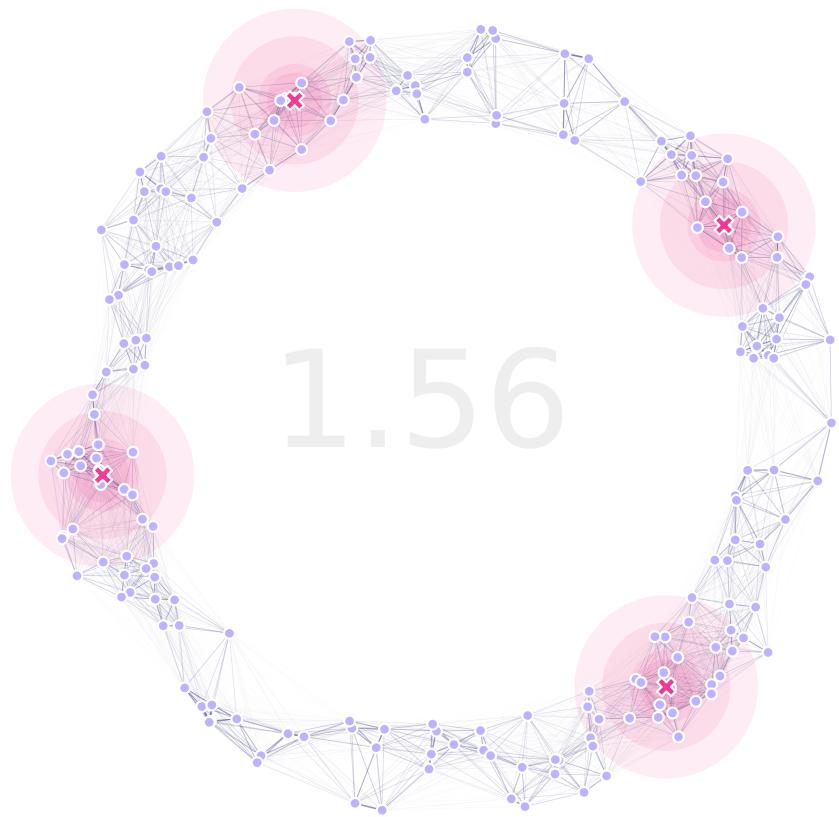


Figura 6.4: Mejor resultado PSO

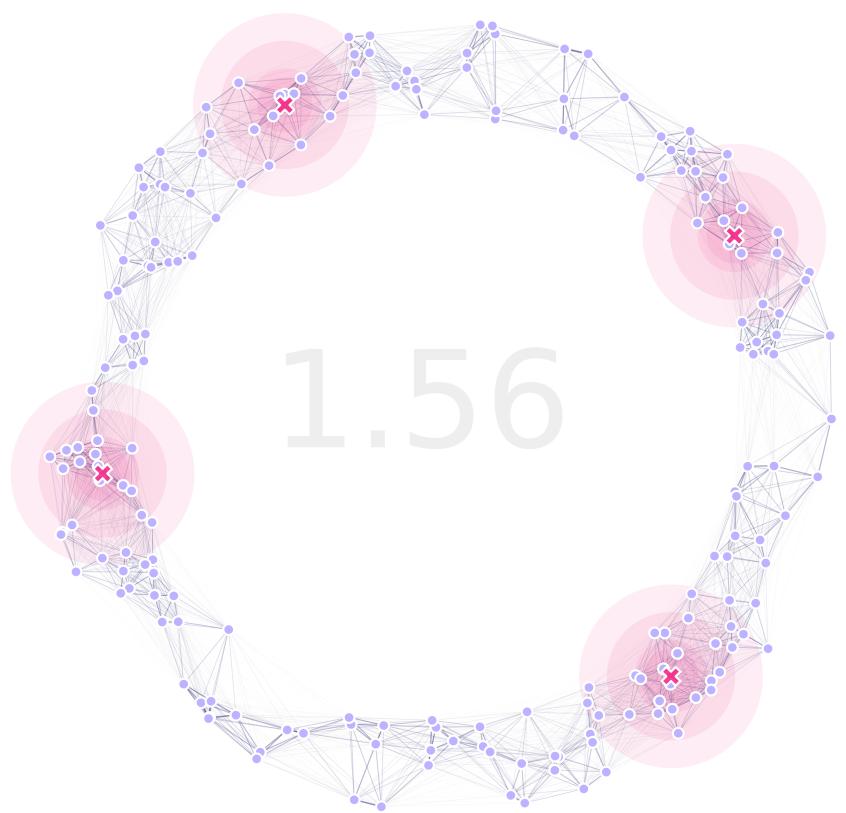


Figura 6.5: Mejor resultado GA

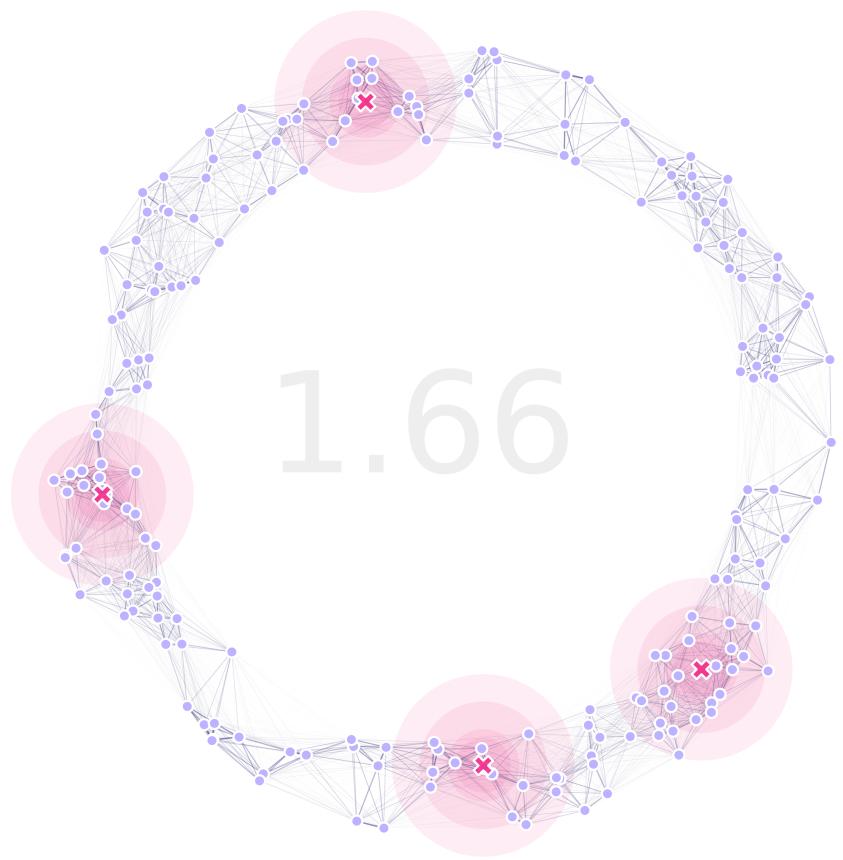


Figura 6.6: Peor resultado PSO

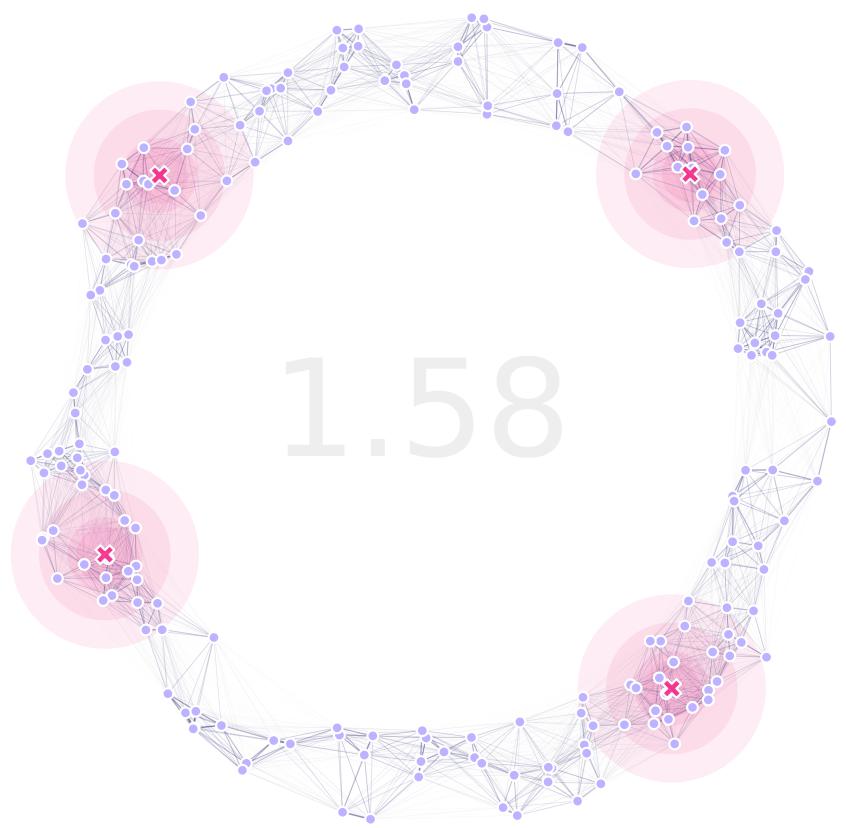


Figura 6.7: Peor resultado GA

Capítulo 7

Conclusiones y trabajo futuro

En este capítulo se presentan las conclusiones obtenidas de este trabajo. Se toman en cuenta los trabajos que componen el estado del arte, la solución propuesta y los resultados obtenidos de la evaluación de la efectividad del algoritmo.

7.1. Conclusiones

En la última década se ha visto un incremento en los artículos publicados en donde se aplican algoritmos de optimización a distintos campos. Sin embargo, las aplicaciones para el control y monitoreo de mosquito son mínimas. El trabajo más relevante es MGSurvE, pues es la única herramienta enfocada totalmente a la optimización en el ámbito de monitoreo de mosquitos.

