

Modelo algorítmico para alta disponibilidad en transporte de volúmenes crecientes de tráfico variable en redes ópticas

Diego Fernando Aguirre Moreno

Universidad Distrital Francisco José de Caldas

Facultad de Ingeniería

Bogotá, Colombia

2022

Modelo algorítmico para alta disponibilidad en transporte de volúmenes crecientes de tráfico variable en redes ópticas

Diego Fernando Aguirre Moreno

Tesis presentada como requisito parcial para optar al título de:

Doctor en Ingeniería

Director:

Ing. Julio Barón Velandia Ph.D

Co director:

Ing. Octavio Salcedo Parra Ph.D

Grupo de Investigación Interoperabilidad Tecnológica y Semántica

Universidad Distrital Francisco José de Caldas

Facultad de Ingeniería

Bogotá, Colombia

2022

Comisión de Doctorado

Esta tesis, titulada "Modelo algorítmico para alta disponibilidad en transporte de volúmenes crecientes de tráfico variable en redes ópticas", escrita por Diego Fernando Aguirre Moreno, ha sido aprobada en cuanto a estilo y contenido intelectual.

Hemos leído esta tesis y la aprobamos,

_____

Ph. D. Dr. Wilfrido Alejandro Moreno – Universidad del Sur de la Florida – USA.
Jurado 1

_____

Ph. D. Evelio Astaiza Hoyos – Universidad del Quindío.
Jurado 2

_____

Ph. D. Gustavo Adolfo Puerto Leguizamón – Docente de Planta Universidad Distrital.
Jurado 3

_____

Ph. D. Julio Barón Velandia
Director

Fecha de la defensa: 28/10/2022

Resumen

Internet ha presentado un crecimiento exponencial en la última década, generando que los usuarios demanden solución a los requisitos de forma instantánea. Estas exigencias constituyen la problemática principal que se encuentra relacionada con el desempeño y la caracterización de la infraestructura de red de acceso y agregación. Entre tecnologías prometedoras se encuentran las redes ópticas, sus características en capacidad, calidad de servicio y desempeño, permiten soportar tráfico generado por aplicaciones y tecnologías futuras como vídeo en alta definición, comunicaciones móviles de quinta generación (5G) y transmisiones en ultra definición.

La implementación de redes totalmente ópticas implica: madurez de dispositivos, buffers de almacenamiento óptico, conmutación de paquetes ópticos, desarrollo de métodos efectivos de administración de red y disminución de costos elevados. Estas situaciones han retrasado el desarrollo de las redes totalmente ópticas, impulsando investigaciones en redes ópticas elásticas (EON), que permiten ajustar dinámicamente sus recursos de acuerdo con los requisitos de cada demanda.

Esta investigación aborda los beneficios en el uso de las redes ópticas elásticas como soporte en el transporte de los crecientes volúmenes de tráfico variable, apoyado en técnicas de aprendizaje automático para afrontar problemas de: enrutamiento, asignación de espectro, formato de modulación y selección de núcleo. Aprovechando una perspectiva de infraestructura tecnológica que asegure los procesos de conmutación y enrutamiento completamente ópticos, para optimizar los recursos en función de las demandas.

Palabras clave: Aprendizaje automático, Aprendizaje por refuerzo, Redes ópticas elásticas, Tráfico dinámico.

Abstract

The Internet has shown exponential growth in the last decade, generating that users demand solutions to the requirements instantly. These requirements constitute the main problem related to the performance and characterization of the access and aggregation network infrastructure. Among the promising technologies are optical networks, their characteristics in capacity, quality of service and performance allow to support the traffic generated by future applications and technologies such as high-definition video, 5G networks and ultra-definition transmissions.

Today, the implementation of all-optical networks presents several challenges such as: device maturity, optical storage buffers, optical packet switching, lack of effective network administration and management methods, and high costs. These limitations have delayed the development of fully optical networks, promoting research into elastic optical networks (EON), which dynamically adjust their resources according to the requirements of each demand.

This research addresses the benefits in the use of elastic optical networks as support in the transport of the increasing volume of traffic, supported with machine learning techniques to face the problems of routing, spectrum allocation, modulation format and core selection. Providing a technological vision for the exploitation of the infrastructure that ensures completely optical switching and routing processes, optimizing resources based on demands.

Keywords: Dynamic traffic, Elastic optical network, Machine learning, Reinforcement learning

Dedicatoria

A mis padres por haberme forjado como la

persona que soy y a Dios Todopoderoso

Soberano del universo

Agradecimientos

Quiero agradecer a mi familia por el apoyo incondicional que siempre me han brindado en todos los proyectos y metas que me he trazado.

Le agradezco a los profesores Julio Barón y Octavio Salcedo por su dedicación y paciencia, sin sus palabras y correcciones precisas no hubiese alcanzado esta instancia.

Agradezco al equipo del doctorado y los docentes por su trabajo y gestión, sin lo cual no estarían las bases ni las condiciones para aprender estos conocimientos.

Contenido

Resumen IV

Abstract V

Introducción 1

I. Descripción del Proyecto 3

1. Formulación del proyecto 3

1.1. Formulación del problema 3

1.2. Interrogante 4

1.3. Justificación 4

1.4. Hipótesis 6

1.5. Grupo y línea de investigación 6

2. Objetivo general y objetivos específicos 7

2.1. Objetivo General 7

2.2. Objetivos Específicos 7

3. Antecedentes 8

II. Marco teórico 11

4. Aprendizaje por Refuerzo 11

4.1. Introducción 11

4.2. Aprendizaje por refuerzo 11

4.3. Q-Learning 12

4.4. Deep Q-Network (DQN) 13
4.5. Extensiones Deep Q-Network 14
4.5.1. Doble DQN (DDQN) 14
4.5.2. Repetición de experiencia priorizada 15
4.5.3. Duelo DQN (Dueling DQN) 15
4.6. Aprendizaje por refuerzo multi agente (MARL) 16
5. Redes Ópticas 18
5.1. Introducción 18
5.2. Generaciones de Redes Ópticas 18
5.3. Conmutación de paquetes ópticos 20
5.4. Conmutación de ráfagas ópticas 21
5.5. Multiplexación por división de frecuencias ortogonales 21
5.6. Multiplexación por división espacial 22
5.7. Tráfico dinámico 23
5.8. Redes ópticas elásticas 24
5.8.1. Segmentación del ancho de banda 24
5.8.2. Agregación de ancho de banda 25
5.8.3. Acomodación eficiente de múltiples velocidades de datos 25
5.8.4. Tasa de línea adaptable al alcance 26
III. Desarrollo 27
6. Modelo Empleado 27
6.1. Problema de asignación de enrutamiento, modulación, espectro y núcleo 27
6.2. Red óptica elástica 28
6.2.1. Unidades de ranuras de frecuencia 29

6.2.2. Fibra multinúcleo 30

6.2.3. Fragmentación 31

6.2.4. Requerimientos de espectro en términos de FSU vs máxima distancia de alcance 32

6.2.5. Topología 34

6.3. Esquema aprendizaje por refuerzo 36

6.3.1. Definiciones de representación 37

6.3.2. Interacción con el entorno 38

7. Propuesta algorítmica 39

7.1. Introducción 39

7.2. Herramientas 39

I. Python 39

II. Visual Studio Code 40

7.3. Entorno 40

7.3.1. Enlaces de fibra 40

7.3.2. Cantidad de núcleos de las fibras troncales 42

7.3.3. Modulación 43

7.3.4. Ancho de banda máximo soportado por las tarjetas de línea 43

7.3.5. Restricciones del entorno según retardo, pérdida de paquetes, probabilidad de bloqueo.43

7.4. DQN RMSCA 44

7.4.1. Agentes 47

7.5. Mecanismo de Entrenamiento 49

8. Resultados experimentales 51

8.1. EON Deep Q Network 52

8.2. EON MARL 54
9. Análisis de resultados 60
Conclusiones, aportes originales, trabajos futuros y socialización de resultados 62
10.1. Conclusiones 62
10.2. Aportes originales 63
10.3. Trabajos futuros 63
10.4. Socialización de resultados 64
Referencias bibliográficas 65

Lista de Figuras

Figura 4-1 Interacción aprendizaje por refuerzo. 12

Figura 4-2 Interacción aprendizaje por refuerzo multi agente. 16

Figura 5-1 Generaciones de redes ópticas. 19

Figura 5-2 Comparación de uso del espectro en DWDM y OFDM.. 22

Figura 5-3 ODFM: a) Dominio de frecuencia, b) Dominio de tiempo..... 23

Figura 6-1 Ejemplo asignación espectral..... 30

Figura 6-2 Fibra de 7 núcleos. 31

Figura 6-3 Fragmentación. 32

Figura 6-4 Topología. 35

Figura 7-1 Entorno. 41

Figura 7-2 Ejemplo de diccionario de enlaces..... 41

Figura 7-3 Ejemplo códigos de enlaces.. 42

Figura 7-4 Ejemplo diccionario de núcleos. 42

Figura 7-5 Ejemplo diccionario de modulación. 43

Figura 7-6 Diagrama de entradas y salidas. 45

Figura 7-7 Algoritmo RMSCA MARL 46

Figura 7-8 Interacción entre el agente y el entorno. 48

Figura 7-9 Seleccionar acción. 48
Figura 7-10 Transición. 49
Figura 7-11 Repetición de experiencia. 49
Figura 7-12 Bucle de entrenamiento.. 50
Figura 7-13 Configuración de hiperparámetros. 50
Figura 8-1 Topología NSFNet. 51
Figura 8-2 Uso de espectro. 53
Figura 8-3 Trafico ofrecido vs caminos ópticos. 53
Figura 8-4 Tráfico ofrecido versus Total tráfico Bloqueado. 54
Figura 8-5 Blocking probability 3 64. 56
Figura 8-6 Episodios de éxito vs Total de Episodios. 56
Figura 8-7 Duration vs Episode. 57
Figura 8-8 Blocking probability 5 128. 57
Figura 8-9 Recompensa de los agentes. 58
Figura 8-10 Multiagente uso de FSU en cada Test. 59

Lista de Ecuaciones

Ecuación 1 Actualización Función Q. 12

Ecuación 2 Pérdida DQ. 13

Ecuación 3 Pérdida DDQN..... 14

Ecuación 4 Error TD..... 15

Ecuación 5 Módulo de agregación. 15

Ecuación 6 Aproximación de BW total. 24

Ecuación 7 Modulación multiportadora. 28

Ecuación 8 Símbolo de información. 28

Ecuación 9 Correlación. 29

Ecuación 10 Diafonía sufrida por el f-ésimo FS. 31

Ecuación 11 SNR. 33

Ecuación 12 Función de valor de acción óptima. 36

Lista de Tablas

Tabla 5.1 Resumen de distintos formatos de señal y ganancia. 25

Tabla 6.1 Requisitos de espectro en términos de ancho de banda, formato de modulación y velocidad. 34

Tabla 6.2 Máxima distancia de alcance (MAR) para formatos de modulación y BER.. 34

Tabla 8.1 Hiperparámetros Q learning. 52

Tabla 8.2 Simulaciones. 54

Tabla 8.3 hiperparámetros Multi Agente. 55

Tabla 8.4 Simulaciones. 58

Lista de Abreviaturas

BP: Blocking Probability: Probabilidad de bloqueo.

DDQN: Double Deep Q-Network: Algoritmo Red Q profunda doble.

DNN: Deep Neural Networks: Redes neuronales profundas.

DQN: Deep Q-Network: Algoritmo Red Q profunda.

DQN RMSCA: Deep Q-Network Routing, Modulation, Spectrum, and Core Allocation: Algoritmo Red Q profunda para Asignación de enrutamiento, modulación, espectro y núcleo

Dueling DQN: Dueling Deep Q-Netwok: Duelo Red Q profunda

EON: Elastic Optical Networks: Redes ópticas elásticas.

FSU: Frequency Slot Unit: Unidades de ranuras de frecuencia

IoT: Internet of things: Internet de las cosas.

IQL: Independent Q-learning: Q-learning independiente

ITU: International Telecommunication union: Unión internacional de telecomunicaciones.

MARL: Multi-agent Reinforcement Learning: Aprendizaje por refuerzo multi agente

ML: Machine Learning: Aprendizaje automático.

M2M: Machine to Machine: Máquina a máquina.

NN: Neural Networks: Redes neuronales.

OFDM: Ortogonal Frequency Division Multiplexing: Multiplexación por división de frecuencias ortogonales.

O-OFDM: Optical Ortogonal Frequency Division Multiplexing: Multiplexación por división de frecuencias ortogonales óptica.

OTN: Optical Transport Network: Red de transporte óptico.

PCE: Path Computation Element: Elemento de cálculo de ruta

QoS: Quality of Service: Calidad de servicio.

Q-Learning: Q-learning Algorithm: Algoritmo Q-Learning

RSA: Routing and Spectrum Allocation: Asignación de enrutamiento y espectro.

RMSA: Routing, Modulation and Spectrum Allocation: Asignación de enrutamiento, modulación y espectro.

RMSCA: Routing, Modulation, Spectrum, and Core Allocation: Asignación de enrutamiento, modulación, espectro y núcleo

RL: Reinforcement Learning. Aprendizaje por refuerzo.

ROADM: Reconfigurable Optical Add/Drop Multiplexer: Multiplexores ópticos reconfigurables de inserción y extracción.

RWA: Routing and wavelength assignment: Enrutamiento de longitudes de onda

SDM: Spatial Division Multiplexing: Multiplexación por división espacial

SDN: Software Defined Networking: Redes definidas por software.

WADM: Wavelength Add/Drop Multiplexer: Multiplexores de adición y extracción de longitud de onda

WDM: Wavelength Division Multiplexing: Multiplexación por división de onda.

5G: Fifth Generation Technology Standard for Cellular Networks: Comunicaciones móviles de quinta generación

6G: Sixth Generation Technology Standard for Cellular Networks, comunicaciones móviles de sexta generación.

Introducción

La constante innovación y desarrollo tecnológico en redes ópticas busca soportar los futuros volúmenes de tráfico que pueden llegar a congestionar la red actual y causar caos en el transporte de información (Cisco, 2019). Las proyecciones de tráfico son generadas por servicios y aplicaciones tales como: automóviles autónomos (Kalyani & Pinky, 2017), pronóstico avanzado del tiempo (Anusuya & Gomathi, 2018), automatización proveniente de las máquinas generadas por la cuarta revolución industrial (Perasso & Mundo, 2016), redes sociales en línea, transmisión de video (Koyama et al., 2019), Internet de las cosas (Chung & Jeong, 2020), ciudades inteligentes (Mehdi & Hayar, 2018) y 5G (Al-Mekhlafi et al., 2020), etc. Los requisitos más importantes para estas aplicaciones son: baja latencia, comunicación ultra confiable y alta velocidad de línea, estos requerimientos han desencadenado una búsqueda exhaustiva de redes con mayor capacidad y rentabilidad.

La mayor exigencia en los actuales sistemas de telecomunicaciones fue la llegada del confinamiento, cambiando las dinámicas mundiales a nivel político, económico, social y sanitario. Lo que ha generado un aumento en el número de usuarios de servicios tecnológicos, recursos de educación y ocio (Király et al., 2020). En el ámbito social se han creado nuevas formas de interacción a través de las redes, por ejemplo, realizar celebraciones familiares y reuniones de amigos por videoconferencia, algo que antes solo sucedía en situaciones muy particulares. La interacción remota prolongada exige el uso de una mayor cantidad de servicios (Alashhab et al., 2020), disponibilidad y condiciones de servicio para poder soportar las necesidades de los usuarios, incrementando la demanda de aplicaciones que requieren recursos de redes de telecomunicaciones.

Según el reporte de tráfico anual de Internet de Cisco para 2023, 5300 millones de usuarios que corresponden al 66% de la población mundial accederá a Internet (Cisco, 2020). La tecnología inteligente se aplicará en la industria del transporte para 2025, con más de 60 millones de vehículos conectados a redes 5G y el 100% de los vehículos nuevos conectados a Internet (Huawei, 2018). Por otro lado, el proveedor de equipos Ciena, estima que habrá 75 mil millones de dispositivos conectados para 2025, así como nuevas aplicaciones críticas que exigirán alta confiabilidad, alta disponibilidad y ultra baja latencia (Ciena Corporation - AnnualReports.Com, s/f). Este proveedor encuestó a personas de todo el mundo y reveló que

el 81% de las personas en los EE. UU. manifiestan que el hogar seguirá siendo una de las localizaciones preferidas para trabajo y entretenimiento, priorizando la velocidad como criterio fundamental de Internet. En el Reino Unido el 43% de las personas planean aumentar su Internet dentro de los próximos 24 meses. En los EAU y Arabia Saudita más del 90% de las personas están evidenciando los beneficios de Internet de las cosas (IoT) y 5G (Ciena Corporation - AnnualReports.Com, s/f). Estas proyecciones y encuestas están aplicando presión en crecimiento y adaptabilidad a las redes actuales y futuras.

La dinámica del mercado ejerce nuevas presiones sobre las redes de larga distancia, metropolitanas y perimetrales, evidenciando la necesidad de evolucionar estas arquitecturas para satisfacer las demandas de los usuarios finales. Se espera que el desarrollo acelerado del tráfico y la existencia de servicios con demandas muy variadas contribuyan a las condiciones de tráfico en constante cambio y a las demandas de gran ancho de banda en las redes 5G (Idowu-Bismark et al., 2019; Zhou et al., 2020). El redimensionamiento en las redes se ha convertido en una actividad vital, por consiguiente, los proveedores de Internet deben revisar y actualizar sus estrategias de capacidad para anticipar y evitar complicaciones de degradación en la experiencia del usuario.

La tecnología con mayor proyección para soportar las velocidades de transmisión de perfiles de tráfico aleatorio emergentes lo proporcionan las redes ópticas elásticas (Velasco et al., 2017), con uso de espectro óptico de manera flexible y eficiente en comparación a su contraparte la multiplexación por división de onda (WDM) (Murthy & Gurusamy, 2002). En esta tarea es imprescindible la aplicación de algoritmia, con marcada tendencia al uso de técnicas de aprendizaje de máquina (ML) sobre las redes de comunicaciones para realizar las funciones de control y operación. Aunque el concepto de ML no es nuevo (O'Shea & Hoydis, 2017), estas iniciativas han ganado impulso gracias a dos tecnologías: Redes definidas por software (SDN) (Li & Kilper, 2018) y analítica de red (NA). Aunque es difícil predecir las aplicaciones y servicios que demandarán los usuarios en los próximos años, el comportamiento del tráfico a nivel de velocidad y aleatoriedad crece aceleradamente, por consiguiente, es imprescindible el desarrollo de una red flexible, confiable, disponible y segura que dé respuesta a las proyecciones que demandan los crecientes perfiles de tráfico variable.

I. Descripción del Proyecto

Capítulo 1

1. Formulación del proyecto

1.1. Formulación del problema

En la actualidad los enlaces de fibra óptica son considerados la mejor opción en capacidad y velocidad para soportar los requisitos de usuario en aplicaciones y contenidos enriquecidos (Wei et al., 2012). En la última década se han implementado las redes WDM y las redes de transporte óptico (OTN) equipadas con una tasa máxima de transmisión de 200 Gbps en rejilla fija con espacio entre canales adyacentes de 50 GHz (Gigahercios) o 100 GHz según los estándares de la Unión Internacional de Telecomunicaciones (ITU, 2012), evidencia la rigidez en ancho de banda de los canales como el principal problema de las redes WDM, provocando un uso ineficiente del espectro. Por otra parte, las redes OTN actuales delegan la reconfiguración de las características de enrutamiento a capas superiores, aumentando la probabilidad de bloqueo, los tiempos de transporte y el procesamiento.

A pesar del crecimiento de las tasas de transmisión de OTN las demandas de tráfico continúan aumentando, hasta 2008, los requerimientos de ancho de banda se incrementaron 10 veces cada 6 años (Heron et al., 2008), en la actualidad cada persona cuenta con 2,2 dispositivos que consumen servicios de Internet y reflejan a nivel mundial 1,5 ZB de tráfico IP. Para finales de 2023 se espera que la cantidad de dispositivos conectados a redes IP sea más del triple de la población mundial, con lo cual el tráfico de Internet llegará a los 4,8 ZB (Cisco, 2020). El tráfico de video aumentará casi dos tercios (66 %) de los televisores instalados serán 4k o Ultra-High-Definition (UHD) (Cisco, 2019), donde también tendrá una participación importante dispositivos 8k (Koyama et al., 2019). IoT se ha convertido en un sistema predominante en el que las personas, los procesos, los datos y las cosas se conectan a Internet y entre sí. A nivel mundial, las conexiones maquina a maquina (M2M) pasarán de

6,1 mil millones en 2018 a 14,7 mil millones en 2023, logrando 1.8 conexiones M2M para cada miembro de la población mundial.

Frente a las proyecciones de tráfico, los actuales esquemas de redes ópticas presentan limitaciones en capacidad, administración y supervivencia. De continuar con el diseño de infraestructura actual y sin tener presentes las proyecciones de tráfico, los servicios de comunicaciones fijas y móviles colapsaran. Para el sistema financiero global la situación puede ser devastadora, afectando la banca electrónica, comercio electrónico y los nuevos ecosistemas financieros descentralizados basados en tecnologías como blockchain. En el sector de la medicina, se verían amenazados los accesos en línea a historiales médicos y el apoyo tecnológico en las salas de cirugía, que cada día exigen un mayor sustento clínico especializado. En lo social se debilitaría la fortaleza en la formación del talento digital, aumentando la brecha de pobreza, educación y equidad. Estos factores y otros obligarían a los usuarios a incurrir en mayores costos, restringiendo el número de personas, con lo cual se perdería el perfil de Internet como activo público.

Este constante crecimiento sostenido de aplicaciones con diferentes requerimientos de calidad de servicio (QoS) exige innovación y constante desarrollo tecnológico que permita la implementación viable de redes completamente ópticas. En respuesta a las futuras necesidades ha emergido la arquitectura EON (Cai et al., 2012) que supera el límite de la red rígida WDM convencional, adaptando dinámicamente sus recursos de acuerdo con los requisitos de cada conexión. Para esta tecnología es necesario establecer nuevos algoritmos que logren explotar las características de asignación de espectro y espacio cuando este tiene un comportamiento elástico.

1.2. Interrogante

¿Cómo asignar recursos físicos en redes de transporte óptico que respondan dinámicamente a las demandas variables que exigen diferentes perfiles de tráfico para mantener alta disponibilidad?

1.3. Justificación

Para asegurar los efectos positivos de la economía digital del siglo XXI es necesario expandir y fortalecer la infraestructura que soporta las autopistas de información, generando nuevas propuestas que permitan mantener el crecimiento de los servicios que se demanden. Tendencias cómo: 5G, IoT, transmisión de contenido en streaming y servicios en la nube, exigen conectividad de alto rendimiento a nivel de núcleo de red móvil y fija, que suministren enlaces directos a plataformas globales de proveedores de contenidos, con gran capacidad de transmisión para satisfacer las demandas actuales y futuras.

Entre los grandes desafíos para la implementación de las redes ópticas elásticas están: consolidación de componentes de hardware que trabajen de manera elástica y, diseño de algoritmos que establezcan, de forma rápida y eficiente una trayectoria espectral para atender cada una de las peticiones. La incorporación de algoritmos de aprendizaje automático facilitará la adaptación a las crecientes tendencias de tráfico aleatorio disminuyendo el procesamiento requerido para identificar, asignar, direccionar y priorizar los recursos con respecto a los diferentes tipos de tráfico, por lo tanto, serán fundamentales para gestionar los recursos a nivel espectral y espacial de acuerdo con las necesidades del usuario.

La clave para el éxito en el desarrollo de los futuros sistemas de comunicación móvil 5G y de sexta generación (6G) será cumplir con los objetivos de rendimiento, tales como una transmisión de velocidad de datos más alta, una latencia de extremo a extremo más baja, una capacidad más alta, un costo más bajo y una calidad de experiencia (QoE) satisfactoria. La gran ventaja de las redes ópticas es su capacidad para soportar cualquier tipo de servicio con independencia de sus requisitos de ancho de banda y calidad de servicio. Por lo tanto, permite la transmisión simultánea de servicios con altas tasas de bits como video en alta definición o bajas tasas como la voz, incluso apuntalará la inserción de nuevos servicios de Internet. Otra ventaja está soportada en la flexibilidad de las redes ópticas elásticas, donde las funciones de agregación y multiplexación estadística pueden ubicarse en capa óptica, lo que genera un espacio de investigación para comparar el desempeño de las diferentes alternativas de flexibilidad.

