



**UNIVERSIDAD NACIONAL DE CHIMBORAZO  
FACULTAD INGENIERÍA  
CARRERA ELECTRÓNICA Y TELECOMUNICACIONES**

**DISEÑO DE UN ALGORITMO ADAPTATIVO DE BACKOFF  
BASADO EN APRENDIZAJE REFORZADO PARA  
INCREMENTAR LA PROBABILIDAD DE ACCESO EXITOSO  
EN UNA RED CELULAR CON COMUNICACIONES MASIVAS.**

**Trabajo de Titulación para optar al título de  
Ingeniero en Electrónica y Telecomunicaciones**

**Autor:**

Guilla Muñoz Jacqueline Cristina

**Tutor:**

PhD. Luis Patricio Tello Oquendo

**Riobamba, Ecuador. 2023**

## DERECHOS DE AUTORÍA

Yo, **Jacqueline Cristina Guilla Muñoz**, con cédula de ciudadanía **060405805-7**, autora del trabajo de investigación titulado: **DISEÑO DE UN ALGORITMO ADAPTATIVO DE BACKOFF BASADO EN APRENDIZAJE REFORZADO PARA INCREMENTAR LA PROBABILIDAD DE ACCESO EXITOSO EN UNA RED CELULAR CON COMUNICACIONES MASIVAS**, certifico que la producción, ideas, opiniones, criterios, contenidos y conclusiones expuestas son de mí exclusiva responsabilidad.

Asimismo, cedo a la Universidad Nacional de Chimborazo, en forma no exclusiva, los derechos para su uso, comunicación pública, distribución, divulgación y/o reproducción total o parcial, por medio físico o digital; en esta cesión se entiende que el cesionario no podrá obtener beneficios económicos. La posible reclamación de terceros respecto de los derechos de autor (a) de la obra referida, será de mi entera responsabilidad; librando a la Universidad Nacional de Chimborazo de posibles obligaciones.

En Riobamba, 10 de noviembre de 2023.



---

Jacqueline Cristina Guilla Muñoz

C.I:060405805-7

**DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL;**

Quienes suscribimos, catedráticos designados Tutor y Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **DISEÑO DE UN ALGORITMO ADAPTATIVO DE BACKOFF BASADO EN APRENDIZAJE REFORZADO PARA INCREMENTAR LA PROBABILIDAD DE ACCESO EXITOSO EN UNA RED CELULAR CON COMUNICACIONES MASIVAS**, presentado por Jacqueline Cristina Guaylla Muñoz, con cédula de identidad número 060405805-7, certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha asesorado durante el desarrollo, revisado y evaluado el trabajo de investigación escrito y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 10 noviembre de 2023.

Alejandra del Pilar Pozo Jara, Msc.  
**PRESIDENTE DEL TRIBUNAL DE GRADO**



---

Leonardo Rentería Bustamante, PhD.  
**MIEMBRO DEL TRIBUNAL DE GRADO**



---

Dr. Klever Hernán Torres Rodríguez, Msc.  
**MIEMBRO DEL TRIBUNAL DE GRADO**



---

Luis Patricio Tello Oquendo, PhD.  
**TUTOR**



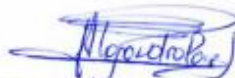
---

## CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **DISEÑO DE UN ALGORITMO ADAPTATIVO DE BACKOFF BASADO EN APRENDIZAJE REFORZADO PARA INCREMENTAR LA PROBABILIDAD DE ACCESO EXITOSO EN UNA RED CELULAR CON COMUNICACIONES MASIVAS**, presentado por Jacqueline Cristina Guaila Muñoz, con cédula de identidad número 060405805-7, bajo la tutoría de PhD. Luis Patricio Tello Oquendo; certificamos que recomendamos la APROBACIÓN de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 10 de noviembre de 2023

Alejandra del Pilar Pozo Jara, Msc.  
**PRESIDENTE DEL TRIBUNAL DE GRADO**



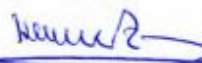
Firma

Leonardo Rentería Bustamante, PhD.  
**MIEMBRO DEL TRIBUNAL DE GRADO**



Firma

Dr. Klever Hernán Torres Rodríguez, Msc  
**MIEMBRO DEL TRIBUNAL DE GRADO**



Firma



Dirección  
Académica  
VICERRECTORADO ACADÉMICO



UNACH-RGF-01-04-02.20  
VERSIÓN 02: 06-09-2021

# CERTIFICACIÓN

Que, **GUAILLA MUÑOZ JACQUELINE CRISTINA** con CC: **0604058057**, estudiante de la Carrera **ELECTRÓNICA Y TELECOMUNICACIONES, NO VIGENTE**, Facultad de **INGENIERÍA**; ha trabajado bajo mi tutoría el trabajo de investigación titulado "**DISEÑO DE UN ALGORITMO ADAPTATIVO DE BACKOFF BASADO EN APRENDIZAJE REFORZADO PARA INCREMENTAR LA PROBABILIDAD DE ACCESO EXITOSO EN UNA RED CELULAR CON COMUNICACIONES MASIVAS**", cumple con el 5 %, de acuerdo al reporte del sistema Anti plagio URKUND, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 08 de noviembre de 2023



PhD. Luis Tello Oquendo  
TUTOR TRABAJO DE INVESTIGACIÓN

## **DEDICATORIA**

Con gratitud y amor, dedico este trabajo a las personas que han sido pilares fundamentales en mi vida. A mis padres, Marina y Julio Cesar quienes, con su amor incondicional, protección y apoyo me guían en cada paso que doy. A mi tía Lilian Muñoz, quien es como una segunda madre y me enseñó lo importante que es superarse y alcanzar mis sueños. A mi sobrina Sofia, quien con su amor y alegría me motiva a seguir adelante y no rendirme. Finalmente, a mi amor, por ayudarme y guiarme en el proceso.

*Jacqueline Cristina Guaila Muñoz*

## **AGRADECIMIENTO**

Mi agradecimiento eterno a Dios, quien me ha brindado salud y vida. A mis padres por darme el regalo más valioso, los estudios y enseñarme el valor del esfuerzo y la perseverancia. A mis abuelitos, tíos y hermanos, quiénes siempre han confiado en mí y en mis capacidades para seguir adelante, ser una mejor persona y nunca rendirme.

Y de manera especial agradecer a mi tutor PhD. Luis Tello, por su compromiso y dedicación en ayudarme a culminar este proceso, sus enseñanzas y guía han sido fundamentales para mi crecimiento académico.

Agradezco a todas las personas que estuvieron presentes en esta etapa de mi vida.

# ÍNDICE GENERAL

DERECHOS DE AUTORÍA	
DICTAMEN FAVORABLE DEL TUTOR Y MIEMBROS DE TRIBUNAL;	
CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL	
CERTIFICADO ANTIPLAGIO	
DEDICATORIA	
AGRADECIMIENTO	
ÍNDICE GENERAL	
ÍNDICE DE FIGURAS	
ÍNDICE DE TABLAS	
RESUMEN	
ABSTRACT	
CAPÍTULO I INTRODUCCIÓN.....	16
1.1 Planteamiento del Problema y Justificación.....	18
1.2 Objetivos.....	19
1.2.1 Objetivo General.....	19
1.2.2 Objetivos Específicos.....	19
CAPÍTULO II MARCO TEÓRICO.....	20
2.1 Estado del Arte.....	20
2.2 LTE-A Procedimiento de acceso aleatorio.....	21
2.3 Procedimiento de acceso aleatorio basado en la contención.....	22
2.4 Procedimiento de <i>BACKOFF</i> .....	24
2.5 Aprendizaje automático en LTE-A.....	25
2.5.1 Aprendizaje reforzado.....	25
2.5.2 Algoritmo de aprendizaje por refuerzo QL.....	27
2.6 Métricas de rendimiento.....	28
2.6.1 Probabilidad de acceso exitoso ( $P_s$ ).....	28
2.6.2 Número de preámbulos transmitidos ( $k$ ).....	29
2.6.3 Retardo en el acceso ( $D$ ).....	30
CAPÍTULO III METODOLOGÍA.....	31