La problemática presentada en este trabajo gira en torno a MGSurvE, pues su algoritmo de optimización, un GA, no está encontrando las mejores posiciones posibles para las trampas de mosquitos, y estos resultados son encontrados en etapas tardías de las búsquedas.

A partir de esto, se propuso la implementación de un algoritmo PSO para reemplazar al GA y mejorar el comportamiento general del paquete MGSurvE.

Para validar que el algoritmo propuesto cumplía con su propósito, se diseñó una metodología basada en la comparación directa de ambos algoritmos, el PSO y el GA.

A partir de las métricas generadas de la comparación, se puede concluir que el algoritmo PSO propuesto en este trabajo cumple su objetivo de mejorar los resultados arrojados por MGSurvE, pues genera mejores posiciones para las trampas la mayoría de las veces y lo hace en una menor cantidad de generaciones.

7.2. Contribuciones

Las contribuciones logradas gracias al algoritmo desarrollado en este trabajo son las siguientes:

- Se mejoraron los resultados arrojados por la optimización de trampas de mosquitos en el paquete MGSurvE.
- Se redujo el tiempo de optimización de trampas para mosquito en el paquete MGSurvE.

7.3. Trabajo futuro

Durante el desarrollo y evaluación del algoritmo PSO propuesto en este trabajo se encontraron algunas áreas de oportunidad para mejora. Estas oportunidades se deben a falta de recursos para el desarrollo. Las posibles mejoras futuras son las siguientes:

- Buscar solución al problema de mínimos locales presente en el algoritmo PSO.
- Graficar el percentil 50 o la media en la Figura 6.3 en lugar del promedio para mejorar la visualización de los datos.
- Realizar una comparación del comportamiento del PSO con diferentes cantidades de trampas a optimizar.

