

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

**“DISEÑO DE UN ALGORITMO BASADO EN APRENDIZAJE REFORZADO PARA AJUSTAR
DINÁMICAMENTE EL TIEMPO DE BACKOFF EN UN ENTORNO MACHINE-TYPE
COMMUNICATION”**

PROYECTO INTEGRADOR

Previo la obtención del Título de:

MAGISTER EN TELECOMUNICACIONES

Presentado por:

PAÚL ENRIQUE GUAMÁN GUAMÁN, Ing.

GUAYAQUIL - ECUADOR

Año: 2023

DEDICATORIA

A mis amadas hijas, Dalia Camila y Paula Catalina quienes no soltaron mi mano en todo este proceso. A mi esposa Jenny (+), quien nunca dejo de confiar en mí.

AGRADECIMIENTOS

Mi más sincero agradecimiento a mis padres, tíos, hermanos y personas quienes me apoyaron e incentivaron en este camino.

A mi tutor de proyecto, PhD. Luis Tello, por su predisposición, guía y soporte durante el desarrollo de la investigación.

Compañeros y amigos de trabajo, quienes supieron entender el tiempo requerido y el apoyo que estuvo presente.

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, me corresponde conforme al reglamento de propiedad intelectual de la institución; Yo, *PAÚL ENRIQUE GUAMÁN GUAMÁN* doy mi consentimiento para que la ESPOL realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

Paúl E. Guamán

EVALUADORES

Luis Patricio Tello O., PhD.

PROFESOR TUTOR

María Antonieta Álvarez V., PhD.

PROFESOR EVALUADOR

RESUMEN

Los dispositivos que requieren un acceso a la red de comunicaciones más grande como el internet va creciendo de forma exponencial, por lo que se requiere disponer de medios los cuales garanticen el acceso a estos dispositivos. Las redes celulares son la mejor opción para garantizar esta conectividad. Los dispositivos de comunicación tipo máquina (Machine-Type Communication, MTC) generan elevadas solicitudes de conexión simultáneas, lo que implica un elevado número de colisiones en el canal de acceso aleatorio de la red celular. El proceso de *backoff* en el procedimiento de acceso aleatorio se utiliza para resolver las colisiones cuando varios usuarios han seleccionado los mismos recursos para el intento de acceso. Este proceso consiste en definir un tiempo de espera previo a una re-transmisión; la estación base determina este tiempo por medio del Indicador de *Backoff* (*BI*), lo difunde en la red y los dispositivos que han colisionado calculan un tiempo de retransmisión aleatoriamente en función del *BI* configurado por la red. En la especificación técnica, el valor de *BI* es fijo y, en ocasiones de elevada carga en el canal de acceso, esto provoca un elevado número de retransmisiones en las solicitudes de conexión que utilizan un mismo slot, así como una baja probabilidad de acceder a la red de comunicaciones. En esta investigación se evalúa el canal de acceso aleatorio y se diseña un algoritmo basado en aprendizaje por refuerzo para ajustar dinámicamente el valor de *BI*. Se utilizó el software MATLAB para desarrollar un modelo de simulación del procedimiento de acceso aleatorio basado en contención; este modelo contempla las especificaciones del organismo de estandarización *Third Generation Partnership Project* (3GPP) para comunicaciones celulares. Se utilizó Q-Learning para ajustar dinámicamente el valor del *BI* y se evaluaron las métricas de rendimiento de red consideradas de mayor relevancia: probabilidad de acceso satisfactorio, número de retransmisiones y retardo en el acceso. Los resultados demuestran que el algoritmo diseñado incrementa la probabilidad de acceso satisfactorio manteniendo un valor superior a 75% incluso en escenarios de elevada carga con un ligero impacto en el número de retransmisiones, pero incrementando el retardo en el acceso de los dispositivos debido a la dispersión del tiempo de retransmisión de las peticiones fallidas. La solución es óptima para aplicaciones tolerantes al retardo en las que la calidad de servicio se mide en base a la probabilidad de acceso satisfactorio de los dispositivos a la red de comunicación.

Palabras Clave: backoff, MTC, RACH, Q_Learning, redes celulares.

ABSTRACT

Devices that require access to the largest communications network, such as the Internet, are growing exponentially, so it is necessary to have means that guarantee access to these devices. Cellular networks are the best option to guarantee this connectivity. Machine-type Communication (MTC) devices generate high simultaneous connection requests, which implies a high number of collisions on the random access channel of the cellular network. The backoff process in the random access procedure is used to resolve collisions when multiple users have selected the same resources for the access attempt. This process consists of defining a waiting time prior to a re-transmission; the base station determines this time through the Backoff Indicator (BI), broadcasts it on the network, and the colliding devices calculate a retransmission time randomly based on the BI configured by the network. In the technical specification, the BI value is fixed, and sometimes, a high load on the access channel causes a high number of retransmissions in connection requests that use the same slot as well as a low probability of accessing the communication network. In this research, the random access channel is evaluated, and an algorithm based on reinforcement learning is designed to adjust the BI value in a dynamic manner. MATLAB software was used to develop a simulation model of the contention-based random access procedure; This model includes the specifications of the Third Generation Partnership Project (3GPP) standardization body for cellular communications. Q-Learning was used to adjust the BI value dynamically, and the network performance metrics considered most relevant were evaluated: Probability of successful access, number of retransmissions, and access delay. The results show that the designed algorithm increases the probability of successful access by maintaining a value greater than 75% even in high load scenarios, with a slight impact on the number of retransmissions but increasing the delay in device access because of dispersion of the retransmission time of the failed requests. The solution is optimal for delay-tolerant applications in which the quality of service is measured based on the probability of successful access of the devices to the communication network.

Keywords: *backoff, MTC, RACH, Q_Learning, cellular networks.*

ÍNDICE GENERAL

RESUMEN	I
ABSTRACT.....	II
ÍNDICE GENERAL.....	III
ÍNDICE DE GRÁFICOS.....	V
ÍNDICE DE TABLAS.....	VII
ABREVIATURAS.....	IX
CAPÍTULO 1.....	1
1. INTRODUCCIÓN	1
1.1 Descripción del Problema.....	2
1.2 Justificación	3
1.3 Objetivos.....	4
1.4 Propuesta	4
1.5 Alcance del trabajo.....	5
CAPÍTULO 2.....	7
2. MARCO TEÓRICO	7
2.1 Actualidad de las Redes de Nueva Generación.	7
2.2 Comunicación masiva tipo máquina.....	13

2.3	Canal de acceso aleatorio y procedimiento de retransmisión	16
2.4	Sistemas dinámicos basados en aprendizaje reforzado	24
<i>CAPÍTULO 3.....</i>		29
3. DISEÑO DE UN ALGORITMO BASADO EN Q-LEARNING PARA AJUSTAR DINÁMICAMENTE EL TIEMPO DE BACKOFF		29
3.1.	Modelo del sistema.....	29
3.2.	Definición y desarrollo del algoritmo Q-Learning	29
3.3.	Métricas de rendimiento a evaluar	34
<i>CAPÍTULO 4.....</i>		39
4. EVALUACIÓN Y DISCUSIÓN DE RESULTADOS		39
4.1	Escenarios de tráfico MTC y configuración de parámetros para simulación.....	39
4.2	Resultados y análisis de las métricas de rendimiento	39
4.3	Comparación y discusión de resultados	46
<i>CAPÍTULO 5.....</i>		49
5. CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS		49
5.1	Conclusiones.....	49
5.2	Recomendaciones	50
5.3	Trabajos Futuros	51
<i>BIBLIOGRAFÍA.....</i>		52
<i>ANEXOS.....</i>		<i>a</i>

ÍNDICE DE GRÁFICOS

CAPÍTULO 1

Figura 1. 1. Diagrama de flujo del algoritmo de simulación.....	5
---	---

CAPÍTULO 2

Figura 2. 1 Componentes de la Arquitectura de una Red NGN. (Castillo Medina, Salcedo Parra, & Forero Rodriguez, 2014)	8
---	---

Figura 2. 2 Evolución de las tecnologías de Red desde 1G hasta 3G (CHEN & ZHANG, 2004).	10
---	----

Figura 2.3 Dispositivos conectados a Nivel Mundial de 2019 a 2030. (Fernandez, 2021).	14
---	----

Figura 2. 4 Integración de Aplicaciones basados en mMTC en una celda celular. (Santander González & Tello Oquendo, 2021)	16
--	----

Figura 2. 5 Procedimiento de Acceso Aleatorio basado en Contención (Tello Oquendo, et al., 2018).....	18
---	----

Figura 2. 6 Procedimiento de retransmisión realizado por un UE. (Tello Oquendo, et al., 2018)	21
---	----

Figura 2. 7 Distribución de tiempo de retransmisión dado diferentes valores de BI para 100 UE's.	22
---	----

Figura 2. 8 Distribución de tiempo de retransmisión dado diferentes valores de BI para 30000 UE's.	23
---	----

Figura 2. 9. Modelo de Aprendizaje Supervisado. (Ruiz Manosalva, 2019).....	25
---	----

Figura 2.10 Modelo de Aprendizaje No Supervisado. (Ruiz Manosalva, 2019).....	26
Figura 2. 11 Modelo de Aprendizaje por Refuerzo (García Pascual, 2021).....	27
CAPÍTULO 3	
Figura 3. 1. Modelo del sistema	29
CAPÍTULO 4	
Figura 4. 1 Comportamiento del algoritmo frente a varias cantidades de usuarios.	40
Figura 4. 2 Comportamiento de la Probabilidad de Acceso Satisfactorio; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.....	41
Figura 4. 3 Comportamiento del número de retransmisiones; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.	43
Figura 4. 4 Comportamiento del Retardo al acceso; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.	45

ÍNDICE DE TABLAS

CAPÍTULO 2

Tabla 2. 1 Características de 5G (Medeiros, Pereira, & Castilho, 2017) (Luis Tello-Oquendo, 2020). 13

Tabla 2. 2 Características de comunicación para M2M y H2H, (Alvi, Abualnaja, Tariq Toor, & Saadi, 2021)..... 15

CAPÍTULO 3

Tabla 3. 1 Valores que puede tomar el Indicador de Retransmisión según 3GPP (LTE Quick Reference , n.d.)..... 31

Tabla 3. 2 Valores de las propiedades de algoritmo de RL..... 32

Tabla 3. 3 .- Parámetros de configuración del RACH 34

Tabla 3. 4 Tabla de la Probabilidad de Acceso Satisfactorio al variar el indicador de Retransmisión BI 35

CAPÍTULO 4

Tabla 4. 1 Condiciones de Escenario de Trafico MTC y HTC..... 39

Tabla 4. 2 Condiciones para el desarrollo del Algoritmo Q-Learning 39

Tabla 4. 3 Probabilidad de Acceso Satisfactorio en un ambiente con BI fijo. 42

Tabla 4. 4 Probabilidad de Acceso Satisfactorio en un ambiente de BI dinámico. 42

Tabla 4. 5 Número de retransmisiones en un ambiente con BI fijo..... 44

Tabla 4. 6 Número de retransmisiones en un ambiente de BI dinámico..... 44

Tabla 4. 7 Retardo en el acceso en un ambiente con BI fijo 46

Tabla 4. 8 Retardo en el acceso en un ambiente de BI dinámico. 46

ABREVIATURAS

Abreviación	Descripción
1G	First Generation – Primera Generación de telefonía móvil
2G	Second Generation – Segunda Generación de telefonía móvil
2.5 G	2.5 Generation – Generación de Transición
3G	Third Generation – Tercera Generación de telefonía móvil
3GPP	3rd Generation Partnership Project - Proyecto de Colaboración de Tercera Generación
4G	Fouth Generation – Cuarta Generación de telefonía móvil
5G	Fifth Generation, Quinta Generación de telefonía móvil
ACB	Access Class Barring - Restricción de Clases de Acceso
BI	Backoff Indicator – Indicador de Backoff
BS	Base Station – Estación Base
C-RNTI	Cell Radio Network Temporary Identifier - Identificador Temporal de Red de Radio Celular
CDF	Cumulative Distribution Function - Función de Distribución Acumulada
eNB	Evolved Node B – Interconexión entre Móvil y LTE
H2H	Human to Human, Comunicación Humano a Humano
HARQ	Hybrid Automatic Retransmission Request – Solicitud de Retransmisión Automática Híbrida
HTC	Human Type Communication - Comunicaciones Tipo Humano
IoT	Internet of Things – Internet de las Cosas
IP	Internet Protocol – Protocolo de Internet

LTE-A	Long Term Evolution – Advanced (4G)
M2M	Machine to Machine – Máquina a Máquina
Mbps	Megabits por segundo
MIB	Master Information Block – Bloque de Información Maestro
MTC	Machine Type Communication – Comunicación Tipo Máquina
mMTC	Massive MTC
NGN	Next Generation Network – Red de nueva Generación
OFDMA	Orthogonal Frequency Division Multiple Access - Acceso Múltiple por División de Frecuencia Ortogonal
PDCCH	Physical Downlink Control Channel - Control de Canal Físico de Enlace Descendente
RACH	Random-Access Channel – Canal de Acceso Aleatorio
P_s	Probability of Successful access – Probabilidad de Acceso Satisfactorio
QoS	Quality of Service – Calidad de Servicio
RA	Random Access – Acceso Aleatorio
RA-RNTI	Random Access Radio Network Temporary Identifier - Identificador Temporal de Red de Radio de Acceso Aleatorio
PRACH	Physical RACH
RAO	Random Access Opportunity – Oportunidad de Acceso Aleatorio
RAP	Random Access Procedure - Procedimiento de Acceso Aleatorio
RAR	Random Access Response – Respuesta de Acceso Aleatorio
RL	Reinforcement Learning - Aprendizaje por Refuerzo

SC-FDMA	Single Carrier Frequency Division Multiple Access - Acceso Múltiple por División de Frecuencia de Portadora Única
SIB	System Information Blocks – Bloques de Información del Sistema
SMS	Short Message Service – Servicio de Mensajes de Texto
TBO	Time Backoff – Tiempo de Retransmisión
UE	User Equipment – Dispositivo del Usuario
UMTS	Universal Mobile Telecommunications System - Sistema Universal de Telecomunicaciones Móviles.
VOD	Video on Demand, Video Bajo Demanda
VoIP	Voice Over Internet Protocol – Voz Sobre el Protocolo de Internet

CAPÍTULO 1

1. INTRODUCCIÓN

La evolución tecnológica ha permitido que el término *Internet of Things* (IoT) sea más estruendoso en los últimos tiempos. Los dispositivos que lo conforman requieren conexión a Internet y qué mejor hacerlo a través de una infraestructura desplegada como es el caso de las redes celulares. Sin embargo, hay que considerar la saturación o colisiones que se pueden dar en el Canal de Acceso Aleatorio (Random Accesos Channel, RACH) cuando existen múltiples solicitudes de acceso por parte de los dispositivos al intentar acceder a la red celular, provocando un sobre consumo energético por las retransmisiones luego de una colisión.

Para que un dispositivo pueda hacer una nueva solicitud de acceso luego de haber sufrido una colisión o pérdida de la misma, es necesario que tenga que esperar un tiempo de backoff (Time Backoff, T_{BO}); este tiempo de backoff va a depender del valor del indicador de backoff (Backoff Indicator, BI), el mismo que por estandarización del Proyecto de Colaboración de Tercera Generación (3rd Generation Partnership Project, 3GPP) puede oscilar entre 0 a 960 mseg; sin embargo, actualmente en las redes celulares el valor de BI es fijo cuyo valor establecido es de 20 mseg.

Con el fin de disminuir el número de retransmisiones que se pueden generar por colisiones en el RACH y aumentar el valor de la Probabilidad de Acceso Satisfactorio (Probability of Successfull access, P_s), se plantea realizar un ajuste dinámico del valor del BI en función de la carga del RACH. Para ello, se pretende utilizar un algoritmo de aprendizaje reforzado como es el caso de Q-Learning.

1.1 Descripción del Problema

Para el diseño de una red de comunicaciones confiable, se debe de considerar características tales como: disponibilidad, seguridad y escalabilidad. En los últimos años, la tecnología ha venido desarrollándose en proporciones gigantescas, lo que conlleva a que cada vez sean más los dispositivos que requiere conectividad y tratan de acceder a una red de comunicaciones (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018), comprometiendo en cierta manera el acceso y la disponibilidad de esta red.