Debido a las características especiales de las redes ópticas elásticas, los mecanismos tradicionales de asignación, enrutamiento, protección y recuperación deben reformularse. Gracias a esta reformulación los algoritmos de aprendizaje aplicados en el campo de las redes de transporte óptico están buscando su consolidación como el hilo conductor que implemente la separación de los servicios, proporcionando nuevos modelos para la predicción de tráfico que se reflejen en rapidez, economía, calidad y seguridad.

El establecimiento de una sociedad digital crea una mayor dependencia a servicios y aplicaciones en Internet, llegando a considerarse un activo público al igual que en su momento sucedió con las carreteras y los ferrocarriles. Por lo tanto, una infraestructura de telecomunicaciones actualizada a los requerimientos de los usuarios es fundamental para que sean parte de esta emergente sociedad. El desarrollo de algoritmos permitirá hacer un mejor uso de los recursos de infraestructura de telecomunicaciones apoyando los nuevos servicios que la sociedad demande, para evitar la relegación y el incremento de la brecha digital, en especial en los países en vía de desarrollo.

1.4. Hipótesis

Considerando lo anterior, se plantea la siguiente hipótesis:

La aplicación de algoritmos de aprendizaje automático sobre capa óptica mejora el desempeño de red en términos de eficiencia espectral, costos y capacidad del canal; aportando elasticidad y reduciendo las pérdidas de recursos físicos de transmisión ocasionadas por la variabilidad en los crecientes volúmenes de tráfico.

1.5. Grupo y línea de investigación

La investigación se centra en el campo de las redes ópticas, aplicando las nuevas técnicas en direccionamiento y encaminamiento sobre capa óptica para respaldar el transporte de altos volúmenes de tráfico. Temática abordada a nivel institucional por los grupos de investigación "Interoperabilidad Tecnológica y Semántica" e "Internet Inteligente" de la Universidad Distrital Francisco José de Caldas.

Capítulo 2

2. Objetivo general y objetivos específicos

2.1. Objetivo General

Desarrollar un modelo algorítmico para proporcionar alta disponibilidad en el transporte de volúmenes crecientes de tráfico variable en redes ópticas.

2.2. Objetivos Específicos

- Caracterizar la capa física de redes ópticas para determinar cómo aprovechar los recursos físicos en situaciones donde los tipos de tráfico varían dinámicamente.
- Diseñar un modelo algorítmico para transportar los crecientes volúmenes de tráfico ajustando dinámicamente los recursos físicos a los requerimientos de cada tipo de usuario.
- Validar el modelo algorítmico propuesto sobre diferentes caracterizaciones aplicando procesos de simulación, que permitan comparar las capacidades de la red óptica en adaptación y soporte de volúmenes de tráfico que desbordan los modelos convencionales.
- Socializar los resultados de investigación mediante publicaciones en revistas indexadas, fortaleciendo el proceso de creación de nuevo conocimiento, la difusión social y la consolidación de la investigación a nivel institucional.

Capítulo 3

3. Antecedentes

Una red totalmente óptica es aquélla que transporta los datos desde el nodo origen al nodo final en formato óptico sin conversiones optoelectrónicas, ni electroópticas intermedias. Son redes transparentes ideales, pero hasta el momento inviables, en la práctica siempre se requiere algún tipo de regeneración electrónica. La tecnología actual de los sistemas WDM permite velocidades de 10 Gbps, 40 Gbps, 100 Gbps, con multiplexación superior a 40 canales de transmisión simultánea bajo asignación de frecuencias de tamaño fijo que garantiza la calidad de la señal y el filtrado en los receptores. Esta capacidad puede ser manejada exclusivamente en la capa óptica cuando se utilizan multiplexores ópticos reconfigurables de inserción y extracción (ROADM) con múltiples grados de conmutación en los nodos de la red.

Los sistemas actuales WDM de slots de frecuencia estáticos denominados de “rejilla fija” brindan escalabilidad y flexibilidad limitada, esta gestión estática de la infraestructura dificulta la adaptación a la variabilidad y heterogeneidad del tráfico (Mukherjee & Mukherjee, 2006; Ramaswami et al., 2010). En los últimos años se ha logrado aliviar algunas de las limitaciones que tienen las redes ópticas mediante control centralizado por medio de SDN y la virtualización de funciones de red (NFV) (Li & Kilper, 2018; Tanha et al., 2018; Vissicchio et al., 2017). El objetivo de NFV es eliminar los dispositivos de hardware propietario y permitir que los servicios de red que ahora son llevados a cabo por enrutadores, firewalls, balanceadores de carga y otros dispositivos de hardware dedicado sean hospedados en máquinas virtuales (VM), esta propiedad permite crear cadenas de servicios sin tener hardware dedicado. SDN tiene el potencial de proporcionar una solución de control y administración unificada, reuniendo los protocolos de capa 3 (IP), capa 2 (Ethernet) y capa 1 (fotónica), su uso en la administración de redes ópticas aumenta la automatización y estandarización, pero no proporciona flexibilidad, ni capacidad de programación en tiempo real.

Las redes EON proponen un acercamiento en la administración eficiente de los recursos espectrales por medio de la tecnología: multiplexación por división de frecuencias ortogonales óptica (O-OFDM) (Minn et al., 2000; Schmidl & Cox, 1997). Las EON logran asignar espectro óptico de forma flexible mediante una gestión ágil del ancho de banda en la capa óptica (Lu et al., 2015), estas características han despertado el interés de los investigadores para mejorar la configuración y el funcionamiento de los dispositivos de red, la supervisión del rendimiento óptico, el reconocimiento del formato de modulación, la mitigación de las no linealidades de fibra y la estimación de la calidad de transmisión (QoT). En las empresas de telecomunicaciones el gran tamaño de la infraestructura es un desafío para el procesamiento, detección, predicción, localización y corrección de cambios de topología, por este motivo crece la necesidad de adquirir la capacidad de recolectar y analizar de manera eficiente la información.

En diversas áreas de la ciencia se están aplicando los algoritmos de aprendizaje automático, en medicina los programas de aprendizaje se utilizan para el diagnóstico médico (Deo, 2015; Khan et al., 2001; Wernick et al., 2010), en biología para el reconocimiento de Fenotipos (Sommer & Gerlich, 2013; Tarca et al., 2007), en la energía para calcular parámetros de rendimiento de los calentadores de agua solares (Kalogirou et al., 1999; Liu et al., 2016) y en telecomunicaciones los patrones de llamada se analizan para la optimización de la red (Clark et al., 2003), existe la necesidad de crear métodos que colaboren automáticamente para realizar predicciones acertadas en diferentes escenarios de aplicación.

Los algoritmos de aprendizaje automático (ML) están impulsado el desarrollo de las redes ópticas en áreas como: predicción y clasificación de tráfico, automatización de procesos en la reducción de costos operativos, apoyo en la inteligencia humana para la planificación y reconfiguración de red, virtualización, detección de anomalías y correlación de alarmas, entre otras. La inteligencia artificial (IA) presenta varias oportunidades para automatizar operaciones e introducir decisiones inteligentes en: planificación de redes, control dinámico y gestión de recursos de red, incluyendo el establecimiento de conexiones, autoconfiguración y auto optimización, a través de la predicción y estimación, utilizando el estado de red en el instante e históricos.

Las técnicas de ML se han utilizado en algunas funciones de optimización de red, mediante la estimación de las matrices de tráfico predictivo se implementan servicios en las redes ópticas dinámicas. Como solución al problema de implementación también se han propuesto arquitecturas basadas en SDN (Yan et al., 2018), aplicando métodos de aprendizaje automático para las matrices predictivas de tráfico (Panayiotou et al., 2018; Alvizu et al., 2017), incluyendo redes neuronales (NN), redes neuronales profundas (DNN) (Chen et al., 2017) y aprendizaje por refuerzo (RL) (Panayiotou et al., 2018). Trabajos recientes han utilizado aprendizaje por refuerzo para abordar problemas relacionados, como el enrutamiento en redes ópticas (Aguirre et al., 2022), el enrutamiento IP (Valadarsky et al., 2017) o el aprovisionamiento de QoS (S. C. Lin et al., 2016).

Las investigaciones que aplican aprendizaje automático en redes ópticas de mayor relevancia son: (I) Estimación y adquisición de parámetros físicos que permite el diagnóstico de la red para tomar medidas contra fallas, existe un método basado en aprendizaje automático que ajusta el punto de funcionamiento de una red EON, logrando reducir la figura de ruido para adaptar la ganancia de cada amplificador en cascada. Este modelo es capaz de definir la ganancia de los amplificadores, considerando la restricción en la cual la potencia de entrada y la potencia de salida del enlace pueden predefinirse (De A Barboza et al., 2013). Un amplio estudio para aplicar redes neuronales artificiales en la administración del rendimiento óptico (OPM) realiza la identificación simultánea de la no linealidad acumulada, relación de señal óptica al ruido (OSNR), dispersión cromática (CD) y la dispersión por modo de polarización (PMD), a partir de los parámetros de histograma de ojo y diagrama de ojo (Xiaoxia Wu et al., 2009). (II) QoT el cual es relevante en el diseño y operación de redes ópticas sensibles al deterioro, mediante el cálculo de factores Q asociados a un conjunto de trayectorias de luz, en la combinación de modelos analíticos y métodos numéricos, para lo cual se emplea la herramienta Q-Tool con base en una topología de referencia(T. Jiménez et al., 2012). Existen otras dos propuestas para la estimación de QoT, la primera propone un clasificador basado en el aprendizaje automático para predecir la probabilidad de BER para que un camino de luz candidato, no exceda el umbral propuesto (Barletta et al., 2017) y la segunda presenta el concepto de "red viviente" donde la red óptica mantiene registros de su desempeño a nivel espacial, aprovechando el monitoreo de la información de BER y un proceso de aprendizaje para estimar el BER de cada nueva solicitud de servicio(Oda et al., 2017b). (III) Predicción de probabilidad de bloqueo (BP) mediante redes neuronales artificiales aplicando regresión no lineal basada en una lista de variables de entrada con base en métricas de topología e información de la capa física (de Araújo et al., 2015). (IV) Predicción de tráfico del próximo periodo con base en el volumen y dirección de tráfico mediante una red neuronal artificial para cada servicio, presentando mejores resultados tanto en ahorros de capital como de gastos operativos (Velasco et al., 2017). (V) Enrutamiento de longitudes de onda (RWA) complementado por medio de la técnica de aprendizaje por refuerzo Q-learning se resuelve el camino y la selección de la longitud de onda, proporcionando mejores resultados que cuando es abordado mediante algoritmos heurísticos (Kiran et al., 2007; Venkatesh et al., 2008).

II. Marco teórico

Capítulo 4

4. Aprendizaje por Refuerzo

4.1. Introducción

En el presente capítulo se revisan los aspectos teóricos que fundamentan el aprendizaje por refuerzo. También se realiza un recuento de las diferentes técnicas que se tienen para la solución de los problemas asociados con el aprendizaje por refuerzo, siendo de importancia el concepto aprendizaje multiagente.

4.2. Aprendizaje por refuerzo

En una configuración de RL (Figura 4-1), totalmente observable, hay un agente que interactúa con un entorno mediante la realización de acciones. En cada tiempo t, el agente observa el estado actual $S_t \in S$ del entorno, elige una acción $U_t \epsilon U$ de acuerdo con una política estocástica π y recibe una señal de recompensa $R_t = (S_t, U_t)$. Luego, el entorno transita a un nuevo estado $S_{t+1} \in S$ de acuerdo con la función de probabilidad de transición $P(S_{t+1}|S_t, U_t)$. El objetivo es encontrar una política π para el agente que maximice la expectativa de rendimiento descontado $G_t = \sum_{k=t}^{\infty} \gamma^{k-t} R_k$, donde R_k es la recompensa recibida en el tiempo k y $\gamma \epsilon [0,1)$ es un factor de descuento. La función acción-valor Q de una política π es $Q(s,u) = E^{\pi}[G_t|S_t = s, U_t = u]$. La función de valor de acción óptima $Q^*(s,u) = max_{\pi} Q^{\pi}(s,u)$ obedece a la ecuación de optimización de Bellman $Q^*(s,u) = r(s,u) + \gamma \sum_{s'} P(s'|s,u)\, max_{u'} Q^*(s',u')$. Dado Q^*, la política óptima π^* para cada estado $s \in S$ se puede recuperar seleccionando la acción con la función Q más alta: $\pi^*(s) = argmax_u Q^*(s,u)$. Esta propiedad ha llevado a varios algoritmos de aprendizaje que estiman directamente la función Q, como Q-Learning (*Learning from Delayed Rewards / Christopher John Cornish Hellaby Watkins. - University of Edinburgh*, n.d.). Esta interacción se puede visualizar gráficamente en la Figura 4-1:

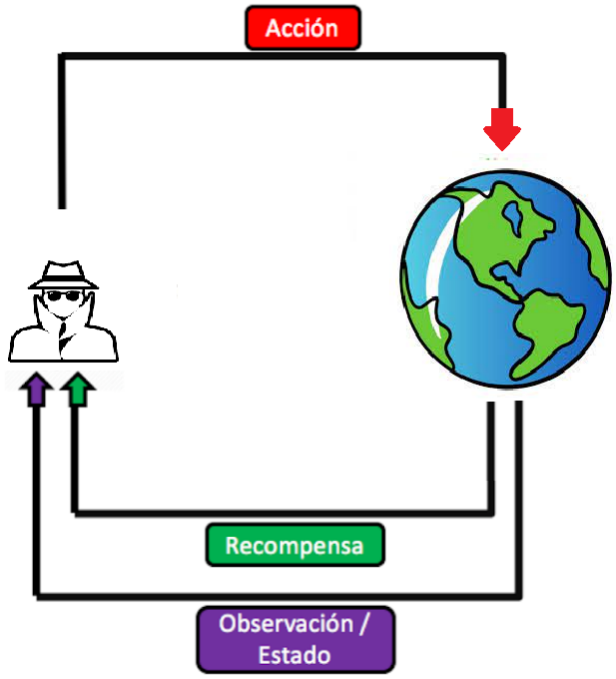


Figura 4-1 Interacción aprendizaje por refuerzo. Elaboración propia.

4.3. Q-Learning

Q-Learning es un método de Aprendizaje por Refuerzo que permite resolver problemas de decisión secuencial en los cuales la utilidad de una acción depende de una secuencia de decisiones y donde además existe incertidumbre en cuanto a las dinámicas del ambiente en que está situado el agente. El algoritmo Q-learning para estimar q hizo de las funciones acción-valor una parte importante del aprendizaje por refuerzo y, en consecuencia, estas funciones a menudo se denominan funciones Q.

En Q-Learning un agente comienza con una estimación arbitraria Q_0 de Q^* y mejora iterativamente su estimación de la siguiente manera: en cada momento t, el agente realiza una acción u_t para el estado s_t de acuerdo con una política, recibe una recompensa $r_t = r(s_t, u_t)$ y observa un nuevo estado S_{t+1}. Basado en la experiencia $< s_t, u_t, r_t, s_{t+1} >$, el agente actualiza la función Q, según:

$$Q_{t+1}(s_t, u_t) = Q_t(s_t, u_t) + \propto_t [r^t + \gamma max_{u_t+1} Q(s_{t+1}, u_{t+1}) - Q(s_t, u_t)]$$

Ecuación 1 Actualización Función Q. Tomado de (Mnih et al., 2015).

Donde $\alpha_t \in [0,1)$, es la tasa de aprendizaje o el tamaño del paso que determina en qué medida la información recién adquirida anula la información anterior. En la Ecuación 1 el término $r^t + \gamma max_{u_t+1} Q(s_{t+1}, u_{t+1})$ se considera objetivo y el término $Q(s_t, u_t)$ se

considera estimación. Una forma común de derivar una nueva política durante las iteraciones es seguir una estrategia codiciosa. Con base en esta estrategia en cada estado s_t, con probabilidad ϵ se selecciona una acción aleatoria y con probabilidad $1-\epsilon$, se selecciona la acción que maximiza Q_t. Este tipo de estrategias se utilizan para introducir una forma de exploración que es un concepto importante en RL: al seleccionar aleatoriamente acciones que son subóptimas de acuerdo con sus estimaciones actuales, el agente puede descubrir y corregir sus estimaciones cuando corresponda.

4.4. Deep Q-Network (DQN)

Los agentes de DQN desarrollan sus políticas con aprendizaje Q, una técnica de aprendizaje por refuerzo que se centra en la idea de los valores Q y una función Q: para cada par único de un estado específico del entorno y una acción específica a tomar, hay un valor Q que representa el valor o la utilidad de tomar esa acción específica cuando ese estado específico está en su lugar.

El algoritmo DQN (Deep Q-Network) fue desarrollado por Mnih en 2015 (Mnih et al., 2015) y es capaz de resolver una amplia gama de juegos de Atari (algunos a nivel sobrehumano) mediante la combinación de aprendizaje por refuerzo y redes neuronales profundas a gran escala. El algoritmo fue desarrollado por la mejora de un algoritmo de RL clásico llamado Q-aprendizaje con redes neuronales profundos y una técnica llamada experiencia de reproducción En el campo de la inteligencia artificial (IA) las redes neuronales son una forma de aproximar funciones no lineales. Por lo tanto, este algoritmo usa una red neuronal para aproximar la función Q, evitando así utilizar una tabla para representar la misma. DQN es una variación del clásico algoritmo Q-Learning con 3 contribuciones principales:

- En lugar de la función Q tabular utilizada en Q-learning, DQN utiliza redes neuronales multicapa para aproximarse a la función Q. En otras palabras, en DQN la función acción-valor $Q(.,.)$ es aproxima por $Q(.,.;\theta_t)$ donde θ_t representa los parámetros de la red neuronal.
- En lugar de usar una función Q para calcular tanto el objetivo como la estimación, DQN usa dos redes neuronales separadas, una red neuronal objetivo $Q(.,.;\overline{\theta_t})$ y una red neuronal en línea $Q(.,.;\theta_t)$. Más específicamente, para cualquier experiencia $< s_t, u_{t,} r_t, s_{t+1} >$, DQN usa la red objetivo para calcular el objetivo $r^t + \gamma max_{u_t+1}\bar{Q}(s_{t+1}, u_{t+1}; \overline{\theta_t})$ y usa la red en línea para calcular la estimación $Q(s_t, u_t; \theta_t)$]. Los valores objetivo y de estimación se utilizan luego para calcular la siguiente pérdida:

$$[r^t + \gamma max_{u_t+1}\bar{Q}(s_{t+1}, u_{t+1}; \overline{Q_t}) - Q(s_t, u_t; \theta_t)]^2$$

Ecuación 2 Pérdida DQ. Tomado de (Mnih et al., 2015).

- En lugar de usar la última experiencia obtenida para actualizar la función Q, DQN almacena las experiencias en una memoria de repetición (L.-J. Lin, 1992) y usa mini lotes de experiencias tomadas uniformemente al azar para actualizar los parámetros de la red en línea usando descenso de gradiente estocástico para minimizar la pérdida mediante la Ecuación 2. La red de destino no se optimiza directamente y solo se actualiza periódicamente a la red en línea. De esta manera, el entrenamiento puede proceder de una manera más estable. El uso de la memoria de repetición de experiencias puede aliviar dos problemas: i) las redes neuronales son propensas a "olvidar" las cosas que han aprendido, ii) Existe una gran cantidad de correlación entre experiencias consecutivas. Al mantener las experiencias anteriores en una memoria de reproducción y tomar muestras aleatoriamente de esta memoria, DQN evita que la red aprenda en función de las experiencias correlacionadas y también le permite volver a aprender de las experiencias pasadas.

4.5. Extensiones Deep Q-Network

Existen varias extensiones del DQN original propuesto por (Mnih et al., 2015). A continuación, se describen un conjunto de extensiones que abordan distintas preocupaciones asociadas con el DQN original.

4.5.1. Doble DQN (DDQN)

Una red Q profunda doble o doble DQN utiliza el aprendizaje Q doble para reducir la sobreestimación al descomponer la operación máxima en el objetivo en selección de acción y evaluación de acción. En el DQN la red objetivo $Q(.,.;\overline{\theta_t})$, se usa tanto para seleccionar la mejor acción del siguiente estado como para calcular el objetivo para evaluar la acción seleccionada (Ecuación 2). Esto hace que sea más probable seleccionar valores sobreestimados, lo que da como resultado estimaciones de valores demasiado optimistas. Para evitar este problema, los autores de Double Deep Q-Learning (Van Hasselt et al., 2015) propusieron una técnica simple para descomponer la selección de acciones de la evaluación de acciones. En DDQN, la red en línea se usa para seleccionar una acción y la red objetivo se usa para generar el valor objetivo para esa acción. Por lo tanto, la pérdida de la Ecuación 2 se puede reescribir de la siguiente manera para el doble DQN.

$$[r^t + \gamma \bar{Q}(s_{t+1}, argmax_{u_t+1} Q(s_{t+1}, u_{t+1}; \theta_t); \overline{Q_t}) - Q(s_t, u_t; \theta_t)]^2$$

Ecuación 3 Pérdida DDQN. Tomado de (Mnih et al., 2015).

Se ha demostrado que este cambio reduce sustancialmente las sobreestimaciones que estaban presentes en DQN y mejora el rendimiento.

4.5.2. *Repetición de experiencia priorizada*

En DQN las muestras se toman uniformemente al azar de la memoria de repetición de la experiencia. Sin embargo, al tratar todas las muestras de la misma manera, estamos ignorando una simple intuición del mundo real, es decir, potencialmente podemos aprender más de las experiencias cuyos resultados difieren más de las expectativas. Para aprovechar este hecho, la repetición de experiencia priorizada (PER) (Schaul et al., 2015) muestrea experiencias con probabilidad p_t proporcional a la diferencia absoluta entre el objetivo y los valores de estimación para esta experiencia, lo que se conoce como error TD.

$$p_t\ \alpha\ (\left|r^t + \gamma max_{u_t+1}\bar{Q}(s_{t+1}, u_{t+1}; \bar{\theta_t}) - Q(s_t, u_t; \theta_t)\right| + \epsilon)^w$$

Ecuación 4 Error TD. Tomado de (Mnih et al., 2015).

Donde ϵ, es una pequeña constante positiva que asegura que ninguna experiencia tiene probabilidad cero de ser muestreada. Además, el w determina cuánta priorización se usa, $w = 0$ corresponde al caso uniforme. Para la suite de Atari (Bellemare et al., 2012) el uso de la repetición de experiencia priorizada ha llevado a un aprendizaje más rápido y a una mejor calidad de la política final en la mayoría de los juegos en comparación con la repetición de experiencia uniforme.

4.5.3. *Duelo DQN (Dueling DQN)*

En la arquitectura de red DQN, aunque existen dos redes, que son la red en línea y la red de destino, en cada una de estas redes, se usa una sola red neuronal para estimar la función acción-valor Q, la cual mide el valor de elegir una acción particular cuando se está en un estado. La idea clave detrás de la arquitectura de red de duelo (Wang et al., 2015), es que para muchos estados no es necesario estimar el valor de cada opción de acción. Con este fin se estiman dos funciones separadas, una función de valor V, que mide qué tan bueno es estar en un estado particular y una función de ventaja A, que mide la importancia relativa de elegir una acción particular cuando se está en un estado particular. Luego, estas dos funciones se agregan para estimar la función de acción-valor Q. Por ejemplo, un módulo de agregación para la red en línea propuesto en (Wang et al., 2015) es el siguiente:

$$Q(s_t, u_t; \theta_t, \alpha_t, \beta_t) = V(\theta_t, \alpha_t, \beta_t) + (A(s_t, u_t; \theta_t, \alpha_t) - \frac{1}{|U|}\sum_{u'\epsilon U} A(s_t, u'_t; \theta_t, \alpha_t))$$

Ecuación 5 Módulo de agregación. Tomado de (Mnih et al., 2015).

Donde θ_t representa los parámetros de capas compartidas (por ejemplo, las capas convolucionales si existen) entre los estimadores para la función de valor V y la función de

ventaja A, mientras que α_t y β_t, son los parámetros de capas separadas para la función de valor V y la función de ventaja A, respectivamente.

