

UNIVERSIDAD NACIONAL DE SAN AGUSTÍN DE
AREQUIPA
FACULTAD DE INGENIERÍA DE PRODUCCIÓN Y SERVICIOS
ESCUELA PROFESIONAL DE INGENIERÍA EN
TELECOMUNICACIONES



Curso
Perspectiva y Enfoque de la Investigación Científica

Docente: Alberth Ronal Tamo Calla

**Tema: Optimización de la Asignación de Recursos en
Redes 5G-IoT Mediante Técnicas de Aprendizaje
Automático**

Integrantes del Grupo:

Turpo Quispe, Darwin Cristhian
Luque Llanqui, Vladimir Williams
Chacca Villagra Royer Paul

1 de octubre de 2025

1. Marco Teórico

1.1 Introducción al Paradigma Convergente 5G-IoT

La quinta generación de redes móviles (5G) y el Internet de las Cosas (IoT) representan dos de las transformaciones tecnológicas más significativas de la era digital. Aunque a menudo se discuten por separado, su verdadera capacidad disruptiva emerge de su convergencia. Esta sinergia crea un nuevo paradigma de conectividad inteligente, pero también introduce desafíos sin precedentes en la gestión de redes, particularmente en el ámbito de la asignación de recursos.

1.2 Fundamentos de las Redes 5G: Más Allá de la Velocidad

Las redes 5G constituyen una evolución paradigmática con respecto a sus predecesoras, trascendiendo la simple mejora en la velocidad de datos para ofrecer una plataforma de comunicación versátil y de alto rendimiento[1]. Su arquitectura está diseñada para ser un “perfecto fundamento para que los ecosistemas IoT prosperen” [2], sustentada en tres categorías de servicio fundamentales definidas por organismos de estandarización como la Unión Internacional de Telecomunicaciones (ITU) y el 3GPP[3]:

- **Banda Ancha Móvil Mejorada (eMBB):** Proporciona velocidades de datos de varios gigabits por segundo, diseñadas para soportar aplicaciones de alto consumo de ancho de banda como el streaming de video en ultra alta definición, la realidad virtual (VR) y la realidad aumentada (AR)[4].
- **Comunicaciones Ultra Confiables y de Baja Latencia (URLLC):** Es el habilitador clave para aplicaciones de misión crítica. Ofrece una fiabilidad del 99.999 % y latencias de extremo a extremo en el rango de 1 milisegundo, condiciones indispensables para vehículos autónomos, cirugía remota, control de redes eléctricas inteligentes y automatización industrial avanzada[3].
- **Comunicaciones Masivas de Tipo Máquina (mMTC):** Está diseñada para soportar una densidad masiva de dispositivos conectados, con capacidad para hasta un millón de dispositivos por kilómetro cuadrado. Estos dispositivos se caracterizan por un bajo consumo energético y transmisiones de datos esporádicas, siendo ideal para despliegues a gran escala en ciudades inteligentes, agricultura de precisión y logística[4].

1.3 El Ecosistema del Internet de las Cosas (IoT): Un Universo de Dispositivos Heterogéneos

El IoT se concibe como una red global de miles de millones de dispositivos físicos interconectados, que van desde simples sensores de temperatura y actuadores hasta maquinaria industrial compleja y wearables avanzados[1]. Cada uno de estos dispositivos recopila y comparte datos, creando un flujo de información sin precedentes. La “rápida proliferación de dispositivos IoT y la diversidad de sus requisitos” es precisamente el catalizador principal del complejo problema de la gestión de recursos de red. Las aplicaciones de IoT abarcan prácticamente todos los sectores de la economía y la sociedad, incluyendo la salud (monitoreo remoto de pacientes), la industria (mantenimiento predictivo), la agricultura

(sensores de humedad del suelo), las ciudades inteligentes (gestión del tráfico y alumbrado público), los hogares inteligentes (automatización) y el transporte (vehículos conectados).