3.1	Tipo de Investigación.....	31
3.1.1	Investigación Exploratoria.....	31
3.2	Diseño de Investigación.....	31
3.2.1	Método Científico.....	31
3.3	Técnicas de Investigación.....	31
3.3.1	Recolección de Datos.....	31
3.3.2	Instrumentos de la Investigación.....	32
3.4	Población de estudio y tamaño de Muestra.....	32
3.4.1	Población.....	32
3.4.2	Muestra.....	32
3.5	Operacionalización de las variables.....	32
3.6	Métodos de Análisis y Procesamiento de Datos.....	33
3.6.1	Modelo del Sistema.....	35
3.6.2	Parámetros de configuración del algoritmo de QL.....	36
3.6.3	Procedimiento de la implementación del algoritmo de QL.....	38
3.6.4	Función de la asignación de Recompensas a las posibles acciones tomadas.....	40
CAPÍTULO IV RESULTADOS Y DISCUSIÓN.....		41
4.1	Análisis de los resultados de las métricas de rendimiento.....	41
4.1.1	Resultados de la probabilidad de acceso exitoso, número promedio de preámbulos transmitidos y retardo en el acceso con un percentil 95 en un entorno tanto de BI estático como con BI dinámico ...	41
4.1.1.1	Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 20 ms y un valor de BI aplicado el algoritmo de QL.....	44
4.1.1.2	Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 160 ms y un valor de BI aplicado el algoritmo de QL.....	45
4.1.1.3	Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 320 ms y un valor de BI aplicado el algoritmo de QL.....	46
4.1.1.4	Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 480 ms y un valor de BI aplicado el algoritmo de QL.....	47
4.1.1.5	Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 960 ms y un valor de BI aplicado el algoritmo de QL.....	48
4.2	Resultados de los mejores valores de BI para diferentes cargas de tráfico.....	50
4.3	Análisis estadístico de las métricas de rendimiento.....	52
4.3.1	Prueba de hipótesis para la probabilidad de acceso exitoso (Ps). .....	53

4.3.1.1	Diagrama de cajas .....	53
4.3.1.2	Análisis Descriptivo.....	53
4.3.1.3	Pruebas de normalidad.....	54
4.3.1.4	Prueba de Mann-Whitney U .....	55
4.3.2	Prueba de hipótesis para el Número promedio de preámbulos transmitidos (k).....	56
4.3.2.1	Diagrama de cajas .....	57
4.3.2.2	Análisis Descriptivo.....	57
4.3.2.3	Pruebas de normalidad.....	58
4.3.2.4	Prueba de Mann-Whitney U .....	59
4.3.3	Prueba de Hipótesis para el Retardo en el acceso (D). .....	60
4.3.3.1	Diagrama de cajas .....	60
4.3.3.2	Análisis Descriptivo.....	61
4.3.3.3	Pruebas de normalidad.....	62
4.3.3.4	Prueba de Mann-Whitney U .....	62
4.4	Discusión de los Resultados.....	64
CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES .....		66
5.1	Conclusiones .....	66
5.2	Recomendaciones .....	67
BIBLIOGRAFÍA.....		68
ANEXOS.....		71

## ÍNDICE DE FIGURAS

<b>Figura 1</b> Procedimiento de acceso aleatorio basado en contención de LTE-A.....	23
<b>Figura 2</b> Procedimiento de backoff realizado por los equipos de usuario .....	24
<b>Figura 3</b> Diagrama del aprendizaje por refuerzo. ....	26
<b>Figura 4</b> Diagrama de Bloques del Procedimiento y análisis de resultados .....	34
<b>Figura 5</b> Modelo del sistema .....	35
<b>Figura 6</b> Diferentes configuraciones del índice de PRACH – Estructura de trama tipo 1.....	36
<b>Figura 7</b> Diagrama de Flujo de la implementación del algoritmo de QL .....	38
<b>Figura 8</b> Probabilidad de acceso exitoso ( $P_s$ ) varía con diferentes cargas de tráfico; (a) Utilizando distintos valores de BI estáticos y (b) Al obtener valores óptimos de BI mediante la implementación del algoritmo QL.....	41
<b>Figura 9</b> Número promedio de retrasmisiones ( $k$ ) cambia en función de las diversas cargas de tráfico; (a) Utilizando distintos valores de BI estáticos y (b) Al obtener valores óptimos de BI con la implementación del algoritmo QL. ....	42
<b>Figura 10</b> Retardo en el Acceso con percentil 95 ( $D_{95}$ ) varía en función de las diversas cargas de tráfico; (a) Utilizando diferentes valores de BI estáticos y (b) Al obtener valores óptimos de BI mediante la implementación de algoritmo QL.....	43
<b>Figura 11</b> Implementación del algoritmo de QL, obtención de los mejores BI para una baja carga de tráfico. ....	50
<b>Figura 12</b> Implementación del algoritmo de QL, obtención de los mejores BI para una carga de tráfico media. ....	51
<b>Figura 13</b> Implementación del algoritmo de QL, obtención de los mejores BI para una alta carga de tráfico. ....	52
<b>Figura 14</b> Diagrama de Cajas de la probabilidad de acceso exitoso en función de las dos configuraciones. .	53
<b>Figura 15</b> Diagrama de Cajas del número de retrasmisiones en función de las dos configuraciones.....	57
<b>Figura 16</b> Diagrama de Cajas del retardo en el acceso en función de dos configuraciones.....	60

## ÍNDICE DE TABLAS

<b>Tabla 1</b> Posibles valores que puede tomar el Indicador de Backoff (BI).....	27
<b>Tabla 2</b> Configuración para la obtención de las posibles recompensas .....	28
<b>Tabla 3</b> Operacionalización de variables .....	33
<b>Tabla 4</b> Parámetros para la configuración del RACH .....	35
<b>Tabla 5</b> Parámetros de configuración para el desarrollo del algoritmo QL .....	37
<b>Tabla 6</b> Asignación de recompensas a diferentes acciones en diversas condiciones .....	40
<b>Tabla 7</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 20 y valor de BI óptimo aplicando el algoritmo QL .....	44
<b>Tabla 8</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 160 y valor de BI óptimo aplicando el algoritmo QL .....	45
<b>Tabla 9</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 320 y valor de BI óptimo aplicando el algoritmo QL .....	46
<b>Tabla 10</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 480 y valor de BI óptimo aplicando el algoritmo QL .....	47
<b>Tabla 11</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 960 y valor de BI óptimo aplicando el algoritmo QL .....	48
<b>Tabla 12</b> Resumen del procesamiento de los casos .....	52
<b>Tabla 13</b> Resultados estadísticos descriptivos de la métrica Ps.....	53
<b>Tabla 14</b> Resultados de las pruebas de Normalidad de la métrica Ps .....	55
<b>Tabla 15</b> Resultado comparativo de la prueba U para cada tipo de algoritmo.....	56
<b>Tabla 16</b> Resultado comparativo de la prueba de Mann-Whitney .....	56
<b>Tabla 17</b> Resultados estadísticos descriptivos de la métrica k.....	57
<b>Tabla 18</b> Resultados de las pruebas de Normalidad de la métrica k.....	58
<b>Tabla 19</b> Resultado comparativo de la prueba U para cada tipo de algoritmo.....	59
<b>Tabla 20</b> Resultado comparativo de la prueba de Mann-Whitney .....	59
<b>Tabla 21</b> Resultados estadísticos descriptivos de la métrica D.....	61
<b>Tabla 22</b> Resultados de las pruebas de Normalidad de la métrica D .....	62
<b>Tabla 23</b> Resultado comparativo de la prueba U para cada tipo de algoritmo.....	63
<b>Tabla 24</b> Resultado comparativo de la prueba de Mann-Whitney .....	63
<b>Tabla 25</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 30 y valor de BI óptimo aplicando el algoritmo QL .....	71
<b>Tabla 26</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 40 y valor de BI óptimo aplicando el algoritmo QL .....	71
<b>Tabla 27</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 60 y valor de BI óptimo aplicando el algoritmo QL .....	72
<b>Tabla 28</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 80 y valor de BI óptimo aplicando el algoritmo QL .....	72

<b>Tabla 29</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 120 y valor de BI óptimo aplicando el algoritmo QL .....	73
<b>Tabla 30</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 240 y valor de BI óptimo aplicando el algoritmo QL .....	73
<b>Tabla 31</b> Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 750 y valor de BI óptimo aplicando el algoritmo QL .....	74
<b>Tabla 32</b> Asignación de recompensas para un valor de BI de 960 en diversas condiciones.....	74
<b>Tabla 33</b> Asignación de recompensas para un valor de BI de $\geq 240$ en diversas condiciones .....	75
<b>Tabla 34</b> Asignación de recompensas para un valor de BI de $\geq 80$ en diversas condiciones .....	76
<b>Tabla 35</b> Asignación de recompensas para un valor de BI de $\geq 20$ en diversas condiciones .....	76

## RESUMEN

En la investigación, se aborda el diseño de un algoritmo de *backoff* adaptativo basado en el aprendizaje por refuerzo, con el propósito de aumentar la probabilidad de acceso exitoso en una red celular que maneja comunicaciones masivas. El enfoque principal de este trabajo es estudiar y modelar tanto el canal de acceso aleatorio como el procedimiento de *backoff* en condiciones de comunicación masiva generada por dispositivos IoT. Este algoritmo ajusta el tiempo de *backoff* en función del indicador de *backoff*, el cual se determina a partir de la carga del canal de acceso aleatorio. La evaluación del rendimiento de la red se llevó a cabo al considerar la probabilidad de acceso exitoso en diferentes condiciones de tráfico masivo. Para llevar a cabo este análisis, se empleó un modelo de simulación implementado en el software de MatLab. Los resultados obtenidos a través de las configuraciones adecuadas de los parámetros del RACH, demostraron que se puede alcanzar una probabilidad de acceso exitoso mayor al 90%. Este logro se acompaña de una considerable reducción en las retransmisiones y el retardo de acceso. La eficiencia de estos resultados se debe a que el algoritmo basado en aprendizaje por refuerzo tiene la capacidad de aprender de sus experiencias y de adaptarse de manera eficiente a las cambiantes condiciones del canal de comunicación.