4.6. Aprendizaje por refuerzo multi agente (MARL)

El aprendizaje por refuerzo de múltiples agentes (MARL) estudia cómo interactúan múltiples agentes en un entorno común. Estos agentes interactúan con el entorno y entre sí, se puede observar en estado de colaboración, coordinación, competición o aprender colectivamente para realizar una tarea en particular. En este problema multiagente cooperativo totalmente observable, existen n agentes donde el conjunto de agentes se denota $A = \{1,2,3 \dots n\}$. En cada tiempo t, cada agente $a \in A$ observa el estado actual $S_t \in S$ del entorno y elige una acción $U_t^a \in U^a$ de acuerdo con una política estocástica π^a. Las acciones de los agentes forman una acción conjunta U_t. Como resultado de esta acción conjunta, todos los agentes reciben una señal de recompensa $R_t = r(S_t, U_t)$ y el entorno transita a un nuevo estado $S_{t+1} \epsilon S$ según la función de probabilidad de transición $P(S_{t+1}|S_t, U_t)$. El objetivo es encontrar un conjunto de políticas $\pi^a, a \in A$, para todos los agentes que maximizan la expectativa de rendimiento descontado $G_t = \sum_{k=t}^{\infty} \gamma^{k-t} R_k$ donde R_K es la recompensa recibida en el tiempo k y $\gamma \in [0,1]$ es un factor de descuento. Esta interacción se puede visualizar gráficamente en la Figura 4-2:

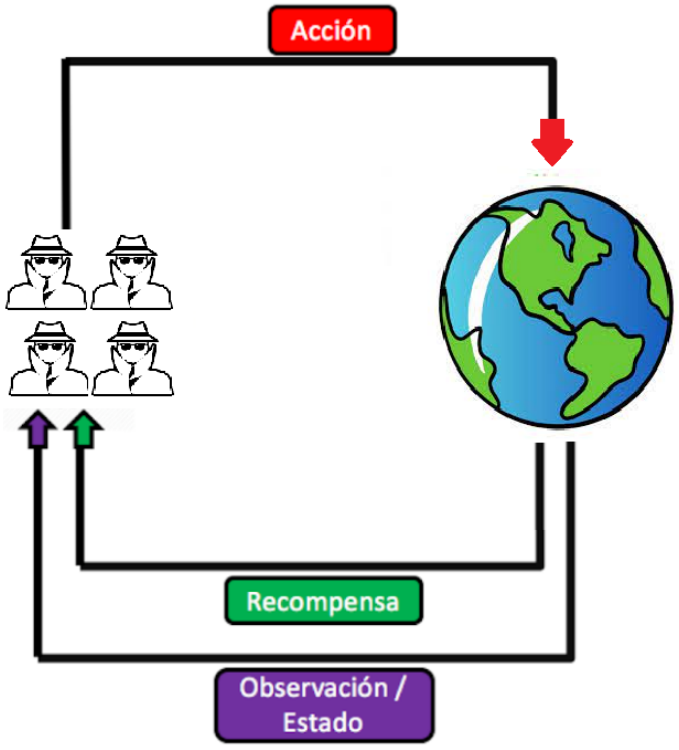


Figura 4-2 Interacción aprendizaje por refuerzo multi agente. Elaboración propia.

Un enfoque natural para un problema de RL multiagente cooperativo y totalmente observable es considerar un "meta-agente", elige la acción conjunta U_t de π según la cual es un vector que incluye políticas $\pi^a, a \epsilon A$, es decir $\pi = (\pi^1, \dots \dots \pi^n)$. Este meta-agente aprende la Q-función $Q(s,u) = E^{\pi}[G_t|S_t = s, U_t = u]$ que condiciona el estado y la acción conjunta de todos los agentes. Entonces, todos los algoritmos antes mencionados para un problema de RL de un solo agente se pueden aplicar a este problema. Sin embargo, el meta-agente necesita elegir la acción conjunta U_t en $\prod_{a \in A} U^a$ en este espacio crece exponencialmente el número de agentes. Por lo tanto, este problema hace que este enfoque no sea escalable y, por lo tanto, poco práctico para problemas con una gran cantidad de agentes.

Otro enfoque que se puede aplicar a problemas de RL multi agente es Q-learning independiente (IQL) (Tan, n.d.). Este enfoque es el más simple y popular para problemas de RL de múltiples agentes, cada agente $a \in A$ ignora la presencia de otros agentes en el entorno y aprende su propia función $Q^a(s,u^a) = E^{\pi^a}[G_t|S_t = s, U_t^a = u^a]$ que condiciona al estado y a su propia acción. Comparado con el primer enfoque, IQL es atractivo porque evita el problema de escalabilidad discutido anteriormente. Como Q-learning puede extenderse a DQN para problemas profundos de RL, IQL puede extenderse a DQN independiente haciendo que cada agente realice DQN para entrenar su propia red neuronal para aproximar su propia función Q.

A pesar de la simplicidad de IQL, adolece de un problema clave: el entorno parece no estacionario desde el punto de vista de cualquier agente, ya que contiene a otros agentes que actualizan sus políticas de forma independiente a medida que avanza el aprendizaje. La no estacionariedad del entorno desde el punto de vista de cada agente descarta las garantías de convergencia que existen para Q-learning. A pesar de este problema, IQL tiene fuertes resultados empíricos (Matignon et al., 2012), sin embargo, tales resultados no implican un aprendizaje profundo.

Hay un desafío clave para aplicar IQL a problemas profundos de RL: mientras que la existencia de la memoria de repetición de experiencia ayuda a estabilizar el entrenamiento de una red neuronal profunda y también mejora la eficiencia de la muestra, ya que puede volver a aprender repetidamente de experiencias pasadas, su combinación con IQL parece problemático. Debido a la no estacionariedad del entorno desde la perspectiva de cada agente introducido por IQL, es posible que la memoria de repetición de la experiencia ya no refleje la dinámica actual del entorno que el agente debe aprender. Por lo tanto, reutilizar experiencias obsoletas puede confundir continuamente a la red neuronal y hacer que el entrenamiento sea inestable.

Capítulo 5

5. Redes Ópticas

5.1. Introducción

En el presente capítulo se revisan los aspectos teóricos que fundamentan las redes ópticas y se realiza un recuento de las diferentes generaciones de redes ópticas, siendo de importancia el concepto de redes ópticas elásticas.

5.2. Generaciones de Redes Ópticas

El campo de las redes de telecomunicaciones en los últimos años ha experimentado un crecimiento acelerado, generado por las nuevas aplicaciones que están poniendo a prueba los límites de las redes actuales. La tecnología WDM fue desarrollada para transportar altos volúmenes de ancho de banda (Mukherjee & Mukherjee, 2006), WDM se componen de fibras ópticas y nodos con todo el equipamiento necesario para establecer, enrutar y eliminar conexiones ópticas a través de la red. Cada enlace de fibra óptica tiene una capacidad útil que se denomina espectro o ancho espectral (S) medido en Gigahercios (GHz) y transporta múltiples canales de comunicación, operando cada canal en una longitud de onda diferente, en total cada fibra óptica bajo WDM tiene el potencial de proporcionar más de 50 Tera bits por segundo (Tbps) de ancho de banda. Las redes de fibra óptica se han desarrollado en tres generaciones dada la necesidad de desarrollar nuevas redes de alta capacidad, que soporten las demandas de ancho de banda y el aumento progresivo en el transporte de tráfico.

La arquitectura de red óptica de primera generación Figura 5-1(a), consta de enlaces WDM punto a punto. En cada nodo se descarga el tráfico que llega, se convierte de óptico a eléctrico, posteriormente se procesa electrónicamente y se convierte de eléctrico a óptico antes de ser enviado al siguiente nodo. Cada procesamiento implica una sobrecarga significativa en términos de complejidad de conmutación y costo de procesamiento electrónico, particularmente si la mayor parte del tráfico es de derivación. Para reducir costos surgió la necesidad de una conmutación totalmente óptica.

La arquitectura de red óptica de segunda generación Figura 5-1(b) consiste en anillos que interconectan cada nodo, ayudados por multiplexores de adición y extracción de longitud de onda (WADM), donde el tráfico se puede agregar o eliminar (Murthy & Gurusamy, 2002). Los WADM permiten la terminación de canales de longitud de onda seleccionados en una fibra, mientras que otras longitudes de onda pasan sin ser manipuladas. La solución cuenta con crosconectoras ópticas (OXCs) que pueden conmutar el tráfico de cualquier puerto de entrada a cualquier otro puerto de salida.

La arquitectura de red óptica de tercera generación Figura 5-1(c) se basa en una red de malla de fibras interconectadas por dispositivos totalmente ópticos, donde los datos pueden evitar los nodos intermedios sin pasar por las conversiones electrónicas, logrando reducir los costos asociados con la provisión de capacidad de enrutamiento y conmutación electrónica. Los dispositivos de interconexión ópticos se dividen en tres categorías: acopladores pasivos en estrella, enrutadores pasivos y conmutadores activos (Mukherjee, 2000). La estrella pasiva es un dispositivo de transmisión, una señal que llegue a una longitud de onda dada en un puerto de fibra de entrada del acoplador en estrella tendrá su potencia dividida entre todos los puertos de salida del acoplador. Un enrutador pasivo puede encaminar por separado cada una de las longitudes de onda que llegan a una fibra de entrada a la misma longitud de onda en diferentes fibras de salida. Un conmutador activo también enruta las longitudes de onda desde las fibras de entrada a las fibras de salida y puede soportar conexiones simultáneas. A diferencia de un enrutador pasivo, el conmutador activo se puede reconfigurar para cambiar el patrón de interconexión de las longitudes de onda de entrada y salida.

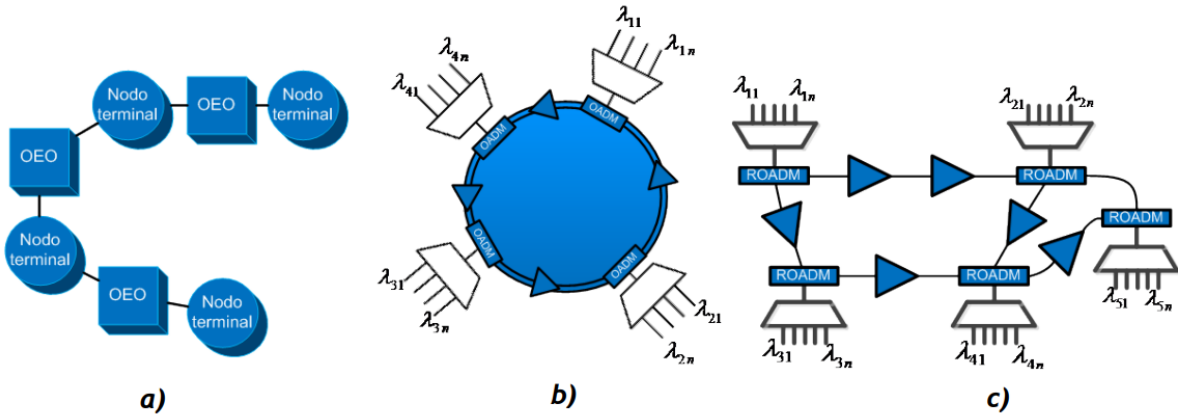


Figura 5-1 Generaciones de redes ópticas. Tomado de (Metconnex, s/f).

Entre los sistemas más populares de núcleo óptico están: conmutación de circuitos ópticos (OCS), conmutación de paquetes ópticos (OPS), conmutación de ráfagas ópticas (OBS) (Murthy & Gurusamy, 2002). En la conmutación de circuitos se establece un camino óptico de extremo a extremo entre origen y destino para toda la sesión, evitando la conversión optoelectrónica en los nodos intermedios. La conmutación de circuitos es estática y presenta dificultades para adaptarse a las ráfagas del tráfico de Internet. En la conmutación de paquetes la red de acceso se conecta directamente en el dominio óptico a lo largo de la ruta que exige funciones de almacenamiento y

procesamiento en el dominio óptico, este tipo de conmutación no es factible hasta un futuro próximo debido a las limitaciones tecnológicas (Tucker, 2008). En la conmutación de ráfagas ópticas los paquetes se concatenan en unidades de transporte denominadas ráfagas y posteriormente se conmutan a través de un núcleo óptico, sin necesidad de dedicar una longitud de onda para toda la sesión entre un par de nodos. Además, OBS es viable ya que la ráfaga de datos no necesita ser almacenada o procesada en los nodos intermedios. OBS es la propuesta para superar las deficiencias de la conmutación de circuitos y conmutación de paquetes, se dirige principalmente como una posible tecnología para apoyar la conmutación todo-óptica en Internet antes de que sea factible la conmutación de paquetes ópticos.

5.3. Conmutación de paquetes ópticos

La conmutación de paquetes ópticos (OPS) permite realizar la conmutación y enrutamiento de los paquetes IP en el dominio óptico sin realizar conversiones al dominio electrónico. Un nodo OPS tiene un esquema de conmutación que es capaz de reconfigurar para cada paquete. Los paquetes ópticos se envían junto con sus encabezados sin ninguna configuración previa en la red. En un nodo central, el paquete está retardado ópticamente utilizando fibras de retardo (FDL), mientras que el encabezado se somete a conversión óptica a electrónica y se procesa electrónicamente. En función de la información del encabezado, un conmutador está configurado para transmitir el paquete óptico desde un puerto de entrada a un puerto de salida, para liberarlo inmediatamente.

Como los recursos de red no están reservados de antemano, algunos paquetes ópticos pueden competir por el mismo puerto de salida que conduce a pérdidas de paquetes. La falta de una adecuada tecnología de almacenamiento intermedio óptico agrava el problema de contención en las redes OPS en comparación con las redes tradicionales, donde la tecnología de almacenamiento intermedio electrónico es madura. El almacenamiento en búfer óptico se realiza mediante el uso de FDL que pueden contener un paquete óptico durante un período de tiempo variable por medio de múltiples líneas de retardo en etapas (Mukherjee & Mukherjee, 2006). Para retrasar un paquete óptico durante unos pocos microsegundos, se requiere un kilómetro de fibra óptica, por esta razón un nodo OPS puede ser muy ineficiente en el manejo de cargas elevadas o tráfico de ráfagas. Otra solución al problema de contención es encaminar los paquetes contendientes a un puerto de salida diferente al previsto. Esta técnica se conoce como enrutamiento de deflexión. El enrutamiento de deflexión puede causar la entrega de paquetes en bucle y fuera de orden, y requiere más estudios para abordar estas limitaciones.

5.4. Conmutación de ráfagas ópticas

La red de conmutación de ráfagas ópticas está compuesta por: nodos de entrada, nodos de salida y un "CORE" de nodos centrales. En el nodo de ingreso se agregan diferentes tipos de tráfico datos del cliente y la red de acceso agrega una ráfaga de datos que se transmite por completo en el dominio óptico. Los paquetes destinados al mismo nodo de salida con el mismo nivel de servicio se agregan en una misma cola de ensamblaje de ráfaga. Para evitar el almacenamiento en búfer y el procesamiento de la ráfaga de datos ópticos en los nodos intermedios llamados nodos centrales, un paquete de control también denominado paquete de cabecera de ráfaga (BHP) viaja por adelantado con la información sobre la longitud y el tiempo de llegada de la ráfaga de datos. Los BHP se transmiten en una longitud de onda de control dedicada, mientras que las ráfagas de datos se envían después de un tiempo en longitudes de onda separadas. El intervalo de tiempo entre un BHP y la ráfaga de datos se denomina tiempo de compensación que se establece lo suficiente como para permitir el procesamiento de BHP y configurar los conmutadores en los nodos centrales. Una vez que una ráfaga alcanza el nodo de salida, se desensambla en paquetes que se enrutan a través de la red de acceso.

Los enrutadores centrales están compuestos por una OXC y una unidad de control de conmutación (SCU). La SCU crea y mantiene una tabla de reenvío y es responsable de configurar el OXC (Vokkarane & Jue, 2003). Cuando la SCU recibe un paquete de encabezado de ráfaga, identifica el destino deseado y consulta al procesador de señalización del enrutador para encontrar el puerto de salida deseado. Si el puerto de salida está disponible cuando llega la ráfaga de datos, la SCU configura el OXC para permitir el paso de la ráfaga de datos. Si el puerto no está disponible, entonces el OXC se configura dependiendo de la política de resolución de contención implementada en la red. En general, la SCU es responsable de la interpretación del encabezado, programación, detección y resolución de colisiones, búsqueda de tabla de reenvío, control de matriz de conmutación, reescritura de encabezado y control de conversión de longitud de onda. En el caso de una ráfaga de datos que ingresa al OXC antes de su paquete de control, la ráfaga simplemente se descarta (se conoce como problema de arribo temprano).

5.5. Multiplexación por división de frecuencias ortogonales

La multiplexación por división de frecuencia Ortogonales (OFDM) es una clase de modulación multiportadora que puede transmitir altas velocidades dividiendo los datos entre un número de canales ortogonales, cada canal ortogonal cuenta con una baja velocidad de datos (Shieh & Djordjevic, 2010). Comparando OFDM con WDM, el espaciamiento en WDM entre canales a nivel de espectro debe estar predefinido en el diseño de la red, OFDM permite que las subportadoras se solapen gracias a las características de ortogonalidad alcanzando eficiencia en la administración de los recursos espectrales.

En la Figura 5-2 se detalla la administración del espectro en OFDM donde es posible contar con una grilla espectral flexible para acomodar tráficos de diferentes tamaños, en la gráfica de WDM se observa una rejilla fija de 50 GHz donde se puede acomodar tráfico hasta de 100 Gbps. La condición de ortogonalidad entre varias subportadoras se logra cuando las frecuencias centrales están espaciadas n/Ts, Donde Ts es la duración del símbolo y n un entero.

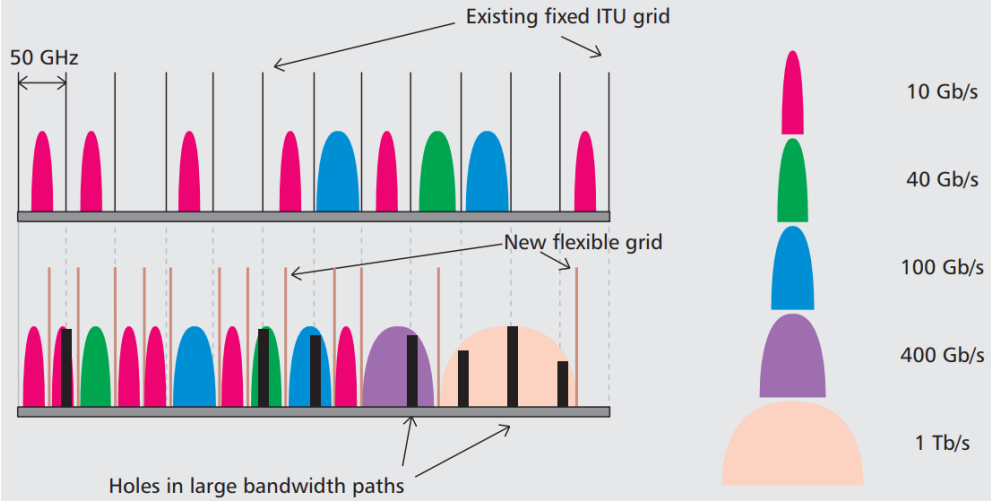


Figura 5-2 Comparación de uso del espectro en DWDM y OFDM. Tomado de (Gerstel et al., 2012).

En la Figura 5-3(a) se observa que el punto máximo en amplitud de una subportadora corresponde al punto cero del espectro de otras subportadoras, por lo tanto, las demás subportadoras no interfieren con la subportadora que se está muestreando. La ortogonalidad conduce al uso eficiente de los recursos espectrales. En la Figura 5-3(b), se observa la señal OFDM en el domino del tiempo es una síntesis de formas de onda de múltiples subportadoras. En las redes totalmente ópticas el transmisor genera múltiples subportadoras desde una fuente de luz continua, cada subportadora es modulada de forma individual y finalmente acoplada en la señal óptica OFDM, la condición de ortogonalidad se satisface a través de una forma de pulso adecuada y bloqueo de fase de la subportadora de frecuencia ortogonal.

5.6. Multiplexación por división espacial

La multiplexación por división espacial (SDM) se ha fortalecido gracias al concepto de copropagación en modos y núcleos espaciales. Este tipo de multiplexación funciona con fibras especiales en las cuales los modos son guiados por múltiples núcleos en una sola fibra o por un solo núcleo más grande con índice de refracción modificado (Morioka et al., 2012). Algunos investigadores se han centrado en la adición de canales espaciales mediante la exploración de ciertas características o combinaciones de estructuras de fibra. En Suecia desarrollaron los conceptos de SDM para lograr tasas de transmisión altas basados en la teoría de la información, el procesamiento de señales digitales y la codificación (*Technologies for*

Spatial-Division Multiplexing: The next Frontier in Optical Communications | Chalmers, s/f). En la búsqueda para aumentar la capacidad de fibra óptica a través de núcleos espaciales, investigadores produjeron, desarrollaron y optimizaron los componentes de SDM para apoyar la implementación de fibras con múltiples núcleos (*Optical Technologies for Spatial Division Multiplexing (SDM) SDMThechs*, s/f).

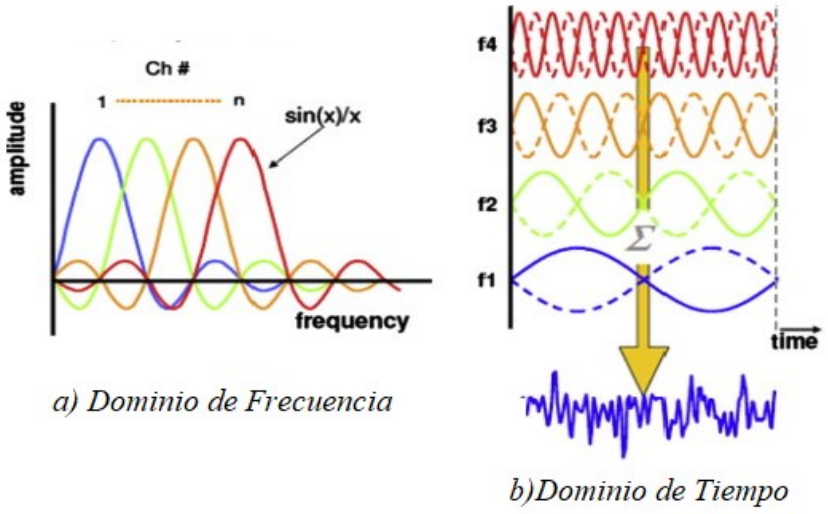


Figura 5-3 ODFM: a) Dominio de frecuencia, b) Dominio de tiempo. Tomado de (Eng & Kocot, 2013)

5.7. Tráfico dinámico

En entornos de tráfico dinámico, el grado de servicio de la red es medida mediante la probabilidad donde una petición de conexión no pueda ser aceptada. Esta probabilidad se denomina probabilidad de bloqueo (BP) y es la figura más usada para evaluar la efectividad de la red. Se puede considerar una BP del 1% como el valor más alto admisible para entregar servicio a los clientes. Aunque la red tenga capacidad sobrante, no se puede permitir más intensidad, decaería el grado de servicio a un nivel no deseable. Nuevamente se evidencia que no es lo mismo rechazar una demanda de 400 Gbps que una de 10 Gbps, ya que la cantidad de ancho de banda rechazado es mucho menor. Así, en redes ópticas elásticas surgen dos medidas de probabilidad de bloqueo: la parcial de cada tipo de cliente y la ponderada, haciendo uso de las proporciones del perfil de tráfico y de sus velocidades. A la probabilidad de bloqueo ponderada se le llama con precisión como la proporción de ancho de banda rechazado. Finalmente, para definir la cantidad media de ancho de banda total que transporta una red ($BWtotal$) en un determinado momento es necesario conocer la intensidad global (I), para cada tipo de cliente i su proporción en el perfil de tráfico (TPi), su ancho de banda demandado (Bwi), y su probabilidad de bloqueo parcial (PBi). Así, el cálculo se puede efectuar como sigue en la Ecuación 6. Una muy buena aproximación de BW total se puede obtener si no se considera la probabilidad de bloqueo. Esto ocurre cuando restringimos ésta a valores pequeños (e.g. 1%).

$$BW_{total} = I \sum_{Vi} Bw_i * TP_i * (1 - PB_i)$$

Ecuación 6 Aproximación de BW total. Tomado de (Jinno et al., 2009).

