

Aplicación de Drones y Aprendizaje Reforzado en el Sistema de Envíos: Un Enfoque para Mejorar la Eficiencia en Rutas en Cali.

Juan Diego Zapateiro Carabali
zapateirojuandiego@gmail.com
Diego Cabezas

Universidad Santiago de Cali, Facultad de Ingeniería, Programa de Tecnología en Sistemas de Información

Resumen

En este trabajo se aborda la aplicación del aprendizaje reforzado en el trazado de ruta de un dron, para su desarrollo en proyectos de entrega (comúnmente conocidos como "delivery"), haciendo comparaciones con distintos métodos de trazado fijo actuales y cómo esto podría mejorar la situación de vuelo de dichos aparatos, proponiendo como escenario de uso la ciudad de Cali. Se aprecia un hallazgo significativo con el desarrollo de este proyecto, el cual es un mejoramiento significativo en el aspecto de lograr una identificación de manera autónoma de diferentes obstáculos por parte de la aplicación del código desarrollado en entornos de aprendizaje reforzado, conllevando a que la maquina en cuestión, bajo un entorno virtual, logre determinar la ruta óptima para llegar al destino establecido, sorteando los obstáculos para poder lograrlo. Teniendo en cuenta que la ciudad de Cali presenta una gran cantidad de obstáculos, los cuales deberían sortearse al momento de la ejecución del vuelo, también se pudo apreciar en la metodología planteada y los resultados obtenidos, que el aprendizaje reforzado para su futura aplicación en el tema de entregas mediante el uso de drones, puede constituirse en una solución acertada para optimizar los tiempos de entrega de diferentes elementos, y solucionar problemas como la entrega de distintos insumos o materiales en lugares de difícil acceso.

Palabras Clave: Drones, Delivery, Aprendizaje reforzado, UAV's, Navegación Autónoma.

Abstract

This work addresses the application of reinforcement learning in the route planning of a drone, for its implementation in delivery projects, commonly known as "delivery." It compares different current fixed-path methods and how this could enhance the flight performance of these devices, proposing the city of Cali as a use case scenario. A significant finding is observed through the development of this project, which results in a notable improvement in autonomously identifying various obstacles using the code developed within reinforcement learning environments. This advancement enables the machine, within a virtual environment, to determine the optimal route to reach the designated destination while navigating around obstacles.

Considering that the city of Cali presents numerous obstacles that must be navigated during flight execution, the methodology and obtained results also highlight that reinforcement learning, when applied to delivery using drones, could be a suitable solution to optimize delivery times for various items. It can address challenges such as delivering different supplies or materials to hard-to-reach locations.

Keywords: Drones, Delivery, Reinforcement Learning, UAVs, Autonomous Navigation.

Keywords:

1. INTRODUCCIÓN

Con este proyecto, se pretende resaltar la importancia de implementar el aprendizaje reforzado en los drones de envío, ya que podría optimizar su funcionamiento y hacerlo más eficiente. Se ha demostrado su relevancia, por ejemplo, en zonas afectadas por desastres naturales, donde abundan obstáculos ambientales (León, Cruz,

Del Rosario, Zambrano Solorsano y Elba, 2022). En este proyecto, los drones funcionarán de tal manera que puedan aprender por sí mismos cómo desplazarse según el conocimiento adquirido en la ruta. Esto se logrará mediante el aprendizaje reforzado, que esencialmente utilizará un plano cartesiano como punto de referencia. El dron se desplazará desde el punto (0,0) hasta el punto (3,3), y se registrará la cantidad de intentos necesarios para encontrar y seleccionar la ruta óptima hasta el objetivo.

En la actualidad, para los vuelos autónomos con drones, se establece una ruta fija con un objetivo o punto de destino. Sin embargo, esta ruta puede ser modificada manualmente según los obstáculos que surjan durante el vuelo. Esto implica que el operador debe estar alerta a posibles cambios para evitar obstáculos. El aprendizaje reforzado se perfila como una solución viable para automatizar la evasión de obstáculos que puedan surgir durante el vuelo, un problema especialmente significativo en la ciudad de Cali. Esta aproximación permite determinar la mejor ruta para llegar al objetivo final.

El aprendizaje de automatización de máquinas tiene como finalidad utilizar la prueba y el error para lograr que la máquina se entrene y aprenda la manera más útil de superar un obstáculo (Kaelbling, Littman y Moore, 1996). En este caso, se desarrolló un código basado en el aprendizaje reforzado para que una máquina, en este caso un dron, pueda llegar a un punto o meta predeterminada. Así, si se parte de una matriz donde la posición inicial es (0,0), la máquina debería ser capaz de alcanzar la posición (3,3), que es la meta (recompensa). Antes de alcanzar esta meta, el código registra la cantidad de intentos necesarios y la ruta más óptima para lograrlo.