1.4 Sinergias y Desafíos de la Integración 5G-IoT

La convergencia de 5G e IoT no es una mera suma de tecnologías, sino una simbiosis que da lugar a un ecosistema de “conectividad inteligente”[4]. La red 5G proporciona la infraestructura de comunicación robusta, de baja latencia y alta capacidad que el IoT necesita para desarrollar todo su potencial, permitiendo aplicaciones que antes eran inviables[6].

Sin embargo, esta poderosa sinergia genera un desafío central: la gestión eficiente de recursos en redes que son, por diseño, de “alta densidad” (debido a mMTC) y extremadamente “dinámicas” (debido a la movilidad de los usuarios y los requisitos fluctuantes de URLLC y eMBB). Una tensión fundamental surge de la propia arquitectura de servicios de 5G. La red debe gestionar simultáneamente, y a menudo en la misma área geográfica, dispositivos con demandas de recursos radicalmente opuestas. Por ejemplo, un sensor de bajo consumo (mMTC) requiere una conectividad mínima y esporádica, mientras que un vehículo autónomo cercano (URLLC) exige una conexión infalible y de latencia casi nula, y una aplicación de realidad aumentada (eMBB) en el mismo sector demanda un ancho de banda masivo. Esto transforma el problema de la asignación de recursos de una simple cuestión de “cantidad” a un complejo problema de optimización multidimensional de “calidad”, “prioridad” y “temporalidad”, que los enfoques de gestión estáticos no pueden resolver.

1.5 El Problema Central de la Asignación de Recursos en Entornos Dinámicos

La asignación de recursos en redes 5G-IoT se define como la orquestación óptima y dinámica de recursos de red finitos —tales como el espectro de radiofrecuencia, la potencia de transmisión, los bloques de recursos computacionales y la conectividad de backhaul— entre una multitud de agentes heterogéneos (dispositivos IoT, usuarios finales, aplicaciones de misión crítica) con demandas variables en el tiempo y en un entorno de red propenso a fluctuaciones rápidas[5]. El objetivo no es la maximización de un único parámetro, como el rendimiento agregado, sino un problema de optimización multiobjetivo que busca un equilibrio entre métricas a menudo contrapuestas[7].

1.6 Métricas Clave de Rendimiento y Calidad de Servicio (QoS)

Cualquier solución viable para la asignación de recursos debe ser evaluada en función de un conjunto de métricas críticas que definen el rendimiento de la red y la experiencia del usuario. El documento de investigación principal identifica la gestión de la Calidad de Servicio (QoS) como un “impedimento” fundamental en la asignación de recursos. Las métricas clave incluyen:

- **Calidad de Servicio (QoS):** Un concepto multifacético que engloba la fiabilidad de la conexión, la tasa de error de bloque (BLER), la disponibilidad del servicio y el cumplimiento de los acuerdos de nivel de servicio (SLA) para diferentes aplicaciones y segmentos de red[8].

- **Latencia:** El tiempo de retardo en la transmisión de datos de extremo a extremo. Es una métrica primordial para las aplicaciones URLLC, donde cada milisegundo es crítico. El objetivo es lograr mejoras sustanciales en la latencia.
- **Ancho de Banda:** La capacidad de transmisión de datos, medida típicamente en bits por segundo. Es esencial para los servicios eMBB que manejan grandes volúmenes de datos. Se buscan mejoras significativas en la utilización del ancho de banda disponible.
- **Eficiencia Energética:** El consumo de energía tanto de la infraestructura de red como de los dispositivos terminales. Esta métrica es crucial para la sostenibilidad económica y medioambiental de los despliegues masivos de IoT (especialmente mMTC) y para prolongar la vida útil de los dispositivos alimentados por batería.

1.7 Estado del Arte en la Optimización de Recursos para Redes 5G

Para contextualizar la contribución de este marco teórico, es esencial realizar un análisis crítico de la investigación previa en el campo de la asignación de recursos para redes 5G.

