EXAMEN FINAL

PAUL MURILLO*, KRITZAN SULLCA*

PROGRAMACIÓN CONCURRENTE

CONTENTS

1	Intro	oducción
2	Mar	co Teórico
	2.1	Algoritmos Genéticos
	2.2	Población
	2.3	Generaciones
	2.4	Crossover
	2.5	Mutación
	2.6	Fitness
3	Met	odología
	3.1	Definición de parámetros
	3.2	Diseño de la arquitectura
	3.3	Implementación del servidor
	3.4	Implementación de los trabajadores
	3.5	Prueba en un clúster
	3.6	Desarrollo del programa
4	Resi	altados y Discusiones

ABSTRACT

En este informe se presenta un estudio detallado sobre la implementación de algoritmos genéticos mediante concurrencia con hilos, paralelismo y distribución. Se define una serie de parámetros, como el tamaño de la población, el número de generaciones, el porcentaje de crossover, mutación y elite, y la función de fitness basada en la ecuación descrita en la el documento brindado por el docente. Se evalúa el tiempo de ejecución utilizando los parámetros iniciales y se realiza una comparación con los resultados esperados. El enfoque propuesto utiliza un servidor en LP1 (Java) y LP2 (Python) para enviar tareas a los trabajadores en LP1 (Java) y LP2 (Python), y se realiza una prueba en imagenes de docker que simulan clústers. Se describe en detalle la arquitectura diseñada, que incluye la interacción entre el servidor y los trabajadores, y se presenta un diagrama de protocolo que muestra las etapas del proceso. Además, se proporciona una explicación exhaustiva del desarrollo del programa. Los resultados obtenidos se presentan y se discuten, analizando el tiempo de ejecución y comparándolo con los resultados esperados, mientras que se exploran las ventajas y desventajas de utilizar concurrencia con hilos, paralelismo y distribución en la implementación de algoritmos genéticos.

^{*} Facultad de Ciencias, Universidad Nacional de Ingeniería, Lima, Perú

1 INTRODUCCIÓN

Los algoritmos genéticos son técnicas de optimización inspiradas en la evolución biológica que han demostrado su eficacia en la resolución de problemas complejos. En este artículo, se presenta un estudio exhaustivo sobre la implementación de algoritmos genéticos utilizando concurrencia con hilos, paralelismo y distribución, con el objetivo de mejorar la eficiencia y el tiempo de ejecución de estos algoritmos. El uso de concurrencia y paralelismo permite realizar cálculos simultáneos en diferentes hilos y procesos, mientras que la distribución permite aprovechar los recursos de un clúster para acelerar el proceso de optimización.

En este estudio, se definen varios parámetros cruciales para la implementación de los algoritmos genéticos, como el tamaño de la población, el número de generaciones, el porcentaje de crossover, mutación y elite. Además, se utiliza una función de fitness.

La implementación propuesta utiliza un enfoque cliente-servidor, donde un servidor en LP1 (Java) envía tareas a los trabajadores, que pueden estar implementados en LP1 (Java) o LP2 (Python). Se realiza una prueba en un clúster para evaluar el rendimiento y la escalabilidad del sistema.

Se proporciona una descripción detallada de los algoritmos genéticos, explicando conceptos fundamentales como la población, las generaciones, el crossover, la mutación y la función de fitness. Además, se discute la relevancia y aplicabilidad de los algoritmos genéticos en la resolución de problemas de optimización.

Además incluye el diseño de una arquitectura específica para la implementación de los algoritmos genéticos en un entorno concurrente y distribuido. Se describe la interacción entre el servidor y los trabajadores, y se presenta un diagrama de protocolo que muestra las etapas del proceso, desde el envío de tareas hasta la consolidación de los resultados.

Finalmente, se presentan los resultados obtenidos a través de la ejecución del programa implementado. Se analiza el tiempo de ejecución utilizando los parámetros iniciales y se compara con los resultados esperados. Además, se exploran las ventajas y desventajas de utilizar concurrencia con hilos, paralelismo y distribución en la implementación de algoritmos genéticos, y se discuten posibles mejoras y áreas de investigación futuras, y se incluye el código fuente utilizado en la implementación, en el anexo. Esto proporciona una visión completa del proceso de implementación y los resultados obtenidos.