En principio, se concluye que este proyecto destaca por proponer un método que podría ser de gran importancia en la actual situación de Cali. Este método permitiría, en el futuro, llevar a cabo desplazamientos y entregas de distintas mercancías mediante desplazamientos aéreos principalmente por empresas, como Amazon. Dado que la ciudad de Cali presenta una gran cantidad de objetos que podrían considerarse obstáculos en el entorno, como antenas, vegetación y edificios, esta solución busca facilitar el desplazamiento de vehículos aéreos no tripulados. Para ello, se utiliza el aprendizaje reforzado, de manera que estos dispositivos puedan determinar de manera más óptima el camino a seguir a través de una ruta predefinida.

Los drones, más comúnmente conocidos como dispositivos no tripulados, son vehículos sin tripulación capaces de mantener una ruta controlada y previamente definida de manera autónoma o manual (Scott y Scott, 2017). Empresas como Amazon han determinado que, actualmente, para el servicio de entregas (delivery), se utilizan drones MK27-2, y en el futuro se prevé la incorporación de los MK3. Estos últimos modelos, presentados hace pocos meses, se espera que estén operativos en 2024. Los MK30 prometen, entre otras mejoras, una mayor capacidad de carga, una detección de obstáculos aún más efectiva, mayor velocidad y menor nivel de ruido (Rivera, 2022).

Los avances tecnológicos actuales en el campo de los drones han permitido su aplicación en diversas áreas. Por ejemplo, en el ámbito de la mensajería, los drones han demostrado su capacidad para realizar entregas de manera rápida y eficiente, cuidando la integridad de los paquetes. Además, esta tecnología podría ser aplicada en el transporte de medicamentos y paquetería a larga distancia, siempre y cuando se realicen las adaptaciones necesarias en el diseño de los drones.

2. MATERIALES Y MÉTODOS/METODOLOGÍA

La metodología usada en este proyecto se presenta a continuación:

2.1. Recopilación de la información

Se llevó a cabo una investigación exhaustiva sobre el lenguaje de aprendizaje reforzado, con el propósito de incorporar este concepto en el código desarrollado. Se consideraron los drones empleados en la actualidad para los sistemas de entregas, y se evaluó el dron que resultaría más adecuado para su implementación fuera de la simulación (MK24-2).

2.2. Estructura del código mediante aprendizaje reforzado

Se empleó el lenguaje de programación Python en conjunto con su librería numpy. Esta librería permitió que el código, basado en la matriz propuesta, pudiese determinar la ruta más eficiente para el desplazamiento y lograr completar el recorrido planteado.

2.3. Simulación del código mediante entornos virtuales

Una vez completada la programación, procedimos a realizar pruebas exhaustivas para obtener un registro de todas las rutas trazadas por el dron, así como el recuento de intentos realizados en cada una de ellas. Este análisis permitió evaluar, en función del aprendizaje del sistema, cuál fue la ruta más eficiente seleccionada para el trayecto en cuestión. Estas pruebas se llevaron a cabo en un entorno virtual simulado a través de una matriz que representa el contexto de desplazamiento del dron. Esta matriz incluyó puntos de inicio y finalización, los cuales el dron debía alcanzar para completar su recorrido.

3. RESULTADOS Y DISCUSIÓN

Se puede observar un código elaborado mediante aprendizaje reforzado. En este enfoque, el sistema utiliza una matriz como su método de aprendizaje. La matriz guiará las decisiones del sistema para determinar la ruta más eficiente con el propósito de alcanzar su destino. Es importante tener en cuenta que el punto de partida es 0,0 y la meta es llegar a la posición 3,3, lo cual se ilustra claramente en la figura 1.

Figura 1. Representación de la matriz.



Fuente: Fuente propia.

La figura 2 muestra claramente cómo se ha configurado el entorno del código para que el dron disponga de una ruta definida por medio de una matriz. En este enfoque, el dron deberá avanzar de celda en celda dentro de la matriz hasta alcanzar su objetivo, que se encuentra en la posición 3,3. Una vez que el dron logre posicionarse en esta ubicación, se considerará que ha completado su recorrido. A través de este proceso, se identificará la ruta más óptima mediante los intentos realizados por el dron en su trayectoria.