**Palabras claves:** Indicador de *Backoff*, aprendizaje reforzado, probabilidad de acceso exitoso, redes celulares.

## ABSTRACT

The research focuses on designing an adaptive Backoff algorithm based on reinforcement learning to increase the probability of successful access in a cellular network handling massive communications. The primary focus of this work is to study and model both the random access channel and the Backoff procedure under conditions of massive communication generated by IoT devices. This algorithm adjusts the Backoff time based on the Backoff indicator, which is determined from the load of the random access channel. The network's performance evaluation was carried out by considering the probability of successful access under various conditions of massive traffic. A simulation model was implemented using Matlab software to conduct this analysis. The results obtained with the appropriate RACH parameter settings demonstrated that a satisfactory access probability greater than 90% can be achieved. A significant reduction in retransmissions and access delays accompanies this achievement. The efficiency of these results is attributed to the reinforcement learning-based algorithm's ability to learn from its experiences and adapt efficiently to changing communication channel conditions.

**Keywords:** Backoff Indicator, Reinforcement Learning, Probability of Successful Access, Cellular Networks.



Reviewed by:  
Mg. Dario Javier Cutiopala Leon  
**ENGLISH PROFESSOR**  
c.c. 0604581066

## CAPÍTULO I INTRODUCCIÓN

Hoy en día cursamos una nueva era tecnológica en la que todo está conectado. En particular, millones de dispositivos heterogéneos de Internet de las cosas (IoT) utilizan conexiones celulares para proveer distintos servicios e información [1]. Las redes celulares son una excelente opción para proporcionar dicha hiperconectividad por la infraestructura ampliamente desplegada dando como resultado numerosas ventajas, como los reducidos costes de despliegue, la seguridad, la gestión y la calidad de servicio (QoS) [2].

La adaptabilidad a los cambios constantes y a los operadores de redes móviles es un desafío continuo, lo que a menudo significa nuevas inversiones en infraestructura [2]. Al mismo tiempo, la alta tasa de éxito de la tecnología móvil y la capacidad de recopilar fácilmente grandes cantidades de información sobre el comportamiento de los usuarios permite una mejor comprensión de los requisitos de la red y por lo tanto brinda nuevas soluciones para optimizar los recursos [3].

La comunicación de tipo máquina (MTC) constituye un servicio dominante para aplicaciones emergentes de IoT. La comunicación masiva tipo máquina (mMTC) es uno de los principales tipos de servicio en 5G [4]. Los servicios de mMTC se caracterizan por un gran número de dispositivos conectados que suelen transmitir un volumen relativamente bajo de datos tolerantes al retardo. En este sentido cuando un equipo de usuario quiere acceder a la red, debe obtener primero algunos parámetros de configuración de la red, como los recursos de tiempo/frecuencia predefinidos donde se permiten los intentos de acceso aleatorio [2]. Entonces cada instancia de estos recursos, en la que se puede realizar un intento de acceso, se denomina oportunidad de acceso aleatorio (RAO) y se realiza utilizando el canal de acceso aleatorio (RACH) [5].

El procedimiento de *backoff* en redes celulares es un proceso que implica tener en cuenta varios factores como la congestión de la red, la tasa de transferencia de datos y la probabilidad de acceso exitoso [5]. El procedimiento de *backoff* se utiliza para evitar colisiones en la transmisión de datos en redes celulares y se basa en la idea de que los dispositivos deben esperar un tiempo aleatorio antes de intentar transmitir datos otra vez después de una colisión, dicho procedimiento implica la selección de un tiempo de espera



adecuado para cada dispositivo, lo que puede ser un desafío debido a la variabilidad en la congestión de la red y la tasa de transferencia de datos.

El aprendizaje por refuerzo es una técnica de aprendizaje automático que se refiere a la capacidad de asignar acciones a situaciones para maximizar la recompensa. Esto se logra interactuando con el entorno y descubriendo que acciones producen la mayor recompensa al probarlas. La estación base almacena una función de valor  $Q(s, a)$ , que mide la recompensa esperada de realizar una acción específica  $a$  en un estado  $s$  y continuar realizando la mejor acción indefinidamente [6].

El problema en las redes celulares es la gestión del número masivo de intentos de conexión por un gran número de dispositivos (UEs), provocando que el RACH se sobrecargue, lo que eventualmente conducirá a posibles sobrecargas. Uno de los mecanismos intrínsecos para resolver la congestión de la red móvil es el procedimiento de *backoff* [7]. Por tanto, es importante investigar el aumento de probabilidad de acceso exitoso para asegurarnos menos congestión de red, ya que hay pocos estudios que abordan este procedimiento a profundidad. En escenarios en los que la red es sometida a una alta carga de tráfico, conviene aumentar el tiempo de *backoff* para dispersar las transmisiones simultáneas. Mientras que en escenarios en los que la red es sometida a bajas cargas de tráfico, conviene disminuir el tiempo de *backoff* para no generar retardos innecesarios [7].

Organismos de estandarización como el Proyecto Asociación de Tercera Generación (3GPP) no proporciona algún algoritmo destinado para el ajuste adecuado de este tiempo de *backoff* en función de la carga de la red [8]. Para lograr esto, el enfoque propuesto consiste en diseñar un algoritmo basado en aprendizaje reforzado que permita ajustar dinámicamente y de forma autónoma el tiempo de *backoff*. Esto con la finalidad de que la red pueda reaccionar rápidamente a los cambios de tráfico utilizando la información local disponible en la estación base. Se pretende que la red pueda superar situaciones de sobrecarga y no generar retardos innecesarios en situaciones de carga ligera, obteniendo así un mejor rendimiento del canal de acceso.

## 1.1 Planteamiento del Problema y Justificación

El tráfico de mMTC implica un gran número de dispositivos, que llegan a generar masivas solicitudes de acceso y de manera simultánea en un intervalo corto de tiempo. En el primer contacto con la red, los dispositivos deben realizar un procedimiento de acceso aleatorio basado en el intercambio de cuatro mensajes [6]. Para esto se utiliza el RACH en el enlace de subida que tiene un limitado número de recursos y por ende se generan colisiones en los intentos de acceso, reduciendo significativamente su rendimiento [5].

Existe un algoritmo intrínseco en la red celular para gestionar las retransmisiones que se conoce como el procedimiento de *backoff* el cual se encarga de distribuir las peticiones de los usuarios que no han tenido éxito en el acceso durante un determinado tiempo. La duración de este tiempo lo determina la estación base y lo comunica a los dispositivos en un mensaje que se conoce como bloque de información del sistema [9]. Al ser un valor estático, existirán ocasiones en las que este procedimiento sea efectivo (escenarios de alta carga) y otras en las que no (escenarios de baja carga), perjudicando el rendimiento del canal de comunicación [5].

Por este motivo se diseñará un algoritmo adaptativo para ajustar el tiempo de *backoff* y disminuir la sobrecarga en el RACH mejorando la eficiencia y el rendimiento de la red, por medio de técnicas de aprendizaje reforzado. El algoritmo se basará en la idea de que un agente aprenda a tomar decisiones óptimas a través de la experiencia, utilizando una función de valor para determinar la acción a elegir en un estado dado [10]. Entonces se determinará la mejor estrategia para elegir el tiempo de *backoff* en función del estado actual de la red y las acciones previas tomadas, en este caso, el estado actual de la red se puede definir en términos de la congestión de tráfico y la tasa de transferencia de datos.

## 1.2 Objetivos

### 1.2.1 Objetivo General

- Diseñar un algoritmo adaptativo de *backoff* basado en aprendizaje reforzado para incrementar la probabilidad de acceso exitoso en una red celular con comunicaciones masivas.

### 1.2.2 Objetivos Específicos

- Estudiar y modelar el canal acceso aleatorio junto con el procedimiento de *backoff* bajo condiciones de comunicación masiva generada por dispositivos IoT.
- Diseñar un algoritmo de *Backoff* adaptativo basado en aprendizaje reforzado.
- Establecer la probabilidad de acceso exitoso como métrica de análisis.
- Evaluar el rendimiento de la red en función de la probabilidad de acceso exitoso bajo diferentes condiciones de tráfico masivo mediante un modelo de simulación y contrastar los resultados con una configuración estándar del procedimiento de *Backoff*.