1.7.1 Análisis Crítico de Enfoques Previos

La literatura existente ha abordado el problema desde diversas perspectivas, cada una con sus propias fortalezas y debilidades:

- **Ukkalli et al.** propusieron DYSOLVE, un enfoque para la partición dinámica de redes utilizando recursos de múltiples operadores, enfocado en escenarios vehiculares. Aunque efectivo para gestionar la fragmentación dinámica, su principal limitación es que requiere “mejor gestión y modificaciones alternativas para mejorar la robustez y eficiencia del sistema en redes dinámicas”.
- **Li y Hu** se centraron en el procesamiento rápido de datos basado en Redes Definidas por Software (SDN) para mejorar la asignación de recursos para eMBB y URLLC. Su trabajo logra reducir la latencia, pero la tecnología “tiene limitaciones que requieren una evaluación y modificaciones exhaustivas de la red de IoT” para ser aplicable a gran escala.
- **Wu et al.** introdujeron un modelo de fragmentación de red inspirado en principios biológicos para aumentar la flexibilidad en la asignación de recursos. Sin embargo, esta técnica de implementación “necesita una integración y supervisión completas” y depende de la actualización continua de grupos de usuarios.
- **Randhava et al.** abordaron la gestión del tráfico 5G mediante la asignación dinámica de espectro basada en comunicación Dispositivo a Dispositivo (D2D). Lograron mejoras notables en el rendimiento del sistema y la relación señal-interferencia-ruido (SINR), pero el método se considera eficaz principalmente para “sistemas pequeños pero requiere recursos adicionales para redes extensas”.

- **Pinto et al.** argumentaron a favor de la segmentación para la asignación de recursos y el aislamiento de usuarios, demostrando una reducción en el tiempo de finalización de flujo y la pérdida de paquetes en simulaciones. No obstante, su trabajo carecía de “información sobre la adaptación al mundo real y la capa de tránsito”, lo que limita la evaluación de su efectividad práctica.

1.7.2 Tabla Comparativa del Estado del Arte

La siguiente tabla sintetiza las contribuciones y limitaciones de los enfoques discutidos, ofreciendo una visión panorámica del panorama de la investigación y destacando un patrón recurrente de desafíos relacionados con la escalabilidad, la robustez y la aplicabilidad en el mundo real.

Autor(es)	Contribución Principal	Metodología Clave	Limitaciones Identificadas
Kukkalli et al.	Gestión de la fragmentación dinámica de la red en escenarios vehiculares.	Partición de red multioperador (DYSOLVE).	Requiere mejor gestión y modificaciones para mejorar la robustez y eficiencia en redes dinámicas.
Li y Hu	Reducción de latencia en redes sensibles al retardo para eMBB y URLLC.	Procesamiento de datos basado en SDN.	Necesita evaluación y modificaciones exhaustivas para redes IoT a gran escala.
Wu et al.	Flexibilidad en la asignación de recursos mediante fragmentación de red.	Modelo de beneficios de red basado en propiedades de partículas.	Requiere integración y supervisión completas; asignación basada en grupos de usuarios actualizados.
Randhava et al.	Reducción de interferencia y mejora del rendimiento en tráfico 5G.	Asignación de espectro dinámico basada en D2D.	Eficaz para sistemas pequeños, pero necesita recursos adicionales para redes extensas.
Pinto et al.	Reducción del tiempo de finalización de flujo y pérdida de paquetes.	Segmentación para asignación de recursos y aislamiento de usuarios.	Falta de información sobre la adaptación al mundo real y la capa de tránsito para ser efectiva.

Cuadro 1: Comparativa del Estado del Arte en Asignación de Recursos 5G-IoT

1.7.3 Identificación de Brechas en la Investigación

El análisis del estado del arte revela que, si bien se han logrado avances significativos en aspectos específicos de la asignación de recursos, la mayoría de los trabajos existentes carecen de un enfoque holístico que sea simultáneamente “escalable, adaptativo y eficiente” para el complejo entorno convergente 5G-IoT. La brecha de investigación principal es la ausencia de soluciones integradas que combinen de manera unificada el análisis de datos en tiempo real, las arquitecturas de computación en el borde y algoritmos de aprendizaje automático avanzados. Los enfoques previos tienden a concentrarse en aplicaciones o contextos específicos, sin ofrecer un marco pragmático y extensible que pueda abordar la diversidad y dinamismo de los servicios 5G-IoT de manera integral.