Debido a los miles de dispositivos intentando formar parte de una red de comunicaciones, se pretende dar cabida a estos haciendo uso de infraestructuras ya desplegadas, como es el caso de la red celular. Esta red dispone de una amplia cobertura y en su mayoría se encuentra desplegada con tecnología Long Term Evolution – Advanced (LTE-A) (HUDA , MOHAMED, NOR, & ZURINA, 2020) que en su momento fue diseñada para la comunicación Humano a Humano (Human to Human, H2H).

Cuando un dispositivo requiere acceder a la red celular a través de una estación base (base station, BS), entre ellos empieza un intercambio de mensajes los cuales transitan por el RACH; sin embargo, puede darse el caso de que varios mensajes enviados por los dispositivos hagan uso de los mismos recursos, ocasionando una colisión de mensajes los cuales no serán escuchados por la BS. Si el intercambio de mensajes ha sido satisfactorio, la BS envía una respuesta con los recursos físicos para establecer el enlace, caso contrario, el dispositivo debe esperar un tiempo previo a iniciar nuevamente el intercambio de mensajes. El tiempo de espera antes de una retransmisión se calcula en función del valor de BI, que es comunicado por la BS a los dispositivos. Los dispositivos cuyos mensajes fueron erróneos toman el valor de BI para establecer un tiempo de retransmisión (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018) (Santander González & Tello Oquendo, 2021).

La BS fija un valor constante de BI, el cual no se adapta a las condiciones del tráfico del sistema. Esta es la motivación para este estudio y se propone diseñar un mecanismo para ajustar dinámicamente el valor de BI basado en Q-Learning, que es un algoritmo de aprendizaje reforzado. Con este mecanismo, el tiempo de retransmisión se adaptará a las condiciones de tráfico de la red.

1.2 Justificación

Las redes celulares han considerado un valor de BI fijo, de acuerdo con lo establecido por la 3GPP, de 20 mseg; cabe indicar que este valor puede oscilar entre 0 a 960 mseg (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018). Cuando el valor de BI aumenta, las peticiones se van a dispersar más en el tiempo y esto favorecerá a la disminución de colisiones en el RACH, pero tendrá un impacto en el retardo experimentado por la comunicación.

Con el afán de dar soluciones a esta problemática, los autores (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018) plantean aplicar mecanismos de control de acceso con el fin de no sobrecargar la red haciendo uso de restricciones de clase de acceso, sin embargo, al tratarse de una red híbrida el impacto de la Ps no es significativo en el tráfico H2H debido a que el retardo de conexión aumenta.

El planteamiento realizado por (Tello Oquendo, et al., 2018) muestra un análisis al tomar un valor aleatorio del indicador de backoff en conjunto la restricción de acceso extendido, mostrando resultados de mejora en la probabilidad de acceso satisfactorio Ps, pero al combinar más esquemas de control compromete el retardo en la conexión.

Es importante evaluar las condiciones del tráfico y que la red actúe en función de estas condiciones. Considerando el caso particular del BI, el hecho de que no se ajuste dinámicamente este valor, ocasiona un incremento en el número de retransmisiones de las solicitudes de conexión que emiten los dispositivos al

intentar acceder a la red debido a las colisiones, repercutiendo en el consumo energético de estos (HUDA , MOHAMED, NOR, & ZURINA, 2020) (Yu & Seet, 2019). Además, en las comunicaciones tipo máquina se considera el análisis de la Ps, un ajuste adecuado del valor de BI permitirá incrementar esta probabilidad.

1.3Objetivos

1.3.1 Objetivo General

- Diseñar un algoritmo basado en aprendizaje reforzado para el ajuste del tiempo de Backoff de forma dinámica en un entorno de comunicación tipo máquina (Machine Type Communication, MTC).

1.3.2 Objetivos Específicos

- Analizar el comportamiento del tiempo de Backoff (T_{Bo}) en condiciones de tráfico MTC para la caracterización de su funcionamiento.
- Diseñar un mecanismo de Q-Learning para el ajuste dinámico del T_{Bo} .
- Analizar los resultados con diversas cargas de tráfico para la confirmación del ajuste dinámico del valor de BI conforme la carga en el RACH.

1.4Propuesta

Frente a la problemática de posibles retransmisiones y colisiones que se pueden dar en el RACH en un ambiente de comunicaciones masivas tipo máquina (Massive Machine-Type Communications, mMTC), se propone desarrollar un modelo de simulación del canal de acceso aleatorio e implementar un algoritmo basado en Q-Learning para que la red aprenda a ajustar dinámicamente el tiempo de backoff en función de la carga de tráfico de la red de comunicaciones para gestionar las colisiones.

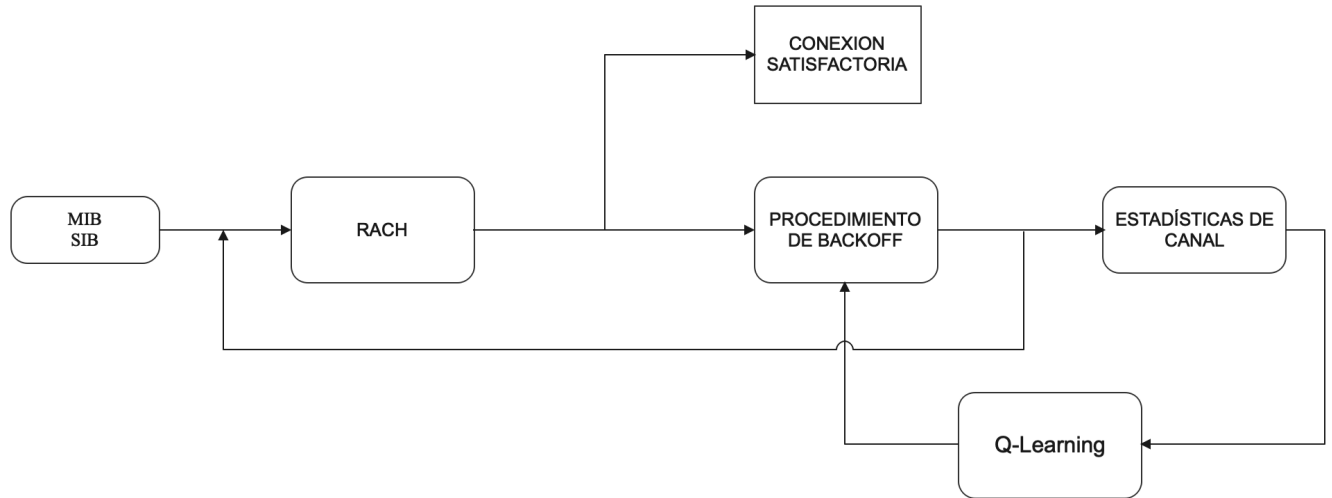


Figura 1. 1. Diagrama de flujo del algoritmo de simulación.

La Figura 1.1, indica un diagrama del algoritmo propuesto, la que comprende diferentes fases, que van desde el modelo hasta el aprendizaje del canal haciendo uso de Q-Learning; con el desarrollo del algoritmo se pretende ajustar el valor del BI considerando las condiciones de tráfico de la red.

Se define una función que permita determinar el estado del sistema, y en base a este configurar el valor de BI de modo que, los dispositivos que requieran el acceso a la red celular luego de un intento fallido lo hagan después de un tiempo adecuado. Esto evitará las colisiones que se pueden generar y como consecuencia se disminuirá las retransmisiones de los dispositivos MTC, aumentará la P_s y se optimiza el consumo energético de los dispositivos móviles.

1.5 Alcance del trabajo

Con el desarrollo de un algoritmo basado en aprendizaje reforzado, se pretende incrementar el valor de la P_s ; el análisis de este incremento está en base del proceso de retransmisión de una solicitud de acceso colisionada en el RACH.

El diseño del RACH se basa en las especificaciones del 3GPP, siguiendo el protocolo de acceso satisfactorio y la señalización correspondiente de una BS por

medio del bloque de información maestro (Master Information Block, MIB) y el bloque de información del sistema (System Information Blocks, SIB).

Con el uso del software MATLAB, se desarrollará un algoritmo dinámico que evaluará el canal y a la vez determinará un valor adecuado de BI, el mismo que permitirá estimar un tiempo óptimo de retransmisión de solicitud de acceso, incrementando la P_s y evitando una nueva colisión en el mismo dispositivo.

Es importante validar el funcionamiento correcto del algoritmo aplicándolo a diferentes escenarios, considerando varias cargas en el canal de acceso aleatorio.

CAPÍTULO 2

2. MARCO TEÓRICO

2.1 Actualidad de las Redes de Nueva Generación.

Las redes *legacy*, redes implementadas que brindan un servicio a la vez como: audio, video o datos, se han tornado obsoletas debido al desarrollo de aplicaciones y a la migración de la infraestructura celular a un sistema de Protocolo de Internet (Internet Protocol, IP), logrando unificar servicios y tener elevadas tasas de transmisión sin comprometer la Calidad de Servicio (Quality of Service, QoS) (Castillo Medina, Salcedo Parra, & Forero Rodriguez, 2014) (Rojas Mendez & Urgiles Peralta, 2011).

Por otro, las redes inalámbricas no se han quedado rezagadas, ya que ha migrado su infraestructura de radio a una infraestructura IP, la cual permite mejoras en su sistema, tales como: control en tiempo real de servicios de voz y multimedia, mejor ruteo al hacer uso de paquetes, calidad de servicio (CHEN & ZHANG, 2004). Esta integración de varios servicios a un esquema netamente IP es conocido como Red de Nueva Generación (Next Generation Network, NGN).

2.1.1 Arquitectura de una Red de Nueva Generación

La arquitectura de una NGN, mostrada en la Figura 2.1, indica la estructura de cuatro capas principales, como son:

- **Servicios de red:** La capa de servicios de red es la encargada de proporcionar los servicios y aplicaciones requeridos por los usuarios.
- **Núcleo de Red:** En la capa de Núcleo de la Red (Core Network) es donde se encuentra representado principalmente por el Backbone, aquí la red se encarga de la transferencia de paquetes, así como de la función de gestión y control.

- **Red de Acceso:** Permite la integración de un usuario hacia la red NGN, que está conformada por diferentes tecnologías de Última Milla.
- **Equipo de Usuario:** En esta capa se encuentra todo el equipo final que requiere el acceso a la red, el cual puede hacer uso de diferentes tecnologías o protocolos (Castillo Medina, Salcedo Parra, & Forero Rodríguez, 2014).

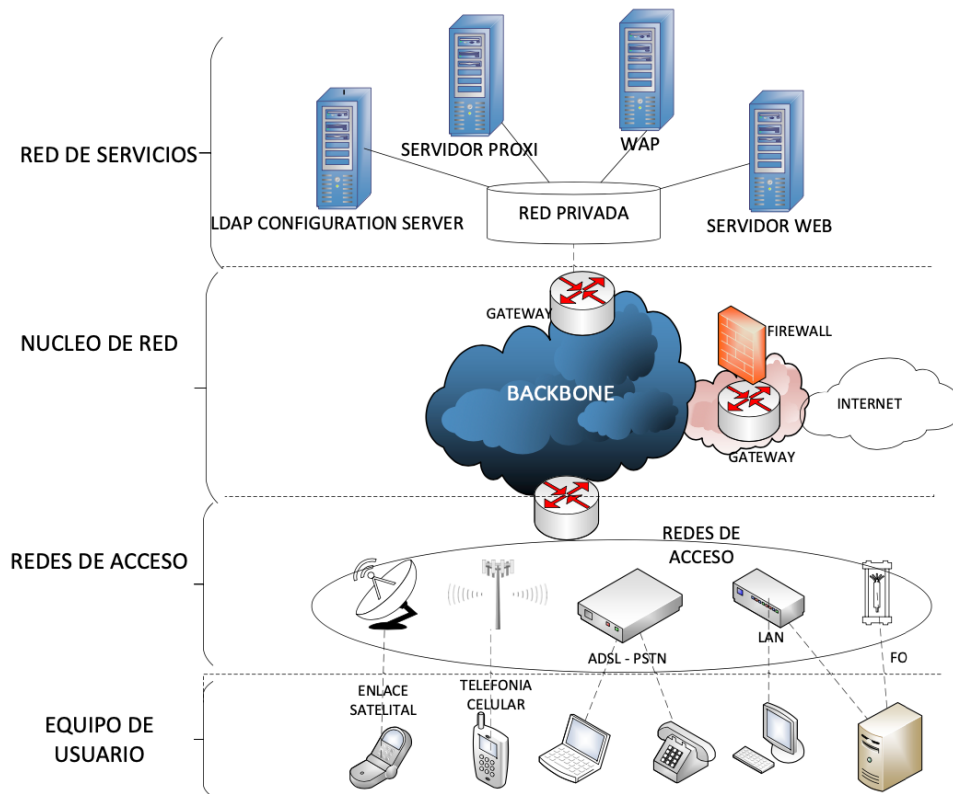


Figura 2. 1 Componentes de la Arquitectura de una Red NGN. (Castillo Medina, Salcedo Parra, & Forero Rodríguez, 2014)

2.1.2 Redes móviles y su evolución.

La telefonía móvil ha bautizado su evolución en términos de la palabra generación, dicha evolución se encuentra plasmada en la Figura 2.2. En su momento, la Primera Generación (First Generation, 1G) fue una red implementada en función de la conmutación de circuitos, cuyo único

servicio ofrecido fue el de voz, esta red de telefonía tuvo su lanzamiento a inicio de los años 80. Así mismo, no era posible realizar roaming entre diferentes operadores de forma automática, por lo que se tenía que llamar y realizar este requerimiento de forma manual; en esta generación de telefonía móvil no fue posible realizar roaming entre operadores de diferentes países (CHEN & ZHANG, 2004).

A inicio de los 90, se lanza la Segunda Generación (Second Generation, 2G) de telefonía móvil, se reemplaza el procesamiento analógico de señales que se realizó en 1G por el procesamiento digital de señales. La utilización de una plataforma que haga uso de tecnología digital trajo consigo el bajo consumo de potencia en los terminales móviles, así como la mejora en la calidad de voz; en esta era de la segunda generación, se agregó un servicio adicional como es el Servicio de Mensajes de Texto (Short Message Service, SMS), se suma la posibilidad de realizar roaming entre diferentes operadoras de red y entre países. Antes de lograr la red de Tercera Generación (Third Generation, 3G), existe la generación 2.5 G, infraestructura que soporta el servicio de datos móviles (CHEN & ZHANG, 2004).

Con el lanzamiento de la 3G a finales del 90, se tiene un incremento de hasta 2 Mbps en la tasa de datos, la infraestructura que soporta esta generación de telefonía móvil es basada en IP, lo que permite agregar servicios como: Internet móvil, servicios de transmisión de video por streaming y transmisión de video en un tiempo no-real, cabe indicar que el QoS no se ve comprometido sino que al contrario, la mejora del QoS es la que permite el poder integrar más servicios a la telefonía móvil (CHEN & ZHANG, 2004).