## CAPÍTULO II MARCO TEÓRICO

### 2.1 Estado del Arte

De la búsqueda y revisión de varias fuentes bibliográficas y artículos de investigación relacionados con el tema, se explican los métodos y algoritmos utilizados.

Los autores del artículo [11], centran su estudio en la resolución de las colisiones en el RACH en escenarios de acceso masivo para las comunicaciones máquina a máquina (M2M) celulares. Se proponen un esquema para resolver las colisiones mediante el procedimiento de *backoff* que ajusta de forma dinámica el Indicador de *Backoff* (BI) en relación con el número de dispositivos atrasados y los recursos disponibles. En el enfoque propuesto se integran tres esquemas de acceso aleatorio: acceso aleatorio estándar (SRA), restricción estática de clase de acceso (ACB) y restricción dinámica de clase de acceso (DBA). Además de presentar un análisis de la probabilidad de acceso exitoso basado en el procedimiento de *backoff* dinámico, se determina un valor óptimo de BI que logra la mayor probabilidad de acceso exitoso para los tres diferentes esquemas.

El estudio en [4], propone un nuevo esquema de canal de acceso aleatorio basado en el método de aprendizaje Q para reducir los problemas de la congestión. De esta manera, se divide de forma adaptativa los preámbulos disponibles entre los dispositivos M2M y entre personas (H2H), de tal manera que se maximice el número de dispositivos M2M activos y se proporcione un servicio aceptable para los dispositivos H2H. Los resultados indican que el método propuesto proporciona una alta probabilidad de éxito del canal de acceso aleatorio para dispositivos H2H, incluso con un gran número de dispositivos M2M.

En [5], se desarrolló un modelo de simulación del procedimiento de acceso aleatorio basado en la contención, el cual utiliza métricas de rendimiento como la probabilidad de acceso exitoso, el retraso del acceso y el número promedio de preámbulos transmitidos para evaluar y caracterizar el impacto de ajustar el valor del Indicador de *Backoff* (BI) en distintos escenarios de tráfico masivo. Los resultados muestran un rango de valores óptimos para el BI que se identificaron mediante el cumplimiento de diferentes condiciones de fiabilidad para cada escenario. Al ajustar adecuadamente los parámetros del RACH, se puede lograr

una alta probabilidad de acceso exitoso con un aumento aceptable en el retardo de acceso, lo que es ideal para aplicaciones tolerantes al retardo en escenarios con tráfico masivo.

En el artículo los autores [6]. En su estudio, realizaron un análisis del rendimiento del canal de acceso aleatorio, buscando mejorar el rendimiento de la evolución a largo plazo avanzada (LTE-A) en escenarios de comunicación de máquina a máquina (M2M) masivos. Se proponen un método de aprendizaje por refuerzo (RL) para ajustar dinámicamente y de forma autónoma la tasa de restricción de la clase de acceso (ACB), de manera que pueda reaccionar rápidamente a los cambios de tráfico. Entre sus resultados se refleja el uso del ACB y la configuración correcta de sus parámetros llevan a un aumento en la probabilidad de éxito del acceso de los dispositivos M2M en los escenarios de prueba extrema.

En el artículo [9], los autores se enfocan en la selección de la estación base en la comunicación M2M mediante un algoritmo de aprendizaje Q en redes LTE-A, que proporciona una forma eficiente de acceso múltiple en la red y minimiza la sobrecarga de la misma. Proponen un algoritmo de aprendizaje no supervisado basado en aprendizaje Q, mediante un esquema de selección de estaciones base en los dispositivos MTC, que se adaptan continuamente a los cambios en el tráfico de la red y decide que estación base se debe seleccionar en función a los parámetros de calidad de servicio. Los resultados muestran que además de mejorar el rendimiento de la comunicación M2M, el algoritmo también puede ayudar a los dispositivos MTC a obtener un mejor rendimiento de la red.

## **2.2 LTE-A Procedimiento de acceso aleatorio**

En LTE-A los dispositivos acceden al RACH cuando inicialmente solicitan el acceso a la red o al cambio de celda. RACH es un canal de enlace ascendente utilizado para transmitir información de control desde dispositivos móviles a la red [12].

Para acceder al RACH, los dispositivos pasan por un proceso conocido como acceso aleatorio. Se define dos modos de acceso aleatorio; libre de contención diseñado para aplicaciones con limitaciones de tiempo y no puede manejar ninguna latencia, tiene 10 preámbulos de los 64 preámbulos LTE reservados para él, mientras que el modo basado en la contención dispone de 54 preámbulos para su uso [5]. El procedimiento consta de 4 pasos:

1. Transmisión de preámbulos de acceso aleatorio;
2. Recepción de respuesta de acceso aleatorio;
3. Solicitud de conexión;
4. Resolución de contención.

Estos 54 preámbulos están disponibles para que los dispositivos M2M y H2H se comuniquen con la estación base. Se hace uso de un ancho de banda de 1.4MHz, donde la ranura de acceso aleatorio (RA) consta de 6 bloques de recurso (RB) de 180KHz cada uno [6].

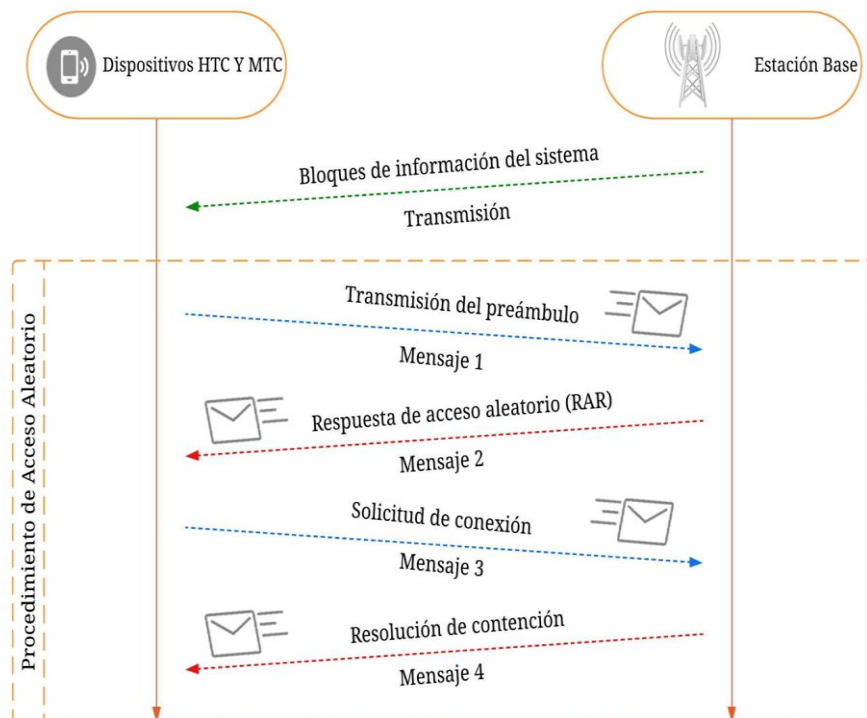
### 2.3 Procedimiento de acceso aleatorio basado en la contención

Las redes celulares se suelen conectar mediante el proceso de acceso aleatorio basado en la contención. A través del bloque de información principal (MIB) y del bloque de información del sistema (SIB), la estación base emite información periódicamente. Una vez que los equipos del usuario han adquirido los ajustes de configuración fundamentales, pueden iniciar los cuatro intercambios de mensajes. Este proceso se efectúa mediante el uso de RACH, como se muestra en la Figura 1.

- **Transmisión del preámbulo:** En el mensaje 1, los equipos eligen aleatoriamente uno de los preámbulos. Dado que los preámbulos son ortogonales entre sí, muchos equipos de usuario pueden acceder a la estación base (EB) en el misma RAO y utilizar diferentes preámbulos. La BS puede estimar el tiempo de transmisión del terminal y decodificar el preámbulo transmitido por un único equipo de usuario [5]. Si más de un equipo de usuario selecciona el mismo preámbulo, se produce una colisión y el acceso se considera invalido.
- **Respuesta de acceso aleatorio (RAR):** Tan pronto como la BS decodifica con éxito cada preámbulo, calcula un identificador y envía el mensaje 2 a través del canal de control de enlace descendente (PDCCH). Este mensaje contiene información sobre el indicador de *backoff* (BI) que se utiliza en el proceso de espera, la identificación del preámbulo detectado (ID), la alineación temporal (TA), los recursos del enlace ascendente y la asignación de un identificador temporal. Los equipos esperan una ventana temporal (WRAR) para recibir una concesión de enlace ascendente por parte de la estación base (BS) a través del mensaje 2 [13]. Para maximizar el rendimiento

de la red, sólo se envía una respuesta por cada subtrama. En consecuencia, hay un límite al número de concesiones de enlaces ascendentes; sólo los equipos que reciben una concesión pueden enviar el mensaje 3.