Figura 2. Definición de las variables y la matriz en el código desarrollado (definición del entorno).

```

1 [1]: # Importamos Las Librerías necesarias
import pandas as pd
import numpy as np

# Formato de Los decimales en Pandas y La semilla del Random
pd.options.display.float_format = '{:,.2f}'.format
np.random.seed(23)

class Environment(object):
    def __init__(self, action_penalty=-1.0):
        """
        Clase que representa y controla en entorno
        :param step_penalty: Factor de descuento del Reward por acción tomada
        """
        self.actions = {'Arriba': [-1, 0],
                        'Abajo': [1, 0],
                        'Izquierda': [0, -1],
                        'Derecha': [0, 1]}
        self.rewards = [[0.0, 0.0, 0.0, 0.0],
                        [0.0, 0.0, 0.0, 0.0],
                        [0.0, 0.0, 0.0, 0.0],
                        [0.0, 0.0, -100.0, 100.0]]

        self.action_penalty = action_penalty # Penalización por cada paso dado
        self.state = [0, 0] # Estado en el que se encuentra el agente
        self.final_state = [3, 3] # Estado final del entorno. Cuando el agente llega, se termina el episodio
        self.total_reward = 0.0 # Contador de recompensas en el episodio
        self.actions_done = [] # Lista en la que se guardan Los pasos (acciones) realizadas en cada episodio

    def reset(self):

```

Fuente: Fuente propia

En la imagen siguiente (figura 3), se puede observar la aplicación del proceso de entrenamiento de la máquina en este proyecto. La imagen muestra el recuento de intentos que la máquina necesitó para alcanzar el punto objetivo (recompensa), así como el número de movimientos que realizó hasta llegar a su posición final. Esta posición se encuentra determinada por la matriz establecida en el código previamente diseñado.

Figura 3. Resultado de la ejecución del código.

```

episodes_list, best_episode = run_agent(learner=Learner,
                                       verbose=True)

print_process_info(episodes_list=episodes_list,
                  best_episode=best_episode)

EPISODIO 1 - Numero de acciones: 264 - Reward: -864.0
EPISODIO 2 - Numero de acciones: 84 - Reward: -184.0
EPISODIO 3 - Numero de acciones: 16 - Reward: -216.0
EPISODIO 4 - Numero de acciones: 42 - Reward: 58.0
EPISODIO 5 - Numero de acciones: 44 - Reward: -44.0
EPISODIO 6 - Numero de acciones: 36 - Reward: 64.0
EPISODIO 7 - Numero de acciones: 29 - Reward: 71.0
EPISODIO 8 - Numero de acciones: 11 - Reward: -11.0
EPISODIO 9 - Numero de acciones: 73 - Reward: -373.0
EPISODIO 10 - Numero de acciones: 37 - Reward: -137.0

MEJOR (ÚLTIMO) EPISODIO:
EPISODIO 7
    Numero de acciones: 29
    Reward: 71.0

```

Fuente: Fuente propia.

En la imagen previa, se muestra el resultado de la ejecución del código, reflejando el número de intentos realizados (en este caso, 10 intentos, ya que en situaciones reales, el dron no tendría mil oportunidades debido a las restricciones temporales de las entregas). Asimismo, se exhibe el número de acciones emprendidas en cada intento, que representan los desplazamientos efectuados en la matriz con el fin de llegar al punto específico. Por último, se presenta el "reward" (recompensa u objetivo), que se suma o resta según si el dron alcanza o toca obstáculos en el proceso. Es importante tener en cuenta que el contacto con un obstáculo, representado por los bordes o espacios fuera de la matriz, es considerado como una elección incorrecta, ya que en un entorno real podría causar daño al dron. Por esta razón, la matriz simula las posibles rutas alrededor de los obstáculos, y la puntuación final se muestra al término del proceso.

En comparación con el aprendizaje por coincidencia (matching learning), el sistema de drones con aprendizaje reforzado resulta más eficiente. Esto se debe a que no es necesario ingresar todas las rutas posibles de antemano para que el dron aprenda a moverse. En su lugar, el sistema comprende las dificultades y, a medida que comete errores, aprende a tomar las acciones más adecuadas en tiempo real.

4. CONCLUSIONES

El empleo del aprendizaje reforzado capacita a los drones para aprender y adaptarse de manera autónoma. Esta capacidad les permite reconocer obstáculos, tomar decisiones en tiempo real y elegir las rutas más eficientes para alcanzar su destino. Esto es especialmente valioso en entornos urbanos que presentan numerosos obstáculos que pueden cambiar o surgir durante el vuelo.