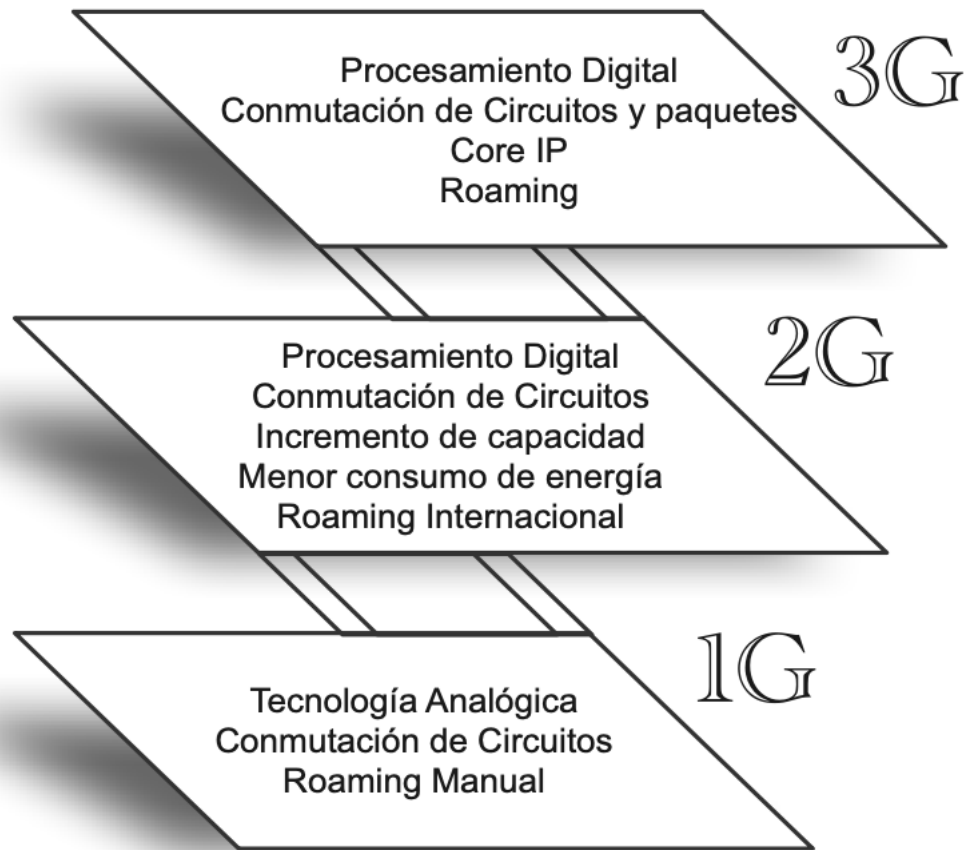


Figura 2. 2 Evolución de las tecnologías de Red desde 1G hasta 3G (CHEN & ZHANG, 2004).

2.1.3 Redes NGN 4G

La exigencia del acceso a la red celular por parte de los usuarios móviles cada vez es mayor, por lo que los operadores de red buscan el integrar nuevos servicios sin comprometer parámetros de alta relevancia como la QoS y el retardo.

A inicios del año 2000 empieza el despliegue de redes de cuarta generación (Fourth Generation, 4G), misma que permite la convivencia con las tecnologías anteriores como son 2G y 3G. Cabe indicar que además de los

servicios ofrecido en la red 3G, a esta red 4G se suman servicios como: Video Bajo Demanda (Video On Demand, VOD), aplicaciones multimedia para video conferencia, Voz sobre IP (Voice Over Internet Protocol, VoIP) y Streaming de datos (Guzman Quintero, 2012).

La red 4G adopta como núcleo de red a la tecnología LTE-Advanced la cual tienen su estructura basada completamente sobre IP, permitiendo tasas de transmisión de hasta 100 Mbps en movimiento y 1 Gbps cuando se está en estado estacionario (Utrilla Salazar, 2014) (Ratana Bhalla & Vardhan Bhalla, August 2010).

La arquitectura del Sistema Universal de Telecomunicaciones Móviles (Universal Mobile Telecommunications System, UMTS) usada en 3G fue desplazada por LTE para dar cabida a los sistemas inalámbricos 4G, basando su dominio sobre un esquema de conmutación de paquetes. Con el fin de lograr una mejor eficiencia del espectro radioeléctrico, se utilizan dos técnicas de acceso múltiple al medio de comunicación, asignando ya sea un mayor o menor número de portadoras en función de la velocidad que requiera y a la vez mitigando las interferencias co-canal que pudieran haberse generado al no usar frecuencias ortogonales. El Acceso Múltiple por División de Frecuencia Ortogonal (Orthogonal Frequency Division Multiple Access, OFDMA) se utiliza en el enlace descendente, mientras que, el Acceso Múltiple por División de Frecuencia de Portadora Única (Single Carrier Frequency Division Multiple Access, SC-FDMA) se utiliza para el enlace ascendente (Utrilla Salazar, 2014).

2.1.4 Redes NGN 5G

El sucesor de la red 4G, se encuentra desplegándose a nivel mundial a partir del 2020 y es la denominada Quinta Generación de telefonía móvil (Fifth Generation, 5G); su sistema tiene un enfoque principal hacia el

usuario más que al operador, implicando mejoras en los servicios ofertados como: Navegación por internet, contenido por *streaming*, altas tasa de transmisión, entre otras (Haard , Patel, Joshi, & Modi, 2014).

Las actividades cotidianas del ser humano, cada vez se van tornando automáticas, en tal punto que es posible poder controlar una máquina desde un teléfono móvil, aunque el operador se encuentre a miles de kilómetros. Entre las aplicaciones más ambiciosas con el despliegue de una red 5G se encuentra el ámbito de salud; se estima pronto poder realizar operaciones remotas, gracias a la confiabilidad y seguridad que tiene esta generación de telefonía móvil (Xia, 2021).

Las tecnologías de telefonía móvil que van desde 1G hasta 4G, se encuentra haciendo uso del espectro en las bandas de 300 MHz a 3 GHz, sin embargo, la 5G de telefonía celular empieza a explotar un espectro superior, como es el caso de 3 GHz hasta los 300 GHz, lo que lo hace propenso a pérdidas de propagación (Tibúrcio Pereira, Sousa Ramos, & Lopes Ramos, 2020).

La 3GPP ha manifestado que la tecnología usada en 5G es LTE Advanced Pro, que aparte de que presenta mejoras en el rendimiento de la red, también tiene como objetivo el de bajar el costo por bit, permitiendo un mayor despliegue de servicios móviles (Medeiros, Pereira, & Castilho, 2017). La Tabla 2.1 presenta las principales características del sistema 5G (Luis Tello-Oquendo, 2020).

Con la implementación de una infraestructura de red 5G, se prevé una densidad de al menos 1 millón de conexiones por kilómetro cuadrado y 100 billones de conexiones en total (Cao, Ma, Li, Fu, & Lio, 2018).

Tabla 2. 1 Características de 5G (Medeiros, Pereira, & Castilho, 2017) (Luis Tello-Oquendo, 2020).

<i>Parámetros del Sistema 5G</i>	
Tráfico de datos	50 exabytes / mes
Máxima tasa de bits (DL)	20 Gbps
Máxima tasa de bits (UL)	10 Gbps
Latencia Máxima	1 ms
Densidad de conexiones	1 millón conexiones / Km ²

2.2 Comunicación masiva tipo máquina

El tiempo diario de uso de dispositivos móviles es cada vez más extenso, se estima que en la actualidad cada persona disponga de 1.5 dispositivos móviles, lo que genera un tráfico mensual de aproximadamente 49 exabytes (Yu & Seet, 2019).

El término IoT, cuyo nombre fue llamado por Kevin Ashton en 1999 (Otero Rojas, 2018), se refiere a la interconexión de dispositivos y/o actuadores con la Internet, haciendo uso de medios como Wifi, Bluetooth, ZigBee entre otros protocolos de comunicación; con el fin de que estos dispositivos se puedan controlar desde cualquier lugar, recolectar o intercambiar datos (Gokhale , Bhat, & Bhat, 2018).

La Figura 2.3 indica una tendencia creciente del número de dispositivos IoT interconectados a Internet, tal es el caso que se estima que en el año 2030 estén conectados más de 25.400 millones de dispositivos (Fernandez, 2021).

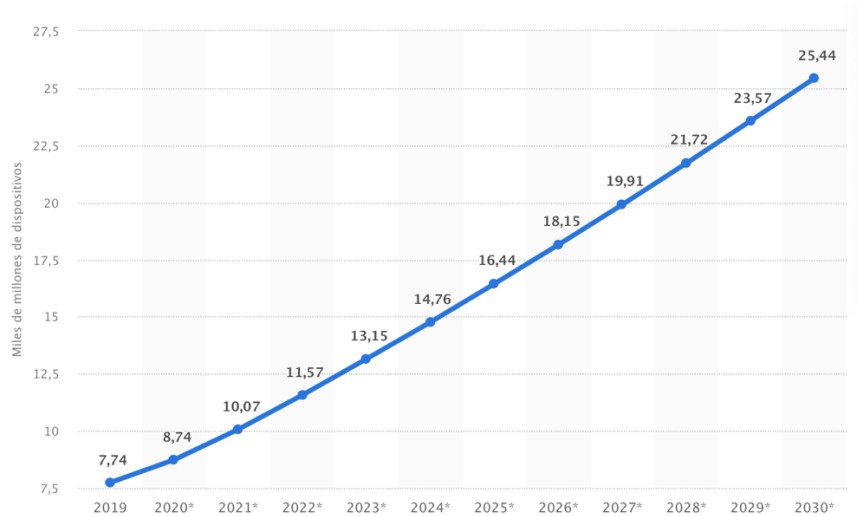


Figura 2.3 Dispositivos conectados a Nivel Mundial de 2019 a 2030. (Fernandez, 2021)

El despliegue ubicuo y la evolución que ha tenido las redes celulares conllevó a disponer de servicios que ofrecen plataformas como LTE-A y 5G, llegando a ser la mejor opción para una conexión concurrente de IoT o mMTC (Tello Oquendo, et al., 2018), en vista de que las redes celulares actuales disponen de características como: escalabilidad, baja latencia, amplia capacidad, garantía de QoS y uso eficiente del espectro (HUDA , MOHAMED, NOR, & ZURINA, 2020).

La tecnología celular en su momento fue desarrollada para un ambiente Humano a Humano (Human-to-Human, H2H); las nuevas infraestructuras de telefonía móvil permiten convivencia de dos tipos de comunicaciones como son H2H y MTC. Las principales características de este tipo de comunicaciones se detallan en la Tabla 2.2 (Alvi, Abualnaja, Tariq Toor, & Saadi, 2021).

La conexión de miles de dispositivos IoT o MTC a una infraestructura de red celular, lleva el nombre de Comunicación Masiva de Tipo Máquina (mMTC), para esto se hace uso de sistemas de acceso de radio heterogéneos, (Cao, Ma, Li, Fu, & Lio, 2018).

Tabla 2. 2 Características de comunicación para M2M y H2H, (Alvi, Abualnaja, Tariq Toor, & Saadi, 2021)

CARACTERÍSTICAS	MÁQUINA A MÁQUINA (M2M) / COMUNICACIÓN TIPO MÁQUINA (MTC)	HUMANO A HUMANO (H2H)
NÚMERO DE DISPOSITIVOS	Cientos de dispositivos por BS	No más de cientos de dispositivos por BS
TRANSMISIÓN DE PAQUETES	Mensajes que contiene pocos bits.	Transmisión en tiempo real ocupa paquetes de gran tamaño
MOVILIDAD	La movilidad no es preocupante	Para estudios se consideran a los humanos estáticos
DIRECCIÓN DE TRÁFICO	Por lo general Ascendente	Por lo general Descendente
PÉRDIDA DE PAQUETES	Alta	Baja
USO DE BATERÍA	Batería dura años y no es reemplazable	Batería recargable.

Este tipo de comunicaciones están pensadas para un enlace confiable sin comprometer la red, teniendo atributos principales como: cobertura, rentabilidad, bajo consumo de energía y disponibilidad a largo plazo (Alvi, Abualnaja, Tariq Toor, & Saadi, 2021) (Rohde & Schwars, 2022). Dispositivos que utilizan entornos mMTC como: sensores, equipos informáticos o actuadores (ver Figura 2.4) enfocan su uso en actividades como: medicina, agricultura, automatización industrial, comercio, ciudades inteligentes.

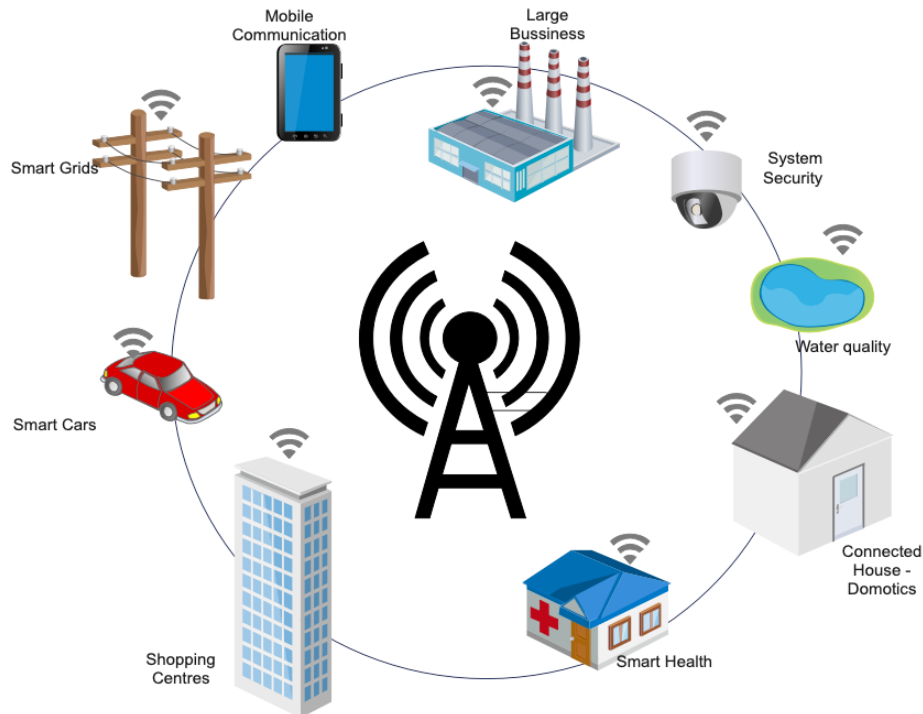


Figura 2. 4 Integración de Aplicaciones basados en mMTC en una celda celular. (Santander González & Tello Oquendo, 2021)

2.3 Canal de acceso aleatorio y procedimiento de retransmisión

Un entorno H2H que tiene convivencia con un entorno MTC y hacer uso de una infraestructura existente como la red celular, le permite tener ventajas como: uso de una infraestructura ubicua existente, ahorro en los costos de despliegue, seguridad, entre otros. (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018). Si embargo, esta convivencia requerirá un acceso simultaneo de miles de dispositivos, que forma un entorno mMTC, lo cual probablemente puede generar problemas de congestión y colisión de paquetes en el RACH (HUDA , MOHAMED, NOR, & ZURINA, 2020); cuando se da una colisión de paquetes de solicitud de conexión a la red celular, los dispositivos empiezan un nuevo proceso de solicitudes de acceso lo que se conoce como un procedimiento de retransmisión.

2.3.1 Canal de acceso aleatorio

En las redes NGN, el acceso por defecto es aleatorio y basado en contención (HUDA , MOHAMED, NOR, & ZURINA, 2020). Un RACH está conformado por slots de acceso aleatorio (Random Access Opportunity, RAO), que son recursos físicos de frecuencia-tiempo asignados para la transmisión de solicitudes de acceso (Laya , Alonso, & Alonso Zarate). Antes de establecer una comunicación con una estación base LTE (envolved NodeB, eNB), el dispositivo (User Equipment, UE) debe configurar unos parámetros que son enviados desde el eNB por broadcast a través de *Master Information Block* (MIB) y *System Information Blocks* (SIBs). Con la información obtenida del eNB (ver Figura 2.5) empieza la fase de negociación de acceso a través del RACH, en un proceso de cuatro mensajes (Tello Oquendo, et al., 2018) (Santander González & Tello Oquendo, 2021).

- Mensaje 1 – Mensaje de Preámbulo: Cuando un dispositivo requiere el acceso a un canal, selecciona un slot de Acceso Aleatorio (Random Access, RA), que no es más que un preámbulo con información del dispositivo, para realizar una solicitud y lo envía al eNB por medio del *Physical RACH* (PRACH). Existen 64 preámbulos ortogonales que se pueden utilizar durante el *Random Access Procedure* (RAP); de estos, son 54 los que están reservados para un RAP basado en contienda, en el caso de que 2 o más dispositivos seleccionen el mismo preámbulo en el mismo Slot del RA, se produce una colisión ya que el eNB no podrá decodificar el preámbulo (Tello Oquendo, et al., 2018) (Laya , Alonso, & Alonso Zarate) (Santander González & Tello Oquendo, 2021).