- **Solicitud de conexión:** Tras recibir el mensaje 2, los equipos de usuario ajustan el tiempo de transmisión de enlace ascendente en base a la alineación temporal recibida y utilizan los recursos del canal físico compartido de enlace ascendente (PUSCH) para enviar un mensaje programado de solicitud de conexión a la BS [11]. Este mensaje incluye el número de identificación del UE y el motivo de la solicitud de acceso.
- **Resolución de contención:** La BS envía un mensaje como respuesta al mensaje anterior y utiliza un proceso de repetición automática híbrida (HARQ) para enviar el mensaje 4 de retorno a los dispositivos. Si un equipo de usuario no recibe este mensaje, debe intentar el RAP de nuevo, utilizando el procedimiento de *backoff* y comenzando desde el primer paso en la siguiente RAO [6]. La cantidad de intentos de transmisión tiene un límite máximo; después de eso, se descarta la solicitud del dispositivo.



**Figura 1** Procedimiento de acceso aleatorio basado en contención de LTE-A.

Fuente: [5]

## 2.4 Procedimiento de *BACKOFF*

El procedimiento de *backoff* es una técnica utilizada para gestionar las colisiones en el RACH. Se aplica a todos los dispositivos que no consiguen conectarse con éxito a la red en su primer intento [5]. Aguardar otro periodo de tiempo antes de realizar un nuevo intento de transmisión de preámbulo en el próximo RAO disponible es necesario cuando un dispositivo falla en su acceso aleatorio (RA). Este periodo de espera se denomina tiempo de *backoff*  $T_{BO}$  (ms) y se elige de manera uniforme y aleatoria en función del indicador de *backoff* (BI) a través de la ecuación 1, como se muestra en la Figura 2:

$$T_{BO} = \mathcal{U}(0, BI) \quad (1)$$

donde  $\mathcal{U}()$  denota una distribución uniforme, el valor de BI tiene un impacto significativo en el rendimiento de la RACH. Los dispositivos pueden intentar un nuevo acceso después de un corto periodo de tiempo si el valor de BI es pequeño, lo que incrementa la probabilidad de colisiones durante las ráfagas de llegada. Por otro lado, un valor alto de BI puede aumentar la probabilidad de éxito durante el acceso a la red a costa de un mayor retraso en el acceso. Los estándares 3GPP estipulan que la estación base selecciona aleatoriamente el valor de BI, que puede oscilar entre 0 y 960ms [11]. Todos los dispositivos que enviaron el preámbulo en el RAO previo leían el valor de BI, que se incluye en el mensaje 2. Para el dispositivo que no obtiene confirmación por medio del RAR, se considera un intento fallido [7]. A pesar de los retos que presenta, el uso del *backoff* ha demostrado ser una herramienta útil para gestionar las colisiones y mejorar el rendimiento de los dispositivos conectados.

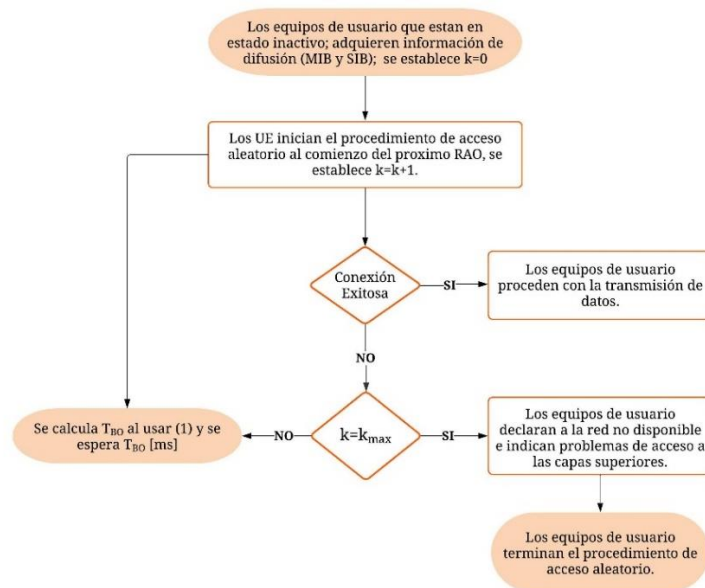


Figura 2 Procedimiento de *backoff* realizado por los equipos de usuario



## 2.5 Aprendizaje automático en LTE-A

La adaptación a nuevos entornos y el aprendizaje de experiencias previas es posible gracias al aprendizaje automático en numerosos entornos, incluidos juegos, optimización, robótica y control de procesos. El aprendizaje reforzado es una técnica de aprendizaje automático que se puede utilizar en el tiempo de *backoff* para mejorar la eficiencia y reducir la congestión en redes de comunicación.

### 2.5.1 Aprendizaje reforzado

El aprendizaje por refuerzo permite que un dispositivo MTC interactúe con el entorno y adquiera conocimientos a partir de experiencias pasadas, incluso si no cuenta con un conjunto de datos de entrenamiento, se basa en como un agente adquiere habilidades para tomar decisiones adecuadas en entornos dinámicos e inciertos [14].

Esta rama de aprendizaje automático se centra en aprender a partir de la retroalimentación que recibe un agente en forma de recompensa o castigo, el cual aprende a elegir la acción de mayor recompensa acumulada en todo ese tiempo, formulando una política que es un mapeo entre estados y acciones [15].

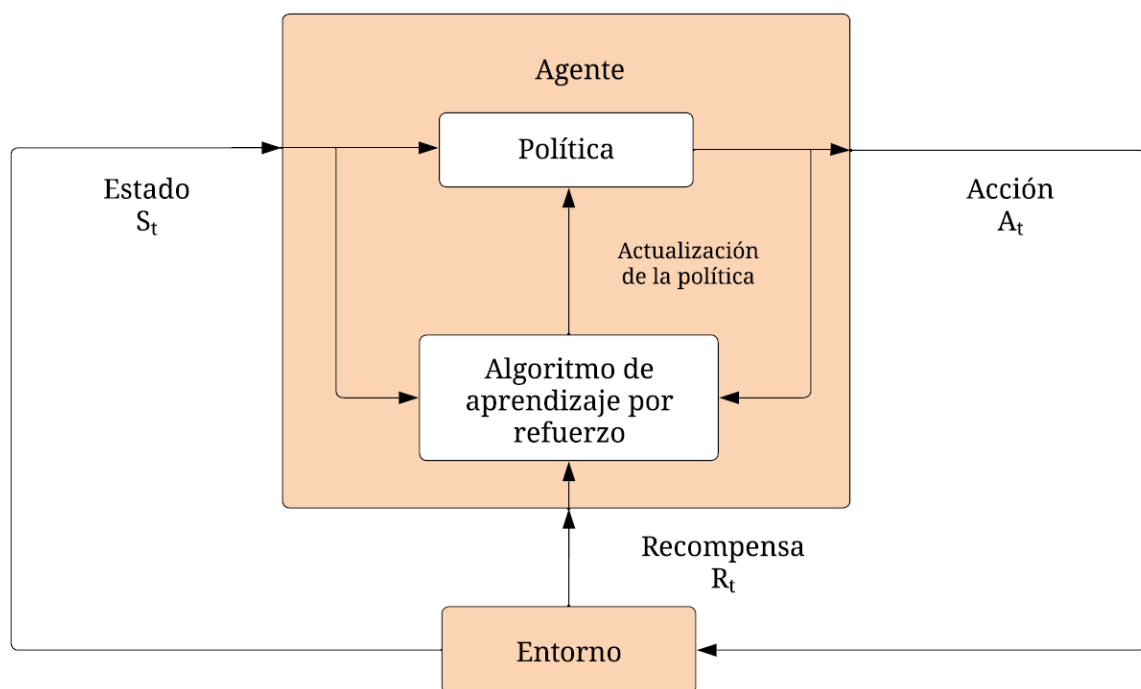
El proceso de aprendizaje por refuerzo se describe cómo [16]:

- **Agente:** Representa el modelo que se busca entrenar para que adquiera la capacidad de tomar decisiones de manera autónoma.
- **Entorno:** El agente interactúa y se mueve en el entorno este incluye las limitaciones y reglas que se puedan aplicar en cada situación.
- **Estado:** En un momento específico, el entorno se encuentra en un estado determinado, y el agente percibe este estado como información relevante para tomar una decisión.
- **Acción:** Según el estado actual el agente elige una acción. La acción puede tener un impacto en el entorno y provoca un cambio de estado.

El ciclo de interacción, recompensas y actualización del modelo es esencial para un aprendizaje exitoso.