1.8 Arquitecturas y Tecnologías Habilitadoras para la Optimización Inteligente

Para implementar una solución de gestión de recursos verdaderamente inteligente y proactiva, es necesario apoyarse en un conjunto de tecnologías habilitadoras que son fundamentales en la arquitectura 5G. Estas tecnologías no son meramente parte del entorno, sino componentes activos e interdependientes que forman la base para la optimización basada en aprendizaje automático.

1.8.1 Computación en el Borde (Edge Computing): Acercando la Inteligencia al Dato

La Computación en el Borde (Edge Computing) es un paradigma de computación distribuida que traslada el procesamiento de datos y la ejecución de aplicaciones desde servidores centralizados en la nube hacia la periferia de la red, más cerca de las fuentes de datos, como los dispositivos IoT[5]. La arquitectura jerárquica muestra cómo los nodos de borde actúan como una capa intermedia entre los dispositivos terminales y el centro de datos, permitiendo un procesamiento local y rápido.

Los beneficios directos de esta arquitectura en el contexto 5G-IoT son múltiples:

- **Reducción drástica de la latencia:** Al procesar los datos localmente, se elimina el tiempo de ida y vuelta (round-trip time) a un servidor en la nube distante. Esto es absolutamente crítico para las aplicaciones URLLC, como los vehículos autónomos o la telecirugía, donde las decisiones deben tomarse en milisegundos[12].
- **Optimización del ancho de banda de la red troncal:** En lugar de transmitir flujos masivos de datos brutos desde millones de sensores a la nube, los nodos de borde pueden pre-procesar, filtrar y agregar esta información, enviando solo los datos relevantes o los resultados del análisis. Esto reduce significativamente la congestión en la red central[13].
- **Mejora de la privacidad y la seguridad:** Los datos sensibles pueden ser procesados y anonimizados en el borde sin necesidad de ser transmitidos a través de redes públicas, lo que reduce la superficie de ataque y ayuda a cumplir con las regulaciones de soberanía de datos[12].

A pesar de sus ventajas, el Edge Computing también presenta desafíos, como la gestión de recursos computacionales en los propios nodos de borde y la necesidad de proteger una infraestructura más distribuida y, por tanto, con más puntos de posible vulnerabilidad[12].

1.8.2 Análisis de Datos en Tiempo Real y Enrutamiento Dinámico

El análisis de datos en tiempo real es el motor que impulsa la toma de decisiones inteligentes en la red. Se considera “crítico para la asignación eficiente de recursos” porque permite que el sistema reaccione instantáneamente a las condiciones cambiantes de la red, como picos de tráfico o degradación del canal[5]. Este análisis continuo alimenta los algoritmos de Enrutamiento Dinámico, que se define formalmente como “un método para seleccionar la forma más eficiente de procesar y entregar datos basándose en el análisis de la carga y el tráfico de la red”. En la práctica, esto significa que las rutas de los paquetes de datos no son fijas, sino que se recalculan constantemente para evitar cuellos de botella,

minimizar la latencia y garantizar que se cumplan los requisitos de QoS de cada flujo de datos[10].

1.8.3 Segmentación de Red (Network Slicing): Virtualización para la Diversidad de Servicios

La Segmentación de Red (Network Slicing) es una tecnología fundamental de la arquitectura 5G que permite la creación de múltiples redes virtuales lógicas, de extremo a extremo y aisladas, sobre una única infraestructura física compartida[9]. Cada “slice” o segmento puede ser configurado y optimizado de forma independiente para satisfacer los requisitos específicos de un servicio o aplicación particular. Por ejemplo, se puede crear un slice con ancho de banda masivo para servicios eMBB, otro con latencia ultra baja y alta fiabilidad para URLLC, y un tercero optimizado para la conectividad de bajo consumo de millones de dispositivos mMTC[14]. El Network Slicing es, por tanto, el mecanismo estructural que habilita la asignación de recursos diferenciada y garantiza el aislamiento del rendimiento entre servicios dispares.