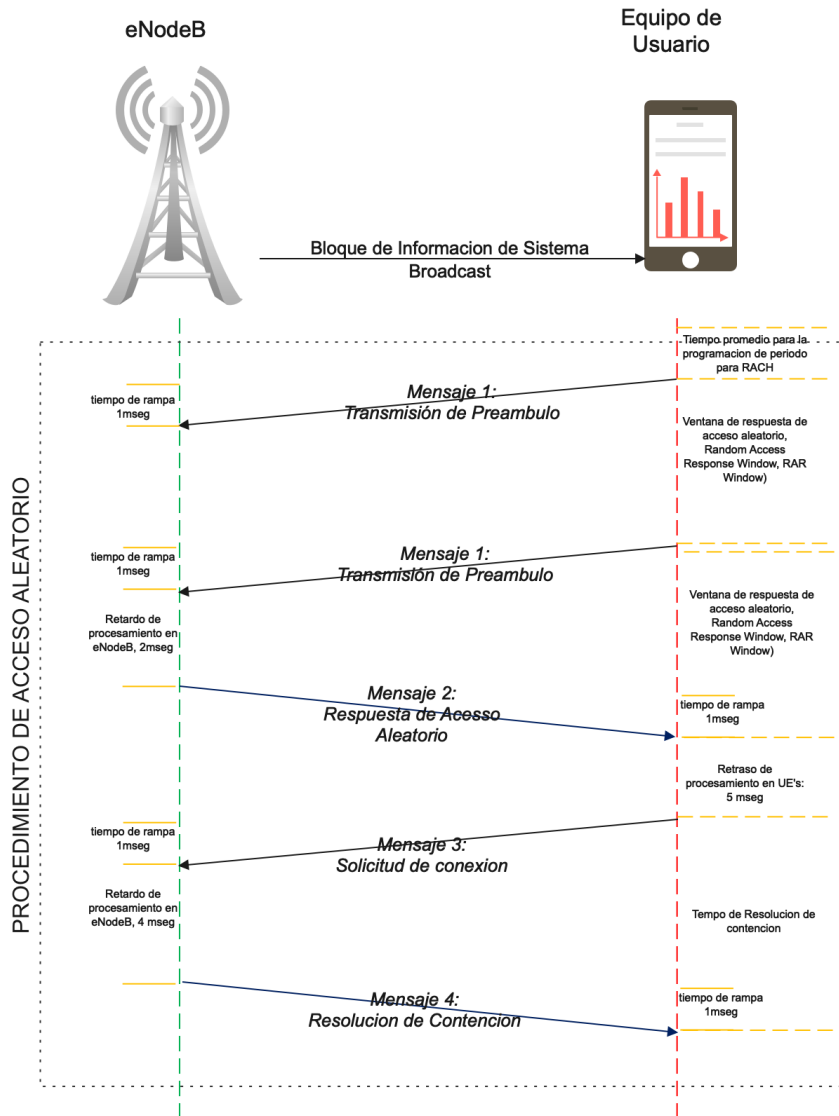


Figura 2. 5 Procedimiento de Acceso Aleatorio basado en Contención (Tello Oquendo, et al., 2018).

- Mensaje 2 – Respuesta de Acceso Aleatorio: Cuando se ha podido decodificar con éxito un preámbulo, eNB calcula un identificador denominado *Random Access Radio Network Temporary Identifier (RA-RNTI)*, el cual está en función del preámbulo y del slot de RA, $ID = f(\text{prambulo}, \text{RAO})$. El eNB haciendo uso del *Physical Downlink Control Channel (PDCCH)*, transmite un mensaje de Respuesta de Acceso

Aleatorio (Random Access Response, RAR), que contiene información como:

- Identificación del preámbulo detectado.
- Instrucciones de alineación de tiempo para sincronizar las transmisiones en el enlace ascendente.
- Asignación de recursos de enlace ascendente que utilizará el dispositivo para transmitir el tercer mensaje del protocolo de enlace.
- Asigna un Identificador Temporal de Red de Radio Celular (Cell Radio Network Temporary Identifier, C-RNTI).
- En caso de falla, un BI.

El RAR está dirigido a un RA-RNTI específico, es decir, a todos los dispositivos que transmitieron un preámbulo en una ranura de RA específica. El RAR contiene diferentes subtítulos asociados a cada preámbulo detectado. Si un dispositivo recibe un mensaje RAR pero no contiene el identificador del preámbulo utilizado, realiza un tiempo de retransmisión aleatorio (según el parámetro BI adjunto al RAR) antes de programar otro intento de transmisión de preámbulo (Mensaje 1) (Tello Oquendo, et al., 2018) (Laya , Alonso, & Alonso Zarate).

- Mensaje 3 – Solicitud de conexión: El UE luego de recibir el RAR, ajusta un tiempo de transmisión de solicitud de conexión al eNB, haciendo uso de los recursos ya asignados. Al mensaje 3, se agrega una petición de retransmisión automática híbrida (Hybrid Automatic Retransmission Request, HARQ), esta es utilizada para proteger la transmisión del mensaje. Hay que considerar que en el caso de que el eNB, durante el Msg2, no detectará el uso de un mismo preámbulo en un mismo RAO, más de un dispositivo utilizará los mismos medios físicos para la solicitud de conexión provocando una colisión en el eNB (Tello Oquendo, et al., 2018) (Laya , Alonso, & Alonso Zarate).

- Mensaje 4 – Resolución de Contención: Al recibir una petición de conexión, el eNB emite un mensaje de resolución de contienda, expresado en el mensaje 4, y a la vez se agrega un proceso HARQ. En el caso de que un UE no reciba un mensaje de resolución de contienda, realiza nuevamente un RAP basado en el BI. El UE realiza un seguimiento de las retransmisiones realizadas, ya que este dispone de un contador de transmisión de preámbulos, en el caso de superar el valor máximo de retransmisiones permitidas, se declara red no disponible finalizando el procedimiento de acceso aleatorio (Tello Oquendo, et al., 2018) (Laya , Alonso, & Alonso Zarate) (Santander González & Tello Oquendo, 2021).

2.3.2 Procedimiento de retransmisión (Backoff)

Cuando se produce una falla durante el intento de acceso al canal de comunicación o una colisión en el canal de acceso a la red, los UEs nuevamente intentan acceder a la red siguiendo el esquema indicado en la Figura 2.6. Para este proceso de retransmisión lo tienen que realizar luego de un tiempo, T_{BO} , cuyo valor está en milisegundos (mseg) y a la vez está en función del valor de BI (Tello Oquendo, et al., 2018) (HUDA , MOHAMED, NOR, & ZURINA, 2020).

Donde:

- $T_{BO} \sim \mathcal{U}(0, BI)$ tiene una distribución uniforme.
- BI representa el valor del Indicador de Retransmisión

EL valor del BI es un valor asignado por un eNB de forma aleatoria en un rango de 0 a 960, aunque por estándar la 3GPP recomienda que el mismo sea de 20 mseg. (HUDA , MOHAMED, NOR, & ZURINA, 2020) (Tello Oquendo, et al., 2018). Hay que considerar que un valor reducido de BI aumentaría la probabilidad de que existan colisiones en un ambiente

masivo de solicitudes de acceso, por otro lado, si se coloca un valor elevado de BI aumentaría la probabilidad de acceso y también aumentaría el tiempo de acceso, originando un retraso en la conexión (HUDA , MOHAMED, NOR, & ZURINA, 2020).

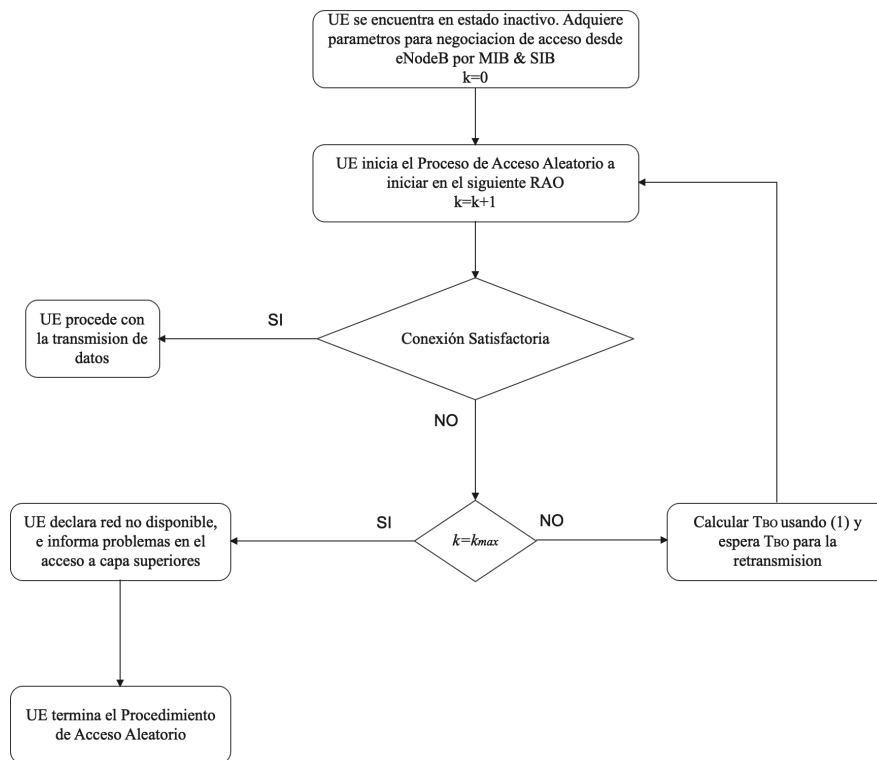


Figura 2. 6 Procedimiento de retransmisión realizado por un UE. (Tello Oquendo, et al., 2018)

2.3.3 Selección de tiempo para la retransmisión, T_{BO}

Aproximar un valor que indique cuántos usuarios lanzan una solicitud de acceso a la BS en un tiempo determinado, permite seleccionar un valor apropiado de BI de manera que menos UE's soliciten acceso en el mismo instante de tiempo, T_{BO} , logrando minimizar nuevamente una colisión de solicitudes de acceso.

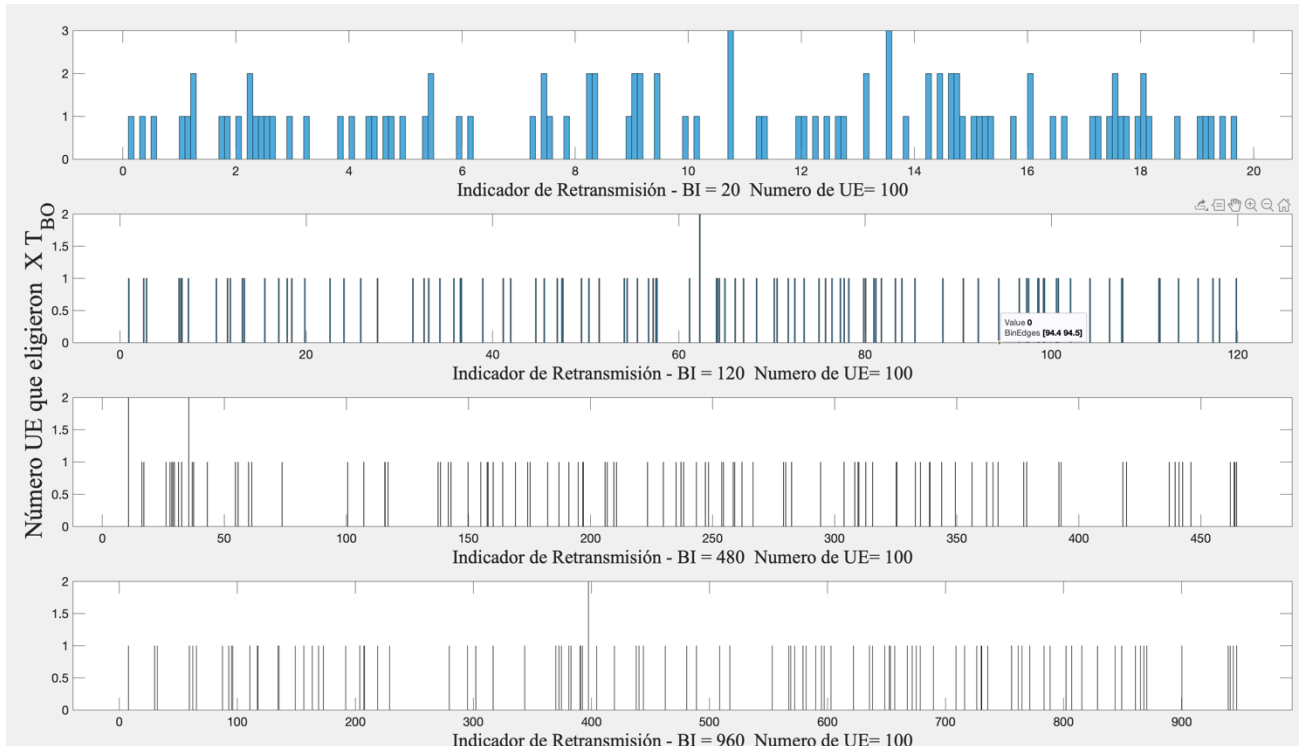


Figura 2. 7 Distribución de tiempo de retransmisión dado diferentes valores de BI para 100 UE's.

La Figura 2.7 así como la Figura 2.8, muestra el impacto que tiene los valores de BI frente a diversas cargas que se puede tener en el RACH. Para el caso de la Figura 2.7 se considera el RACH con carga baja, i.e., una cantidad de 100 usuarios intentando acceder a la red. En el primer caso con un BI=20 se puede ver que cuando se tiene un $T_{BO}=2$, un dispositivo va a empezar el procedimiento de solicitud de acceso, mientras que en un tiempo aproximado de $T_{BO}=10.8$ msec, son hasta tres los dispositivos que realizarán el procedimiento para solicitud de acceso en ese instante de tiempo, lo que ocasionará que en este caso existan 3 dispositivos que empiezan un nuevo procedimiento de retransmisión. Por otro lado, cuando el valor de BI=960, en un tiempo aproximado de $T_{BO} \cong 400$ msec existe 2 UE's que solicitarán un proceso de retransmisión. Para el caso de la Figura 2.8, se realiza un análisis con 30000 usuarios, i.e., carga elevada en el RACH. Cuando se tiene un BI=20 msec, en el valor de $T_{BO} = 0$ msec se

tiene alrededor de 150 usuarios que realizarán un proceso de retransmisión, teniendo el peor de los casos en un $T_{BO} \cong 12.5 \text{ mseg}$ con aproximado de 190 usuarios que realizarían un proceso de retransmisión; considerando un valor de BI= 960 para el caso de un $T_{BO} \cong 0 \text{ mseg}$ se tiene alrededor de 5 UE's que realizaron un proceso de retransmisión en ese instante, mientras que el peor de los casos sucede en un $T_{BO} \cong 490 \text{ mseg}$ con un total de 10 UE's que realizarán un proceso de retransmisión.

Para ambas situaciones, ya sea cuando existen 100 o 30000 usuarios intentando acceder, se puede evidenciar que, cuando el valor de BI incrementa, el número de colisiones disminuye debido a la distribución de los accesos en el tiempo.