- **Recompensa:** El entorno envía una señal al agente como recompensa cada vez que realiza una acción. El tipo de señal enviada informa al agente sobre la eficacia de su acción.
- **Actualización del modelo:** El agente actualiza su modelo interno del entorno y aprende asociar estados y acciones como las recompensas esperadas. Este proceso de aprendizaje se basa en técnicas como método de valoración de acciones como el algoritmo de aprendizaje Q o aproximaciones de funciones de valor como las redes neuronales.
- **Ciclo de interacción:** Interactuar con su entorno es una parte crucial del proceso de aprendizaje del agente. Al seleccionar acciones y actualizar su modelo, finalmente forma una política óptima que conduce a la acumulación de recompensas a lo largo del tiempo.

En general, el entorno percibido por un dispositivo IoT se describe mediante un proceso de decisión de Markov (MDP), como se representa en la Figura 3. [17]. En este modelo MDP, en cada tiempo, el dispositivo puede cambiar su estado actual  $S_t$  al siguiente estado a traves de una acci3n  $A_t$  basada en una funci3n de probabilidad de transici3n y durante esta transici3n, el dispositivo recibe una recompensa  $R_t$  instantanea.



*Figura 3 Diagrama del aprendizaje por refuerzo.*

## 2.5.2 Algoritmo de aprendizaje por refuerzo QL

Existen diferentes tipos de algoritmos de aprendizaje por refuerzo, el aprendizaje Q (QL) es un ejemplo de algoritmo de aprendizaje por refuerzo donde un agente interactúa con el entorno para lograr un objetivo específico [18], y encontrar la política óptima que seleccione el mejor procedimiento para interacciones futuras, lo que resulta en altos rendimientos a largo plazo [19]. Usando este método la estación base almacena una función de valor  $Q(s, a)$  donde  $s$  es el conjunto de todos los estados posibles y  $a$  mide la recompensa esperada de tomar una acción, continuando de manera indefinida tomando acciones de forma óptima.

*Tabla 1 Posibles valores que puede tomar el Indicador de Backoff (BI)*

Acción	Valores del Indicador de Backoff (BI)	Acción	Valores del Indicador de Backoff (BI)
1	20	7	160
2	30	8	240
3	40	9	320
4	60	10	480
5	80	11	750
6	120	12	960

Los parámetros de  $Q(s, a)$  tal y como se establecen en la Tabla 1 el conjunto de estados se determina por los valores de BI proporcionados por la estación base a través de un mensaje SIB2, el cual se envía por difusión cada  $T_{SIB2}$  este mensaje se envía cada 80ms y se estandariza en un valor  $T_{SIB2} = 16RAO$  [20], mientras que las acciones se refieren al conjunto de posibles valores que puede asumir el BI, que estará representados por  $\mathcal{A} = \{1,2,3, \dots, 12\}$ . Por lo que las acciones están diseñadas para aumentar o disminuir el valor de BI según el estado del RACH [21].

El aprendizaje por refuerzo se define mediante la ecuación (2):

$$Q(s, a) = Q(s, a) + \alpha \left[ \mathcal{R} + \gamma \max_{a' \in \mathcal{A}} [Q(s', a')] - Q(s, a) \right]; \quad (2)$$

Donde  $\alpha$  representa la tasa de aprendizaje que controla la cantidad de actualizaciones,  $s'$  denota el estado siguiente,  $a'$  se refiere a la acción tomada en un estado  $s'$  [22]. El parámetro

$\mathcal{R}$  corresponde a la recompensa asociada a tomar la acción  $a$  en el estado  $s$  definida por  $\mathcal{R} = (NpsMcur, NpsCVcur, DeltaNpscur, BI)$ , donde  $NpsMcur = \frac{1}{T_{SIB2}} (\sum_{k=1}^{T_{SIB2}} Nps)$  es el promedio del número de preámbulos transmitidos por RAO que la BS detecto en un  $T_{SIB2}$  completo,  $NpsCVcur = \left[ \frac{1}{T_{SIB2}^{-1}} (\sum_{k=1}^{T_{SIB2}} |N_{psk} - NpsMcur|^2) \right]^{\frac{1}{2}} / NpsMcur$ , es el coeficiente de variación de  $Nps$  para el mismo periodo entre 0 y 1, un valor bajo indica que la transmisión de paquetes es relativamente uniforme.  $DeltaNpscur$  es la diferencia entre el número promedio de preámbulos transmitidos en el período actual y el anterior el cual puede tomar cualquier valor real, como se muestra en la Tabla 2. En este sentido si  $\mathcal{R}$  indica que el  $NpsMcur$  actual es baja y el coeficiente de variación  $NpsCVcur$  actual es alto el agente puede aumentar el valor de BI, caso contrario si  $\mathcal{R}$  indica que el  $NpsMcur$  actual es alto y el coeficiente de variación  $NpsCVcur$  actual es bajo el agente puede disminuir el valor de BI.

*Tabla 2 Configuración para la obtención de las posibles recompensas*

	<b>NpsMcur</b>	<b>NpsCVcur</b>	<b>DeltaNpscur</b>
Configuración 1	3	0,4	1 o 3
Configuración 2	7	0,4	1 o 3
Configuración 3	10	0,2	1 o 3

El factor de descuento está representado por  $\gamma$  el cual controla el impacto de las futuras recompensas. La presencia del factor  $\max_{a' \in \mathcal{A}} [Q(s', a')]$  en el aprendizaje Q presenta dos resultados [21], [6]. El primero, las recompensas a futuro se basan en una estimación que podría estar inclinada positivamente, lo que puede resultar en un rendimiento subóptimo. En segundo lugar, las recompensas futuras no están relacionadas con la política actual, lo que se definen por fuera de dicha política [23], esto posibilita una implementación en tiempo real. En conjunto  $\mathcal{R} + \gamma \max_{a' \in \mathcal{A}} [Q(s', a')]$  se refiere al objetivo buscado. Mediante el uso de QL se busca cambiar el valor de BI antes de que el mensaje SIB2 sea transmitido.

## 2.6 Métricas de rendimiento

### 2.6.1 Probabilidad de acceso exitoso (Ps)

La variable Ps, representa la probabilidad de éxito del acceso, es decir la probabilidad de que un equipo de usuario (UE), ya sea M2M o H2H, tenga éxito en su proceso de acceso a la red. Esta métrica mide la eficiencia del sistema en términos de cuantas UEs lograron acceder

exitosamente en relación con el número total de UEs en la simulación. En la ecuación (3) se define  $P_s$  como:

$$P_s = \frac{\text{totalSuccessfulUEs}}{\text{totalUEs}} \quad (3)$$

Donde *totalSuccessfulUEs* representa el número total de dispositivos MTC con acceso exitoso y *totalUEs* representa el número total de dispositivos MTC que ingresaron a la estación base.

### 2.6.2 Número de preámbulos transmitidos (k)

Se refiere a la cantidad de transmisiones que un equipo de usuario necesita realizar utilizando preámbulos antes de lograr con éxito el procedimiento del RAP después de un intento inicial fallido. Para evaluar este indicador, se emplea una función de distribución acumulada (CDF), como se describe en la ecuación (4), para cada una de las transmisiones de preámbulos realizadas hasta alcanzar el número máximo de intentos [24].

$$p(k_n) = \frac{\sum M_A(k_n)}{\text{totalSuccessfulUEs}} \quad (4)$$

Donde  $p(k_n)$  representa la probabilidad acumulada de “n” transmisiones de preámbulo, y  $\sum M_A(k_n)$  se define como el conjunto de dispositivos lograron el acceso a la red en “n” transmisiones de preámbulo.

En el contexto de la simulación, la variable  $k$  se convierte en un vector que representa la distribución del número de preámbulos transmitidos necesarios para que un UE logre acceder a la red. Esta cantidad puede variar entre diferentes UEs debido a factores como la carga de la red. Durante la simulación, *RACHConfig.maxNumPreambleTxAttempts* representa el número máximo de intentos de preámbulos transmitidos antes de que una UE falle en su intento de acceso. A medida que las UEs intentan acceder a la red, se actualiza el vector  $k$  para registrar cuantas UEs lograron un acceso exitoso después de varios intentos.

La distribución de  $k$  proporciona información sobre cuantos intentos de transmisión de preámbulos necesarios en promedio para que las UEs logren un acceso exitoso. La variable  $E_k$  se calcula como el promedio ponderado de  $k$ , lo que representa el número promedio de preámbulos transmitidos necesarios para un acceso exitoso en la simulación. Esta métrica es importante para evaluar el rendimiento en el RAP para dispositivos MTC un valor bajo de

EK indica que el RAP es eficiente y permite que la mayoría de los dispositivos MTC accedan a la red en un número mínimo de preámbulo.