Los enfoques de tráfico utilizados por los algoritmos de asignación de espectro se pueden clasificar según el tipo de aplicación. El enfoque de tráfico estático se ocupa del enrutamiento y la asignación de recursos durante la fase de planificación de la red, donde una matriz de tráfico proporcionada y las operaciones de asignación y enrutamiento del espectro se realizan fuera de línea (Jinno et al., 2009). El tráfico variable aborda dinámicamente el enrutamiento y la asignación de demandas a medida que las solicitudes de tráfico llegan a la red, realizando operaciones de enrutamiento en línea y asignación de espectro (Zhang et al., 2013).

5.8. Redes ópticas elásticas

Las redes ópticas elásticas tienen como motivación la transmisión eficiente de canales de baja capacidad (10/40 Gbps) y la transmisión de señales de alta capacidad (100/400 Gbps) (Yonenaga et al., 2008). Para conseguir la transmisión de canales de baja capacidad se necesita sustituir la rejilla espectral actual por una nueva de granularidad, permitiendo que formatos de modulación con alta eficiencia espectral (EE), sean explotados en su totalidad. En la Tabla 5.1, se expone un resumen de distintos formatos de señal y la ganancia en EE que se podría obtener aplicando rejilla flexible con respecto a los sistemas actuales basados en rejilla de 50GHz (Sano et al., 2009).

La reconfiguración del espectro por la rejilla flexible posibilita asignar las subportadoras de acuerdo con la demanda del usuario, permitiendo que las longitudes de onda no utilizadas se asignen para conformar otro Super canales. Esto da como resultado un uso eficiente del ancho de banda de la fibra óptica y una disminución en el consumo de energía de la red, optimizando la eficiencia energética, lo que a su vez permite la reducción del costo por bit transmitido

Las principales características de una red óptica elástica son la segmentación del ancho de banda, la agregación del ancho de banda, la acomodación eficiente de múltiples velocidades de datos, la variación elástica de los recursos asignados, la velocidad de línea adaptable al alcance, etc.

5.8.1. Segmentación del ancho de banda

Las redes ópticas tradicionales requieren la asignación completa de la capacidad de longitud de onda a una ruta óptica entre un par de nodos finales. Sin embargo, las redes ópticas elásticas proporcionan un mecanismo de segmentación de ancho de banda eficiente en el espectro que proporciona un servicio de conectividad de ancho de banda fraccional. Si solo

se requiere un ancho de banda parcial, la red óptica elástica puede asignar el ancho de banda óptico suficiente para acomodar el tráfico del cliente. Al mismo tiempo, cada nodo de la ruta óptica asigna una conexión cruzada con el ancho de banda de espectro apropiado para crear una ruta óptica de extremo a extremo de tamaño apropiado. El uso eficiente de los recursos de la red permitirá el aprovisionamiento rentable del servicio de ancho de banda fraccional con diferentes números de bits modulados por símbolo. Para establecer una conexión, cada conexión cruzada variable de ancho de banda en la ruta asigna una conexión cruzada con espectro suficiente para crear una ruta óptica de extremo a extremo del tamaño adecuado.

Tabla 5.1 Resumen de distintos formatos de señal y ganancia. Tomado de (Gerstel et al., 2012).

Demand bit rate (Gb/s)	Modulation format	Channel bandwidth (GHz)	Fixed grid solution	Efficiency increase for EON
40	DP-QPSK	25+10	1 50 GHz channel	35 GHz vs. 50 = 43%
100	DP-QPSK	37.5+10	1 50 GHz channel	47.5 GHz vs. 50 = 5%
100	DP-16QAM	25+10	1 50 GHz channel	35 GHz vs. 50 = 43%
400	DP-QPSK	75+10	4 100 Gb/s in 4 50 GHz channels	85 GHz vs. 200 = 135%
400	DP-16QAM	75+10	2 200Gb/s in 2 50 GHz channels	85 GHz vs. 100 = 17%
1000	DP-QPSK	190+10	10 100G in 10 50 GHz channels	200 GHz vs. 500 = 150%
1000	DP-16QAM	190+10	5 200Gb/s in 5 50 GHz channels	200 GHz vs. 250 = 25%

5.8.2. Agregación de ancho de banda

La agregación de enlaces es una tecnología de red de paquetes estandarizada en IEEE 802.3. Combina múltiples puertos/enlaces físicos en un conmutador/enrutador en un solo puerto/enlace lógico para permitir un crecimiento incremental de la velocidad del enlace a medida que la demanda de tráfico aumenta más allá de los límites de un solo puerto/enlace. De manera similar, la red óptica elástica habilita la función de agregación de ancho de banda y puede crear una ruta óptica de súper longitud de onda combinada de forma contigua en el dominio óptico, lo que garantiza una alta utilización de los recursos espectrales.

5.8.3. Acomodación eficiente de múltiples velocidades de datos

La red óptica elástica tiene la capacidad de proporcionar la acomodación directa espectralmente eficiente de velocidades de bits de datos mixtas en el dominio óptico debido a su asignación de espectro flexible. Las redes ópticas tradicionales con rejilla fija conducen al despilfarro del ancho de banda óptico debido a la excesiva separación de frecuencias para señales de baja tasa de bits.

5.8.4. Tasa de línea adaptable al alcance

La red óptica elástica tiene la capacidad de admitir una tasa de línea adaptable al alcance, así como la expansión y contracción del ancho de banda dinámico, al alterar el número de subportadoras y formatos de modulación.

III. Desarrollo

Capítulo 6

6. Modelo Empleado

Consolidado el estado del arte, se plantea un modelo que permita utilizar de manera efectiva los recursos físicos de una red óptica elástica mediante el aprendizaje por refuerzo.

6.1. Problema de asignación de enrutamiento, modulación, espectro y núcleo

Las redes ópticas elásticas son la tecnología más prometedora para la próxima generación de redes troncales ópticas, para proporcionar ancho de banda de manera flexible a las solicitudes de conexión en la capa óptica es imperativo proporcionar la tecnología que brindará la flexibilidad. Posterior a seleccionar la tecnología para brindar la flexibilidad se debe dividir el espectro de acuerdo con la norma de la ITU, esta división de espectro aportará anchos de banda finos, los cuales se pueden agrupar para solucionar diferentes demandas de tráfico dinámico. Hasta este punto y contando con un algoritmo que enrute todas las demandas desde un origen a un destino, se solucionaría el problema de enrutamiento y la asignación de espectro (RSA). Para mejorar la eficiencia espectral se pueden adoptar diferentes formatos de modulación lo cual convierte el problema de RSA en un problema de enrutamiento, modulación y asignación de espectro (RMSA). Gracias al desarrollo de la tecnología SDM, donde cada fibra contiene varios núcleos con los mismos recursos espectrales, la capacidad de transmisión aumenta, convirtiendo el problema de RSMA en un problema de enrutamiento, modulación, espectro y asignación de núcleos (RMSCA).

El problema RSA contiene tanto la decisión de enrutamiento para cada nodo como la decisión de asignar subportadoras a la matriz de demandas. El problema RSA ya tiene características

de un problema NP-completo cuando se basa en la optimización de asignación del espectro, por lo tanto, el problema RMSCA al contar con ítem más que es el de asignación espacial también se cataloga como NP-completo.

6.2. Red óptica elástica

Como solución al problema RMSCA la implementación de elasticidad en las redes ópticas se soporta en la tecnología O-OFDM. Por medio de este esquema especial de modulación multiportadora se transmite un flujo de datos a alta velocidad dividido dentro de cierto número de canales ortogonales (subportadoras). Como se ilustra en la Ecuación 7, una señal S_t bajo modulación multiportadora se representa como:

$$\mathrm{s}(t) = \sum_{i=-\infty}^{+\infty} \sum_{k=1}^{N_{sc}} c_{ki} s_k (t - iT_s)$$

$$s_k(t) = \prod(t) e^{j2\pi f_k t}$$

$$\prod(t) = \begin{cases} 1, (0 \leq t \leq T_s) \\ 0, (t \leq 0, t > T_s) \end{cases}$$

Ecuación 7 Modulación multiportadora. Elaboración propia.

Donde c_{ki} es el i-ésimo símbolo de información en la k-ésima subportadora, s_k es la forma de onda de la k-ésima subportadora, N_{sc} es el número de subportadoras, f_k es la frecuencia de la subportadora, T_s es el período del símbolo y $\prod(t)$, es la función pulso- moldeadora. El detector de cada subportadora utiliza un filtro que coincide con la forma de onda de la subportadora. Por lo tanto, como se presenta en la Ecuación 8 el símbolo de información detectado c'_{ki} viene dado por:

$$c'_{ki} = \int_0^{T_s} r(t - iT_s) s_k^* dt = \frac{1}{T_s} \int_0^{T_s} r(t - iT_s)\, e^{-j2\pi f_k} dt$$

Ecuación 8 Símbolo de información. Tomado de (Talebi et al., 2014).

Donde $r(t)$, es la señal recibida en el dominio del tiempo. La condición ortogonal de una señal OFDM se origina a partir de una correlación presentado en la ecuación 9, entre dos subportadoras cualesquiera, dada por:

$$d_{kl} = \frac{1}{T_s}\int_0^{T_s} s_k s_l^* dt = \frac{1}{T_s}\int_0^{T_s} e^{j2\pi(f_k - f_l)^t} dt = e^{j\pi(f_k - f_l)T_s}\frac{\sin(\pi((f_k - f_l)T_s))}{\pi((f_k - f_l)T_s)}$$

Ecuación 9 Correlación. Tomado de (Talebi et al., 2014).

Si la condición:

$$f_k - f_l = m\frac{1}{T_s}$$

Si se satisface la condición, las dos subportadoras son ortogonales entre sí, puede verse que estos conjuntos de subportadoras ortogonales, con sus frecuencias espaciadas en múltiplos del inverso de los períodos del símbolo, pueden recuperarse con los filtros adaptados en la Ecuación 9 sin interferencia entre portadoras, a pesar de la fuerte superposición espectral. La modulación y demodulación OFDM se pueden implementar utilizando la transformada de Fourier discreta inversa (IDFT) y la transformada de Fourier discreta (DFT), respectivamente. El valor discreto de la señal OFDM transmitida $s(t)$ es una IDFT de N puntos del símbolo de información c_k y el símbolo de información recibido $c_k^{'}$ es una DFT de N puntos de la señal muestreada recibida $r(t)$. Para reducir la complejidad computacional de DFT/IDFT, las funciones eficientes de transformada rápida de Fourier y transformada rápida inversa de Fourier (FFT/IFFT) se utilizan normalmente en los sistemas OFDM para implementar la modulación y demodulación OFDM.

6.2.1. Unidades de ranuras de frecuencia

Para generar los recursos espectrales a asignar en el problema RMSCA y contar con una red óptica de velocidad de datos variable y ancho de banda flexible, se necesita una especificación de ranuras de espectro flexible que pueda acomodar eficientemente volúmenes de tráfico de longitud de onda inferior, longitud de onda superior y velocidad mixta. Como solución a esta necesidad se presenta la recomendación de la ITU G.694.1, esta recomendación divide el espectro en subsecciones (*G.694.1 : Planes Espectrales Para Las Aplicaciones de Multiplexación Por División de Longitud de Onda: Plan de Frecuencias Con Multiplexación Por División de Longitud de Onda Densa*, n.d.). Las subsecciones se denominan "unidades de ranuras de frecuencia" (FSU), las cuales se pueden agrupar para proporcionar espectro suficiente para cada usuario de acuerdo con sus necesidades.

Los FSU asignados a cada conexión deben satisfacer tres restricciones del espectro: la no superposición, la continuidad y la contigüidad. Estas limitaciones conducen al problema de la fragmentación del espectro, que afecta el uso de los recursos espectrales. Una alta fragmentación del espectro conlleva a aumentar la probabilidad de bloqueo de las demandas de tráfico.

En la Figura 6-1 se presenta el ejemplo de asignación de FSU donde con una frecuencia nominal central de 193.1 THz y al recibir una solicitud de ancho de banda espectral de 50 GHz, se resuelve separando 4 FSU de 12.5 GHz donde n= 0 y m=4, solucionando el requerimiento de recursos espectrales.

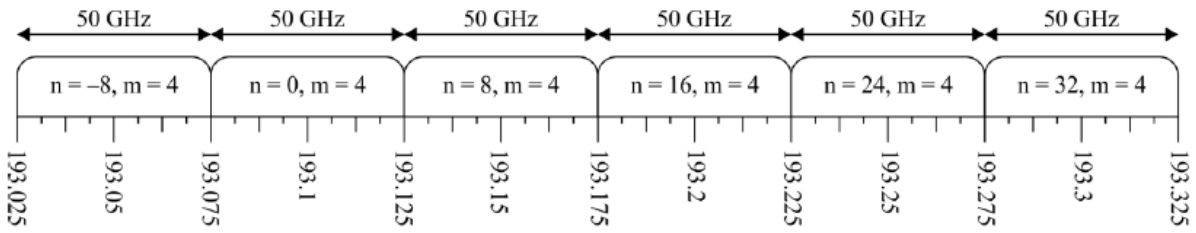


Figura 6-1 Ejemplo asignación espectral. Tomado de (G.694.1 : Planes Espectrales Para Las Aplicaciones de Multiplexación Por División de Longitud de Onda: Plan de Frecuencias Con Multiplexación Por División de Longitud de Onda Densa, n.d.).

6.2.2. Fibra multinúcleo

Los sistemas de transmisión óptica basado en SDM aumentan significativamente la capacidad de transporte de la red ópticas, donde cada fibra contiene varios núcleos con los mismos recursos de espectro, debido a este nuevo espacio RSMA se convierte en RMSCA.

La fibra de 7 núcleos es la configuración más común en SDM-EON (Brasileiro et al., 2020), la estructura de la fibra de 7 núcleos se muestra en la Figura 6-2. Esta fibra incluye seis núcleos marginales y un núcleo central, cada uno de los núcleos marginales tiene tres vecinos y el núcleo central tiene seis. Por ejemplo, los vecinos del núcleo 1 son los núcleos {2, 6, 7} y los vecinos del núcleo 7 son los núcleos {1,2,3,4,5,6}.

Para la implementación del diseño de la red óptica elástica el algoritmo deberá contar con la configuración de 2 a 7 núcleos. Dado que las señales que se transmiten en núcleos adyacentes de la misma fibra interfieren entre sí, la diafonía entre núcleos (XT) debe considerarse en el problema RMSCA (Li et al., 2022). La diafonía entre núcleos hace que el problema de RMSCA sea más complicado que el problema RMSA.

La diafonía que afecta una conexión está determinada por factores como la longitud del camino de luz que transporta la conexión, el número de núcleos adyacentes ocupados en cada enlace del camino de luz y el coeficiente de acoplamiento de potencia de las fibras. Adicionalmente, para el problema de RMSCA en línea, la configuración dinámica y la eliminación de conexiones hacen que la diafonía de cada conexión cambie dinámicamente, por lo que es difícil registrar la diafonía de cada conexión en tiempo real.

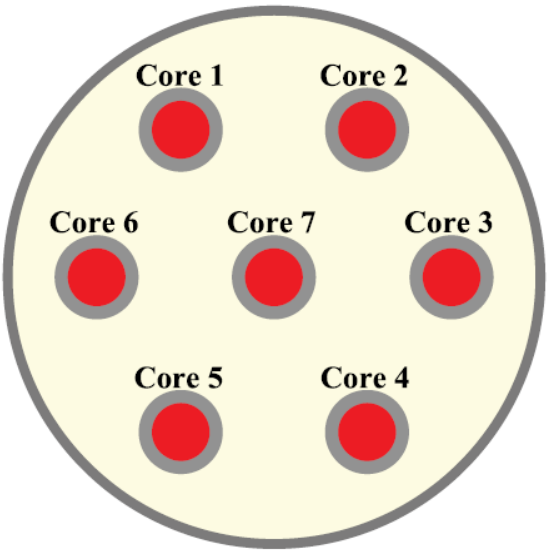


Figura 6-2 Fibra de 7 núcleos. Tomado de (Li et al., 2022).

Se calcula la diafonía en un camino de luz de acuerdo al método propuesto por Klinkowski y Walkowiak (Klinkowski & Walkowiak, 2018), la diafonía d del f-ésimo FSU en el núcleo c del camino de luz P está dada por:

$$XT_P(c) = \sum_{e \epsilon E(P)} A(e,c).h(e).L(e) \leq XT_{th}$$

Ecuación 10 Diafonía sufrida por el f-ésimo FS. Elaboración propia.

Donde, $E(P)$ es el conjunto de enlaces de la ruta P, $L(e)$ es la longitud del enlace e, $h(e)$ es el coeficiente de acoplamiento de potencia del enlace e, $A(e,c)$ es el número de núcleos adyacentes ocupados del núcleo c y XT_{th}, es el umbral de diafonía, es decir, la diafonía máxima que puede tolerar una conexión.

6.2.3. Fragmentación

La fragmentación de espectro hace referencia a los espacios espectrales que quedan sin asignar debido a la asignación dinámica del espectro. Las ranuras de espectro asignadas se dispersan por asignación dinámica de caminos de luz, debido a la eliminación de los caminos de las demandas de tráfico caducadas. Los FSU disponibles en los enlaces se desalinean, provocando el bloqueo de las demandas de tráfico posteriores a pesar de tener suficientes FSU disponibles pero dispersos. Por lo tanto, la fragmentación del espectro es un problema importante en las EON, lo que deteriora la eficiencia del espectro y el rendimiento de la red. Como ejemplo, la Figura 6-3 ilustra un camino de luz candidato, este camino se origina en el nodo a y termina en el nodo e,al asignar las ranuras de espectro con índices 4 al 6. Esto

generará una penalización por fragmentación, ya que en la ranura 3 hay disponibilidad entre el camino de luz candidato y sus conexiones adyacentes en ambos enlaces.



Figura 6-3 Fragmentación. Elaboración propia.

La asignación de recursos en SDM debe satisfacer todas las restricciones de asignación de recursos, estas restricciones son: contigüidad del espectro, continuidad del espectro, continuidad del núcleo y diafonía. Las formas de solucionar la desfragmentación evolucionan a medida que la red se reconfigura continuamente, el objetivo es empaquetar lo más cerca posible las conexiones establecidas y mejorar el rendimiento de la red.

Para mejorar esta situación, los núcleos se clasifican en función del tamaño de los intervalos de frecuencia para transmitir diferentes tipos de servicios, reduciendo así la generación de fragmentos de espectro y mejorando la utilización de los recursos de la red. Para cualquier demanda de un origen A, un destino B y haciendo un único tránsito por C, no se presenta restricción de cambio de canal, es decir, si en el camino de A-C para la demanda se usó la frecuencia central 193.1 THz con n=0 y m=4, en el camino C-B puede usar la misma posición espectral n=0 y m=4 o cualquier otra posición por ejemplo n=8 y m=4.

La asignación de recursos espaciales se realiza por orden de arribo y teniendo en cuenta las prioridades de cada núcleo, en el núcleo $c1$ se transportan las demandas correspondientes al segmento hogares que son de hasta 10 Gbps, en el núcleo $c2$ se transportan las demandas que corresponden al segmento corporativo que están desde los 10 Gbps hasta los 100 Gbps y finalmente, en el núcleo $c3$ se transportan las demandas correspondientes al tráfico entre Datacenter o de alta capacidad que están entre los 100 Gbps y los 1000 Gbps. En caso de contar con fibras multinúcleo de 7 núcleos la distribución será: 2 núcleos para servicios de hogar, 2 núcleos para servicios corporativos y 3 núcleos para tráfico de alta capacidad. De esta forma, al asignar servicios en cada núcleo, se disminuirán los servicios bloqueados por fragmentación del espectro.

6.2.4. Requerimientos de espectro en términos de FSU vs máxima distancia de alcance

El conjunto de la máxima distancia de alcance (MAR) en cada formato de modulación está compuesto por la velocidad de bits disponible en los transpondedores de cada umbral BER. Para calcular el alcance máximo de cada combinación de formato de modulación y BER se usa el modelo de ruido gaussiano (GN) para estimar una relación señal/ruido (SNR) recibido después de la transmisión (Poggiolini et al., 2014). El modelo GN describe la propagación

no lineal en sistemas de transmisión coherentes sin compensación de dispersión y asume que las no linealidades de la fibra se pueden modelar como ruido gaussiano blanco aditivo que se combina de manera incoherente con el ruido de emisión espontánea amplificada (ASE).

Sobre la potencia de lanzamiento óptima se obtiene el MAR que se define como el número máximo de tramos que se pueden transmitir para un umbral de BER con una potencia de lanzamiento óptima. La SNR está relacionada con BER, donde los valores más altos de SNR permiten valores más bajos de BER. Por ejemplo, la modulación por cambio de fase en cuadratura (QPSK) tiene un requisito de SNR bajo, lo que implica que es posible transmitir a larga distancia. Por otro lado, las constelaciones de modulación de amplitud en cuadratura (QAM) más altas pueden transportar más tráfico con requisitos de SNR más altos, obteniendo distancias más cortas.

Como se asume que todos los componentes están cuidadosamente diseñados para evitar cualquier fuente adicional de ruido, degradación de la señal y diafonía, la SNR recibida se usa como una métrica de rendimiento y se calcula según la Ecuación 10.

$$SNR = \frac{P_{in}}{\eta_{ASE} + \eta_{NLI}}$$

Ecuación 11 SNR. Tomado de (Poggiolini et al., 2014).

Donde P_{in} es la potencia de entrada promedio, η_{ASE} es la potencia de ruido ASE y η_{NLI}, es la potencia de ruido de interferencia no lineal (NLI).

Se considera un sistema coherente de ondas de luz que transmite información utilizando una sola polarización, donde la dispersión cromática y la dispersión del modo de polarización se compensan idealmente mediante el procesamiento de señales digitales. Además, la modulación de fase cruzada y la modulación de fase propia son los principales efectos no lineales. Los formatos de modulación considerados son BPSK, QPSK, 8-QAM, 16-QAM, 32-QAM y 64-QAM, cada uno con un umbral de BER de 10−6, 10−9 y 10−12.

Todos los usuarios de red tienen diferentes demandas de velocidad de bits, cada solicitud se empaqueta entre valores de 10, 40, 100, 400 y 1000 Gbps dependiendo de la demanda y utilización. Para traducir las demandas de tasa de bits en requisitos de espectro, la Tabla 6.1 detalla los requisitos de espectro en términos del número de FSU y GHz para cada combinación de tasa de bits y formato de modulación. Posteriormente, para una solicitud de transmisión determinada, se calcula la demanda de ancho de banda en función de la solicitud de tasa de bits del usuario y el mejor formato de modulación posible donde la distancia de comunicación no supere la máxima distancia de alcance (MAR) como se especifica en la Tabla 6.2.

Tabla 6.1 Requisitos de espectro en términos de ancho de banda, formato de modulación y velocidad. Tomado de (Calderón et al., n.d.).

Bit rate (Gbps)	Modulation format	Bandwidth (FSU)	Bandwidth (GHz)
10	BPSK	1	12.5
40	BPSK	4	50
40	QPSK	2	25
40	8-QAM	2	25
40	16-QAM	1	12.5
40	32-QAM	1	12.5
40	64-QAM	1	12.5
100	BPSK	8	100
100	QPSK	4	50
100	8-QAM	3	37.5
100	16-QAM	2	25
100	32-QAM	2	25
100	64-QAM	2	25
400	BPSK	32	400
400	QPSK	16	200
400	8-QAM	11	137.5
400	16-QAM	8	100
400	32-QAM	7	87.5
400	64-QAM	6	75
1000	BPSK	80	1000
1000	QPSK	40	500
1000	8-QAM	27	337.5
1000	16-QAM	20	250
1000	32-QAM	16	200
1000	64-QAM	14	175

Tabla 6.2 Máxima distancia de alcance (MAR) para formatos de modulación y BER. Tomado de (Calderón et al., n.d.).