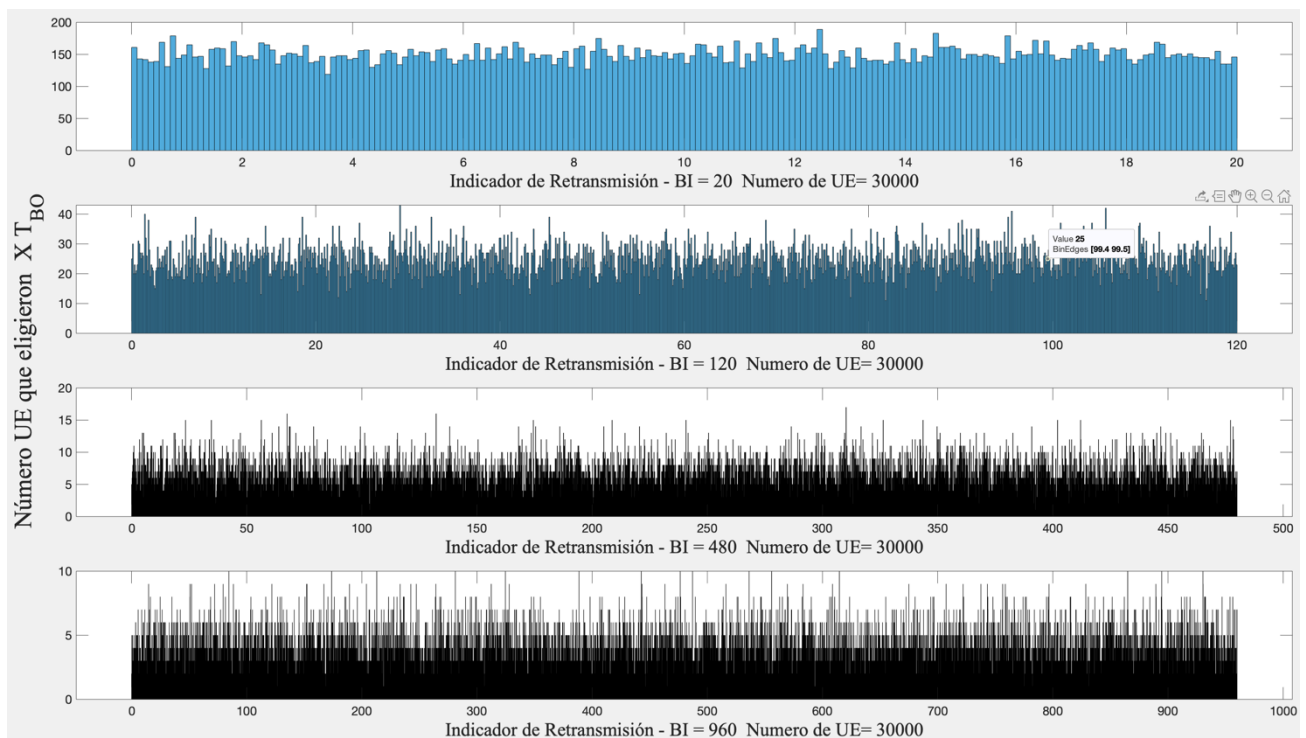


Figura 2. 8 Distribución de tiempo de retransmisión dado diferentes valores de BI para 30000 UE's.

2.4 Sistemas dinámicos basados en aprendizaje reforzado

En el entorno existe sistemas de tipo estáticos y sistemas de tipo dinámicos. Los sistemas estáticos son aquellos donde tienen una entrada conocida y su comportamiento es predecible (Vilchez Valenzuela, 2021) . Por otro lado, un sistema dinámico es aquel sistema que el valor de su estado tiene dependencia de las variables que lo componen, considerando a estas variables como el medio en donde se da el sistema o al tiempo o momento en el cual se está ejecutando dicho sistema. En el caso de un sistema de telefonía celular se toma como base a un sistema dinámico, debido a que existe variables las cuales hacen poco probable el poder predecir con exactitud tal comportamiento.

Cuando se requiere establecer una nueva conexión entre un UE y un eNB, se hace uso del RACH, este canal tiene un comportamiento dinámico y aquí es donde existen colisiones en las solicitudes de conexión (Tello Oquendo, et al., 2018), minimizar el número de colisiones no solo permitirá aumentar la Ps, sino que permite tener un beneficio al presentar un mejor rendimiento de la batería del dispositivo. Con el fin de lograr estos objetivos, el conocer el estado del RACH antes de iniciar una solicitud de acceso, permitirá minimizar el número de retransmisiones; utilizando herramientas de aprendizaje automático es posible conocer en qué estado está el canal para poder indicar un tiempo adecuado en el cual los dispositivos puedan lanzar una solicitud de acceso.

El aprendizaje automático, usualmente conocido como Machine Learning, se clasifica en tres formas de aprendizaje como son (García Pascual, 2021):

- Aprendizaje supervisado;
- Aprendizaje no supervisado;
- Aprendizaje por refuerzo.

2.4.1 Aprendizaje Supervisado

A un sistema de aprendizaje supervisado, se lo conoce como un sistema de aprendizaje predictivo, ya que para una(s) entrada(s) es posible predecir la(s) salida(s) del sistema (Carmona, Pulgar, García, González , & del Jesus, 2015) (Ruiz Manosalva, 2019).

La denominación de aprendizaje supervisado nace de la idea de que, al mostrarle los resultados deseados al algoritmo, supervisamos el aprendizaje, es decir, como es presentado en la Figura 2.9, el algoritmo de aprendizaje supervisado obtiene un modelo predictivo a raíz de la comparación entre datos de entrenamiento (textos, documentos, imágenes, entre otros) y etiquetas que se desean obtener (Aguirre Ascona, 2019) (Barrionuevo, Iersche, & Sattolo, 2020).

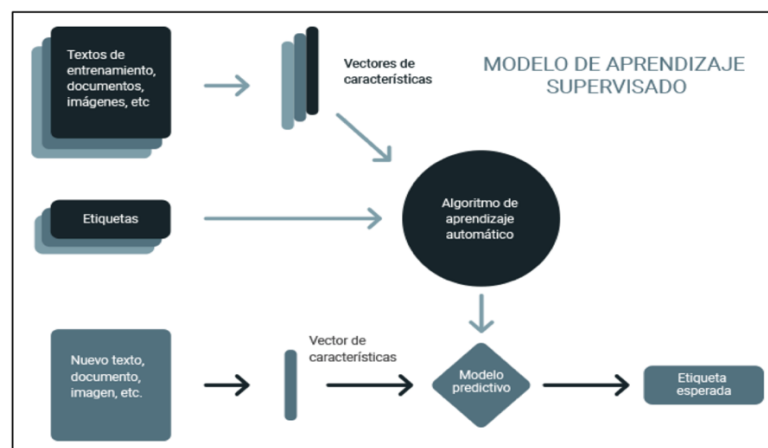


Figura 2. 9. Modelo de Aprendizaje Supervisado. (Ruiz Manosalva, 2019)

2.4.2 Aprendizaje no Supervisado

Un sistema de Aprendizaje no Supervisado es también conocido como sistema de aprendizaje descriptivo (Carmona, Pulgar, García, González , & del Jesus, 2015); la Figura 2.10 muestra el modelamiento de un sistema de

aprendizaje no supervisado, el cual obtiene su modelo predictivo solamente a partir de sus datos de entrenamiento, al desconocer los resultados que debe obtener, este sistema agrupa los datos al encontrar ciertas similitudes entre ellos (Ruiz Manosalva, 2019) (Brusil Cruz, 2020).

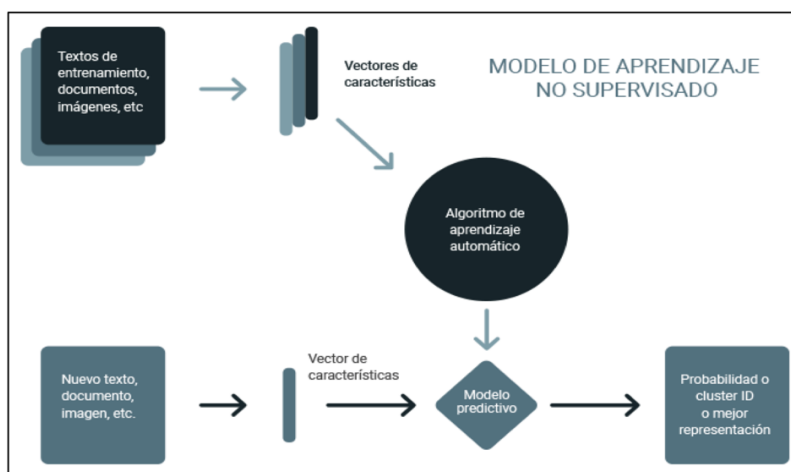


Figura 2.10 Modelo de Aprendizaje No Supervisado. (Ruiz Manosalva, 2019)

2.4.3 Aprendizaje por Refuerzo

El aprendizaje por refuerzo (Reinforcement Learning, RL) tiene como objetivo principal el encontrar un trayecto óptimo que va a mover un agente por el entorno, por lo general este entorno puede modelarse como un Proceso de Decisión de Markov (García Polo, 2012). Los algoritmos de aprendizaje reforzado son muy usados en ambientes dinámicos y usualmente desconocidos, de manera que es poco probable determinar una acción absoluta, pero si modificar las acciones en función de los resultados (Fernández Vizcaíno & Gellego Durán, 2016; Hernández Zaragoza, 2009). El aprendizaje por refuerzo tiene como objetivo obtener la mejor recompensa, en función de las acciones que se han desarrollado en un entorno (Ruiz Manosalva, 2019).

El modelo de aprendizaje por refuerzo, expresado en la Figura 2.11, se define como un agente interconectado a un entorno mediante sensores los cuales le dan información del estado del ambiente (García Polo, 2012), con el uso de esta información el agente toma acciones y recibe los resultados de dichas acciones (Hernández Zaragoza, 2009). En base a esos resultados o también conocidos como recompensas, se genera una política óptima de toma de decisiones, que es la suma de todas las recompensas (García Polo, 2012).

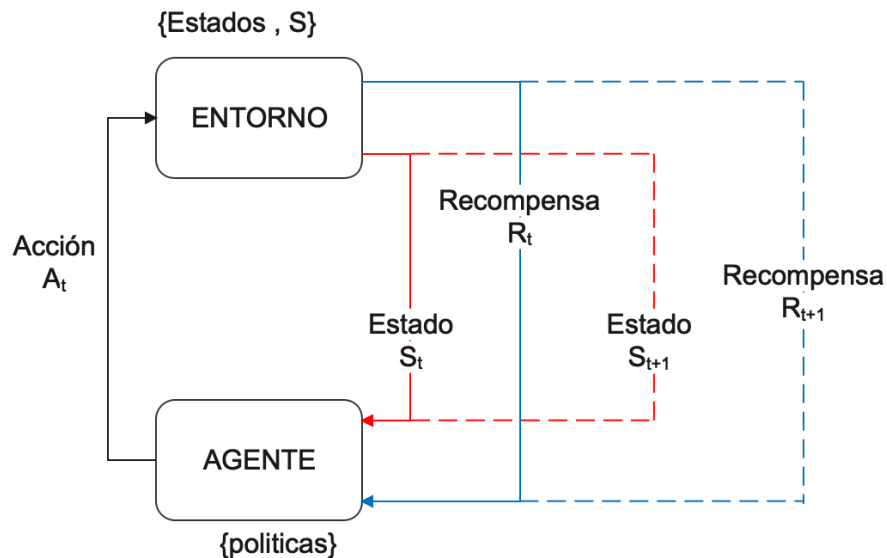


Figura 2. 11 Modelo de Aprendizaje por Refuerzo (García Pascual, 2021).

El modelo de aprendizaje por refuerzo se conforma de:

- A acciones o movimientos a ejecutar un agente dentro de un entorno.
- S estados o posiciones posibles del agente en el entorno.
- R la recompensa es el valor que dado en un estado S cuando se ejecuta una acción A.

Dado un tiempo t el agente se encuentra en un estado S_t , para lo cual se ejecuta una acción A_t llevando al agente a un nuevo estado S_{t+1} y a la vez recibe una recompensa R_t . La recompensa R_t informa al agente qué tan

factible fue el realizar la acción \mathbf{A}_t ; en el nuevo estado \mathbf{S}_{t+1} se ejecuta la acción \mathbf{A}_{t+1} , el agente se posiciona en el estado \mathbf{S}_{t+2} y recibe la recompensa \mathbf{R}_{t+1} , al repetir este proceso el agente obtiene un conjunto de recompensas ($\mathbf{R}_t, \mathbf{R}_{t+1}, \mathbf{R}_{t+2}, \mathbf{R}_{t+3}, \dots$). Obtenido el conjunto de recompensas, el agente busca la política adecuada que está en función de las recompensas obtenidas (García Pascual, 2021) (Hernández Zaragoza, 2009).

CAPÍTULO 3

3. DISEÑO DE UN ALGORITMO BASADO EN Q-LEARNING PARA AJUSTAR DINÁMICAMENTE EL TIEMPO DE BACKOFF

3.1. Modelo del sistema

Para el desarrollo del algoritmo se plantea un esquema como indica la Figura 3.1, un entorno que permite la convivencia entre dispositivos que generan tráfico HTC, así como tráfico MTC, considerando además un solo eNB, que es donde se pretende que algoritmo, que está en base a un sistema de aprendizaje por refuerzo, realice el análisis al momento que un UE intenta obtener un acceso a la red; se prevé validar el algoritmo haciendo hincapié a tres métricas de evaluación, las cuales indican la factibilidad de implementar el algoritmo en el entorno indicado.



Figura 3. 1. Modelo del sistema

3.2. Definición y desarrollo del algoritmo Q-Learning

Existen diversos algoritmos de aprendizaje por refuerzo; sin embargo, en esta investigación utilizamos el algoritmo *Q-Learning* para que un agente (eNB) aprenda a interactuar con el entorno y adquiera políticas de ejecución de acciones que debe realizar hasta conseguir la meta con la recompensa más alta (TELLO OQUENDO,

PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018) (García Polo, 2012) (Hernández Zaragoza, 2009) (Pacheco-Paramo & Tello-Oquendo, 2020).

El algoritmo de *Q-Learning*, representado por la función $Q(s, a)$, representa el valor de la recompensa esperada que se obtiene al ejecutar una acción a en un estado determinado s , para el caso de estudio este valor es almacenado en el eNB (TELLO OQUENDO, PACHECO PARAMO, PLA, & MARTINEZ BAUSET, 2018) (Hernández Zaragoza, 2009).

La función Q , usada en un algoritmo de *Q-Learning*, está representada en la ecuación 3.1:

3. 1

$$Q(s, a) = Q(s, a) + \alpha \left[\mathcal{R} + \gamma \max_{a' \in \mathcal{A}} [Q(s', a')] - Q(s, a) \right];$$

donde:

- α es la tasa de aprendizaje;
- \mathcal{R} recompensa que se obtiene al ejecutar una acción;
- γ conocido como el factor de descuento, que tiene una afectación directa sobre las recompensas futuras.

Dado el presente caso, donde se pretende la integración de un algoritmo de RL como es *Q-Learning* con la asignación del valor BI por parte de una BS, las **acciones** son el conjunto de posiciones de valor que puede tomar el BI, $\mathcal{A} = \{1,2,3,4,5, \dots, 12,13\}$, valores de BI representados en la Tabla 3.1 que a la vez llegan a formar el conjunto de **estados** del algoritmo de RL.

El valor de BI es dado por el eNB en un mensaje SIB2 que es enviado por broadcast cada T_{SIB2} , por estandarización tiene un valor $T_{SIB2} = 16 \text{ RAO}$ (80 *mseg*); con el uso de *Q-Learning* se pretende cambiar el valor de BI antes de que el mensaje SIB2

sea transmitido. Por otro lado, la Recompensa \mathcal{R} está representada en la ecuación 3.2:

3. 2

$$\mathcal{R} = f(\overline{N_{PT}}, CV_{N_{PT}}, \Delta N_{PT}, BI)$$

donde:

- $\overline{N_{PT}} = \frac{1}{T_{SIB2}} (\sum_{k=1}^{T_{SIB2}} N_{PT_k})$, es el número medio de preámbulos transmitidos en un RAO que el eNB detecta en un T_{SIB2} ;
- $CV_{N_{PT}} = \left[\frac{1}{T_{SIB2}-1} (\sum_{k=1}^{T_{SIB2}} |N_{PT_k} - \overline{N_{PT}}|^2) \right]^{1/2} / \overline{N_{PT}}$, es el coeficiente de variación de N_{PT} para un mismo periodo;
- ΔN_{PT} es la diferencia del valor medio de los preámbulos transmitidos entre el periodo actual de transmisión y el anterior.

Tabla 3. 1 Valores que puede tomar el Indicador de Retransmisión según 3GPP (LTE Quick Reference , n.d.).

ÍNDICE	VALOR DEL INDICADOR DE RETRANSMISIÓN
1	10
2	20
3	30
4	40
5	60
6	80
7	120
8	160
9	240
10	320
11	480
12	960

Por otro lado, la Tabla 3.2 muestra los valores de la tasa de aprendizaje α , la probabilidad de exploración ε y el factor de descuento γ , configurados para la ejecución del algoritmo; estos valores fueron considerados al realizar una prueba variando estos parámetros y observando el comportamiento del algoritmo en los resultados obtenidos de las métricas de evaluación.