### 2.6.3 Retardo en el acceso (D)

Desde el primer intento de acceso hasta la culminación exitosa del proceso RAP, se mide el tiempo transcurrido. Esto involucra considerar la acumulación de todos los retrasos generados en las diversas etapas del procedimiento, incluido el retraso causado por BI, en relación con la cantidad de intentos necesarios para lograr un acceso exitoso [24]. La ecuación (5) se utiliza para evaluar este parámetro.

$$p(D_n) = \frac{\sum M_A(D_n)}{totalSuccessfulUEs} \quad (5)$$

La probabilidad de que un dispositivo requiera  $n$  milisegundos, representada como  $p(D_n)$  se calcula sumando los valores de  $\sum M_A(D_n)$  y dividiéndolos por el total de dispositivos MTC que lograron un acceso exitoso (ecuación 5). La ecuación se basa en el conjunto de dispositivos MTC que completaron el proceso RAP de manera satisfactoria. Para cada dispositivo, se calcula el retraso de acceso, y estos retrasos se suman para cada valor de  $n$ . Luego, la suma se divide por el número total de dispositivos MTC que completaron el proceso RAP con éxito, proporcionando así la probabilidad acumulada de que un dispositivo necesite  $n$  milisegundos de retraso.

Al obtener la distribución de probabilidad para cada uno de los retardos generados durante el RAP, se procederá a utilizar interpolación para obtener resultados más precisos. Luego, se organizan los datos de menor a mayor para poder utilizar medidas de posición, como el percentil 95 y determinar de manera precisa el retardo máximo de acceso [5].

En la simulación  $D$  se define y se utiliza para representar la distribución de los retrasos de acceso, con una longitud igual a  $maxNumPreambleTxAttempts$ , que es el número máximo de intentos de preámbulos transmitidos permitidos. A medida que se actualiza  $D$ , acumula datos sobre accesos exitosos en cada intento de preámbulos transmitidos. Luego, se utiliza esta distribución para calcular estadísticas como el promedio del retardo de acceso, percentiles y otras métricas de rendimiento relacionadas con el acceso de la red.

## **CAPÍTULO III METODOLOGÍA**

Se estudiarán varias fuentes bibliográficas para identificar las variables que intervienen en este trabajo. El modelo por utilizar es un algoritmo adaptativo de *backoff* basado en aprendizaje reforzado que permitirá establecer una serie de procedimientos a seguir en cada una de las etapas que intervienen en el proyecto.

### **3.1 Tipo de Investigación**

#### **3.1.1 Investigación Exploratoria**

Para el desarrollo del proyecto se recopila información de estudios científicos sobre el procedimiento de *backoff* para asegurar menos congestión y aumentar la probabilidad de acceso éxito, además diseñar un algoritmo basado en aprendizaje reforzado que permita ajustar dinámicamente y de forma autónoma el tiempo de *backoff*, donde la red pueda superar situaciones de sobrecarga y no generar retardos innecesarios en situaciones de carga ligera, obteniendo así un mejor rendimiento del canal de acceso.

### **3.2 Diseño de Investigación**

#### **3.2.1 Método Científico**

En la investigación se determinará los estándares para un modelo analítico y una estrategia metodológica con el propósito de mejorar el rendimiento de la red mediante el uso de un algoritmo adaptativo de *backoff* basado en aprendizaje reforzado.

### **3.3 Técnicas de Investigación**

#### **3.3.1 Recolección de Datos**

La observación se empleará como una técnica de evaluación, siguiendo un conjunto de directrices diseñadas para examinar un fenómeno específico y comprender su naturaleza. Esto facilitará un análisis sistemático y preciso en relación con los objetivos planteados en la investigación, lo que permite un estudio detallado de los parámetros que se va a evaluar y visualizar de forma ordenada. El número de dispositivos que puedan acceder a la red y las UEs que puedan transmitir de forma exitosa serán los parámetros a tener en cuenta.

### 3.3.2 Instrumentos de la Investigación

Para el desarrollo del documento los instrumentos que se utilizarán son:

- 3GPP TR 37.868 Mejoras de RAN para comunicaciones de tipo máquina.
- 3GPP TR 36.912 Estudio de viabilidad para avances adicionales para LTE-A.
- Artículos científicos relacionados con el esquema de acceso RACH aprendizaje Q.
- Para implementar el modelo de simulación utilizaremos el software de Matlab.

### 3.4 Población de estudio y tamaño de Muestra

#### 3.4.1 Población

La población se define considerando el número de dispositivos que intentan acceder a la red y este conjunto de dispositivos varía en un rango que abarca desde 10000 y 30000, lo que representa el tráfico masivo mMTC. En particular, los intervalos de llegada de los dispositivos siguen una distribución beta de parámetros (3,4) conforme al modelo de tráfico MTC sugerido por los estándares 3GPP [8]. Para replicar el comportamiento real del acceso a la red, esta población se someterá al modelo de simulación del RACH.

#### 3.4.2 Muestra

La muestra será definida en función del número de dispositivos que hayan logrado completar con éxito el acceso a la red. Este criterio se establecerá utilizando el modelo de situación del algoritmo adaptativo de *backoff* basado en aprendizaje reforzado, una vez que hayan concluido el proceso de comunicación en el canal contemplado por dicho el modelo de simulación.

### 3.5 Operacionalización de las variables

En la Tabla 3 se detalla la operacionalización de variables a utilizar en el proyecto de investigación, se determina como variable dependiente al rendimiento de la red y como variable independiente al tiempo de *backoff*.



Tabla 3 Operacionalización de variables

VARIABLE	CONCEPTO	INDICADORES	INSTRUMENTO
Rendimiento de la red	Establecido por el número de usuarios que lograron acceder a la red.	<ul style="list-style-type: none"> <li>▪ Probabilidad de acceso aleatorio (Ps).</li> <li>▪ Número de preámbulos transmitidos (k).</li> <li>▪ Retardo en el acceso (D)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Observación directa a través de un modelo de simulación del algoritmo adaptativo de <i>backoff</i> basado en aprendizaje reforzado (MatLab).</li> </ul>
Tiempo de <i>Backoff</i>	Se produce un retraso aleatorio en el intento de retransmisión después de que se detecta la colisión. El aumento de la latencia también ayuda a reducir las colisiones, especialmente si la red está muy cargada.	<ul style="list-style-type: none"> <li>▪ Indicador de <i>backoff</i> (BI)</li> </ul>	

### 3.6 Métodos de Análisis y Procesamiento de Datos

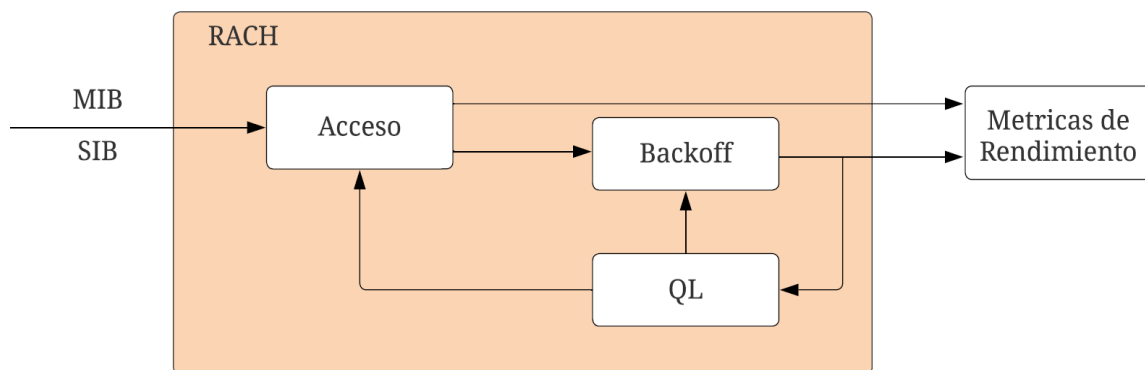
Se ha desarrollado un mecanismo adaptativo del procedimiento de *backoff* basado en aprendizaje reforzado QL utilizando el software de MatLab para incrementar la probabilidad de acceso exitoso en una red celular con comunicaciones mMTC. La integración de un algoritmo de aprendizaje RL como el aprendizaje Q permite ajustar dinámicamente los valores BI en función de las recompensas reales que recibe de las acciones realizadas [22]. El modelo funciona mediante la asignación de un valor BI por parte de una BS a través de

un mensaje SIB2 que se envía por difusión, cada  $T_{SIB2}$  teniendo un valor  $T_{SIB2} = 16RAO$  (80ms) [20].

Se realiza la simulación en el software de MatLab utilizando los parámetros definidos. Se evalúa el rendimiento del RACH en condiciones de congestión, para la obtención de la capacidad de acceso exitoso y otras métricas clave de rendimiento (KPIs). Estos KPIs se utilizan para evaluar la eficiencia del sistema en términos de su capacidad para manejar eficazmente las solicitudes de acceso de los equipos HTC y MTC [8].