2 MARCO TEÓRICO

En este apartado se presenta el marco teórico necesario para comprender los fundamentos y conceptos clave relacionados con la implementación de algoritmos genéticos mediante concurrencia con hilos, paralelismo y distribución. A continuación, se describen los principales elementos teóricos:

2.1 Algoritmos Genéticos

Los algoritmos genéticos son técnicas de optimización inspiradas en la evolución biológica. Estos algoritmos simulan la evolución de una población de posibles soluciones a través de mecanismos de selección, reproducción, crossover y mutación. A través de iteraciones sucesivas, los algoritmos genéticos convergen hacia soluciones óptimas o subóptimas del problema en cuestión.

2.2 Población

La población en un algoritmo genético consiste en un conjunto de individuos, donde cada individuo representa una posible solución al problema. La diversidad y el tamaño de la población son factores importantes que influyen en la calidad de las soluciones obtenidas.

2.3 Generaciones

Las generaciones en un algoritmo genético representan las etapas sucesivas de evolución de la población. En cada generación, se aplican los operadores genéticos (crossover y mutación) para crear una nueva generación de individuos, que reemplaza a la generación anterior.

2.4 Crossover

El operador de crossover implica la combinación de información genética de dos individuos seleccionados de la población para crear nuevos individuos. Este proceso imita la reproducción sexual y busca intercambiar y combinar características favorables de los padres para producir descendencia potencialmente mejorada.

2.5 Mutación

La mutación es un operador genético que introduce cambios aleatorios en los individuos de la población. Este proceso permite la exploración del espacio de búsqueda y evita que el algoritmo genético quede atrapado en óptimos locales.

2.6 Fitness

La función de fitness es una medida utilizada para evaluar la calidad de un individuo en términos de su capacidad para resolver el problema objetivo. Cuanto mayor sea el valor de fitness, mejor será la solución representada por el individuo.

El pseudocódigo para la implementación de algoritmos genéticos mediante concurrencia con hilos, paralelismo y distribución se presenta a continuación:

```
Función AlgoritmoGenético(P, G, Crossover, Mutación, Elite):
    // Inicialización de la población
    población = InicializarPoblación(P)
    // Evaluación inicial de la población
    EvaluarPoblación(población)

Para cada generación g en 1 hasta G:
    // Operadores genéticos
    nuevaGeneración = SeleccionarElite(población, Elite)

Mientras la nuevaGeneración no esté completa:
```

```
padre1 = SeleccionarPadre(población)
    padre2 = SeleccionarPadre(población)
    descendiente = AplicarCrossover(padre1, padre2, Crossover)
    AplicarMutación(descendiente, Mutación)
    nuevaGeneración.AgregarIndividuo(descendiente)

población = nuevaGeneración

// Evaluación de la población
    EvaluarPoblación(población)

// Devolver la mejor solución encontrada
    mejorSolución = ObtenerMejorSolución(población)

Devolver mejorSolución
```

En este pseudocódigo se muestra la estructura general de un algoritmo genético, incluyendo la inicialización de la población, los operadores genéticos (crossover y mutación), la evaluación de la población y la selección de la mejor solución encontrada. El pseudocódigo puede ser adaptado y complementado según las necesidades específicas de la implementación en un entorno concurrente y distribuido.

Este marco teórico proporciona los conceptos y fundamentos necesarios para comprender la implementación de algoritmos genéticos mediante concurrencia con hilos, paralelismo y distribución. Estos elementos teóricos serán aplicados en la metodología presentada en el siguiente apartado.

3 METODOLOGÍA

3.1 Definición de parámetros

En esta etapa, se realiza una cuidadosa definición de los parámetros necesarios para la implementación de los algoritmos genéticos. Se determina el tamaño de la población (P), el número de generaciones (G) y se ajusta el porcentaje de crossover, mutación y elite de acuerdo con las características del problema y los requisitos de optimización. Además, se especifica la función de fitness utilizada para evaluar la calidad de las soluciones generadas. Por ejemplo, se puede emplear la función de fitness basada en la ecuación:

$$1.23 \cdot \sin \frac{x^2 + y^2}{1.0563} \left(\left(\frac{0.4}{1 + 0.02 \cdot \left((x + 20)^2 + (y + 20)^2 \right)} \right) + \left(\frac{0.2}{1 + 0.05 \cdot \left((x + 5)^2 + (y + 25)^2 \right)} \right) + \left(\frac{0.7}{1 + 0.01 \cdot \left((x)^2 + (y - 30)^2 \right)} \right) + \left(\frac{1}{1 + 2 \cdot \left((x + 30)^2 + (y)^2 \right)} \right) + \left(\frac{0.05}{1 + 0.1 \cdot \left((x - 30)^2 + (y + 30)^2 \right)} \right)$$

para evaluar la calidad de las soluciones generadas por los algoritmos genéticos.

donde x e y representan las variables a optimizar. La elección adecuada de estos parámetros es fundamental para lograr un equilibrio entre la exploración y explotación del espacio de búsqueda.