En el Anexo 1 se muestra resultados utilizando valores diferentes para α , ε y γ para observar el impacto de la configuración de estos parámetros durante la puesta en marcha del algoritmo.

Tabla 3. 2 Valores de las propiedades de algoritmo de RL.

Ítem	Denominación	Símbolo	Valor Colocado	Rango
1	Tasa de Aprendizaje	α	0.2	0-1
2	Factor de descuento	γ	0.9	0-1
3	Probabilidad de exploración	ε	0.1	0-1

La **Tasa de aprendizaje " α "**, es el valor que va a impulsar los nuevos valores óptimos en comparación con los valores anteriores. El rango de este valor se encuentra entre 0 y 1; al colocar un valor de $\alpha = 0$ el algoritmo se demorará más en actualizar los valores óptimos, por otro lado cuando $\alpha = 1$ se pueden dar errores en la selección de los nuevos valores.

El **Factor de descuento" γ "**, es un valor que va a incentivar al agente llegar a la meta de una forma más rápida y efectiva, el cual es medido en un valor de retorno (ver ecuación 3.3), para esto se multiplica la recompensa obtenida por el factor de descuento elevado al instante de tiempo en el que se dio esa recompensa menos 1, el valor puede variar entre 0 y 1.

$$G_0 = \gamma^0 R_1 + \gamma^1 R_2 + \gamma^2 R_3 + \dots \\ + \gamma^{n-1} R_n$$

donde:

- G_0 , valor de retorno;
- γ , Factor de descuento;
- R_x , recompensa al ejecutar una acción.

Cuando el valor el factor de descuento es 0, las recompensas futuras salvo la primera serán 0, por lo que el algoritmo no tomará en cuenta a las acciones futuras. Por otro lado, cuando el factor de descuento tiene un valor de 1, aumentará el tiempo para que el algoritmo pueda tomar la decisión con la mejor estrategia (Labs, 2022).

La **Probabilidad de exploración " ε "**, indica la probabilidad que tiene una acción óptima de ser elegida. Esta probabilidad de exploración forma parte de una política denominado ε avaricioso, (ver ecuación 3.4), esta política define que todas las acciones van a ser escogidas con una probabilidad > 0 . Por otro lado, la probabilidad de tomar una acción aleatoria ε_r , (ver ecuación 3.5), mientras que con probabilidad $1 - \varepsilon$ el algoritmo selecciona la acción con máximo $Q(s, a)$.

3.4

$$\pi(a|s) = \begin{cases} 1 - \varepsilon + \varepsilon_r & a = a^* \\ \varepsilon_r & a \neq a^* \end{cases}$$

3.5

$$\varepsilon_r = \frac{\varepsilon}{|A|}$$

donde:

- A , número de acciones;

- a^* , acción con el valor Q óptimo;
- ε , probabilidad de exploración.

3.3. Métricas de rendimiento a evaluar

Para el presente estudio, el tener métricas de rendimiento permiten evaluar el desempeño de la implementación de un algoritmo de aprendizaje por refuerzo en un entorno mMTC, para lo cual se plantea realizar esta evaluación haciendo un enfoque en tres parámetros como son: probabilidad de acceso satisfactorio, el retardo en el acceso y el número de retransmisiones.

3.3.1 Probabilidad de Acceso Satisfactorio (Ps)

Al tener una red celular que permite la convivencia de tráfico tanto HTC con MTC, es necesario estimar un valor que indique qué tan probable es acceder a dicha red celular, dadas ciertas condiciones del canal.

La Tabla 3.3, indica los valores configurados en el RACH del eNB, con los cuales se realiza el análisis de rendimiento de la red, cabe mencionar que estos se encuentran como valores estándares establecidos por la 3GPP

Tabla 3. 3 .- Parámetros de configuración del RACH

Parámetros	Valor Configurado
Número de Preámbulos	54
Periodicidad del RAO	5
Número de otorgamiento de enlace ascendente por subtrama	3
Número máximo de retransmisiones del preámbulo	10
Indicador de retransmisión	20
Temporizador de resolución de contienda	48
Retraso en el procesamiento del preámbulo	2
Número máximo de retransmisiones entre msg3 y msg4	5

Número total de dispositivos	30000
------------------------------	-------

Dado los parámetros indicados en la Tabla 3.3, se hace una variación del total de dispositivos que se intentaran conectar a la red celular; cuyos resultados son indicados en la Tabla 3.4, que a la vez muestran la Ps al variar el BI.

Tabla 3. 4 Tabla de la Probabilidad de Acceso Satisfactorio al variar el indicador de Retransmisión BI

Numero de Dispositivos	Probabilidad de Acceso Satisfactorio, Ps	Backoff Indicator BI
10000	100%	20
15000	98.26%	
20000	69.46%	
25000	44.48%	
30000	31.07%	
10000	100%	50
15000	98.58%	
20000	70.38%	
25000	45.79%	
30000	32.16%	
10000	100%	100
15000	98.97%	
20000	73.48%	
25000	47.75%	
30000	33.68%	
10000	100%	300
15000	99.32%	
20000	87.53%	
25000	58.05%	
30000	40.80%	
10000	100%	600
15000	99.95%	
20000	96.46%	
25000	78.83%	
30000	55.43%	

El valor numérico de la P_s , (ver ecuación 3.6), está definido por la relación entre el total de dispositivos que ingresan al eNB con el total de dispositivos que completan un RAP satisfactorio (Santander González & Tello Oquendo, 2021):

$$P_s = \frac{n_A}{n_T}$$

3. 6

donde:

- n_A , definido como el total de dispositivos MTC con acceso satisfactorio;
- n_T , definido como el total de dispositivos MTC que ingresaron a la BS.

3.3.2 Número medio de retransmisiones.

Conocer el valor de la métrica, número medio de retransmisiones, permite saber las veces que un UE hace uso de un preámbulo hasta lograr una conexión satisfactoria luego de un intento fallido. Este valor parte de una Función de Distribución Acumulada (Cumulative Distribution Function, CDF), con la cual se obtiene una distribución de probabilidad, (ver ecuación 3.7), que está en función del número de retransmisiones hasta lograr un acceso satisfactorio y el número de UE's que han logrado acceder satisfactoriamente a la red:

$$p(k_n) = \frac{\sum M_A(k_n)}{n_A}$$

3. 7

donde:

- $p(k_n)$, es la probabilidad acumulada de “n” transmisiones de preámbulo;
- $\sum M_A(k_n)$, es el conjunto de dispositivos que accedieron a la red en “n” transmisiones de preámbulo;
- n_A , es el total de dispositivos MTC con acceso satisfactorio.

El valor medio de retransmisiones de accesos satisfactorio, (ver ecuación 3.8), se obtiene a partir de la media de distribución discreta:

$$\mu = \frac{\sum k_n \cdot p(k_n)}{N}$$

3. 8

donde:

- μ , es la media aritmética ponderada;
- k_n , es el número de preámbulos transmitidos;
- $p(k_n)$, es la probabilidad acumulada “n” preámbulo;
- N , es la sumatoria de las probabilidades acumuladas en los diferentes preámbulos.

3.3.3 Retardo en el acceso.

El retardo de acceso se define como el tiempo que toma una solicitud de acceso entre el primer intento hasta conseguir un acceso satisfactorio. El valor de retardo de acceso se lo define por la CDF del retardo de cada uno de los dispositivos, (ver ecuación 3.9), hasta lograr un acceso satisfactorio y conocer el número de UEs que han logrado acceder satisfactoriamente a la red:

$$p(D_{n[mseg]}) = \frac{\sum M_A(D_{n[mseg]})}{n_A}$$

donde:

- $p(D_{n[mseg]})$, es la probabilidad acumulada de “n” retardo;
- $\sum M_A(D_{n[mseg]})$, es el conjunto de dispositivos MTC con un acceso satisfactorio en el retardo “n”;
- n_A , es el número total de dispositivos MTC que han tenido un acceso satisfactorio.

Para lograr una mayor precisión en los datos se realiza una interpolación de estos, organizándoles de menor a mayor con el fin de aplicar medidas de posición –percentil 95– y determinar exactamente el retardo de acceso correspondiente a cada intento de acceso (Santander González & Tello Oquendo, 2021) (Tello Oquendo, et al., 2018).

CAPÍTULO 4

4. EVALUACIÓN Y DISCUSIÓN DE RESULTADOS

En este capítulo se presenta a detalle los resultados obtenidos al aplicar un algoritmo de aprendizaje reforzado, y el impacto que este provoca sobre el parámetro denominado Indicador de Backoff, frente a una ráfaga de solicitudes de conexión en un entorno mMTC, haciendo énfasis en las métricas de evaluación.

4.1 Escenarios de tráfico MTC y configuración de parámetros para simulación

La Tabla 4.1, muestra los valores de tráfico HTC y MTC considerados para las simulaciones utilizando Q-learning, los cuales son los máximos valores que soporta un eNB. Durante las simulaciones se probarán diversos valores de BI en el proceso de solicitud de conexión de un UE al eNB.

Tabla 4. 1 Condiciones de Escenario de Trafico MTC y HTC

Ítem	Parámetro	Unidad	Valor
1	Conexiones MTC	u	30000
2	Conexiones HTC	u	1300

Por otro lado, en base lo indicado en el apartado 3.2, la Tabla 4.2, muestra los parámetros considerados para el desarrollo del algoritmo por refuerzo.

Tabla 4. 2 Condiciones para el desarrollo del Algoritmo Q-Learning

Ítem	Parámetro	Unidad	Valor
1	Tasa de aprendizaje	u	0.2
2	Factor de descuento	u	0.9
3	Probabilidad de exploración	u	0.1

4.2 Resultados y análisis de las métricas de rendimiento

Una vez implementado el algoritmo Q-Learning, la Figura 4.1 muestra la evolución del valor del indicador de Backoff en comparación con las diferentes cargas que se

presentan en el RACH en ciertos instantes de tiempo; las cargas tomadas en cuenta para este análisis son: 10000, 20000 y 30000 solicitudes de conexión en un tiempo determinado. Particularmente, en el caso de 30000 solicitudes de conexión, se observa que la red requiere configurar altos valores para BI con el fin de aliviar la congestión, ya que mientras mayor es el indicador de Backoff, disminuye la probabilidad de que más de un dispositivo transmita en un mismo instante de tiempo.

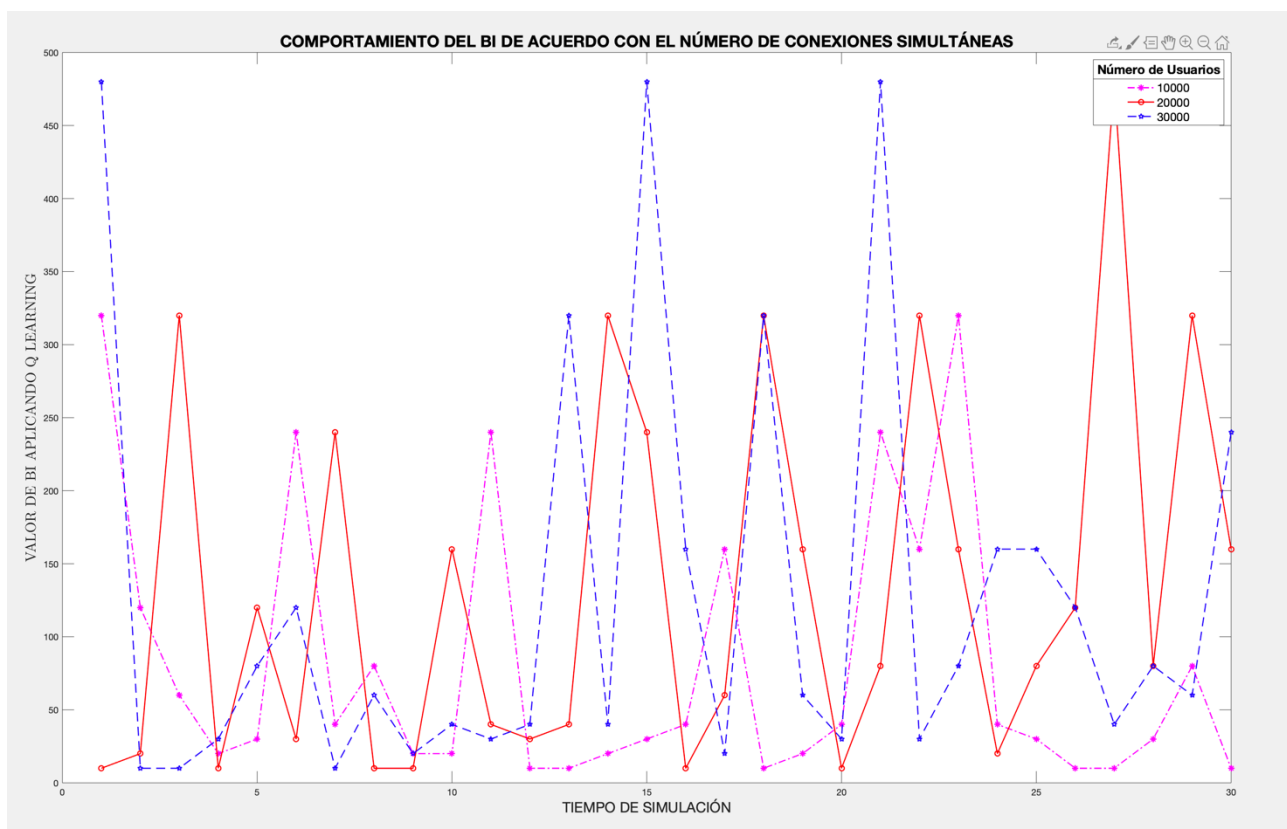


Figura 4. 1 Comportamiento del algoritmo frente a varias cantidades de usuarios.

4.2.1. Probabilidad de acceso satisfactorio.

La Figura 4.2 ilustra el comportamiento de la métrica **Probabilidad de acceso satisfactorio, P_s** . La Figura 4.2 (a) toma como base el comportamiento del sistema usando un BI fijo de 20mseg, según lo establecido por la 3GPP, mientras que la Figura 4.2 (b), representa la P_s

cuando se tiene un BI elegido por un sistema de aprendizaje por refuerzo, como es el caso de Q_Learning.