En contraste con las rutas fijas preprogramadas, el aprendizaje reforzado confiere flexibilidad y adaptabilidad a la elección de la mejor ruta para el dron. Esto asegura entregas más seguras y eficientes. Los drones son capaces de aprender de su experiencia, recibir retroalimentación y recompensas basadas en sus acciones, lo que les permite mejorar de manera constante su rendimiento y tomar decisiones más informadas en futuros vuelos.

En relación a las decisiones tomadas por el dron al ejecutar el código, se observa que se generan diversas rutas posibles. Estas rutas se presentan al final de la ejecución del código, permitiendo identificar la más óptima. La determinación de la ruta óptima no solo se basa en la cantidad mínima de intentos requeridos para alcanzar el objetivo (reward), sino también en consideración de las preferencias del usuario. Esto posibilita establecer rutas alternativas que se ajusten a diferentes factores de aplicación, como se ilustra en la Figura 3: Resultado de la ejecución del código.

REFERENCIAS BIBLIOGRÁFICAS

- Ambrojo, J. C. (2013). *Los drones 'se alistan' al servicio civil*. REPORTAJE.
- Botija, F. G. (2021). *DRONES Y CARGA: PRESENTE Y FUTURO*. CONGRESO CET PONENCIA DRONES.
- D. E. Moriarty, A. C. Schultz, & J. J. Grefenstette. (1999). *Evolutionary Algorithms for Reinforcement Learning*. AI Access Foundation, Inc.
- Gabriel, D.-A., Mankowitz, D., & Todd, H. (2019). *Challenges of Real-World Reinforcement Learning*. Cornell University.
- Hassanalian, M., & Abdelkefi, A. (2017). Classifications, applications, and design challenges of drones: A review. *Progress in Aerospace Sciences*, 99-131.
- Jun Morimoto, & Kenji Doya. (2005). *Robust Reinforcement Learning*. MIT Press.
- Kaelbling, L., Littman, M., & Moore, A. (1996). *Reinforcement Learning: A Survey*. JAIR.
- Leon, M., Cruz, F., Del Rosario, M., Zambrano Solorsano, & Elba, T. (2022). *Revisión de estudios sobre sistemas FANET y drones para emergencias o desastres naturales*. Dialnet.
- Lisette Sandra Toledo, L. M. (2010). *Proyectos de aprendizaje de servicio-comunitario y su influencia en las conductas prosociales de estudiantes universitarios*. Revista de Pedagogía.
- Machicao, J. (2020). *El aprendizaje reforzado: un punto de encuentro entre las políticas públicas y la inteligencia artificial*. Content uploaded by Jose Machicao.
- Marín Meza, A. F. (2021). *Desarrollo de un servicio logístico de entrega de medicamentos urgentes de peso liviano con drones en la ciudad de Barranquilla*. Corporación Universidad de la Costa.
- Matthew Ayamga a, S. A. (2021). *Multifaceted applicability of drones: A review*. EL SEVIER.
- Mena Pruneda, A. (2019). *Instalación fotovoltaica como punto de recarga para el reparto de mercancías a través de drones en zonas rurales*. Atribución-NoComercial-SinDerivadas 3.0 España.
- Molina Garcia, A., Espinoza Matapuncho, L., & Figueroa Vargas, A. (2021). *Proyecto de prefactibilidad para creación de microempresa de abastecimiento de alimentos de primera necesidad que brinde servicio con drones de carga a domicilio, cantón El Triunfo*. UNIVERSIDAD SAN GREGORIO DE PORTOVIEJO.
- Orellana Granja, A. S. (2020). *Plan de negocios para la creación de una empresa de servicio de mensajería mediante la utilización de drones en la ciudad de Quito*. Quito: Universidad de las Américas, 2020.
- Rivera, N. (2022). *Hola, Prime Air: Amazon empieza a entregar paquetes con drones*. hipertextual.
- Scott, J., & Scott, C. (2017). Drone Delivery Models for Healthcare. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 3297-3304.

Sewak, M. (2019). *Deep Reinforcement Learning*. SpringerLink.

Suárez, J., López Morales, V., & Juárez Vargas, C. (2019). *Conocimiento" en un Sistema de Drones Levantando una Carga*. Padi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI.

Tovar Ospitia, A. (2020). *Estructura para un dron de carga elevada*. Atribución-NoComercial-SinDerivadas 3.0 España.

Yash Chandak, Georgios Theocharous, James Kostas, Scott Jordan, & Philip Thomas. (2019). *Learning Action Representations for Reinforcement Learning*. Proceedings of the 36th International Conference on Machine Learning,.