Modulation Format	Maximum achievable reach (km)		
	$BER = 10^{-6}$	$BER = 10^{-9}$	$BER = 10^{-12}$
BPSK	5520	3440	2720
QPSK	2720	1680	1360
8-QAM	1360	720	560
16-QAM	560	320	240
32-QAM	240	160	80
64-QAM	80	80	0

6.2.5. Topología

En la Figura 6-4 se representa la topología de un segmento de red óptica elástica, donde se observa cómo cambian los conjuntos de longitudes de onda señalizadas por diferentes colores entre los nodos origen y destino. Esto posibilita que la red óptica reutilice longitudes de onda en algunas trayectorias donde no se utilizan y asignarlas de manera eficiente en la

implementación de otra demanda o en la conmutación de demandas provocadas por alguna falla en hardware o corte de fibra óptica.

Los dispositivos físicos que componen la topología incluyen fibras ópticas multi núcleo, amplificadores ópticos, regeneradores ópticos, divisores, crosconectoras ópticas, tarjetas de línea y elementos de conmutación. Cada dispositivo puede tener una ganancia (o pérdida) inherente, una potencia de ruido generada y una forma de filtro óptico. Cada nodo cuenta con conversores de onda totalmente ópticos para optimizar los recursos físicos en espacio y espectro. Estos conversores tienen la capacidad de enrutar longitudes de onda sin recurrir a la conversión óptica-eléctrica-óptica y además tienen la capacidad de reconfigurar el espaciamiento entre canales. Esta reconfiguración posibilita asignar las subportadoras de acuerdo con la demanda del usuario, permitiendo que las longitudes de onda no utilizadas se asignen para conformar otros canales. Esto da como resultado un uso eficiente del ancho de banda de la fibra óptica y una disminución en el consumo de energía de la red, optimizando la eficiencia energética, lo que a su vez permite la reducción del costo por bit transmitido.

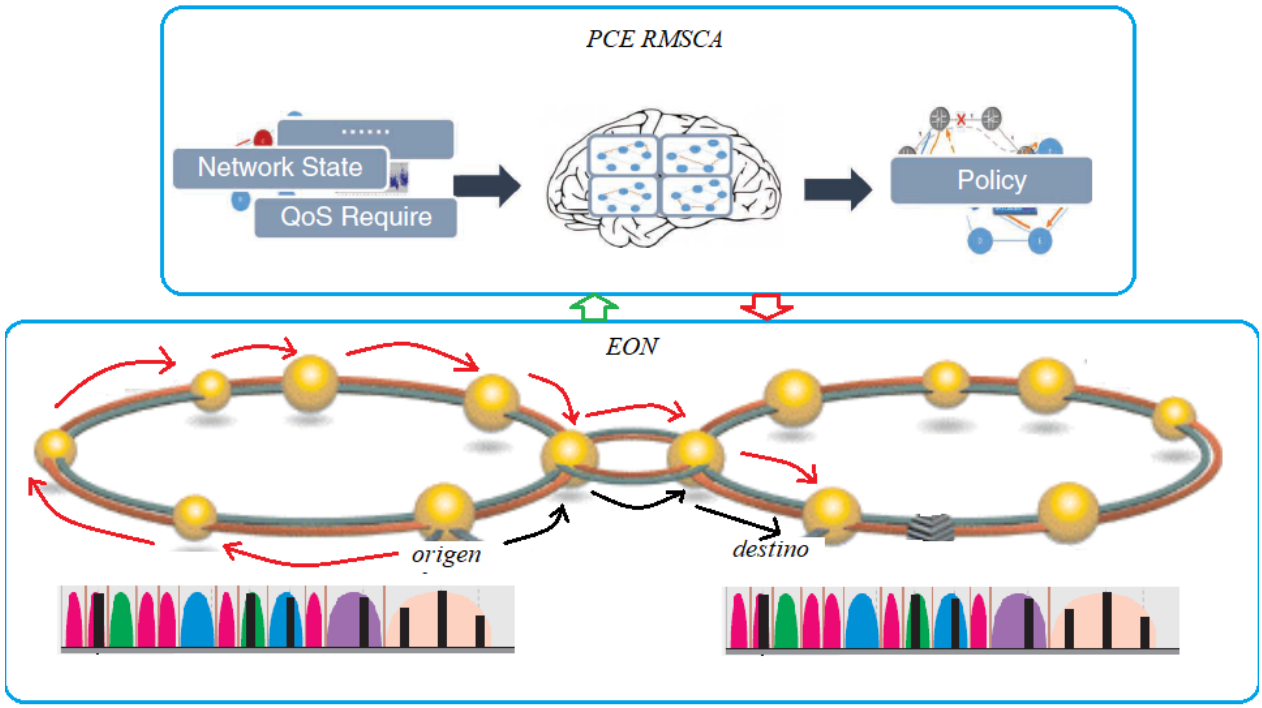


Figura 6-4 Topología. Elaboración propia.

El servidor de elemento de cálculo de ruta (PCE) proporciona una función de cálculo de ruta centralizada, este se conecta con los equipos de red óptica (NE) por medio de la controladora óptica de cada NE y se ejecutan los cambios solicitados para cada demanda. La centralización del cálculo de rutas permite un control de las políticas y algoritmos de enrutamiento desde

una única ubicación. La parte exitosa y madura de la arquitectura permanece en los NE. Esta arquitectura permite: enrutamiento sin problemas entre dominios, cálculo de ruta flexible/personalizable, mejor precio/rendimiento y operaciones simplificadas.

Cada demanda de tráfico es recibida por el PCE, este reporta a cada NE con la señalización para la separación de los recursos, la tarjeta de línea implementa el ancho de banda recibido y lo entrega a la crosconectora óptica de ancho de banda variable quien asignar una conexión cruzada de tamaño apropiado y la modulación indicada desde el PCE. Como paso siguiente selecciona el núcleo de salida por medio de un multiplexor de núcleos ópticos y finalmente se amplifica para llegar al siguiente NE. En el NE (t+1) se desagrega la demanda sobre la crosconectora óptica, si es el NE final se desagrega sobre la tarjeta de línea y se entrega la capacidad por el tiempo requerido, en caso contrario se crosconecta el servicio sobre la crosconectora y amplificadora del siguiente salto.

6.3. Esquema aprendizaje por refuerzo

El modelo seleccionado como referencia tiene como base la arquitectura DQN propuesta por Mnih omitiendo el conjunto de capas convolucionales ya que la entrada a la red neuronal no es una imagen. La entrada a la red neuronal es una representación de estado y hay una unidad de salida separada para cada acción posible. Para lograrlo Mnih desarrolló como agente novedoso, una red Q profunda (DQN), que es capaz de combinar aprendizaje por refuerzo con redes neuronales profundas. Los avances recientes en las redes neuronales profundas utilizan varias capas de nodos para construir representaciones progresivamente más abstractas de los datos, esto genera que las redes neuronales artificiales aprendan conceptos como categorías de objetos directamente a partir de datos sin procesar.

Durante las tareas el agente interactúa con un entorno a través de una secuencia de observaciones, acciones y recompensas. El objetivo del agente es seleccionar acciones de una manera que maximice la recompensa futura acumulada. Más formalmente como se ilustra en la Ecuación 12, se usa una red neuronal convolucional profunda para aproximar la función de valor de acción óptima.

$$Q^*(s,a) = \max_{\pi} E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots .. | s_t = s, a_t = a, \pi]$$

Ecuación 12 Función de valor de acción óptima. Tomado de (Mnih et al., 2015).

Se considera una red neuronal de 3 capas donde las dos primeras capas están completamente conectadas, cada una consta de 512 unidades rectificadoras. La capa de salida es una capa lineal completamente conectada con una salida única para cada acción posible.

Para Duelo DQN, se sigue la arquitectura Wang que está conformada por dos secuencias de capas totalmente conectadas, la primer y tercer capa tiene una sola salida para cada acción posible, mientras que la segunda capa tiene una sola salida (Wang et al., 2015), estos dos flujos de salida de las dos capas se combinan usando la Ecuación 5. La memoria PER se implementa mediante la priorización proporcional de Schaul muestreando un mini lote de tamaño k (Schaul et al., 2015).

En el esquema de DDQN se tienen dos redes neuronales a lo largo del proceso de aprendizaje, una se denomina red en línea y la otra se red de destino. La red en línea se utiliza para seleccionar la acción para la red de destino. La red de destino calcula los valores q de la acción que especifica la red en línea. Se aplica la ecuación de Bellman para estos valores de q. Los valores de Q son predichos por el modelo en línea, pero el valor de q de la acción actual se actualiza en la salida de la ecuación de Bellman. El estado actual se pone como x, y los valores q actualizados como y y se entrena el modelo en línea. Para actualizar los pesos de la red de destino se utiliza el método de actualización de la red de destino. La ventaja es que el entrenamiento se vuelve más estable que DQN y proporciona un mejor rendimiento. Sin embargo, la desventaja es que, en cada paso de tiempo, DDQN necesita actualizar dos modelos, por lo que iterar muchos episodios no es práctico.

6.3.1. Definiciones de representación

A continuación, se presentan las definiciones de representación de estado, espacio de acción y recompensa:

- Estado: Posibles combinaciones de origen-destino con la información espacial y espectral, representadas mediante una matriz. La información de los enlaces y esquemas de modulación se encuentran contenidos en diccionarios para la toma de decisiones en el espacio y espectro. Para el camino entre dos nodos adyacentes se proponen hasta cuatro vértices de conexión y la decisión depende de los pesos tomados de acuerdo con la exploración o explotación. Para cada camino k, se calcula el tamaño y el índice inicial del primer bloque FSU disponible en cada uno de los núcleos, el número requerido de FSU se toma de acuerdo con el formato de modulación seleccionado y la distancia del enlace.

- Acción: Espacio constituido por las k rutas candidatas y la asignación de bloques FSU de acuerdo con la modulación seleccionada.

- Recompensa: Estimulo entregado al finalizar cada episodio, se recibe una recompensa inmediata R_t cerca de 0 si R_t logra la conexión con éxito. De lo contrario, R_t<= -1. Esta recompensa se calcula de acuerdo con el retraso total obtenido en el enrutamiento de la demanda.

- DNN: Es la combinación entre algoritmo Q-Learning con redes neuronales profundas (Deep Neural Networks).

- FSU: Unidades de ranuras de frecuencia, se calculan por medio de la Tabla 6.1 y la Tabla 6.2, obteniendo la cantidad de FSU necesarios para cada demanda de acuerdo con el alcance máximo factible (MAR) en función del formato de modulación y la tasa de bit disponibles en los transpondedores.

6.3.2. Interacción con el entorno

DQN resuelve de manera efectiva problemas complejos de toma de decisiones que pueden aplicarse para el aprovisionamiento de recursos en EON. DQN utiliza una red neuronal profunda (DNN) que aproxima la función Q para evaluar qué tan buena es una acción para cada estado. Se introduce la repetición de la experiencia para superar los problemas de correlación y distribución no estacionaria de los datos empíricos. Para superar la sobreestimación de los valores Q, se introduce el doble DQN (DDQN), que incluye dos redes neuronales denominadas red neuronal en línea y red neuronal de destino, respectivamente.

El agente DQN debe interactuar con el medio que en este caso la topología de red a través del intercambio de las tres señales estado, acción y recompensa. En este problema multiagente cooperativo totalmente observable, existen n agentes donde $A = \{1,2,3 \dots n\}$ denota el conjunto de agentes.

Para cada episodio se entrega un par de nodos (origen, destino), el ancho de banda de la solicitud y tiempo de la solicitud, denotando el grafo de la siguiente forma $G(N, L, C)$. Donde N representa los nodos de la red, L todos los enlaces y C representa la capacidad espacial - espectral de cada enlace. El estado se representa por una solicitud de recursos entre cada par de nodos origen s y destino t, la acción es una tupla de pesos de enlace que determina las rutas para todos los pares origen-destino, y la recompensa se basa en el retardo de la red.

Para calcular cada solicitud de ruta R_t se necesita el camino de extremo a extremo $P_{S,T}$, el formato de modulación m, el ancho de banda b y a nivel espacial el núcleo c. De esta forma se garantizan los recursos solicitados por cada demanda, asignando un número de ranuras espectrales FSU contiguas en cada enlace a lo largo de $P_{S,T}$. En cada enlace de fibra se debe respetar la contigüidad en el espectro, pero no la misma localización, ni la misma modulación, teniendo prelación las políticas para evitar la diafonía. Cada enlace de fibra entre dos nodos tiene de 1 a 7 núcleos para utilizar ordenadamente de acuerdo con la cantidad de FSU asignadas respetando las políticas de fragmentación.

Capítulo 7

7. Propuesta algorítmica

7.1. Introducción

En este capítulo se describe el algoritmo propuesto el cual se soporta en el modelo seleccionado. La estrategia de búsqueda de recursos ópticos se basa en la convergencia como resultado del aprendizaje por refuerzo profundo.

7.2. Herramientas

La investigación se desarrolló con la ayuda de las siguientes herramientas:

I. Python

Python es un lenguaje de programación con estructuras de datos de alto nivel eficientes y un simple pero efectivo sistema de programación orientado a objetos. La elegante sintaxis de Python y su tipado dinámico, junto a su naturaleza interpretada lo convierten en un lenguaje ideal para scripting y desarrollo rápido de aplicaciones en muchas áreas, para la mayoría de las plataformas (*El Tutorial de Python — Documentación de Python - 3.10.7*, n.d.).

La mayoría de las empresas están implementando algoritmos de Machine Learning basados en Python para optimizar sus productos. Aunque hay otros lenguajes de programación que se pueden utilizar en Machine Learning, Python ha resultado ser el mejor lenguaje de programación para esta área. Ofrece grandes librerías y marcos para todas las áreas de la Inteligencia Artificial. Para el desarrollo de la propuesta se ejecutó bajo Python versión 3.9.12 64bits con las siguientes librerías:

- ***TensorFlow:*** Permite construir y entrenar redes neuronales para detectar patrones y razonamientos usados por las personas. Además de trabajar con redes neuronales,

TensorFlow es multiplataforma. Trabaja con GPUs y CPUs e incluso con las unidades de procesamiento de tensores (TPUs) (*TensorFlow*, n.d.).

- ***matplotlib.pyplot:*** Colección de funciones de estilo de comando similar a MATLAB. Cada función de pyplot realiza algún cambio en una figura: por ejemplo, crea una figura, crea un área de trazado en una figura, traza algunas líneas en un área de trazado, decora el trazado con etiquetas, etc. En matplotlib.pyplot se conservan varios estados en la función llamadas, de modo que realiza un seguimiento de cosas como la figura actual y el área de trazado, y las funciones de trazado se dirigen a los ejes que conforman la figura (*Matplotlib — Visualization with Python*, n.d.).
- ***Pytorch:*** Computación de tensores con una aceleración fuerte a través de unidades de procesamientos gráficos (GPU). Redes neuronales profundas construidas en un sistema de diferenciación automática de bases de datos (*PyTorch*, n.d.).
- ***Numpy:*** Librería estable y muy rápida escrita en Python para representar matrices multidimensionales, similar a las matrices en R.

II. Visual Studio Code

Editor de código fuente ligero soporta: JavaScript, TypeScript y Node.js y cuenta con un ecosistema de extensiones para otros lenguajes y tiempos de ejecución (como C++, C#, Java, Python, PHP, Go, .NET) (*Documentation for Visual Studio Code*, n.d.). Para el desarrollo de la propuesta se utilizó bajo la versión 1.72.0 64bits.

7.3. Entorno

En la Figura 7-1 se observan los métodos de la clase entorno, los agentes inician cada episodio mediante el reinicio, sin afectar el estado de ocupación del espectro de los episodios anteriores. En cada paso se realiza la validación de la ruta, espectro, espacio y al finalizar el episodio se realiza la asignación. Por cada episodio finalizado se calcula la probabilidad de bloqueo de la red, el uso del espectro, el tiempo de ejecución y se actualizan las gráficas.

El entorno está representado por la topología de red óptica, compuesta por elementos activos y pasivos con las siguientes características:

7.3.1. Enlaces de fibra

Representa la conexión entre dos NEs, cada enlace está constituido por dos fibras ópticas, las cuales cumplen en cada extremo las funciones de transmisión y recepción. Cada enlace cargado en el entorno es de fibra monomodo (SSM) y debe contar con la información de longitud de enlaces dado en km. La implementación en Python se realiza mediante una clase diccionario que se utiliza para almacenar valores de datos como un mapa. A diferencia de otros tipos de datos que contienen solo un valor único como elemento, el diccionario contiene un par clave-valor. En el diccionario la clave debe ser única e inmutable, la clave para cada

enlace es un código de enlace generado por la ubicación del origen y el destino. En la Figura 7-2 se presenta un ejemplo de diccionario para los enlaces de fibra donde el nodo 1 tiene adyacencias con los nodos: 2 a una distancia de 1050 km, 3 a una distancia de 1500 km y 8 a una distancia de 2400 km.

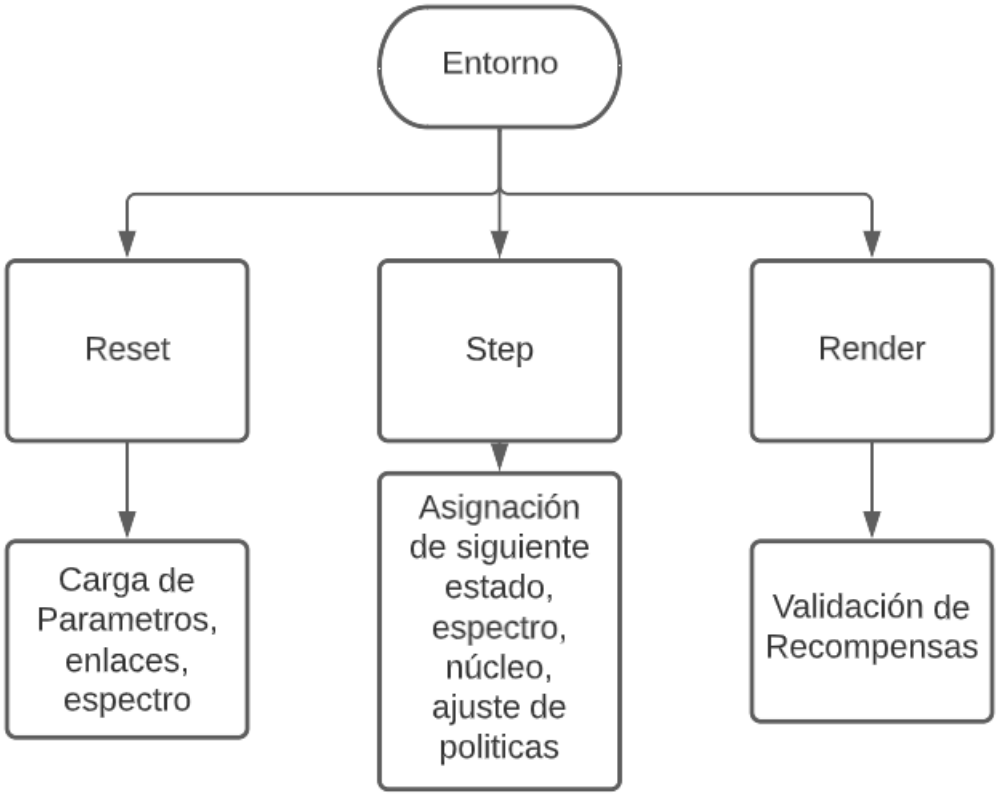


Figura 7-1 Entorno. Elaboración propia.

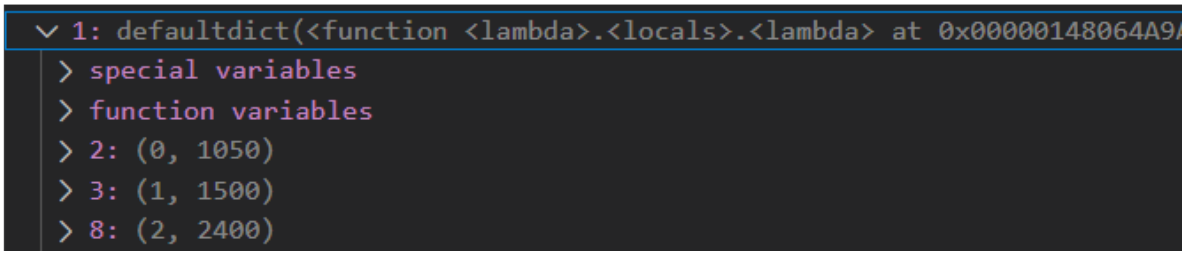


Figura 7-2 Ejemplo de diccionario de enlaces. Elaboración propia.

La Figura 7-3 ilustra cómo se asignan código de enlace de acuerdo con el orden de origen destino de los enlaces, el código de enlace 00 hace referencia al enlace de fibra entre el NE1 y el NE2, el código de enlace 01 hace referencia al enlace de fibra entre NE1 y el NE3.

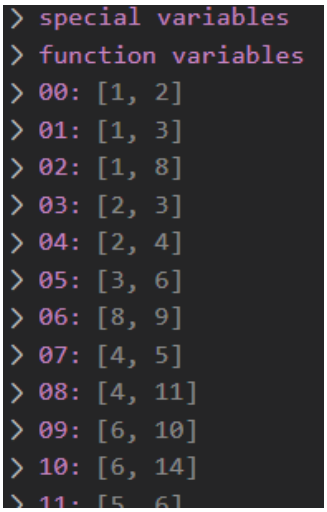


Figura 7-3 Ejemplo códigos de enlaces. Elaboración propia.

7.3.2. Cantidad de núcleos de las fibras troncales

Se implementa mediante fibra multinúcleo entre NEs constituida de 1 a 7 núcleos, como valor por defecto se toman fibras de 3 núcleos. La implementación en Python se realiza mediante una clase diccionario que se utiliza para almacenar valores de datos como un mapa. A diferencia de otros tipos de datos que contienen solo un valor único como elemento, el diccionario contiene un par clave-valor. En el diccionario la clave debe ser única e inmutable, la clave en el diccionario es el código del enlace de fibra y el id del núcleo, cada id de núcleo se compone de 100 FSU.

En la Figura 7-4 se presenta un ejemplo de diccionario de núcleos, donde para cada núcleo se tienen 22 enlaces identificados por su código de enlace y cada uno con 100 FSU

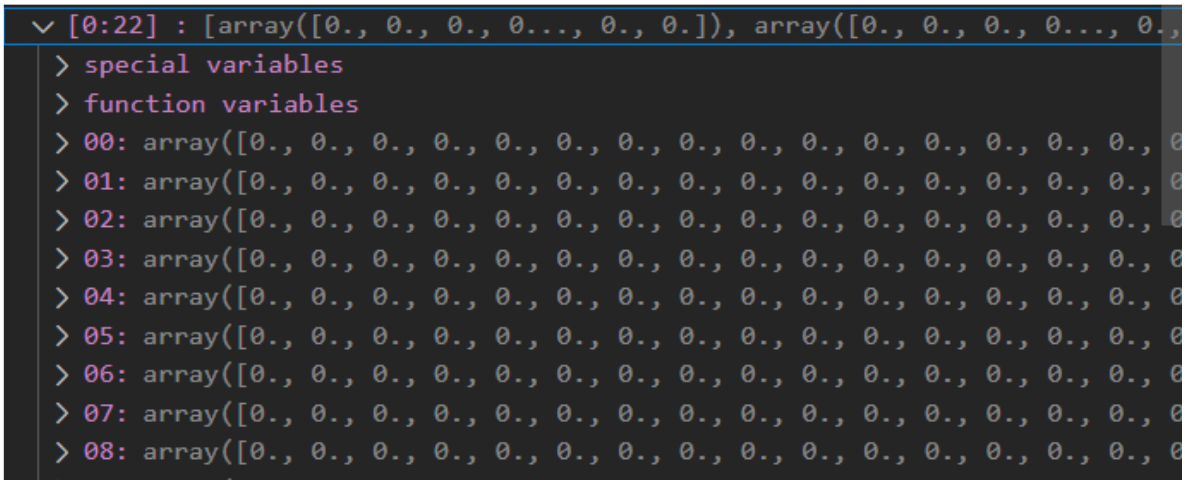


Figura 7-4 Ejemplo diccionario de núcleos. Elaboración propia.

7.3.3. *Modulación*

La modulación es asignada de acuerdo con la relación de los parámetros de la Tabla 6.1y la Tabla 6.2. La implementación en Python se realiza mediante una clase diccionario que se utiliza para almacenar valores de datos como un mapa. A diferencia de otros tipos de datos que contienen solo un valor único como elemento, el diccionario contiene un par clave-valor. En el diccionario la clave debe ser única e inmutable, la clave en el diccionario es la distancia máxima de alcance. Como ejemplo en la Figura 7-5 se presenta un diccionario de modulación, donde el valor de 4000 representa la distancia máxima alcanzable en kilómetros y contiene 5 velocidades de bit posibles {10, 40, 100, 400, 1000} Gbps, cada velocidad tienen la cantidad de FSUs a asignar con el esquema de modulación aplicado.