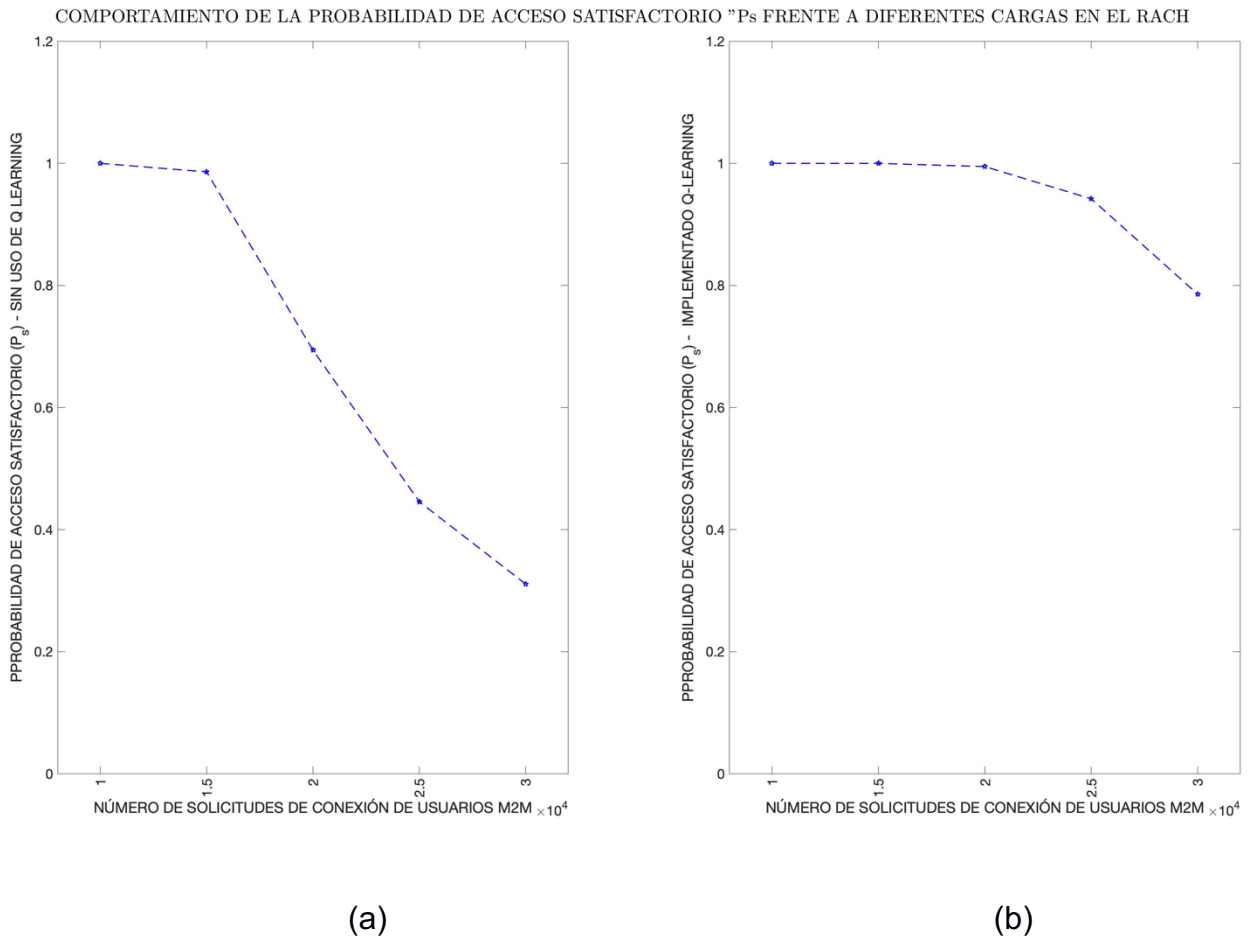


Figura 4. 2 Comportamiento de la Probabilidad de Acceso Satisfactorio; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.

La Tabla 4.3, muestra de manera numérica el comportamiento de la P_s , frente a la variación de las solicitudes de acceso o dispositivos MTC que intentan acceder de manera simultánea a una red celular, teniendo en cuenta un valor de BI fijo.

Tabla 4. 3 Probabilidad de Acceso Satisfactorio en un ambiente con BI fijo.

Dispositivos MTC	Ps	BI
10000	1	20
15000	0.9827	20
20000	0.6916	20
25000	0.4475	20
30000	0.3131	20

Por otro lado, la Tabla 4.4, muestra de manera numérica el comportamiento de la Ps, frente a la variación de las solicitudes de acceso o dispositivos MTC que intentan acceder de manera simultánea a una red celular, teniendo en cuenta un valor de BI que se acopla de acuerdo con el estado del RACH, este acoplamiento está en función de la respuesta que entrega el algoritmo Q-Learning implementado.

Tabla 4. 4 Probabilidad de Acceso Satisfactorio en un ambiente de BI dinámico.

Dispositivos MTC	Ps	BI
10000	1	10
15000	0.9999	80
20000	0.9947	30
25000	0.9430	80
30000	0.7763	80

4.2.2. Número medio de retransmisiones.

Otra de las métricas considerada de alta relevancia para esta investigación es el número de las retransmisiones que llega a realizar un UE para tener acceso exitoso. El sistema se considera para analizar una ráfaga de solicitudes de acceso que puede llegar a tener un entorno MTC, considerando como base el valor del BI. La Figura 4.3, muestra el comportamiento de los dispositivos MTC que realizan retransmisiones en el proceso de conseguir acceso a la red haciendo uso de una BS, la Figura 4.3 (a) indica un comportamiento con un valor de BI constante de 20 mseg,

mientras que la Figura 4.3 (b) muestra un comportamiento con un valor de BI acoplado de acuerdo con el estado del canal.

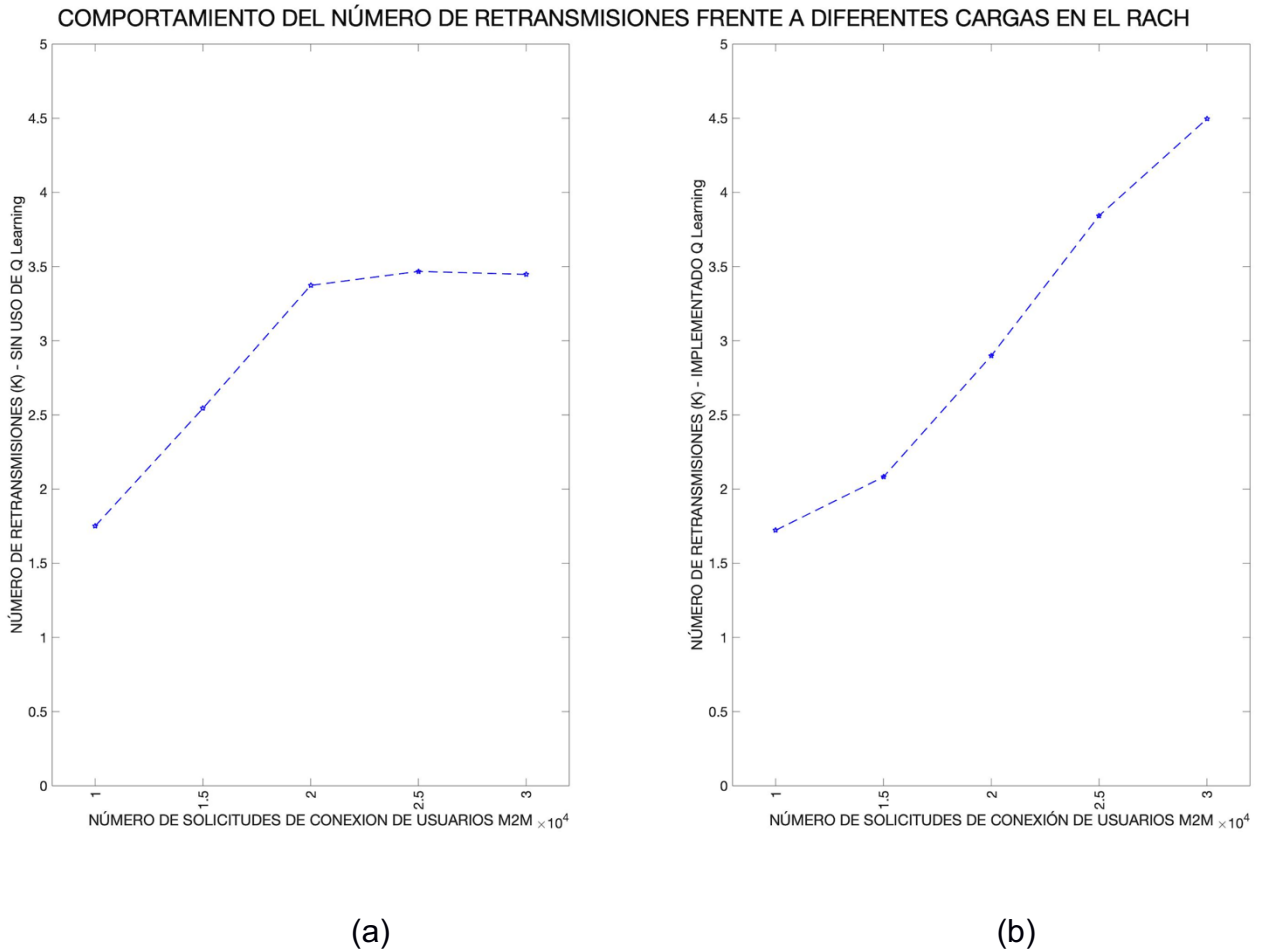


Figura 4. 3 Comportamiento del número de retransmisiones; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.

La Tabla 4.5 indica el número de retransmisiones k , frente a la variación de las solicitudes de acceso o dispositivos MTC que intentan acceder de manera simultánea a una red celular, teniendo en cuenta un valor de BI fijo.

Tabla 4. 5 Número de retransmisiones en un ambiente con BI fijo

Dispositivos MTC	K	BI
10000	1.7553	20
15000	2.5639	20
20000	3.4028	20
25000	3.4937	20
30000	3.4482	20

La Tabla 4.6, muestra de igual manera el número de retransmisiones K frente a la variación de las solicitudes de acceso o dispositivos MTC que intentan acceder de manera simultánea a una red celular, teniendo en cuenta un valor de BI que se acopla de acuerdo con el estado del RACH, este acoplamiento está en función de la respuesta que entrega el algoritmo Q-Learning implementado.

Tabla 4. 6 Número de retransmisiones en un ambiente de BI dinámico.

Dispositivos MTC	K	BI
10000	1.7281	20
15000	2.0823	120
20000	2.8897	40
25000	3.8297	160
30000	4.5310	80

4.2.3. Retardo en el acceso.

La Figura 4.4, hace referencia a la métrica de retardo en el acceso, la cual se ha considerado evaluar en función del percentil 95, la Figura 4.4 (a) hace referencia al análisis del retardo cuando se ha configurado un BI=20 mseg, mientras que la Figura 4.4 (b) muestran los resultados de la métrica retardo en el acceso con un valor de BI que se acopla a la carga de tráfico del RACH.

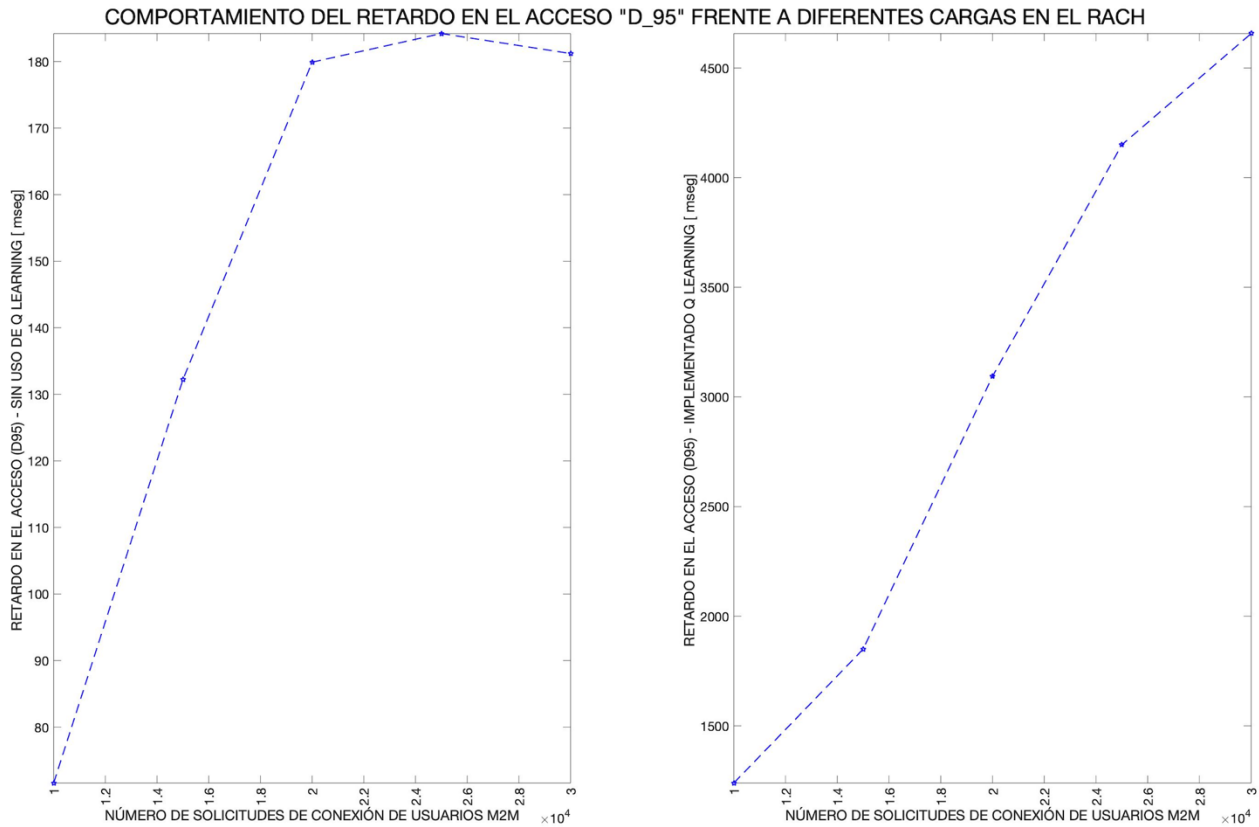


Figura 4. 4 Comportamiento del Retardo al acceso; (a) Ambiente con un valor de BI fijo, (b) Ambiente con un valor de BI óptimo implementando un algoritmo de aprendizaje por refuerzo.

Por otro lado, la Tabla 4.7, indica los valores para D_{95} , considerando un valor fijo del BI y a la vez una masiva solicitud de acceso de dispositivos MTC al intentar acceder a una red de comunicaciones.

Tabla 4. 7 Retardo en el acceso en un ambiente con BI fijo

Dispositivos MTC	D₉₅ (mseg)	BI
10000	71.1	20
15000	132.5	20
20000	180.2	20
25000	182.6	20
30000	182.4	20

Luego de implementar un algoritmo basado en aprendizaje reforzado, la Tabla 4.8 indica los datos obtenidos con respecto a la métrica de retardo en el acceso.

Tabla 4. 8 Retardo en el acceso en un ambiente de BI dinámico.

Dispositivos MTC	D₉₅ (mseg)	BI
10000	1234.8	80
15000	1862.3	40
20000	3095.2	40
25000	4148.5	160
30000	4740.4	320

4.3 Comparación y discusión de resultados

Para el análisis del comportamiento de las métricas, se ha considerado un tráfico que varía desde los 10000 dispositivos hasta los 30000 dispositivos MTC, para términos de simulación se ha realizado en pasos de 5000 dispositivos, debido a que son valores los cuales nos darán una tendencia más exacta del comportamiento del tráfico real.

La Figura 4.2 muestra la variación de la probabilidad de acceso satisfactorio en los escenarios configurados. La Figura 4.2 a) así como la Tabla 4.3 muestran la Ps al tener un valor de retransmisión fijo; cuando existen 10000 dispositivos MTC que requieren acceso a un sistema de red, se tiene un 100% de Probabilidad de acceso satisfactorio; pero, al saturar la capacidad de un eNB con número de 30000

dispositivos solicitando acceso a la red, cae esta métrica a un 31.31 % de dispositivos MTC que logran establecer una comunicación. Hay que considerar que una conexión satisfactoria se da cuando el dispositivo ha cursado con éxito los 4 pasos del intercambio de mensajes en el RACH basado en contención. Por otro lado, el escenario de la Figura 4.2 b) y la Tabla 4.4 ilustran el comportamiento de la Ps al tener un valor que se acopla de acuerdo con la carga de tráfico en el RACH, haciendo uso de Q-Learning; las simulaciones indican que el valor de Ps aumenta gradualmente cuando se tiene un escenario con 30000 dispositivos MTC intentando acceder simultáneamente a la red, en comparación con los valores obtenidos cuando se tiene un valor de BI fijo.

Así mismo, para el análisis del comportamiento de la métrica número de retransmisiones se hace en base a los resultados obtenidos en la Figura 4.3. El primer análisis se los puede observar en la gráfica de la Figura 4.3 a) como en los valores indicados en la Tabla 4.5; estos valores se obtienen cuando el valor del BI es un valor fijo, donde se puede deducir que a mayor cantidad de solicitudes de conexión, mayor se tiene el número de retransmisiones, cuando 10000 dispositivos MTC intentan conectarse existe una media de 1.7553 retransmisiones que ocurre para que un dispositivo se integre a una red de comunicaciones, en cambio cuando se tienen 30000 dispositivos MTC intentando acceder a la red, se tiene 3.44 retransmisiones que tendrá que realizar un dispositivo para formar parte de la red a la cual requiere integrarse; mientras que la gráfica de la Figura 4.3 b) y los valores indicados en la Tabla 4.6, indican los valores obtenidos cuando existen solicitudes de conexión con la diferencia de que el valor del BI es un valor que se acopla al estado del canal, haciendo uso de un sistema de aprendizaje por refuerzo, al haber 10000 solicitudes se tiene un numero de 1.7281 retransmisiones que realiza un dispositivo para integrarse a la red que solicita conexión, sin embargo, cuando se tienen 30000 dispositivos MTC que requieren conexión se da un total de 4.5310 solicitudes de conexión que tiene que realizar un dispositivos para poder integrarse a la red de comunicación . Hay que tener en cuenta que el número de

retransmisiones se da en valores enteros y se debe elevar a su valor inmediato superior.