3.2 Diseño de la arquitectura

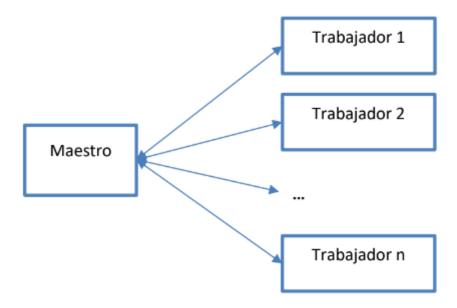
En esta fase, se diseña una arquitectura cliente-servidor que permita la implementación de los algoritmos genéticos en un entorno concurrente y distribuido. El servidor, implementado en LP1 (Java), actúa como punto central de control y gestiona la comunicación entre los clientes y los trabajadores. Se establecen canales de comunicación utilizando sockets u otros mecanismos adecuados para el intercambio de información entre las diferentes partes. Además, se considera la escalabilidad del sistema, permitiendo que nuevos trabajadores se unan al proceso de optimización de manera dinámica.

3.3 Implementación del servidor

En esta etapa, se desarrolla el servidor en LP1 (Java) con el objetivo de gestionar las tareas de los algoritmos genéticos. El servidor recibe las tareas generadas por los clientes, distribuye estas tareas entre los clientes y recopila los resultados obtenidos. Se emplea la concurrencia con hilos para gestionar múltiples solicitudes de tareas simultáneamente y asegurar una adecuada utilización de los recursos del sistema. La implementación del servidor se realiza de manera modular y escalable, permitiendo la expansión de la capacidad de procesamiento de acuerdo con las necesidades.

3.4 Implementación de los trabajadores

En esta etapa, se implementan los trabajadores, los cuales pueden estar desarrollados en LP1 (Java) o LP2 (Python). Los trabajadores son responsables de ejecutar los algoritmos genéticos asignados por el servidor. Utilizan la concurrencia y el paralelismo para ejecutar los cálculos en múltiples hilos o procesos, aprovechando al máximo los recursos disponibles. Cada trabajador recibe una tarea específica del servidor, procesa los cálculos necesarios y devuelve los resultados al servidor. Se aplican técnicas de programación concurrente y paralela para optimizar el rendimiento y la eficiencia de los cálculos realizados por los trabajadores.



3.5 Prueba en un clúster

Para evaluar el rendimiento y la escalabilidad del sistema implementado, se realiza una prueba en un clúster. El clúster consta de un conjunto de máquinas interconectadas que se utilizan para distribuir y acelerar el proceso de optimización. Se analiza el rendimiento en función del tamaño del clúster y la carga de trabajo asignada a cada máquina. Se recopilan métricas de rendimiento, como el tiempo de ejecución y la eficiencia del sistema, para evaluar la capacidad de escalabilidad y la utilización efectiva de los recursos del clúster.

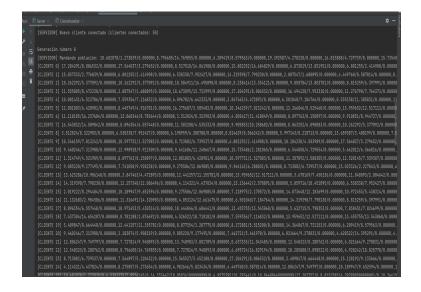
3.6 Desarrollo del programa

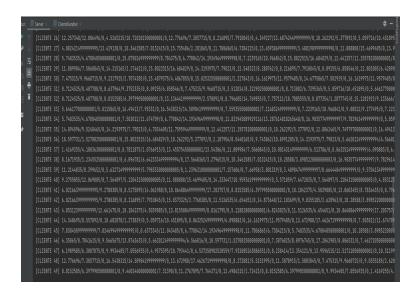
En esta etapa, se explica en detalle el desarrollo del programa basándose en las explicaciones y el código proporcionado en clase. Se describen los pasos necesarios para implementar la lógica de los algoritmos genéticos utilizando las capacidades de concurrencia, paralelismo y distribución según lo definido en la arquitectura diseñada. Se pueden utilizar librerías específicas, como la Java Concurrency API o herramientas de programación paralela en Python, para facilitar la implementación y la gestión de la concurrencia y el paralelismo en el programa. Se realizan pruebas exhaustivas y se ajustan los parámetros y la arquitectura según sea necesario para optimizar el rendimiento y la eficacia del sistema.