Estas tres tecnologías —Edge Computing, Análisis en Tiempo Real y Network Slicing— no operan de forma aislada. Forman un “triángulo virtuoso” interdependiente que es crucial para la gestión inteligente de la red. El Network Slicing particiona los recursos físicos para crear contenedores virtuales adaptados a cada servicio. El Edge Computing aloja las funciones de red virtualizadas (VNFs) que componen estos slices y proporciona la capacidad de cómputo distribuida para procesar los datos localmente con baja latencia. Finalmente, el Análisis en Tiempo Real, impulsado por el aprendizaje automático, actúa como el “cerebro” del sistema, dirigiendo la asignación dinámica de recursos dentro y entre los slices en función de la demanda prevista y las condiciones de la red. Ninguna de estas tecnologías puede resolver el complejo problema de la asignación de recursos 5G-IoT de manera eficiente por sí sola; su poder reside en su integración sinérgica.

1.9 El Aprendizaje Automático como Solución Estratégica y Proactiva

Dada la complejidad, dinamismo y escala de las redes 5G-IoT, el Aprendizaje Automático (Machine Learning, ML) emerge no solo como una herramienta de optimización, sino como un componente estratégico indispensable para una gestión de red proactiva y autónoma.

1.9.1 Justificación del Uso de Machine Learning (ML)

El ML es fundamental para la asignación de recursos porque proporciona “evaluaciones inteligentes y flexibles” en sistemas complejos y en constante evolución, superando las limitaciones de los enfoques algorítmicos tradicionales. Sus atributos clave en este contexto son:

- **Adaptabilidad y Escalabilidad:** Los modelos de ML pueden aprender y adaptarse continuamente a las condiciones cambiantes de la red, como patrones de tráfico, movilidad de usuarios y condiciones del canal de radio, y pueden escalar para gestionar eficientemente millones de dispositivos y flujos de datos.
- **Toma de Decisiones en Tiempo Real:** A diferencia de los métodos estáticos que requieren una reconfiguración manual o se basan en reglas predefinidas, el ML puede

“prever las condiciones de la red y asignar recursos en tiempo real”, permitiendo ajustes dinámicos en escalas de tiempo de milisegundos[5].

- **Optimización Proactiva:** La capacidad predictiva del ML es su mayor fortaleza. Permite anticipar futuras demandas de recursos y posibles cuellos de botella, posibilitando la mitigación de la congestión antes de que afecte la QoS, cumpliendo así con el requisito de una gestión proactiva[5].
- **Mejora del Rendimiento Global:** Al distribuir los recursos de manera inteligente basándose en la demanda real y prevista, los algoritmos de ML mejoran la estabilidad, la eficiencia y el rendimiento general de la red, optimizando simultáneamente la latencia, el ancho de banda y el consumo de energía[5].

1.9.2 Tipologías de Algoritmos de ML Aplicables

Las tres principales categorías de aprendizaje automático tienen aplicaciones directas en la gestión de redes 5G-IoT[15]:

- **Aprendizaje Supervisado:** Utiliza conjuntos de datos históricos y etiquetados para entrenar modelos de clasificación y regresión. Se aplica para predecir la demanda futura de ancho de banda o clasificar el estado de la red (por ejemplo, congestionado o normal). Algoritmos como Redes Neuronales, Máquinas de Vectores de Soporte (SVM) y Árboles de Decisión son altamente relevantes[16].
- **Aprendizaje No Supervisado:** Trabaja con datos no etiquetados para descubrir patrones y estructuras ocultas[15]. Esto permite una gestión de recursos más personalizada y eficiente.
- **Aprendizaje por Refuerzo (RL):** Este es un paradigma en el que un agente aprende a tomar decisiones óptimas de forma autónoma a través de la interacción continua con un entorno dinámico. El Aprendizaje por Refuerzo Profundo (DRL), que integra RL con redes neuronales profundas, es particularmente poderoso para manejar los enormes espacios de estado y acción de las redes 5G[11].