Los valores de la métrica Retardo en el Acceso es mostrada en la Figura 4.4, al igual que el análisis indicados de las métricas anteriores, se han creado dos escenarios; el escenario indicado por la Figura 4.4 a) cuyos valores se encuentran identificados en la Tabla 4.7, muestran el tiempo de conexión que se tardaría un preámbulo el lograr integrarse a una red de comunicaciones, cuando se tienen 10000 dispositivos que solicitan una conexión de forma simultánea, le toma alrededor de 71.1 mseg en lograr formar parte de la red, en cambio al tener 30000 dispositivos MTC intentando acceder a una red de comunicaciones, a este le tomaría alrededor de 182.4 mseg formar parte de la red de comunicaciones; sin embargo, al aplicar un análisis del canal previo a la asignación del BI, haciendo uso de Q_Learning, y teniendo 10000 usuarios intentando formar parte de una red, el tiempo estimado es de 1234.8 mseg que le tomaría a un preámbulo lograr establecer una comunicación, mientras que al tener 30000 usuarios MTC que requieren un acceso a una red de comunicaciones, le tomaría un tiempo estimado de 4740.4 mseg.

CAPÍTULO 5

5. CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS

5.1 Conclusiones

En esta investigación se realizó un modelamiento del canal de acceso aleatorio basado en contención utilizando el software MATLAB; haciendo un enfoque principal el proceso de retransmisión que realiza un UE, y al objeto de estudio que es la selección del valor óptimo del indicador de backoff, que va a permitir descongestionar el tráfico que se puede generar en el RACH en un ambiente mMTC.

El indicador de backoff BI, tiene sus rangos de valores que varían desde 10 a 960 msec; sin embargo, para la selección de este valor adecuado, se requiere conocer o evaluar el estado del Canal de Acceso Aleatorio. Para la evaluación del RACH se ha utilizado el algoritmo de aprendizaje por refuerzo Q-Learning. Utilizando el algoritmo es posible obtener un valor de BI óptimo, ya que va a seleccionar el mejor valor sin comprometer la recompensa, que es un indicador en el algoritmo Q-Learning, que está en base al número de preámbulos que son transmitidos en un RAO detectados por el eNB en un T_{SIB2} .

Las métricas que fueron consideradas para la evaluación de la hipótesis planteada son: probabilidad de acceso satisfactorio P_s , número medio de retransmisiones y el retardo de acceso. La Probabilidad de Acceso Satisfactorio, P_s , indicado en el apartado 4.2.1, nos muestra los resultados al tener dos escenarios distintos, el primer escenario es evaluado cuando existe una convivencia de dispositivos HTC con dispositivos MTC, con un valor de retransmisión BI fijo, considerando las recomendaciones emitidas por la 3GPP que es de 20 msec, en el peor de los casos que es el tener 30000 dispositivos MTC la P_s es demasiado baja, 31.31%, pero cuando se hace uso de un algoritmo que nos ayude con la selección de un valor

óptimo del BI, este valor sube al 77.63 %. Analizando la siguiente métrica, número de retransmisiones (apartado 4.2.2), considerando los mismos escenarios antes indicados, en el primer caso se puede observar que existe un número de retransmisiones de 3.4482 mientras que en el segundo caso este valor sube a 4.531, hay que recalcar que el número de retransmisión siempre será un valor entero por lo que se deberá considerar subir a su inmediato superior. Por otro lado, para el análisis de la métrica Retardo de Acceso (apartado 4.2.3), considerando los escenarios evaluados en las métricas anteriores, se ha notado que cuando existe un un valor de BI fijo el tiempo de conexión estimado es de 182.4 mseg, mientras que al hacer uso del algoritmo de aprendizaje por refuerzo este valor sube a 4740.4 mseg, que es el valor que toma un dispositivo en integrarse a la red.

La métrica de retardo en el acceso tiende a elevar su valor cuando se analiza en un ambiente con un valor de BI que se acopla al estado del canal, el valor considerado está en función de los retardo más elevado que nos entrega el modelamiento del canal, este valor elevado está de acuerdo al número de retransmisiones y a la vez está ligado propiamente con el valor del Indicador de retransmisión, ya que en si cuando sucede una retransmisión el tiempo que espera el dispositivo para lanzar un nuevo preámbulo es directamente proporcional con el valor de BI, hay que tomar en cuenta que el algoritmo selecciona el valor en función de la Ps.

5.2 Recomendaciones

El retardo es un factor importante en las comunicaciones móviles y se requiere siempre un menor tiempo de conexión. Se validó que el variar el tiempo de retransmisión afecta directamente el retardo en el acceso, por lo que se recomienda que los estudios que consideran aplicaciones sensibles al retardo o críticas en el tiempo de respuesta se utilice un umbral máximo en la variación del Indicador de backoff (BI) para minimizar el impacto en el tiempo de acceso.

5.3 Trabajos Futuros

El despliegue de las comunicaciones ha crecido exponencialmente con el desarrollo de la tecnología, por lo que, para buscar mejoras en el acceso en un ambiente MTC, se debe considerar el acople dinámico de los medios físicos de acuerdo con el estado del RACH. Además, se pretende utilizar la métrica antigüedad de la información para abarcar aplicaciones sensibles al retardo.

BIBLIOGRAFÍA

- TELLO OQUENDO, L., PACHECO PARAMO, D., PLA, V., & MARTINEZ BAUSET, J. (2018). Reinforcement Learning-Based ACB in LTE-A Networks for Handling Massive M2M and H2H Communications. *2018 IEEE International Conference on Communications*, 1-7.
- Santander González, D. A., & Tello Oquendo, L. (2021). *EVALUACIÓN Y CARACTERIZACIÓN DEL IMPACTO DEL TIEMPO DE BACKOFF EN LA RED DE ACCESO RADIO DE IOT CELULAR CON TRÁFICO MASIVO*. RIOBAMBA: UNIVERSIDAD NACIONAL DE CHIMBORAZO .
- Tello Oquendo, L., Leyva-Mayorga, I., Pla, V., Martinez-Bauset, J., Ramon Vidal, J., Casares-Giner, V., & Guijarro, L. (2018). Performance Analysis and Optimal Access Class Barring Parameter Configuration in LTE-A Networks With Massive M2M Traffic. *IEEE(IEEE)*.
- Yu , Z., & Seet, B.-C. (2019). Performance Analysis of Random Backoff based D2D Discovery in Cellular Networks. *International Telecommunication Networks and Applications Conference* , 29.
- Park, H. (2019). Adaptive Backoff enabled WUR on non-cellular local IoT for extreme low power operation. *ELSEVIER*, 108(2020), 62-67.
- HUDA , D. A., MOHAMED, O., NOR, K. N., & ZURINA, M. H. (2020). Dynamic Backoff Collision Resolution for Massive M2M Random Access in Cellular IoT Networks. *IEEE* , 8(2020).
- Camargo, M., Osma, G., & Ortega, H. (2017). Plataforma de servicios basada en IoT para el monitoreo y control de un sistema de generación fotovoltaica (FV) integrado con vegetación. *SICEL*, 9(2017).

- Martínez, E. (2001). La evolución de le telefonía móvil . *RED*.
- Rojas Mendez, H. E., & Urgiles Peralta, F. P. (2011). *ESTUDIO Y DISEÑO PARA LA AMPLIACION DE LA RED NGN DE LA CORPORACION NACIONAL DE TELECOMUNICACIONES EN LA PARROQUIA BORRERO DEL CANTON AZOGUES PROVINCIA DEL CANAR*. CUENCA: UNIVERSIDAD POLITÉCNICA SALESIANA.
- Castillo Medina, C. A., Salcedo Parra, O. J., & Forero Rodriguez, F. (2014). A brief survey of new generation networks. En *Electronic Vision* (págs. 194-205). Jose Caldas, Colombia: Universidad distrital Francisco José de Caldas.
- CHEN, J.-C., & ZHANG, T. (2004). *IP-BASED NEXT-GENERATION WIRELESS NETWORK*. New Jersey: John Wiley & Son, Inc.
- Guzman Quintero, Y. (2012). GESTION DE REDES Y SERVICIOS NGN/4G. *Revista tecnica de la Empresa de Telecomunicaciones de Cuba S.A.*, 9, 26-35.
- Utrilla Salazar, D. (2014). QoS. *EN REDES MOVILES DE CUARTA GENERACION*. Callao : Universidad Nacional de Callao Escuela de Postgrado.
- Ratana Bhalla, M., & Vardhan Bhalla, A. (August 2010). Generations of Mobile Wireless Technology: A Survey. *International Journal of Computer Applications*, 5(4), 26-32.
- Haard , M., Patel, D., Joshi, B., & Modi, H. (2014). 0G to 5G Mobile Technology: A Survey. *Journal of Basic and Applied Engineering Research*, 1(6), 56-60.
- Tibúrcio Pereira, P., Sousa Ramos, F., & Lopes Ramos, G. (2020). Study of Attenuation of Centimetric Waves of the Fifth Generation of Cellular Telephony in Vegetation Areas using the Kriging Model. *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT)*. Sakheer, Bahrain.

- Xia, Y. (2021). The debates on the advent of the fifth generation of mobile telephony (5G). *15th INTERNATIONAL CONFERENCE NooJ 2021*. Virtual conference Besançon, France.
- Medeiros, A. M., Pereira, C. G., & Castilho, J. V. (2017). The fifth generation of mobile communication and its applications on the Internet of Things (IoT). *CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*. Pucon, Chile.
- Otero Rojas, M. A. (2018). *Evaluación del desempeño de protocolos de Control de Acceso al Medio para comunicaciones Máquina a Máquina (M2M)*. Santa Clara: Universidad Central Marta Abreu de las Villas.
- Fernandez, R. (23 de Nov de 2021). *Internet de las Cosas (IoT): dispositivos conectados en el mundo 2019-2030*. (STATISTA) Recuperado el 02 de 02 de 2022, de <https://es.statista.com/estadisticas/517654/prevision-de-la-evolucion-de-los-dispositivos-conectados-para-el-internet-de-las-cosas-en-el-mundo/>
- Gokhale , P., Bhat, O., & Bhat, S. (2018). Introduction to IOT. *International Advanced Research Journal in Science, Engineering and Technology*, 5(1), 40-44.
- Cao, J., Ma, M., Li, H., Fu, Y., & Lio, X. (2018). EGHR: Efficient group-based handover authentication protocols for mMTC in 5G wireless networks. *Journal of Network and Computer Applications* , 102(1), 1-16.
- Alvi, M., Abualnaja, K. M., Tariq Toor, W., & Saadi, M. (2021). Performance analysis of access class barring for next generation IoT devices. *Alexandria Engineering Journal*, 60, 615–627.
- Rohde & Schwars. (2022). *Medidas 5G | Casos de uso de mMTC*. (Rohde & Schwars) Recuperado el 3 de 02 de 2022, de <https://www.rohde-schwarz.com/es/soluciones/test-and-measurement/wireless->

communication/estandares-de-telefonía-movil/5g-test-and-measurement/mmtc/mmtc_233752.html

Laya , A., Alonso, L., & Alonso Zarate, J. (s.f.). *Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives*. Catalunya: Universitat Politècnica de Catalunya.

Vilchez Valenzuela, E. E. (2021). *ALGORITMOS DE APRENDIZAJE REFORZADO PARA EL PROBLEMA DE RUTEO DE VEHÍCULOS DINÁMICO*. Santiago de Chile: UNIVERSIDAD DE CHILE FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS DEPARTAMENTO DE INGENIERÍA MATEMÁTICA.

García Pascual, M. (2021). *Aprendizaje por refuerzo profundo con OpenAI Gym*. Madrid: Universidad Autónoma de Madrid.

Aguirre Ascona, Y. D. (2019). *Métodos de aprendizaje supervisado para la predicción de día- betes: una revisión sistemática de la literatura*. Lima: UNIVERSIDAD PERUANA UNIÓN.

Barrionuevo, C., Iersche, J., & Sattolo, I. (2020). *Reconocimiento de emociones a través de expresiones faciales con el empleo de aprendizaje supervisado aplicando regresión logística*. Buenos Aires: Universidad de Morón.

Carmona, C., Pulgar, F., García, A., González , P., & del Jesus, M. (2015). *Análisis descriptivo mediante aprendizaje supervisado basado en patrones emergentes* . Burgos : Universidad de Burgos.

Ruiz Manosalva, G. (2019). *Modelo de análisis de datos utilizando técnicas de aprendizaje supervisado y no supervisado, para identificar patrones en la información generada por los pacientes, sometidos a juegos diseñados como un instrumento de apoyo terapéutico* . Colombia: Universidad Jorge Tadeo Lozano.

Brusil Cruz, C. A. (2020). *ANÁLISIS COMPARATIVO ENTRE APRENDIZAJE SUPERVISADO Y APRENDIZAJE SEMI-SUPERVISADO PARA LA CLASIFICACIÓN DE SEÑALES SÍSMICAS VULCANOLÓGICAS DEL VOLCÁN COTOPAXI*. Quito: ESCUELA POLITÉCNICA NACIONAL .

García Polo, F. J. (2012). *Aprendizaje por Refuerzo para la Toma de Decisiones Segura en Dominios con Espacios de Estados y Acciones Continuo*. Madrid: UNIVERSIDAD CARLOS III DE MADRID.

Fernández Vizcaíno, G., & Gellego Durán, F. J. (2016). *Ajustando Q-Learning para generar jugadores automáticos: un ejemplo basado en Atari Breakout*. San Vicente del Raspeig: Universidad de Alicante.

Hernández Zaragoza, J. C. (2009). *Aprendizaje por Refuerzo Relacional con Acciones Continuas*. Puebla: Instituto Nacional de Astrofísica, Óptica y Electrónica .

LTE Quick Reference . (s.f.). Obtenido de https://www.sharetechnote.com/html/Handbook_LTE_BackOffIndicator.html

Labs, E. V. (2022). *Reinforcement Learning de Cero a Maestro - IA en Python*. https://www.udemy.com/course/reinforcement_learning_principiante_maestro_1/.

Luis Tello-Oquendo, P. (2020). *5G cellular system: A brief review of architecture, use cases, and enabling technologies*. GUAYAQUIL: ESCUELA SUPERIOR POLITECNICA DEL LITORAL.

Pacheco-Paramo, D., & Tello-Oquendo, L. (2020). Delay-aware dynamic access control for mMTC in wireless networks using deep reinforcement learning. *Computer Networks*, 182(107493).

ANEXOS

ANEXO 1

RESULTADOS VARIANDO LOS PARÁMETROS DEL ALGORITMO DE Q_LEARNING

Usuarios MTC=10000

Usuarios HTC=1300

Tiempo=10

α	γ	ϵ	Cálculos sin Q_L			Cálculos con Q_L			T_ejec (seg)	BI_QL
			Ps(%)	k	d_95(mseg)	Ps(%)	k	d_95(mseg)		
0.2	0.9	0.1	31.13	3.45	182.8	75.65	4.42	4698.5	17.28	80
0.8	0.9	0.1	31.44	3.45	182.1	78.33	4.51	4744.5	16.32	40
0.01	0.9	0.1	31.34	3.44	182.9	78.47	4.51	4748.5	16.04	20
0.2	0.5	0.1	31.46	3.46	183.6	78.18	4.51	4749.5	16.36	80
0.2	0.1	0.1	31.51	3.43	182	78.34	4.52	4759.6	16.07	20
0.2	0.9	0.5	31.16	3.44	183.4	72.51	4.27	4649.3	16.39	160
0.2	0.9	0.9	31.12	3.44	181.9	78.31	4.52	4749.5	16.55	120