4 RESULTADOS Y DISCUSIONES

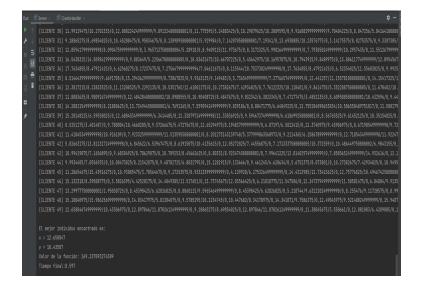
La implementación de algoritmos genéticos mediante concurrencia con hilos, paralelismo y distribución representa un enfoque prometedor para mejorar la eficiencia y el tiempo de ejecución de estos algoritmos. La metodología presentada en este artículo brinda una guía detallada para implementar de manera efectiva estos algoritmos en entornos concurrentes y distribuidos. Sin embargo, se debe tener en cuenta que cada problema puede requerir ajustes y optimizaciones específicas, por lo que se recomienda un análisis exhaustivo y una adaptación adecuada para obtener los mejores resultados en cada caso.

• Para 50 clientes, 25 en java y 25 en python con dos generaciones se obtiene lo siguiente

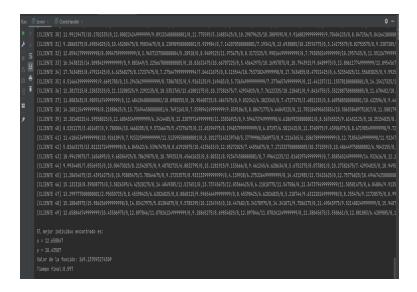


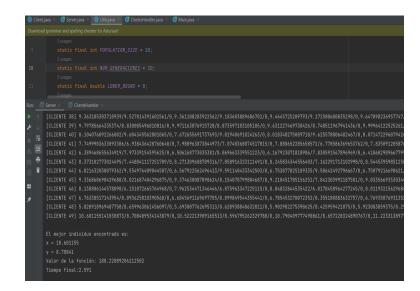


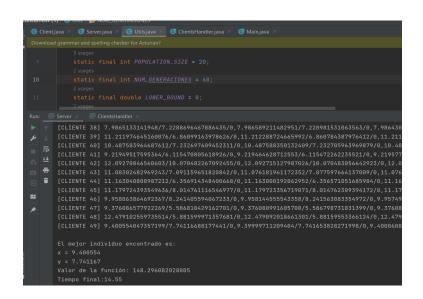
| Colore | 1.0 | Colo



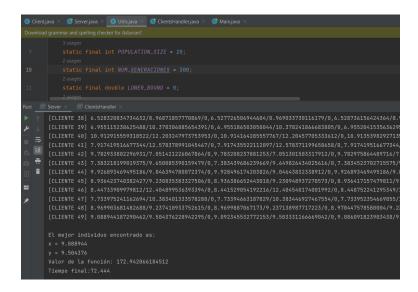
Ahora tomaremos diferentes NumGeneraciones = 2,10,30,60,100,150,300 y mediremos los tiempos







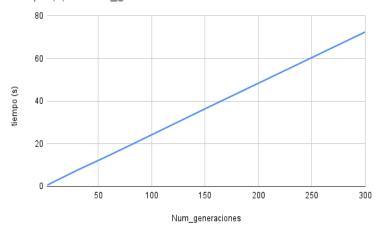
```
Clientyava × Serverjava × Cultisjava × Culti
```



• Tabla que recolecta NumGeneraciones y su tiempo

Num_generaciones	tiempo (s)
2	0.597
10	2.591
30	7.563
60	14.55
100	24.229
150	36.416
300	72.444

tiempo (s) vs Num_generaciones



REFERENCES

[1] Códigos y teoría de las clases. Nuñez, Y. [2] El codigo se puede encontar en Enlace a GitHub