1.9.3 Modelos Específicos para la Asignación de Recursos y Predicción de Tráfico

Una solución de ML óptima para este dominio no consiste en un único algoritmo monolítico, sino en un sistema híbrido y jerárquico de modelos que operan en diferentes escalas de tiempo y niveles de abstracción. Existe una separación funcional clara entre los modelos de predicción y los de control:

- **Predicción de Tráfico (Capa Estratégica):** La asignación proactiva de recursos requiere, por definición, una previsión precisa de la demanda futura. Los modelos de redes neuronales recurrentes, como Long Short-Term Memory (LSTM) y Gated Recurrent Unit (GRU), son especialmente eficaces para analizar series temporales y capturar las complejas dependencias temporales en los datos de tráfico de la red. Estos modelos pueden operar a una escala de tiempo de mediano plazo (ej. predecir la carga de la red para los próximos segundos o minutos), estableciendo los objetivos y restricciones para la capa de control[18].

- **Asignación de Recursos (Capa Táctica):** Una vez que se tiene una predicción de la demanda, se necesita un mecanismo para tomar decisiones de asignación de recursos en tiempo real (milisegundos) que se alinee con esa predicción. Los algoritmos de DRL, y específicamente las Deep Q-Networks (DQN), son ideales para esta tarea de control táctico. En este marco, el “estado” del entorno puede incluir la demanda predicha por el modelo LSTM/GRU, y la “acción” es la asignación concreta de recursos (ej. ancho de banda, bloques de frecuencia). El agente DQN aprende una política para maximizar la recompensa a largo plazo, que se define en términos de cumplimiento de la QoS y eficiencia de la red[17].

Esta arquitectura de IA de dos capas, con una capa de predicción estratégica y una capa de control táctico, permite una gestión de recursos que es a la vez previsor y reactiva, alineando las decisiones a largo plazo con las acciones instantáneas necesarias para mantener un rendimiento óptimo de la red.

1.10 Formulaciones Teóricas del Modelo Propuesto

Para establecer un puente entre la teoría conceptual y la implementación práctica, el enfoque propuesto se fundamenta en un conjunto de definiciones, hipótesis y teoremas formales. Estas formulaciones matemáticas no son meramente ilustrativas, sino que representan la abstracción de principios de diseño complejos en modelos cuantificables, proporcionando una base rigurosa y verificable para el marco de optimización.

1.10.1 Análisis de Datos en Tiempo Real y Enrutamiento Dinámico: Fundamentos

El marco comienza por formalizar el impacto de la inteligencia en tiempo real en la red:

- **Definición 1 (Enrutamiento Dinámico):** Se establece una definición precisa del concepto: “El enrutamiento dinámico es un método para seleccionar la forma más eficiente de procesar y entregar datos basándose en el análisis de la carga y el tráfico de la red”.
- **Hipótesis 1 (Aumento del Rendimiento):** Se postula una predicción cuantitativa sobre el beneficio de integrar sistemas de análisis en tiempo real y enrutamiento dinámico, afirmando que “el rendimiento general de la red aumentará hasta en un 30 %”.
- **Prueba Matemática (Ecuación 1):** Esta hipótesis se formaliza a través de un modelo matemático simplificado que cuantifica el impacto esperado:

$$P_{total} = P_n \cdot (1 + 0,3) \quad (1)$$

Donde P_{total} representa el rendimiento total de la red con los sistemas inteligentes integrados, y P_n es el rendimiento de la red sin ellos. Esta ecuación sirve como un modelo base para evaluar el beneficio de la inteligencia en la red.