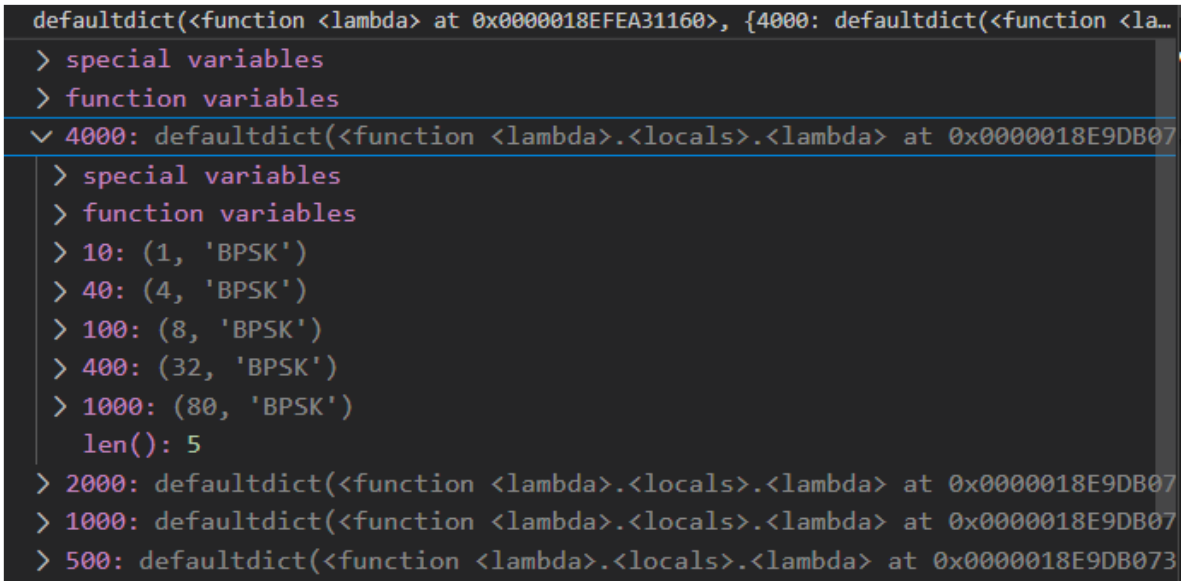


Figura 7-5 Ejemplo diccionario de modulación. Elaboración propia.

7.3.4. *Ancho de banda máximo soportado por las tarjetas de línea*

Las demandas de tráfico están encapsuladas en {10, 40, 100, 400, 1000} Gbps soportadas con la capacidad máxima de modulación. Al tener definida la ruta Rt y la distancia l, se puede calcular la cantidad de FSU a asignar de acuerdo con la Tabla 6.1 y la Tabla 6.2. En este caso si se tiene una demanda de tráfico (origen NE1, destino NE2) con ancho de banda de 9 Gbps y tomando el camino corto con distancia de 1050 km, de acuerdo con la Tabla 6.1 y la Tabla 6.2 la modulación recomendada será QPSK y con una asignación de $FSU = 1$.

7.3.5. *Restricciones del entorno según retardo, pérdida de paquetes, probabilidad de bloqueo.*

Retardo: El retardo de transmisión es el tiempo que toma el paquete en desplazarse en el medio, este tipo de retardo es directamente proporcional al tipo de red por la que se esté

desplazando el paquete, es decir, en enlaces de acceso de baja velocidad este parámetro puede llegar a ser muy alto, caso contrario sucede en redes que cuentan con enlaces de acceso de alta velocidad, en donde el retardo de transmisión puede llegar a ser ínfimo. El retardo en el entorno es calculado de acuerdo con los parámetros de los enlaces de fibra, este cálculo es importante debido a que la suma de los retardos de una demanda corresponde a la ganancia del episodio.

Probabilidad de bloqueo: Para evaluar la probabilidad de bloqueo de las conexiones que operan con varias rutas alternativas, se obtiene una expresión dependiente del bloqueo de cada ruta alternativa. Luego, el bloqueo de cada ruta se calcula como función del bloqueo de cada enlace que lo compone. Para evaluar el bloqueo del enlace, las rutas que pasan a través de este se clasifican de acuerdo con el orden en que estas son utilizadas por el algoritmo, y la carga de tráfico que ofrecen se expresa en función del bloqueo de las rutas exploradas. En base a la información de tipo de ruta y carga de tráfico ofrecida al enlace, se calcula el bloqueo del enlace.

7.4. DQN RMSCA

Los resultados del algoritmo de Red Q profunda para Asignación de enrutamiento, modulación, espectro y núcleo (DQN RMSCA) son la ruta R, los formatos de modulación M, las FSU y el núcleo asignado para cada demanda. La propuesta se separó en dos subsecciones para ser más detallada.

- ***Estrategia de enrutamiento:*** La estrategia de enrutamiento consiste en encontrar una ruta óptica entre los nodos de origen y destino para cada demanda de tráfico en la red. Para solucionar el enrutamiento los agentes DQN usan el criterio del camino con mejor recompensa para solucionar cada demanda.
- ***Formato de Modulación y número de FSU:*** En esta sección se selecciona el formato de modulación para cada usuario, basado en la demanda de ancho de banda y la longitud (km) de cada enlace. Se determina la cantidad de FSU a asignar para cada demanda y el núcleo disponible para la cantidad de FSU asignados.

La Figura 7-6 expone la estrategia de asignación para solucionar el problema RMSCA, la entrada G corresponde al Grafo y D a las demandas. La demanda $\mathrm{D}(\mathrm{S}, \mathrm{T}, \mathrm{b}, \tau)$ contiene la información del origen S, destino T, ancho de banda b y duración τ. El primer bloque calcula la ruta Rt de origen S al destino T, esta ruta debe respetar las políticas de la red y de la demanda. En cada paso de la ruta se determina la modulación m y la cantidad de ranuras de Frecuencia FSU a separar, teniendo en cuenta los parámetros de asignación de las Tabla 6.1 y Tabla 6.2. El tercer bloque realiza la asignación espacial para cada demanda y finalmente se entregan las salidas que corresponden a la recompensa R asignada por el retardo, modulación, espectro, y núcleo asignado.

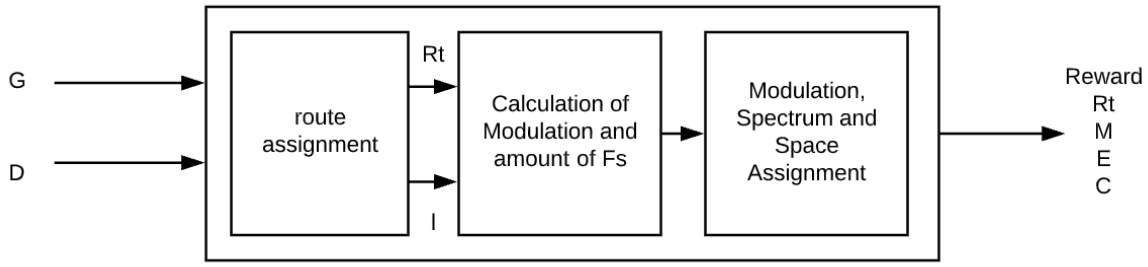


Figura 7-6 Diagrama de entradas y salidas. Elaboración propia.

El Algoritmo 1 representado en Figura 7-7 soluciona el problema RMSCA mediante aprendizaje por refuerzo multiagente, ejecuta la estrategia de bloques explicados anteriormente. La línea 1 carga la topología de red con todas sus características, Las líneas 2-3 cargan las entradas que corresponden al grafo e inicializa los episodios. La línea 4 ejecuta un reinicio del entorno, posteriormente en la línea 5 se actualiza el estado de EON liberando los recursos asignados de los FSU que han caducado. En la línea 6, se carga cada demanda compuesta por ancho de banda y duración. En la línea 7 se determina el estado inicial de los agentes.

Algoritmo 1 DQN RMSCA MARL

```
1: Environment initialization;
2: Inputs (G, D);
3: for episode_num in range(self.episodes_number):
4:       Environment reset;
5:       Release of expired FSU;
6:       Load of demands d;
7:       create randomness in initial state
8:       while not done and listsum==0:
9:                for agent in agents:
10:                       actions.append(agent.greedy_actor(state))
11:                       next_state, reward, done = self.env.step(actions)
13:               end
14:               Rt=routing(G,d);
15:               Algorithm 2
16:               {M,C,E}:= RL(l, D);
17:                D:= Classification (D,l,Rt,E,C);
18:                storage (state, action, next_state, reward, done)
19:          Model optimization
20:      Total reward calculation
21: end
```

La línea 8 determina la condición de finalización del episodio. En las líneas 9 a 13 se selecciona la acción por exploración o explotación de cada agente. En las líneas 14 al 18 se realiza: cálculo de la ruta, asignación de enlaces, núcleos (mediante el Algoritmo 2) y modulación. En la línea 19 se invoca la política y valora DQN para generar una política de RMSCA. Por último, en la línea 20 se calcula la recompensa total del episodio.



Figura 7-7 Algoritmo RMSCA MARL. Elaboración propia.

En el Algoritmo 2, el nivel de modulación m se determina primero de acuerdo con la Tabla 6.1 y Tabla 6.2 bajo la condición de restricción $\mathrm{Reach[m][A_{max}]} \geqslant \mathrm{L_{sd}} > \mathrm{Reach[m+1][A_{max}]}$. Luego se calcula el número de FSU necesarios bajo el nivel de modulación m. Posteriormente los tipos seleccionables de super bloques se comprueban por turnos. Para

cada tipo seleccionable, la disponibilidad de todos los super bloques se verifica uno por uno. Si se dispone de un super bloque, se calculan su N_{NB}y N_{ctp}y se comparan con $N^*{}_{NB}$ y $N^*{}_{ctp}$.

En este contexto, NNB representa el número de bloques vecinos ocupados en los enlaces del camino. Para un super bloque disponible en la ruta de origen-destino, si sus bloques vecinos en los enlaces de la ruta han sido ocupados por otras conexiones, el uso del super bloque no disminuiría la contigüidad del espectro. Por lo tanto, cuanto mayor sea el valor de NNB, menor será el impacto en el espectro contigüidad. NCTP corresponde al número de homólogos ocupados en enlaces adyacentes al trayecto. Para un super bloque disponible en el trayecto fuente-destino, si sus contrapartes en enlaces adyacentes del trayecto han sido ocupadas por otras conexiones, el uso del super bloque no disminuiría la continuidad del espectro. Por lo tanto, cuanto mayor sea el valor de NCTP, menor será el impacto en la continuidad del espectro.

Algoritmo 2 FSU vs Núcleo

1: Core {1,2,3,4,5,6} or {1,2} # For edge cores
2: Core {7} or {3} # For the core
3: Determine m according to Tabla 6.1 y Tabla 6.2
4: If type is 1 to 3 do
5: For each $core \in CoreSet$ do
6: Check availability of the Core block
7: $Bs = B(Core, prime, numbre, start)$
8: If Bs is Available Then
9: Calculate $N_{NB}\ y\ N_{ctp}$
10: If $(N_{NB} > N^*{}_{NB})\ or\ (N_{NB} = N^*{}_{NB})\ or\ (N_{ctp} > N^*{}_{ctp})$ then
11: $N^*{}_{NB} = N_{NB}$
12: $N^*{}_{ctp} = N_{ctp}$
13: $B_s^* = B_s$
14: end
15: end
16: end
17: end

7.4.1. Agentes

DQN toma un estado del entorno como entrada y devuelve una aproximación de valor Q para cada acción posible. Después, elige la acción con el mayor valor Q, tal como lo hace Q-Learning, pero tratando de reconocer patrones en lugar de asignar cada estado a su mejor acción. La red neuronal se entrena con esos datos para aproximar la salida en función de la entrada actualizando los parámetros de forma iterativa. La Interacción entre el agente y el entorno se presenta en la Figura 7-8.

El agente usa dos redes neuronales para aprender y predecir qué acción tomar en cada paso, la red Q o red en línea se utiliza para predecir qué hacer cuando el agente se encuentra con

un nuevo estado. Esta red toma el estado como entrada y emite valores Q para las posibles acciones que se podrían seleccionar. La red en línea toma un tensor con los valores de los estados del entorno como entrada y genera un vector de cuatro valores Q, cada valor para moverse a la adyacencia del NE actual. El agente selecciona la acción que tenga el mayor valor de Q correspondiente emitido por la red en línea. En la Figura 7-9 se presenta el método que usa la red en línea para elegir una acción:

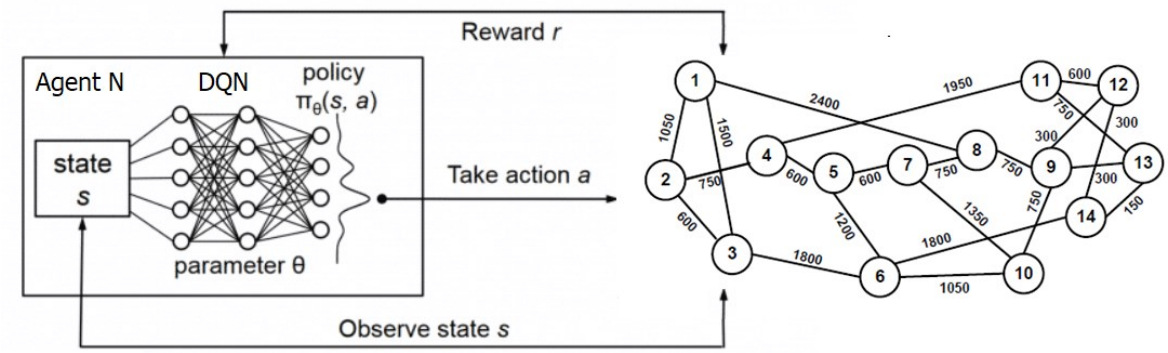


Figura 7-8 Interacción entre el agente y el entorno. Elaboración propia.

```
def get_action(state, eps):
    if np.random.rand() < eps:
        return np.random.choice(N_ACTIONS)
    else:
        policy_net.train(mode=False)
        scores = get_Q(policy_net, state)
        _, argmax = torch.max(scores.data, 1)
        return int(argmax.cpu().numpy())
```

Figura 7-9 Seleccionar acción. Elaboración propia.

Los agentes aprenden y se mejoran a través del método repetición de experiencia, que es donde trabaja la segunda red neuronal denominada red de destino, para llevar a cabo la repetición de la experiencia, el agente "recuerda" cada paso del entorno después de que ocurre, cada recuerdo consiste en la acción realizada, el estado en el que se encontraba cuando se realizó la acción, la recompensa otorgada por realizar la acción y el estado que resultó de tomar la acción. Después de dar cada paso en el episodio, se lleva a cabo el procedimiento de repetición de experiencias si y sólo si después de almacenar suficientes recuerdos. La repetición desarrolla los siguientes pasos:

- Se toma un conjunto aleatorio de experiencias llamado mini lote de la memoria del agente.

- Para cada experiencia en el mini lote se calculan nuevos valores Q de cada par de estado y acción: si la acción finalizó el episodio, el valor Q será negativo entregado por la latencia sumado con la penalización, y si la acción no finalizó el juego, es decir, mantuvo el agente activo durante al menos un turno más, el valor de Q será cercano a cero y se predice mediante la ecuación de Bellman.
- La red neuronal está preparada para asociar cada estado del mini lote con los nuevos valores Q calculados para las acciones realizadas.

Para esto, vamos a necesitar dos clases:

- Transición: una tupla con nombre que representa una sola transición en el entorno. Esencialmente, asigna pares (estado, acción) a su resultado (siguiente_estado, recompensa) como se observa en la Figura 7-10.

```
Transition = namedtuple('Transition',
                        ('state', 'action', 'next_state', 'reward', 'done'))
```

Figura 7-10 Transición. Elaboración propia.

- ReplayMemory: Presentado en la Figura 7-11 es un búfer cíclico de tamaño limitado que contiene las transiciones observadas recientemente. También implementa un método .sample() para seleccionar un lote aleatorio de transiciones para el entrenamiento.

```
class ReplayMemory(object):

    def __init__(self, capacity):
        self.capacity = capacity
        self.memory = []
        self.position = 0


    def push(self, *args):

        if len(self.memory) < self.capacity:
            self.memory.append(None)
        self.memory[self.position] = Transition(*args)
        self.position = (self.position + 1) % self.capacity
```

Figura 7-11 Repetición de experiencia. Elaboración propia.

7.5. Mecanismo de Entrenamiento

Para entrenar el modelo la función optimizar modelo inicia muestreando un lote, luego concatena todos los tensores en uno solo, calcula $Q(s_t, a_t)$ y $V(s_{t+1}) = \max_a Q(s_{t+1}, a)$ y los combina. Por definición se establece que $V(s) = 0$, si s es un estado terminal, se utiliza la red objetivo para calcular $V(s_{t+1})$ aportando mayor estabilidad. La red de destino tiene

sus pesos congelados la mayor parte del tiempo, pero se actualiza con los pesos de la red en línea cada 3 episodios.

En la Figura 7-12 se encuentra el bucle de entrenamiento principal, para iniciar se reinicia el entorno, se inicializa el tensor de estados, luego, se prueba una acción, se ejecuta, se observa el siguiente estado y la recompensa, posteriormente se optimiza el modelo y al finalizar el episodio se reinicia el bucle.

```
for i_episode in range(NUM_EPISODES):
    clear_screen(0)
    state, Bw, Tao = env.reset()
    total_reward = 0
    eps = epsilon_annealing(i_episode, MAX_EPISODE, EPS_END, EPS_START)
    done = False
    t = 0
    print("////////////////////////////////////////////")
    log_progress(env, delay=0.5, message=init_message(i_episode, perf), eps=eps,PblockNet=PblockNet)

    while not done:
        action = get_action(state, eps)
        next_state, reward, done, Block, Trks, _ = env.step(action)
        total_reward += reward

        if i_episode >= 250:
            log_progress(env, reward=reward, total_reward=total_reward, delay=0.5,message=perf_message(i_episode, perf), eps=eps, P

        memory.push(state, action, next_state, reward, done)
```

Figura 7-12 Bucle de entrenamiento. Elaboración propia.

En la Figura 7-13 se detalla la configuración de los hiperparámetros para el agente DQN, los parámetros EPS operan en la selección de acciones para la política codiciosa de épsilon. En ocasiones para elegir la acción aleatoria y en otras seleccionara una de acuerdo con la experiencia. La probabilidad de elegir una acción aleatoria comenzará en EPS_START y disminuirá exponencialmente hacia EPS_END, EPS_DECAY controla la tasa de decaimiento.

```
LEARNING_RATE = 0.01
NUM_EPISODES = 5000
MEMORY = 200000
BATCH_SIZE = 1024
GAMMA = 0.999
EPS_START = 0.9
EPS_END = 0.001
MAX_EPISODE = 150
TARGET_UPDATE = 3
HIDDEN_DIM = 64
N_STATES , N_ACTIONS =env._load_config_params()
```

Figura 7-13 Configuración de hiperparámetros. Elaboración propia.

Capítulo 8

8. Resultados experimentales

Teniendo en cuenta la aplicabilidad y la generalidad del algoritmo propuesto, se evalúa el rendimiento con la topología de red troncal NSFNET de 14 nodos y 22 enlaces, como se muestra en la en la Figura 8-1, en cada uno de estos enlaces se detalla la distancia en km, la cual es cargada en el entorno para calcular la latencia y demás parámetros de configuración. Cada enlace de red es bifilar y cuenta con 3 núcleos, a nivel de espectro se administra una porción de espectro con frecuencia central inicial de 195.9 THz y finalizando en una frecuencia central de 194.6625 THz. Todo esto con el fin de obtener 100 ranuras espectrales con una separación de 12.5 GHz para cada uno de los núcleos de los enlaces de la topología.

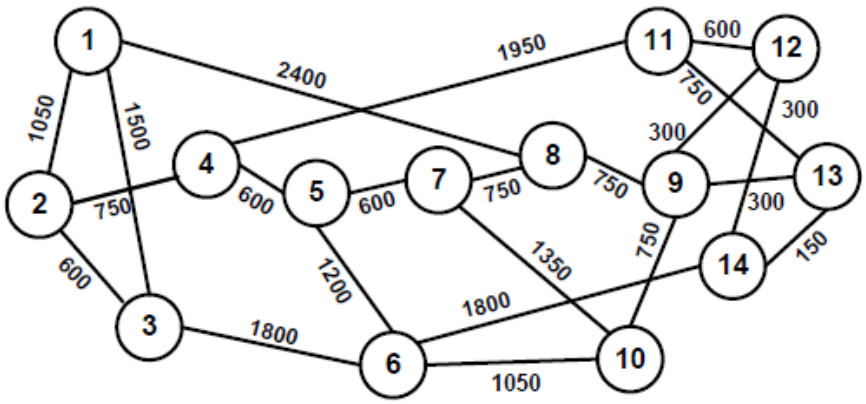


Figura 8-1 Topología NSFNet. Tomado de (Danilo R. B. Araújo, n.d.).

Todos los nodos de la topología son de interconexión, por consiguiente, en cada extremo se entregan y reciben todo tipos de solicitudes (ejemplo: servicios desde 100 Mbps hasta 100Gps), estos servicios administrados de baja capacidad son perfiles de hogares y los de alta capacidad transporte de servicios corporativos. El número de solicitudes se distribuye aleatoriamente, las solicitudes de ancho de banda iniciales son aleatorias dependiendo la

modulación seleccionada y el ancho de banda requerido, la solicitud de ancho de banda aumentará o disminuirá según su tipo.

8.1. EON Deep Q Network

Sin usar una red neuronal de capa oculta con un espacio de estado de codificación en caliente, se aplicó una tasa de aprendizaje de 0,00025 en el optimizador de Adam, tamaño de lote de 32, gamma de 0,95 en la ecuación de Bellman, épsilon en el algoritmo ávido de épsilon comienza desde 1,0, multiplicado por 0,9972 en cada episodio, decae a 0.1 (Tabla 8.1), después de varios episodios se fija en 0.1.

Tabla 8.1 Hiperparámetros Q learning. Elaboración propia.

Hiperparámetros	Valor	Descripción
Alpha	0.1	La tasa de aprendizaje en Q-learning. Cuanto más pequeño es, más influyente es el valor q actual
Gamma	0.6	El factor de descuento. Cuanto más grande es, más considera el agente la recompensa futura
épsilon	0.1	El parámetro en el algoritmo codicioso épsilon. Es 0.1, 1 de cada 10 episodios no sigue el valor q estimado y realiza acciones aleatorias

Implementando DQN RMSCA sobre la topología de 14 nodos y 22 enlaces obtuvo los siguientes resultados:

- Mediante DQN con un solo agente sin desfragmentación de espectro, en configuración espacial de 3 núcleos por troncal. La red comienza a saturarse después de 5000 episodios con una configuración de 3 capas 64 neuronas. Esto se representa en la Figura 8-2, donde cada color representa un núcleo de la fibra troncal, azul para el núcleo $c1$, verde para el núcleo $c2$ y rojo para el núcleo $c3$.
- Con DQN sin fragmentación antes de los 2000 episodios la mayor parte de las recompensas están por debajo de -2, lo que significa que los agentes no han aprendido a realizar las acciones que maximicen su recompensa. Sin embargo, los agentes sufren penalizaciones por repetir tránsito por troncales de red ya visitadas en el mismo episodio, esto también reduce la recompensa. Después de 3000 episodios, los agentes mejoraron este comportamiento y comenzaron a obtener recompensas positivas (cercanas a 0). Sin embargo, el problema es que incluso después de obtener recompensas deseables positivas, los agentes aún cometieron algunos errores y produjeron recompensas negativas.