Bibliografía

- [1] S. Russell and P. Norving, “Artificial Intelligence: A modern approach”, 3rd ed. Prentice Hall, New Jersey: Pearson Education, 2009.
- [2] J. M. Spector and S. Ma, “Inquiry and critical thinking skills for the next generation: From Artificial Intelligence back to human intelligence,” Smart Learning Environments, vol. 6, no. 1, Sep. 2019.
- [3] E. K. Burke and G. Kendall, “Search methodologies: Introductory tutorials in optimization and decision support techniques”. New York, New York: Springer, 2014.
- [4] V. J. Rayward-Smith, “Modern Heuristic Search Methods”. Chichester, New York: Wiley, 1996.
- [5] F. Glover and M. Laguna, “Tabu Search”. Boston, Massachusetts: Kluwer Academic Publishers, 1997.
- [6] C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen, “Evolutionary algorithms for solving multi-objective problems,” Genetic and Evolutionary Computation Series, 2007.
- [7] A. S. Fraser, “Simulation of genetic systems by Automatic Digital Computers II. effects of linkage on rates of advance under selection,” Australian Journal of Biological Sciences, vol. 10, no. 4, p. 492, 1957.
- [8] J. Kennedy and R. Eberhart, “Particle swarm optimization,” Proceedings of ICNN’95 - International Conference on Neural Networks, 1995.
- [9] Y. Shi and R. Eberhart, “A modified particle swarm optimizer,” 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360), 1998.

- [10] S. Cheg, H. Lu, X. Lei, and Y. Shi, “A quarter century of particle swarm optimization,” Complex and Intelligent Systems, vol. 4, no. 3, pp. 227–239, 2018.
- [11] S. Li, C. Zhang, and J. Qu, “Location optimization of wireless sensor network in Intelligent Workshop based on the three-dimensional adaptive fruit fly optimization algorithm,” International Journal of Online Engineering (iJOE), vol. 14, no. 11, p. 202, Nov. 2018.
- [12] N. Hantash, T. Khatib, and M. Khammash, “An improved particle swarm optimization algorithm for optimal allocation of distributed generation units in Radial Power Systems,” Applied Computational Intelligence and Soft Computing, vol. 2020, pp. 1-8.
- [13] H. Liang and F. H. Kang, “Adaptive mutation particle swarm algorithm with dynamic nonlinear changed inertia weight,” Optik-International Journal for Light and Electron Optics, vol. 127, no. 19, pp. 8036-8042, 2016.
- [14] H. M. Sanchez, “MGSurvE’s documentation,” MGSurvE’s documentation! - MGSurvE documentation, 2021. [Online]. Available: <https://chipdelmal.github.io/MGSurvE/build/html/index.html>. [Accessed: 31-Mar-2022].
- [15] DEAP Project, “DEAP documentation,” DEAP documentation - DEAP 1.3.1 documentation, 2022. [Online]. Available: <https://deap.readthedocs.io/en/master/>. [Accessed: 01-Apr-2022].
- [16] A. P. Piotrowski and A. E. Piotrowska, “Differential Evolution and particle swarm optimization against covid-19,” Artificial Intelligence Review, Jul. 2021.
- [17] A. P. Piotrowski, M. J. Napiorkowski, J. J. Napiorkowski, and P. M. Rowlinski, “Swarm intelligence and Evolutionary Algorithms: Performance versus speed,” Information Sciences, vol. 384, pp. 34-85, Apr. 2017.
- [18] A. P. Piotrowski, J. J. Napiorkowski, and A. E. Piotrowska, “Population size in particle swarm optimization,” Swarm and Evolutionary Computation, vol. 58, May 2020.
- [19] The Marshall Lab at UC Berkeley. (s. f.). THE MARSHALL LAB. <https://www.marshalllab.com/>

- [20] R. Singer, “Shape up V 1.8, 2019 edition,” Shape Up: Stop Running in Circles and Ship Work that Matters, 2019. [Online]. Available: <https://basecamp.com/shapeup/webbook>. [Accessed: 03-May-2022].
- [21] “Pandas,” pandas, 02-Apr-2022. [Online]. Available: <https://pandas.pydata.org/>. [Accessed: 06-May-2022].
- [22] NumPy, 2022. [Online]. Available: <https://numpy.org/>. [Accessed: 06-May-2022].
- [23] “DEAP documentation,” DEAP documentation - DEAP 1.3.1 documentation, 22-Jan-2022. [Online]. Available: <https://deap.readthedocs.io/en/master/>. [Accessed: 06-May-2022].