1.10.2 Integración de Edge Computing: Cuantificación del Impacto

A continuación, se cuantifica el beneficio de acercar el cómputo a la fuente de datos:

- **Teorema 1 (Reducción del Tiempo de Procesamiento):** Se enuncia el siguiente teorema: “La implementación de tecnologías de edge computing reduce el tiempo de procesamiento en un 20 %, lo que permite una asignación eficiente de recursos en los dispositivos IoT”.
- **Prueba Matemática (Ecuaciones 2 y 3):** El teorema se sustenta en un modelo que relaciona el tiempo de procesamiento y la eficiencia de los recursos:

$$t_{p.e} = t_{p.c} \cdot (1 - 0,2) \quad (2)$$

Aquí, $t_{p.e}$ es el tiempo de procesamiento con Edge Computing y $t_{p.c}$ es el tiempo de procesamiento basado en la nube. La reducción del 20 % se modela con el factor $(1 - 0,2)$. Esta reducción en el tiempo de procesamiento tiene un impacto directo en la eficiencia de la asignación de recursos ($R_{a.eff}$), que se expresa como una función j que depende del tiempo de procesamiento reducido y los recursos disponibles (r_a):

$$R_{a.eff} = j(t_{p.e}, r_a) \quad (3)$$

Este modelo vincula directamente la mejora en la latencia con la optimización de los recursos.

1.10.3 Fundamento Matemático del Algoritmo de Optimización de Recursos

Finalmente, se proporciona un modelo matemático que describe cómo el aprendizaje automático logra una asignación eficiente:

- **Lema 1 (Eficiencia del ML):** Se establece el lema que afirma que el algoritmo de optimización propuesto utiliza ML para “proporcionar una asignación eficiente de recursos para los dispositivos en las redes 5G-IoT”.
- **Prueba Matemática (Ecuación 4):** Para demostrar este lema, se presenta un modelo integral que captura la eficiencia de la asignación de recursos ($A_{r.eff}$) como una función de la eficiencia optimizada por ML, la prioridad de los recursos y otros parámetros de la red:

$$A_{r.eff} = \int_0^{R_t} \frac{M_l(r) \cdot w_p(r)}{1 + e^{-\gamma \cdot (r - t_h)}} dr \quad (4)$$

En esta ecuación:

- R_t es el recurso total de la red.
- $M_l(r)$ es la eficiencia optimizada del r -ésimo recurso utilizando ML.
- $w_p(r)$ es el peso de prioridad de cada recurso, que depende de las necesidades de los dispositivos.
- γ es un parámetro de sensibilidad que determina la pendiente de la asignación.

- t_h es un umbral que maximiza la eficiencia de la asignación.

Este modelo integral no solo postula un resultado, sino que describe matemáticamente cómo el ML logra una asignación de recursos inteligente y ponderada, incorporando conceptos clave como la prioridad y los umbrales de decisión. Estas formulaciones constituyen la columna vertebral teórica que permite que el marco propuesto se construya sobre una base rigurosa y verificable, conectando la visión conceptual con una implementación cuantificable.