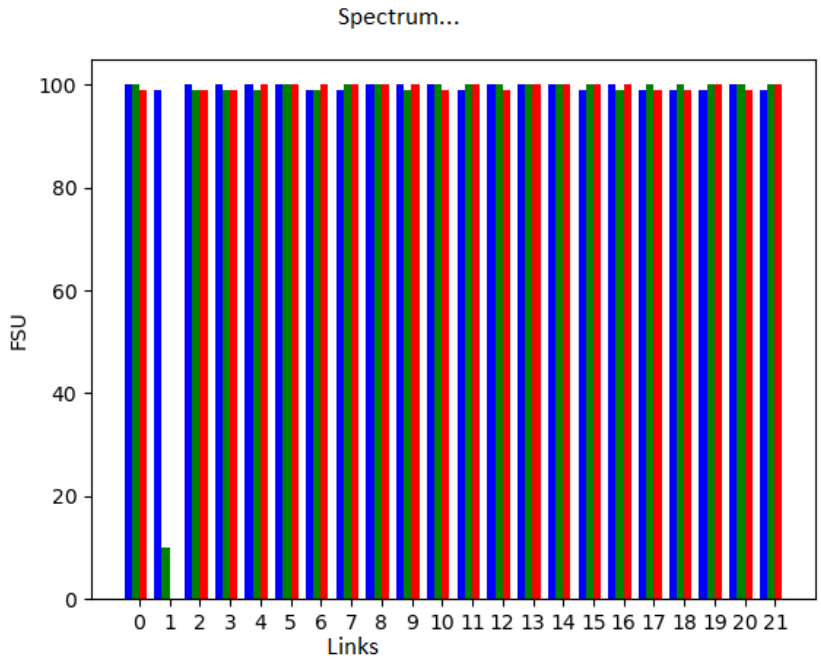


Figura 8-2 Uso de espectro. Elaboración propia

- La probabilidad de bloqueo es mínima para tráfico ofrecido por debajo de los 12.000 Gbps, con tráfico superior a 15.000 Gbps presenta dificultad en la asignación de los recursos espectrales FSU. La administración de red por medio de estos escenarios de simulación sobre topología de 14 nodos permite un máximo de tráfico ofrecido por la red de 10 Tbps, para un promedio de 460 canales implementados antes de presentar bloqueo notable sobre las demandas (Figura 8-3, Figura 8-4,). Se ejecutaron 5 simulaciones presentes en la Tabla 8.2 con diferentes cargas de tráfico para evaluar la asignación.

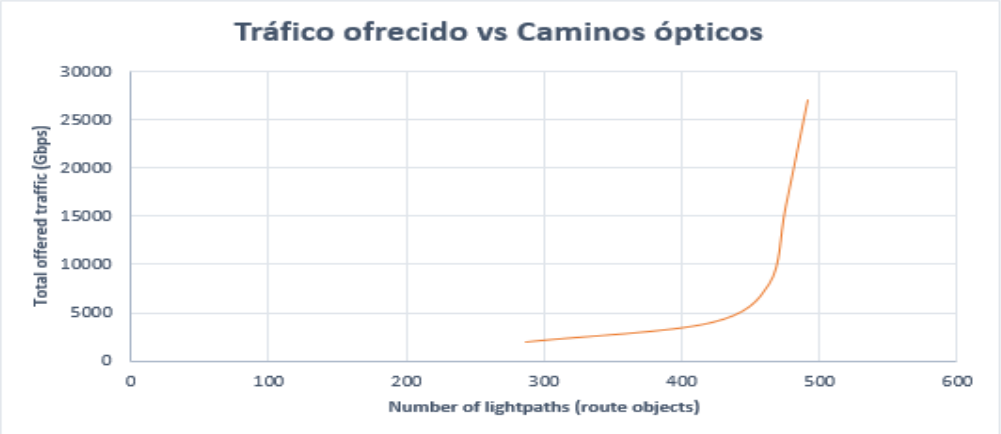


Figura 8-3 Trafico ofrecido vs caminos ópticos. Elaboración propia.

Tabla 8.2 Simulaciones. Elaboración propia.

Test	Number of lightpaths (route objects)	Total offered traffic (Gbps)	Total blocked traffic
1	*287*	2000	0
2	*421*	4000	0.08
3	*463*	8000	0.09
4	*475*	10999	0.31
5	*491*	27089	0.85



Figura 8-4 Tráfico ofrecido versus Total tráfico Bloqueado. Elaboración propia.

8.2. EON MARL

El agente se capacita en el entorno descrito anteriormente utilizando el Q-learning profundo independiente (IDQN). En el entorno, se considera que el número máximo de pasos está dado por la repetición de una troncal de fibra, ósea que la demanda repita el paso por un NE o que no se encuentre recursos disponibles en espacio o espectro. Más específicamente, cada agente está equipado con dos redes neuronales, en línea y objetivo, para aproximar su propia función Q de estado-acción y una memoria para almacenar las experiencias obtenidas a través de la interacción con el entorno. Los agentes están entrenados con la memoria Uniform Experience Replay (UER) y la memoria Prioritized Experience Replay (PER).

El espacio combinatorio de hiperparámetros es demasiado grande para una búsqueda exhaustiva. No se realiza una búsqueda sistemática en topología debido al alto costo computacional, en cambio, se realiza una búsqueda informal. Los valores de todos los hiperparámetros se proporcionan en la Tabla 8.3.

Tabla 8.3 hiperparámetros Multi Agente. Elaboración propia.

Parámetro	Valor
RMSProp learning rate	0.00005
Adam learning rate	0.0000625
Initial ∈ for exploration	1.00
Final ∈ for exploration	0.01
Replay memory size	100
Minibatch size	32
Steps between online network updates	4
Discount factor	0.95
Prioritization type	Proportional
Prioritization exponent ω	0.5
Prioritization importance sampling β	0.4 →1.00

El algoritmo RMSCA se usa con una tasa de aprendizaje de 0.00005. Además, el factor de descuento se configura en 0,99 y se realiza una actualización de aprendizaje cada 4 pasos de tiempo usando mini lotes de 32 transiciones. Para la priorización de repeticiones, el exponente de prioridad es de 0,5 y aumenta linealmente el exponente de muestreo de importancia (IS) de 0,4 a 1 durante el transcurso de la fase de exploración.

Implementando DQN RMSCA MARL sobre la topología de 14 nodos y 22 enlaces con desfragmentación de espectro obtuvo los siguientes resultados:

- En la Figura 8-5 se detalla los episodios versus la probabilidad de bloqueo acumulada sobre una configuración de dos agentes de 3 capas de 64 neuronas, donde ese observa que posterior al episodio 1000 de entrenamiento se comienza a reducir la probabilidad de bloqueo, los agentes comienzan a explotar la información y los pesos en la red se estabilizan.

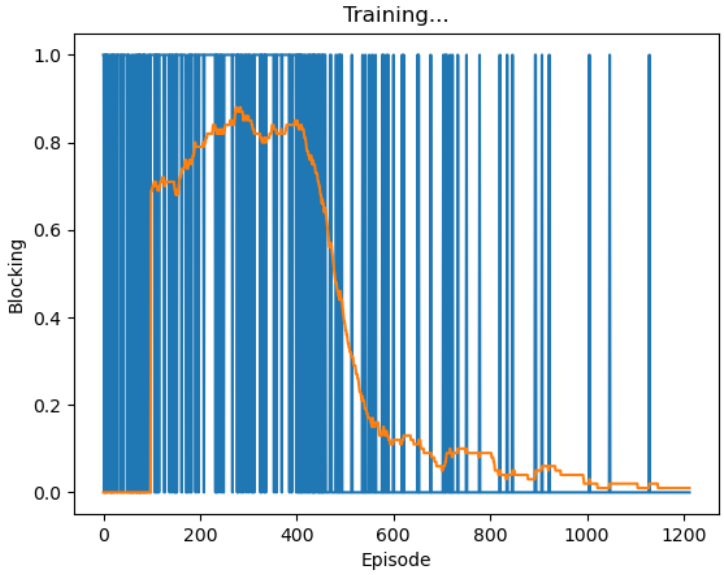


Figura 8-5 Blocking probability 3 64. Elaboración propia.

- La Figura 8-6 representa el éxito de los episodios de los agentes vs el total de los episodios, configurado para 2 agentes en 3 capas de 64 neuronas. Se observa el comportamiento contrario al bloqueo donde el está cerca de 1 después del episodio 1000.

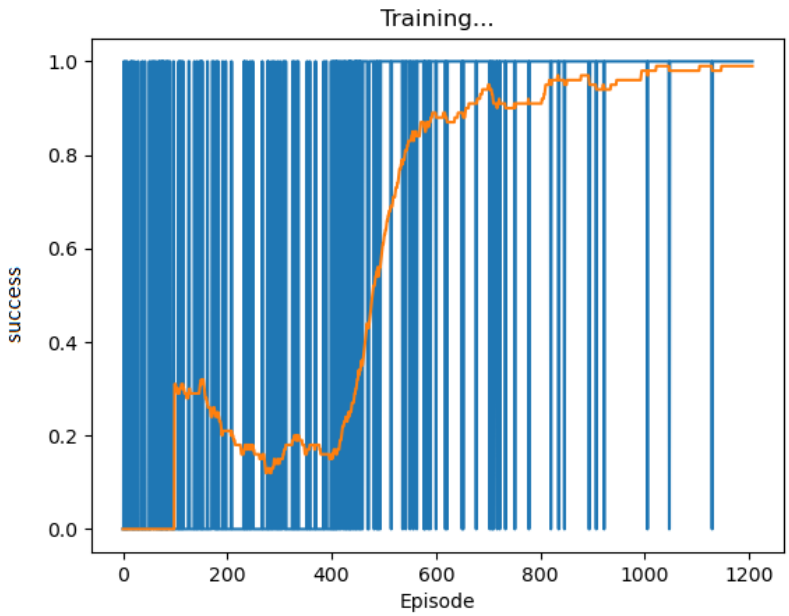


Figura 8-6 Episodios de éxito vs Total de Episodios. Elaboración propia.

- La Figura 8-7 representa la duración de cada episodio en tiempo de ejecución para 2 agentes en 3 capas de 64 neuronas.

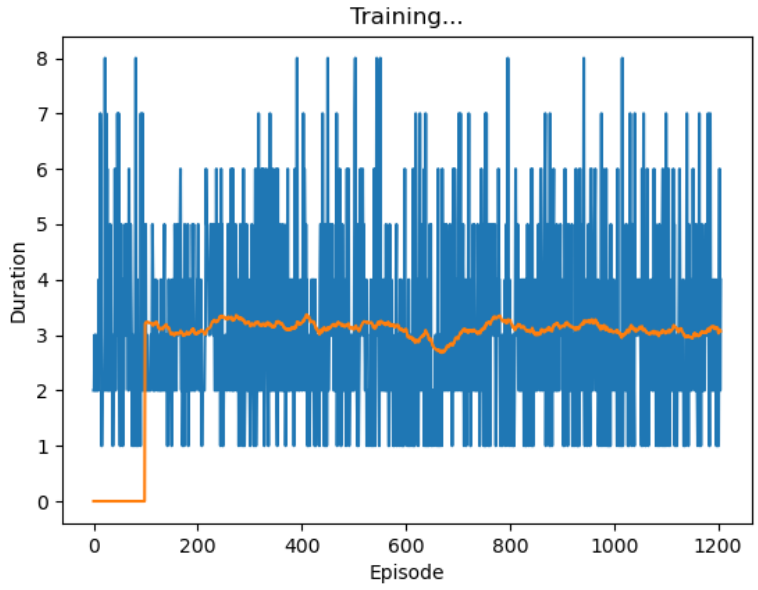


Figura 8-7 Duration vs Episode. Elaboración propia.

- En la Figura 8-8 se detalla los episodios versus la probabilidad de bloqueo acumulada sobre una configuración de dos agentes de 5 capas con 128 neuronas.

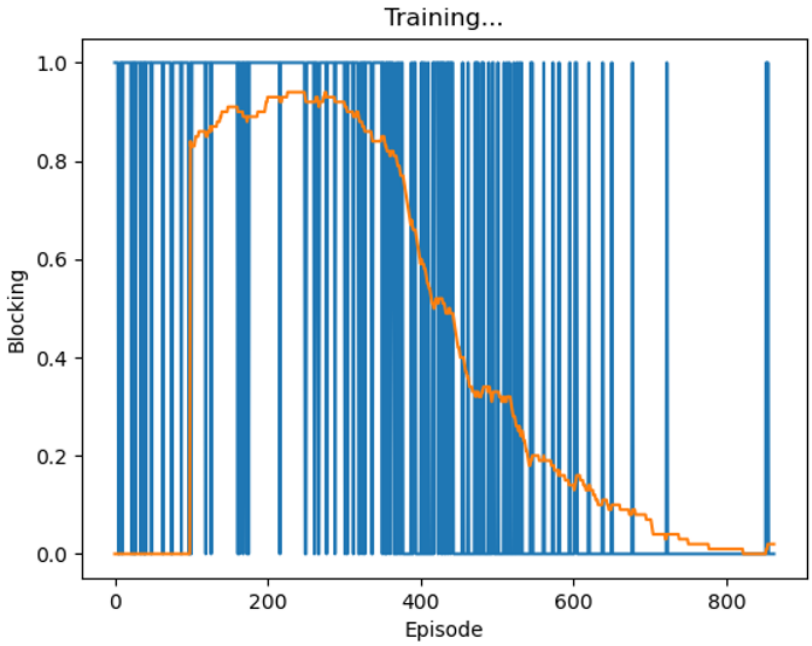


Figura 8-8 Blocking probability 5 128. Elaboración propia.

- La Figura 8-9 relaciona la ganancia de los agentes versus la cantidad de episodios en una configuración de 5 capas de 128 neuronas para dos agentes. Donde se confirma la mejora en la recompensa durante el entrenamiento posterior al episodio 400.

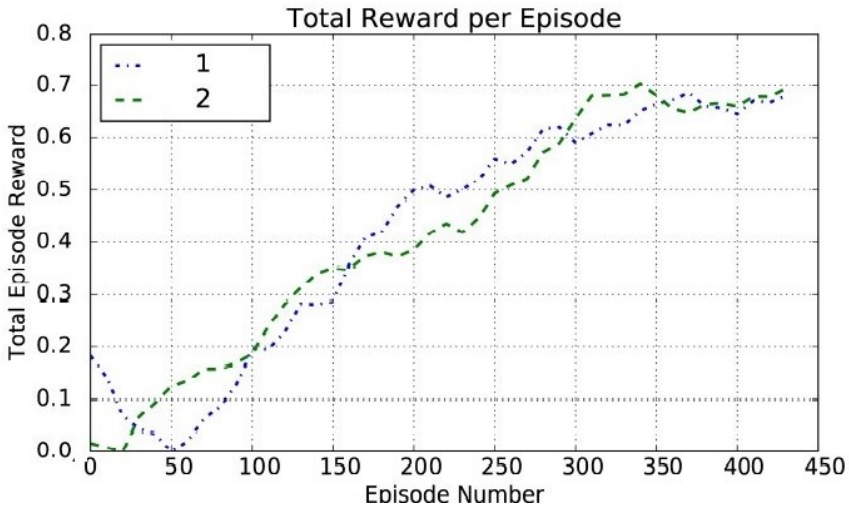


Figura 8-9 Recompensa de los agentes. Elaboración propia.

- En la Tabla 8.4 se presentan las 15 simulaciones más relevantes que se ejecutaron bajo la configuración inicial con desfragmentación del espectro.

Tabla 8.4 Simulaciones. Elaboración propia.

Test	Number of lightpaths (route objects)	Total offered traffic (Gbps)	Total blocked traffic
1	*290*	20000	0
2	*587*	35000	0.05
3	*890*	44200	0.059
4	*1008*	55000	0.076
5	*1250*	87065	0.085
6	*1460*	100200	0.097
7	*1648*	135000	0.106
8	*1890*	168700	0.12
9	*2150*	189000	0.15
10	*2480*	205000	0.164
11	*2900*	227900	0.178
12	*3250*	248000	0.1896
13	*3380*	280000	0.21
14	3500	300500	0.225
15	3800	319000	0.24

- La de asignación de recursos espectrales ejecutado por los algoritmos 1 y 2 se refleja en la Figura 8-10, se exponen 15 simulaciones ejecutadas vs la ocupación de la red para una configuración de dos agentes, el azul de las barras hace relación a los FSU asignados, el color naranja refleja los FSU que se encuentran fragmentados y finalmente el color verde refleja las FSU disponibles.
  Como la asignación espectral para la simulación por enlace es de 100 FSU con una configuración espacial de 3 núcleos y 22 enlaces de troncales, la capacidad total de la red en FSU es de 6600. Se observa que en las simulaciones aún quedan espacio disponible para asignar gracias a la desfragmentación del espacio espectral.

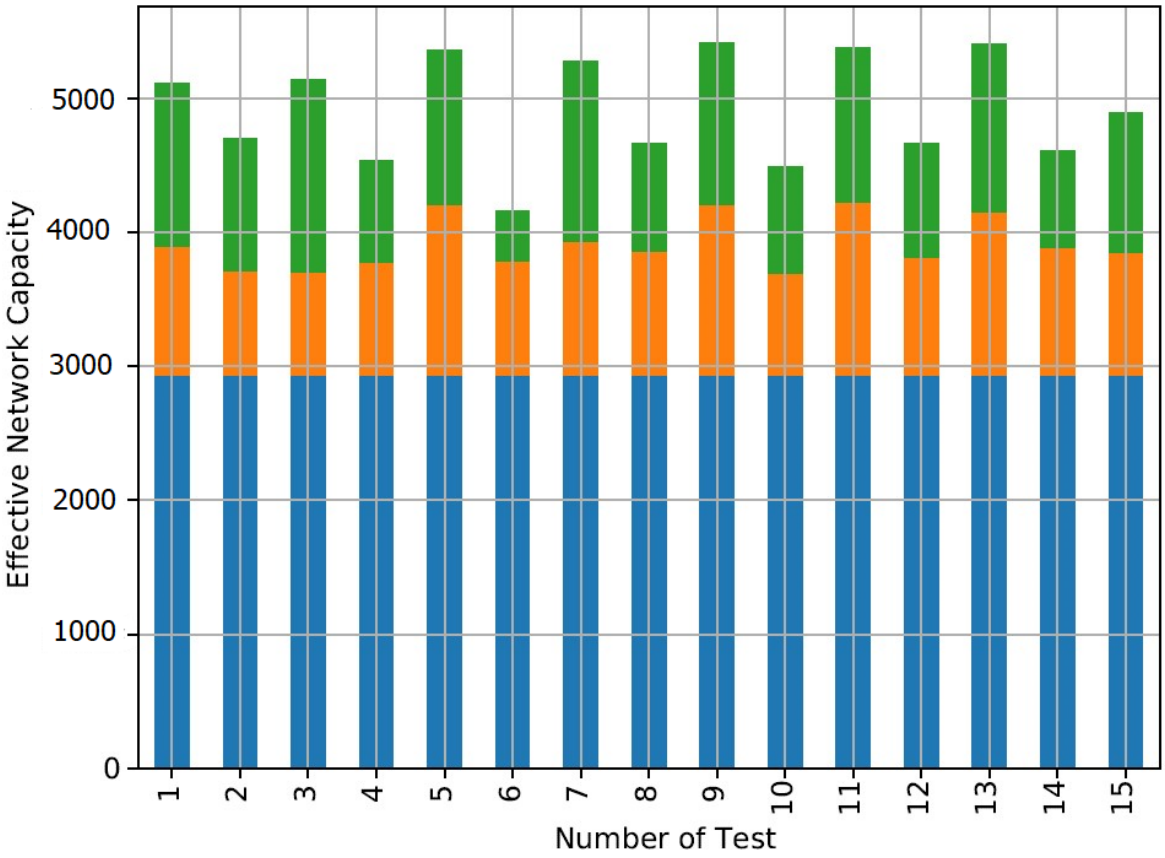


Figura 8-10 Multiagente uso de FSU en cada Test. Elaboración propia.

Capítulo 9

9. Análisis de resultados

Al ejecutar la repetición de la experiencia cada vez que el agente realiza una acción y al actualizar los parámetros de la red en línea, la red en línea comenzará a asociar ciertos pares de estado/acción con valores Q apropiados: mayor será la promesa de realizar una determinada acción en un determinado estado, el modelo comienza a predecir valores de Q más altos y comienza a sobrevivir durante más tiempo a medida que los agente sigan enrutando tráfico por la topología.

La Figura 8-2 refleja el uso de las 22 troncales bajo DQN en una configuración de 3 capas y 64 neuronas, donde cada troncal cuenta con 3 núcleos y 100 slots de frecuencia, cada que se agota el tiempo de separación el algoritmo de almacenamiento libera espacios espectrales. Se observa un estado de saturación sobre las 22 troncales de la red generado por demandas de ancho de banda posterior a los 5000 episodios y la fragmentación del espectro.

Los resultados de la configuración DQN sin desfragmentación son prometedores dado que se alcanza a administrar un tráfico total ofrecido por la red de 8 Tbps con una probabilidad de bloqueo baja 0.09 que permite que más del 99% de las demandas llegue a su destino. Dado que por troncal se administra una porción espectral de 100 FSU al aumentar esta capacidad la fragmentación causara una gran pérdida de recursos.

La Figura 8-5 traza la evolución de la probabilidad de bloqueo de solicitudes versus cada episodio, donde se puede detallar que DQN RMSCA multiagente después de un período de entrenamiento superior a 1000 episodios disminuye la probabilidad de bloqueo a un valor inferior a 1. En la contraparte de la probabilidad de bloqueo, la cual se representa como las demandas exitosas que se detalla en la Figura 8-6 se observa que su valor tiende a 1 después de los 1000 episodios de entrenamiento.

En un inicio el algoritmo explora el entorno buscando los mejores recursos para llevar el tráfico basándose en el retardo, después de 600 episodios la propuesta algorítmica comienza a explotar la información para asegurar la ruta con los recursos disponibles y maximizar la recompensa, esto se refleja en la Figura 8-5 donde es posible observar por el comportamiento de la gráfica que después del episodio 700 en configuración 3 64 el algoritmo realiza más explotación que exploración.

La Figura 8-7 detalla el comportamiento de la duración versus los episodios totales transcurridos, donde se tiene un comportamiento estable en la solución del algoritmo con un promedio de 3 periodos en tiempo de simulación, el comportamiento es similar al realizar cambio en la cantidad de capas a 5 y neuronas a128. Con este aumento de capas y neuronas se tiene un comportamiento de ganancia a nivel de probabilidad de bloqueo ya que se estabiliza con menos episodios ejecutados como se observa en la Figura 8-8.

Cada vez que los agentes operan cooperativamente para solucionar una demanda de tráfico van ajustando los pesos de la red de acuerdo con la explotación versus la exploración, en la Figura 8-9 se refleja el aumento en la efectividad de los agentes después de 350 episodios.

El trabajo realizado por el Algoritmo 2 es reflejado en la Figura 8-10, donde se visualiza la técnica de administración del espacio espectral atacando la diafonía y la fragmentación. Al contar con la asignación de super bloques y desfragmentar el espectro se obtienen mejores resultados que en las simulaciones sin fragmentación, alcanzando una administración de tráfico total ofrecido por la red de 189 Tbps antes de presentar una probabilidad de bloqueo considerable, generada principalmente por las demandas de alta capacidad.

Cuando la probabilidad de bloqueo se incrementa por la falta de recursos físicos el algoritmo da prioridad a las demandas de bajo recursos físicos y bloquea todas las demandas de alta capacidad, esto ocurre poque usa los espacios espectrales disponibles que quedan contiguos a los super bloques ya asignados.

Capítulo 10

Conclusiones, aportes originales, trabajos futuros y socialización de resultados

10.1. Conclusiones

La propuesta realizada permite abordar con otro enfoque el problema de la asignación de recursos físicos en una red óptica elástica mediante aprendizaje por refuerzo bajo la presencia de volúmenes de tráfico de alta variabilidad aleatoria. De forma puntual las conclusiones del proyecto son:

- Como resultado de esta investigación se establece un algoritmo de administración de recursos físicos basado en el comportamiento aleatorio de los volúmenes de tráfico generados por las actuales y futuras aplicaciones de Internet. Logrando la provisión de canales de acuerdo con cada demanda y garantizando alta disponibilidad. Las pruebas experimentales proporcionaron varias soluciones para administrar los recursos espectrales y espaciales sobre una topología de red elástica.
- Los resultados de la simulación evidencian que DQN RMSCA reduce la probabilidad de bloqueo en un 40% después de entrenada, en comparación con las redes WDM actuales. DQN en configuración 5 128 y 8 256 facilita el entrenamiento en comparación con la configuración 3 64. En promedio, DQN RMSCA toma 1000 episodios de entrenamiento para converger en configuración 3 64, 5 128, 8 256, respectivamente. Finalmente, la funcionalidad de cambio de canal en los enlaces de cada NE, proporciona estabilidad a nivel de administración de ancho de banda y recursos físicos.
- El algoritmo RMSCA propuesto además de seleccionar la ruta, el espectro y el formato de modulación, también ayuda a proveer el mejor esquema de canal según la variabilidad del tráfico.