2. Referencias Bibliográficas

Referencias

- [1] MAXIMIZING THE EFFICIENCY OF AI BASED IOT DEVICES WITH UPGRADED 802.11 AND 802.16 BASED NETWORKS TO OPTIMIZE NETWORK CONGESTIONS. *JATIT*, vol. 103, no. 6, 2025. [Online]. Disponible: <http://www.jatit.org/volumes/Vol103No6/22Vol103No6.pdf>
- [2] Artificial Intelligence in Networking Research in the Arab World. *Communications of the ACM*, 2025. [Online]. Disponible: <https://cacm.acm.org/arab-world-regional-special-section/artificial-intelligence-in-networking-research-in-the-arab-world/>
- [3] AN OVERVIEW ON 5G COMMUNICATION TECHNOLOGY. *IJSDR*, 2024. [Online]. Disponible: <https://ijsdr.org/papers/IJSDR2407046.pdf>
- [4] The Integration of the Internet of Things (IoT) Applications into 5G Networks: A Review and Analysis. *MDPI*, vol. 14, no. 7, 2025. [Online]. Disponible: <https://www.mdpi.com/2073-431X/14/7/250>
- [5] Optimizing Resource Allocation for 5G Internet-of-Things Networks Using Machine Learning Techniques, 2025.
- [6] 5G and IoT: Towards a new era of communications and measurements. *ResearchGate*, 2025. [Online]. Disponible: https://www.researchgate.net/publication/337668857_5G_and_IoT_Towards_a_new_era_of_commun
- [7] Efficient Resource Allocation in 5G Massive MIMO-NOMA Networks: Comparative Analysis of SINR-Aware Power Allocation and Spatial Correlation-Based Clustering. *arXiv*, 2025. [Online]. Disponible: <https://arxiv.org/html/2503.08466v1>
- [8] Resource Allocation Schemes for 5G Network: A Systematic Review. *MDPI Sensors*, vol. 21, no. 19, 2021. [Online]. Disponible: <https://www.mdpi.com/1424-8220/21/19/6588>
- [9] ML-Based 5G Network Slicing Security: A Comprehensive Survey. *Semantic Scholar*, 2025. [Online]. Disponible: <https://pdfs.semanticscholar.org/4624/5a2484c3e9e2425dee8df9458a7a40d3199b.pdf>
- [10] Dynamic routing optimization in software-defined networking based on a metaheuristic algorithm. *ResearchGate*, 2024. [Online]. Disponible: https://www.researchgate.net/publication/378180891_Dynamic_routing_optimization_in_software-defined_networking_based_on_a_metaheuristic_algorithm
- [11] A Hybrid Machine Learning Framework for Dynamic Resource Optimization in 5G Networks. *ResearchGate*, 2024. [Online]. Disponible: https://www.researchgate.net/publication/391740172_A_Hybrid_Machine_Learning_Framework_for
- [12] Edge Computing in 5G: A Review. *ResearchGate*, 2019. [Online]. Disponible: https://www.researchgate.net/publication/335510509_Edge_Computing_in_5G_A_Review

-
- [13] IMPACT OF IOT AND EDGE COMPUTING ON 5G NETWORK PERFORMANCE FOR DATA COMMUNICATION. *ResearchGate*, 2024. [Online]. Disponible: https://www.researchgate.net/publication/387005591_IMPACT_OF_IOT_AND_EDGE_COMPUTING_ON_5G_NETWORK_PERFORMANCE_FOR_DATA_COMMUNICATION
- [14] Fisher Market Model based Resource Allocation for 5G Network Slicing. *arXiv*, 2023. [Online]. Disponible: <https://arxiv.org/pdf/2307.16585>
- [15] Survey on Machine Learning in 5G. *International Journal of Engineering Research & Technology*, vol. 9, no. 1, 2020. [Online]. Disponible: <https://www.ijert.org/research/survey-on-machine-learning-in-5g-IJERTV9IS010326.pdf>
- [16] Machine Learning-Based Optimization for 5G Resource Allocation Using Classification and Regression Techniques. *ResearchGate*, 2024. [Online]. Disponible: https://www.researchgate.net/publication/391172513_Machine_Learning-Based_Optimization_for_5G_Resource_Allocation_Using_Classification_and_Regression_Techniques
- [17] Energy-Efficient Radio Resource Allocation in 5G Using Deep Q-Networks. *Universidad de Granada*, 2024. [Online]. Disponible: <https://digibug.ugr.es/bitstream/handle/10481/106424/Radio%20Resource%20Allocation%20in%205G%20Using%20Deep%20Q%20Networks.pdf>
- [18] A Framework for User Traffic Prediction and Resource Allocation in 5G Networks. *MDPI Applied Sciences*, vol. 15, no. 13, 2025. [Online]. Disponible: <https://www.mdpi.com/2076-3417/15/13/7603>