- El establecimiento de una sociedad digital exige mayor disponibilidad a servicios y aplicaciones de Internet, llegando a considerarse un activo público al igual que en su momento sucedió con las carreteras y los ferrocarriles. Por lo tanto, una infraestructura de telecomunicaciones actualizada a los requerimientos de los usuarios es fundamental para que sean parte de esta emergente sociedad. Mediante RMSCA se alcanzará un mejor uso de los recursos de infraestructura apoyando los nuevos servicios que la sociedad demanda con requerimientos especiales de tráfico.
- Esta investigación propone una visión tecnológica para la administración de recursos ópticos apoyado en el aprendizaje automático mediante una colección de técnicas para adaptarse al comportamiento dinámico de la red.

10.2. Aportes originales

El algoritmo establecido permite abordar el proceso de asignación de recursos a las solicitudes de demandas de volúmenes de tráfico variable mejorando la explotación de los recursos físicos en una red óptica elástica mediante Deep reinforcement learning.

- Se establece una estrategia de asignación de espectro y espacio que aborda el problema de fragmentación en los sistemas multinúcleo.
- La estrategia establecida hace frente a la variabilidad de tráfico demandado por las aplicaciones emergentes, administrando los recursos para proveer de elasticidad a la red.

10.3. Trabajos futuros

Doctoral

Promover la continuidad en el desarrollo e implementación de nuevos modelos algorítmicos para la administración de SDM multiagente soportado en la elasticidad de la red óptica y la aleatoriedad del tráfico de internet.

Desarrollo de un modelo para la predicción de tráfico en redes ópticas elásticas.

Maestría

Para aprovechar las nuevas propuestas de aprendizaje por refuerzo multiagente es prioritario desarrollar investigaciones que permitan mejorar: el rendimiento de los algoritmos y los cálculos detallados de las bandas de protección necesarias para reducir la interferencia.

Implementar un sistema de red óptica elástica integrada con redes definidas por software para acceder hasta la capa de aplicación y agrupar demandas de tráfico según el tipo de tecnología.

10.4. Socialización de resultados

Como producto del proceso de investigación se han socializado mediante artículos y ponencias de carácter científico los resultados obtenidos que evidencian las ventajas de aplicar técnicas de RL de asignación de recursos físicos en una red óptica elástica mediante aprendizaje por refuerzo bajo la presencia de volúmenes de tráfico de alta variabilidad aleatoria.

Publicaciones

Título del artículo: Routing in elastic optical networks based on deep reinforcement learning for multi-agent systems, International Review on Modelling and Simulations (IREMOS). 2022.

Autores: Diego Fernando Aguirre, Julio Barón Velandia, Octavio Salcedo Parra

Título del artículo: Routing in Elastic Optical Networks Based on Deep Reinforcement Learning. International Review on Modelling and Simulations (IREMOS). 2022

Autores: Diego Fernando Aguirre, Julio Barón Velandia, Octavio Salcedo Parra

Título del artículo: Heuristic algorithm for flexible optical networks OTN. Springer, Cham International Conference on Smart Computing and Communication. 2018.

Autores: Diego Fernando Aguirre Moreno, Octavio José Salcedo Parra, Danilo Alfonso López Sarmiento

Ponencia

Título: Machine learning in optical networks, El Congreso Internacional en Ingeniería, Tecnología e Innovación – CIITI2022

Autores: Diego Fernando Aguirre, Julio Barón Velandia, Eduardo Avendaño

Referencias bibliográficas

Alashhab, Z. R., Anbar, M., Singh, M. M., Leau, Y.-B., Al-Sai, Z. A., & Alhayja'a, S. A. (2020). Impact of Coronavirus Pandemic Crisis on Technologies and Cloud Computing Applications. *Journal of Electronic Science and Technology*, *19*(1), 25–40. https://doi.org/10.1016/J.JNLEST.2020.100059

Al-Mekhlafi, A. A., Al-Kamali, O. M., Shaddad, R. Q., & Abduljalil, M. A. (2020). *Dense 5G Network Planning for Coastal Area at Aden City, Yemen*. 1–7. https://doi.org/10.1109/icoice48418.2019.9035140

Anusuya, V. V., & Gomathi, V. (2018, noviembre 27). Framework for advanced weather forecasting technique using DMaaS in Cloud. *Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies, ICCTCT 2018*. https://doi.org/10.1109/ICCTCT.2018.8550961

Ayoubi, S., Limam, N., Salahuddin, M. A., Shahriar, N., Boutaba, R., Estrada-Solano, F., & Caicedo, O. M. (2018). Machine Learning for Cognitive Network Management. *IEEE Communications Magazine*, *56*(1), 158–165. https://doi.org/10.1109/MCOM.2018.1700560

Barletta, L., Giusti, A., Rottondi, C., & Tornatore, M. (2017). QoT Estimation for Unestablished Lighpaths using Machine Learning. *Optical Fiber Communication Conference*, Th1J.1. https://doi.org/10.1364/OFC.2017.Th1J.1

Cai, J. X., Davidson, C. R., Lucero, A., Zhang, H., Foursa, D. G., Sinkin, O. V., Patterson, W. W., Pilipetskii, A. N., Mohs, G., & Bergano, N. S. (2012). 20 Tbit/s transmission over 6860 km with sub-nyquist channel spacing. *Journal of Lightwave Technology*, *30*(4), 651–657. https://doi.org/10.1109/JLT.2011.2179975

Chung, C., & Jeong, J. P. (2020). A Design of IoT Device Configuration Translator for Intent-Based IoT-Cloud Services. *2020 22nd International Conference on Advanced Communication Technology (ICACT)*, 52–56. https://doi.org/10.23919/ICACT48636.2020.9061282

Ciena Corporation - AnnualReports.com. (s/f). Recuperado el 9 de septiembre de 2022, a partir de https://www.annualreports.com/Company/ciena-corporation

Cisco. (2019). *Cisco Visual Networking Index: Forecast and Trends, 2017–2022 White Paper*.

Cisco. (2020). *Cisco Annual Internet Report (2018–2023) White Paper*. Cisco.

Clark, D. D., Clark, D. D., Partridge, C., & Ramming, J. C. (2003). A knowledge plane for the Internet. *IN SIGCOMM*, 3--10.

De A Barboza, E., Bastos-Filho, C. J. A., Martins-Filho, J. F., de Moura, U. C., & de Oliveira, J. R. F. (2013). Self-adaptive erbium-doped fiber amplifiers using machine learning. *2013 SBMO/IEEE MTT-S International Microwave & Optoelectronics Conference (IMOC)*, 1–5. https://doi.org/10.1109/IMOC.2013.6646588

de Araújo, D. R. B., Bastos-Filho, C. J. A., & Martins-Filho, J. F. (2015). Methodology to Obtain a Fast and Accurate Estimator for Blocking Probability of Optical Networks. *Journal of Optical Communications and Networking*, *7*(5), 380. https://doi.org/10.1364/JOCN.7.000380

Deo, R. C. (2015). Machine Learning in Medicine. *Circulation*, *132*(20), 1920–1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

Dutta-Roy, A. (2014). Prolog to, "A Tutorial on the Flexible Optical Networking Paradigm: State of the Art, Trends, and Research Challenges." *Proceedings of the IEEE*, *102*(9), 1312–1316. https://doi.org/10.1109/JPROC.2014.2341314

Gerstel, O., Jinno, M., Lord, A., & Yoo, S. J. (2012). Elastic optical networking: a new dawn for the optical layer? *IEEE Communications Magazine*, *50*(2), s12–s20. https://doi.org/10.1109/MCOM.2012.6146481

Heron, R. W., Pfeiffer, T., Veen, D. T. van, Smith, J., & Patel, S. S. (2008). Technology innovations and architecture solutions for the next-generation optical access network. *Bell Labs Technical Journal*, *13*(1), 163–181. https://doi.org/10.1002/bltj.20289

Hillerkuss, D., Schmogrow, R., Schellinger, T., Jordan, M., Winter, M., Huber, G., Vallaitis, T., Bonk, R., Kleinow, P., Frey, F., Roeger, M., Koenig, S., Ludwig, A., Marculescu, A., Li, J., Hoh, M., Dreschmann, M., Meyer, J., Ben Ezra, S., … Leuthold, J. (2011). 26 Tbit s−1 line-rate super-channel transmission utilizing all-optical fast Fourier transform processing. *Nature Photonics*, *5*(6), 364–371. https://doi.org/10.1038/nphoton.2011.74

Hillerkuss, D., Winter, M., Teschke, M., Marculescu, A., Li, J., Sigurdsson, G., Worms, K., Ben Ezra, S., Narkiss, N., Freude, W., & Leuthold, J. (2010). Simple all-optical FFT scheme enabling Tbit/s real-time signal processing. *Optics Express*, *18*(9), 9324. https://doi.org/10.1364/OE.18.009324

Hsu, D.-Z., Wei, C.-C., Chen, H.-Y., Li, W.-Y., & Chen, J. (2011). Cost-effective 33-Gbps intensity modulation direct detection multi-band OFDM LR-PON system employing a 10-GHz-based transceiver. *Optics Express*, *19*(18), 17546. https://doi.org/10.1364/OE.19.017546

Huawei. (2018). *Intelligent Word 2030*.

Hyun, J., & Hong, J. W. K. (2017). Knowledge-defined networking using in-band network telemetry. *19th Asia-Pacific Network Operations and Management Symposium: Managing a World of Things, APNOMS 2017*, 54–57. https://doi.org/10.1109/APNOMS.2017.8094178

Idowu-Bismark, O., Kennedy, O., Husbands, R., & Adedokun, M. (2019). 5G wireless communication network architecture and its key enabling technologies. *International Review of Aerospace Engineering*, *12*(2), 70–82. https://doi.org/10.15866/irease.v12i2.15461

ITU. (2012). *G.694.1 : Planes espectrales para las aplicaciones de multiplexación por división de longitud de onda: Plan de frecuencias con multiplexación por división de longitud de onda densa*. ITU. https://www.itu.int/rec/T-REC-G.694.1-201202-I/es

Jinno, M., Kozicki, B., Takara, H., Watanabe, A., Sone, Y., Tanaka, T., & Hirano, A. (2010). Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network [Topics in Optical Communications. *IEEE Communications Magazine*, *48*(8), 138–145. https://doi.org/10.1109/MCOM.2010.5534599

Jinno, M., Takara, H., Kozicki, B., Tsukishima, Y., Sone, Y., & Matsuoka, S. (2009). Spectrum-efficient and scalable elastic optical path network: Architecture, benefits, and enabling technologies. *IEEE Communications Magazine*, *47*(11), 66–73. https://doi.org/10.1109/MCOM.2009.5307468

Kalogirou, S. A., Panteliou, S., & Dentsoras, A. (1999). Artificial neural networks used for the performance prediction of a thermosiphon solar water heater. *Renewable Energy*, *18*(1), 87–99. https://doi.org/10.1016/S0960-1481(98)00787-3

Kalyani, V. L., & Pinky. (2017). *The Study of Self Driven Cars With Special Reference to Google Drive And Maruti Celerio In High Speed Internet Connectivity*. https://doi.org/10.5281/ZENODO.1134249

Khan, J., Wei, J. S., Ringnér, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu, C. R., Peterson, C., & Meltzer, P. S. (2001). Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature Medicine*, *7*(6), 673–679. https://doi.org/10.1038/89044

Király, O., Potenza, M. N., Stein, D. J., King, D. L., Hodgins, D. C., Saunders, J. B., Griffiths, M. D., Gjoneska, B., Billieux, J., Brand, M., Abbott, M. W., Chamberlain, S. R., Corazza, O., Burkauskas, J., Sales, C. M. D., Montag, C., Lochner, C., Grünblatt, E., Wegmann, E., … Demetrovics, Z. (2020). Preventing problematic internet use during the COVID-19 pandemic: Consensus guidance. *Comprehensive Psychiatry*, *100*, 152180. https://doi.org/10.1016/J.COMPPSYCH.2020.152180

Kiran, Y., Venkatesh, T., & Ram Murthy, C. (2007). A Reinforcement Learning Framework for Path Selection and Wavelength Selection in Optical Burst Switched Networks. *IEEE Journal on Selected Areas in Communications*, *25*(9), 18–26. https://doi.org/10.1109/JSAC-OCN.2007.028806

Koyama, T., Kawamoto, J., Kawaragi, M., Kurakake, T., & Saito, K. (2019). Implementing 8K Vision Mixer for Cloud-Based Production System. *SMPTE Motion Imaging Journal*, *128*(6), 30–37. https://doi.org/10.5594/JMI.2019.2910705

Langley, P., & Simon, H. A. (1995). Applications of machine learning and rule induction. *Communications of the ACM*, *38*(11), 54–64. https://doi.org/10.1145/219717.219768

Lee, S. C. J., Breyer, F., Randel, S., Ziemann, O., van den Boom, H. P. A., & Koonen, A. M. J. (2008). Low-Cost and Robust 1-Gbit/s Plastic Optical Fiber Link Based on Light-Emitting Diode Technology. *OFC/NFOEC 2008 - 2008 Conference on Optical Fiber Communication/National Fiber Optic Engineers Conference*, 1–3. https://doi.org/10.1109/OFC.2008.4528667

Li, Y., & Kilper, D. C. (2018). Optical Physical Layer SDN [Invited]. *Journal of Optical Communications and Networking*, *10*(1), A110. https://doi.org/10.1364/JOCN.10.00A110

Liu, Z., Li, H., Tang, X., Zhang, X., Lin, F., & Cheng, K. (2016). Extreme learning machine: a new alternative for measuring heat collection rate and heat loss coefficient of water-

in-glass evacuated tube solar water heaters. *SpringerPlus*, *5*(1), 626. https://doi.org/10.1186/s40064-016-2242-1

Lu, P., Zhang, L., Liu, X., Yao, J., & Zhu, Z. (2015). Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks. *IEEE Network*, *29*(5), 36–42. https://doi.org/10.1109/MNET.2015.7293303

Mehdi, L., & Hayar, A. (2018). New mathematical model based on OFDM technique for smart lighting systems in smart city. *International Review of Electrical Engineering*, *13*(3), 213–228. https://doi.org/10.15866/iree.v13i3.14388

Mestres, A., Rodriguez-Natal, A., Carner, J., Barlet-Ros, P., Alarcón, E., Solé, M., Muntés-Mulero, V., Meyer, D., Barkai, S., Hibbett, M. J., Estrada, G., Ma'ruf, K., Coras, F., Ermagan, V., Latapie, H., Cassar, C., Evans, J., Maino, F., Walrand, J., & Cabellos, A. (2017). Knowledge-defined networking. *Computer Communication Review*, *47*(3), 1–10. https://doi.org/10.1145/3138808.3138810

Metconnex. (s/f). Wavelength Selective Switching in Metro DWDM Networks. *Metconnex*, *Applicatio*.

Minn, H., Zeng, M., & Bhargava, V. K. (2000). On timing offset estimation for OFDM systems. *IEEE Communications Letters*, *4*(7), 242–244. https://doi.org/10.1109/4234.852929

Moharrami, M., Fallahpour, A., Beyranvand, H., & Salehi, J. A. (2017). Resource Allocation and Multicast Routing in Elastic Optical Networks. *IEEE Transactions on Communications*, *65*(5), 2101–2113. https://doi.org/10.1109/TCOMM.2017.2667664

Moreno, D. F. A., Parra, O. J. S., & Sarmiento, D. A. L. (2018). Heuristic algorithm for flexible optical networks OTN. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10699 LNCS*, 163–172. https://doi.org/10.1007/978-3-319-73830-7_17

Morioka, T., Awaji, Y., Ryf, R., Winzer, P., Richardson, D., & Poletti, F. (2012). Enhancing optical communications with brand new fibers. En *IEEE Communications Magazine* (Vol. 50, Issue 2). https://doi.org/10.1109/MCOM.2012.6146483

Mukherjee, B. (2000). WDM optical communication networks: progress and challenges. *IEEE Journal on Selected Areas in Communications*, *18*(10), 1810–1824. https://doi.org/10.1109/49.887904

Mukherjee, Biswanath., & Mukherjee, Biswanath. (2006). *Optical WDM networks*. Springer.

Murthy, C. S. Ram., & Gurusamy, Mohan. (2002). *WDM optical networks : concepts, design, and algorithms*. Prentice Hall PTR.

Oda, S., Miyabe, M., Yoshida, S., Katagiri, T., Aoki, Y., Hoshida, T., Rasmussen, J. C., Birk, M., & Tse, K. (2017a). A Learning Living Network With Open ROADMs. *Journal of Lightwave Technology*, *35*(8), 1350–1356. https://doi.org/10.1109/JLT.2017.2660540

Oda, S., Miyabe, M., Yoshida, S., Katagiri, T., Aoki, Y., Hoshida, T., Rasmussen, J. C., Birk, M., & Tse, K. (2017b). A Learning Living Network With Open ROADMs. *Journal of Lightwave Technology*, *35*(8), 1350–1356. https://doi.org/10.1109/JLT.2017.2660540

Optical Technologies for Spatial Division Multiplexing (SDM) SDMThechs. (s/f). Recuperado el 21 de octubre de 2019, a partir de https://www.it.pt/Projects/Index/4534

O'Shea, T. J., & Hoydis, J. (2017). An Introduction to Machine Learning Communications Systems. En *Machine Learning Methods for Ecological Applications*. https://doi.org/10.1007/979-20010-1

Perasso, V., & Mundo, B. (2016). *Noticias América Latina Internacional Economía Tecnología Ciencia Qué es la cuarta revolución industrial (y por qué debería preocuparnos)*.

Ramaswami, Rajiv., Sivarajan, K. N., & Sasaki, G. H. (Galen H. (2010). *Optical networks : a practical perspective*. Elsevier/Morgan Kaufmann.

Sanjoh, H., Yamada, E., & Yoshikuni, Y. (s/f). Optical orthogonal frequency division multiplexing using frequency/time domain filtering for high spectral efficiency up to 1 bit/s/Hz. *Optical Fiber Communication Conference and Exhibit*, 401–402. https://doi.org/10.1109/OFC.2002.1036444

Sano, A., Yamada, E., Masuda, H., Yamazaki, E., Kobayashi, T., Yoshida, E., Miyamoto, Y., Kudo, R., Ishihara, K., & Takatori, Y. (2009). No-Guard-Interval Coherent Optical OFDM for 100-Gb/s Long-Haul WDM Transmission. *Journal of Lightwave Technology*, *27*(16), 3705–3713. https://doi.org/10.1109/JLT.2009.2023369

Schmidl, T. M., & Cox, D. C. (1997). Robust frequency and timing synchronization for OFDM. *IEEE Transactions on Communications*, *45*(12), 1613–1621. https://doi.org/10.1109/26.650240

Shieh, William., & Djordjevic, Ivan. (2010). *OFDM for optical communications*. Academic Press/Elsevier.

Sommer, C., & Gerlich, D. W. (2013). Machine learning in cell biology – teaching computers to recognize phenotypes. *Journal of Cell Science*, *126*(24), 5529–5539. https://doi.org/10.1242/jcs.123604

T. Jiménez, Aguado, J. C., Miguel, I. de, Durán, R. J., Fernández, N., Angelou, M., Sánchez, D., Merayo, N., Fernández, P., Atallah, N., Lorenzo, R. M., Tomkos, I., & Abril, E. J. (2012). A cognitive system for fast Quality of Transmission estimation in core optical networks. *IEEE Network*.

Tanha, M., Sajjadi, D., Ruby, R., & Pan, J. (2018). Traffic Engineering Enhancement by Progressive Migration to SDN. *IEEE Communications Letters*, *22*(3), 438–441. https://doi.org/10.1109/LCOMM.2018.2789419

Tarca, A. L., Carey, V. J., Chen, X., Romero, R., & Drăghici, S. (2007). Machine Learning and Its Applications to Biology. *PLoS Computational Biology*, *3*(6), e116. https://doi.org/10.1371/journal.pcbi.0030116

Technologies for spatial-division multiplexing: The next frontier in optical communications | Chalmers. (s/f). Recuperado el 21 de octubre de 2019, a partir de https://www.chalmers.se/en/projects/Pages/Technologies-for-spatial-division-multiplexingQ-The-next.aspx

Tucker, R. S. (2008). Optical packet switching: A reality check. *Optical Switching and Networking*, *5*(1), 2–9. https://doi.org/10.1016/j.osn.2007.08.001

Velasco, L., Vela, A. P., Morales, F., & Ruiz, M. (2017). Designing, operating, and reoptimizing elastic optical networks. *Journal of Lightwave Technology*, *35*(3), 513–526. https://doi.org/10.1109/JLT.2016.2593986

Venkatesh, T., Sankar, A., Jayaraj, A., & Siva Ram Murthy, C. (2008). A Complete Framework to Support Controlled Burst Retransmission in Optical Burst Switching Networks. *IEEE Journal on Selected Areas in Communications*, *26*(3), 65–73. https://doi.org/10.1109/JSAC-OCN.2008.029507

Vissicchio, S., Vanbever, L., Cittadini, L., Xie, G. G., & Bonaventure, O. (2017). Safe Update of Hybrid SDN Networks. *IEEE/ACM Transactions on Networking*, *25*(3), 1649–1662. https://doi.org/10.1109/TNET.2016.2642586

Vokkarane, V., & Jue, J. P. (2003). *Practical Priority Contention Resolution for Slotted Optical Burst Switching*.

Wang, H., Zhao, J., Li, H., & Ji, Y. (2016). Opaque virtual network mapping algorithms based on available spectrum adjacency for elastic optical networks. *Science China Information Sciences*, *59*(4). https://doi.org/10.1007/s11432-016-5525-9

Wei, W., Wang, C., & Yu, J. (2012). Cognitive optical networks: Key drivers, enabling techniques, and adaptive bandwidth services. *IEEE Communications Magazine*, *50*(1), 106–113. https://doi.org/10.1109/MCOM.2012.6122540

Wernick, M., Yang, Y., Brankov, J., Yourganov, G., & Strother, S. (2010). Machine Learning in Medical Imaging. *IEEE Signal Processing Magazine*, *27*(4), 25–38. https://doi.org/10.1109/MSP.2010.936730

Winzer, P. J. (2014). Spatial multiplexing in fiber optics: The scaling of metro/core capacities. *Bell Labs Technical Journal*, *19*, 22–30. https://doi.org/10.15325/BLTJ.2014.2347431

Xiaoxia Wu, Jargon, J. A., Skoog, R. A., Paraschis, L., & Willner, A. E. (2009). Applications of Artificial Neural Networks in Optical Performance Monitoring. *Journal of Lightwave Technology*, *27*(16), 3580–3589. https://doi.org/10.1109/JLT.2009.2024435

Yonenaga, K., Sano, A., Yamazaki, E., Inuzuka, F., Miyamoto, Y., Takada, A., & Yamada, T. (2008). 100 Gbit/s All-Optical OFDM Transmission Using 4 × 25 Gbit/s Optical Duobinary Signals with Phase-Controlled Optical Sub-Carriers. *OFC/NFOEC 2008 - 2008 Conference on Optical Fiber Communication/National Fiber Optic Engineers Conference*, 1–3. https://doi.org/10.1109/OFC.2008.4528078

Zhang, G., De Leenheer, M., Morea, A., & Mukherjee, B. (2013). A survey on OFDM-based elastic core optical networking. *IEEE Communications Surveys and Tutorials*, *15*(1), 65–87. https://doi.org/10.1109/SURV.2012.010912.00123

Zhou, Y., Liu, L., Wang, L., Hui, N., Cui, X., Wu, J., Peng, Y., Qi, Y., & Xing, C. (2020). Service-aware 6G: An intelligent and open network based on the convergence of communication, computing and caching. *Digital Communications and Networks*, *6*(3), 253–260. https://doi.org/10.1016/J.DCAN.2020.05.003