El procesamiento de datos implica la configuración y simulación de un entorno de red celular para evaluar el rendimiento del procedimiento de acceso aleatorio considerando un entorno de celda única que aloja tanto a los equipos HTC como a los equipos MTC. Se establecen parámetros clave tanto del canal como del retardo, representados en la Tabla 4, la configuración del índice del PRACH 14, donde se establece la longitud de la subtrama de 1ms.

Durante la simulación, el algoritmo QL aprende gradualmente a tomar decisiones óptimas que maximizan la probabilidad de acceso exitoso a medida que los dispositivos interactúan repetidamente con el entorno y experimentan diferentes situaciones de acceso, tomando en cuenta la tasa de aprendizaje  $\alpha$  que controla el impacto de los nuevos valores óptimos en relación con los anteriores, mientras que el factor de descuento  $\gamma$  ajusta las recompensas en función del tiempo. Además, la probabilidad de exploración  $\epsilon$  influye en la elección de acciones óptimas [5], [17].



*Figura 4 Diagrama de Bloques del Procedimiento y análisis de resultados*

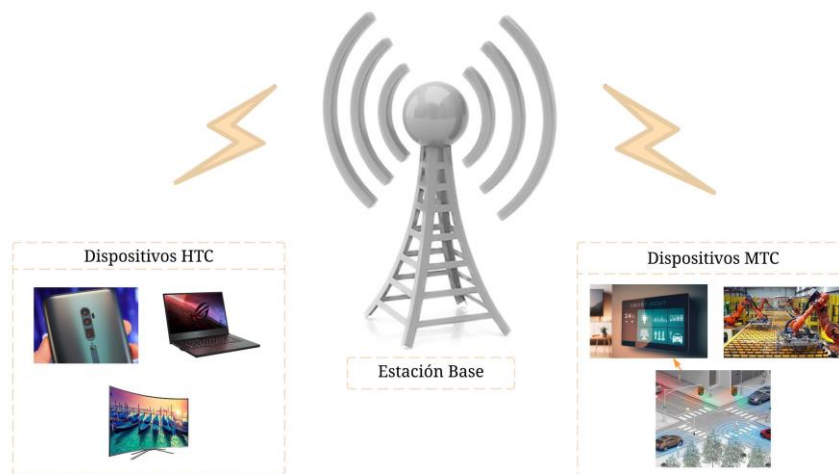
El desarrollo para el análisis del procedimiento de acceso aleatorio basado en la contención se muestra en la Figura 4. Se contemplan los parámetros de configuración del RACH, el

cálculo y la generación de resultados que representan las métricas de rendimiento. Además, del comportamiento del BI debido al algoritmo de QL.

*Tabla 4 Parámetros para la configuración del RACH*

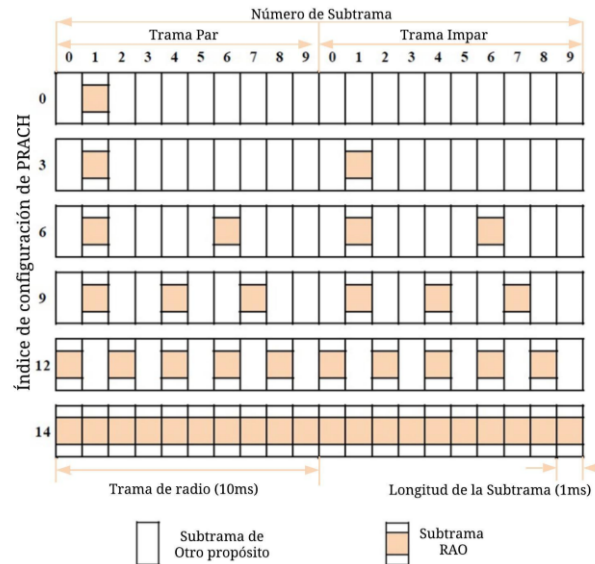
Parámetros de configuración	Configuración
<b>Parámetros del Canal</b>	
Número total de dispositivos MTC	10000-30000
Número de Preámbulos disponibles	54
Configuración del Índice del PRACH	14
Periodicidad RAO	1
Número de concesiones de enlace ascendente por subtrama	3
Máximo número de intentos de transmisiones de preámbulos	10
Indicador de <i>Backoff</i>	Variable (20 a 960 ms)
<b>Parámetros de retardo</b>	
Temporizador de resolución de contención	48
Retardo de procesamiento del preámbulo	2
Retardo de procesamiento del RAR	5
Retardo de procesamiento de la solicitud de conexión	4
Número de intentos de transmisión de Mgs3 y Mgs4	5
Probabilidad de retransmisión HARQ para Mgs3 y Mgs4	0,1
Tiempo de ida y vuelta (RTT) para el Mgs3	8
RTT para el Mgs4	5

### 3.6.1 Modelo del Sistema



*Figura 5 Modelo del sistema*

Para el estudio se considera una sola estación base que atiende a N dispositivos las solicitudes de acceso de usuarios (UEs) que se registraron previamente en la estación base, como se muestra en la Figura 5. En este contexto, cada dispositivo está equipado con una tarjeta SIM (Módulo de identificación de abonado) que le permite conectarse directamente a la estación base sin necesidad de una puerta de enlace.



**Figura 6** Diferentes configuraciones del índice de PRACH – Estructura de trama tipo 1

*Fuente:* [25]

La configuración RACH que se utiliza corresponde al tipo 1 de estructura de trama, donde la longitud de la subtrama es de 1ms, se tiene un total de 64 configuraciones PRACH disponible para cada celda, como se observa en la Figura 6. Cada subtrama tiene un máximo de una oportunidad de acceso aleatorio (RAO) cada 20ms. La periodicidad de RAO se determina mediante el índice de configuración de PRACH en el estudio es igual a 14, entonces la periodicidad de RAOs es de 10ms RAOs por cada trama.

### 3.6.2 Parámetros de configuración del algoritmo de QL

Los parámetros como la tasa de aprendizaje  $\alpha$ , el factor de descuento  $\gamma$  y la probabilidad de exploración  $\epsilon$  en el algoritmo de QL, se establecen para ajustar el proceso de aprendizaje y mejorar la eficiencia del algoritmo.

Estos parámetros actúan como los elementos de control que determinan cómo el agente interactúa con su entorno y aprende a tomar decisiones óptimas. El algoritmo busca maximizar las recompensas acumuladas a lo largo del tiempo mediante de la interacción continua con el entorno y la exploración de diferentes acciones.

*Tabla 5 Parámetros de configuración para el desarrollo del algoritmo QL*

<b>Parámetros de configuración</b>	<b>Configuración</b>
Tasa de aprendizaje (Alpha)	0,2
Factor de Descuento (Gamma)	0,9
Probabilidad de exploración (Épsilon)	0,1

Los tres parámetros de configuración que se establecieron para el desarrollo del algoritmo QL están detallados en la Tabla 5. La tasa de aprendizaje, el primer parámetro regula la rapidez con la que el agente actualiza su función de valor Q en cada RAO al interactuar con el entorno. Un valor alto de alfa indica un aprendizaje rápido, lo que hace que el algoritmo sea más sensible a los errores y variabilidad en los datos, mientras que un valor bajo de alfa indica un aprendizaje lento lo que hace que el algoritmo sea más estable y menos susceptible a los errores [26], por tal razón en la investigación se decidió ajustar este valor a 0,2 para lograr un equilibrio adecuado.

La importancia relativa de las recompensas futuras en comparación con las recompensas inmediatas se determina por el segundo parámetro,  $\gamma$ . Un valor alto de gamma, cerca de 1, indica que el agente valora las recompensas a largo plazo, mientras que un valor bajo de gamma, cerca de 0 implica el agente valora las recompensas inmediatas [26]. Para la investigación se considera un valor típico de gamma a 0,9 lo que indica que el agente tiene en cuenta las recompensas futuras, pero no con la misma importancia que las recompensas inmediatas.

La probabilidad de exploración  $\epsilon$  determina las veces que el agente realiza una acción aleatoria en lugar de la acción óptima basada en su conocimiento actual. Un valor alto de épsilon indica que el agente está enfocado en la exploración, aumentando la probabilidad de elegir acciones aleatorias. En cambio, un valor bajo de épsilon significa que el agente se inclina hacia la explotación reduciendo la probabilidad de elegir acciones aleatorias [6]. El valor de 0,1 establecido en la investigación asegura que el agente aproveche al máximo su conocimiento actual mientras sigue explorando para descubrir nuevas estrategias. En un 10% de las ocasiones, el agente optará una acción aleatoria para explorar nuevas opciones, mientras que el 90% restante, se centrará en explotar la mejor acción actual estimada por la función Q.

### 3.6.3 Procedimiento de la implementación del algoritmo de QL

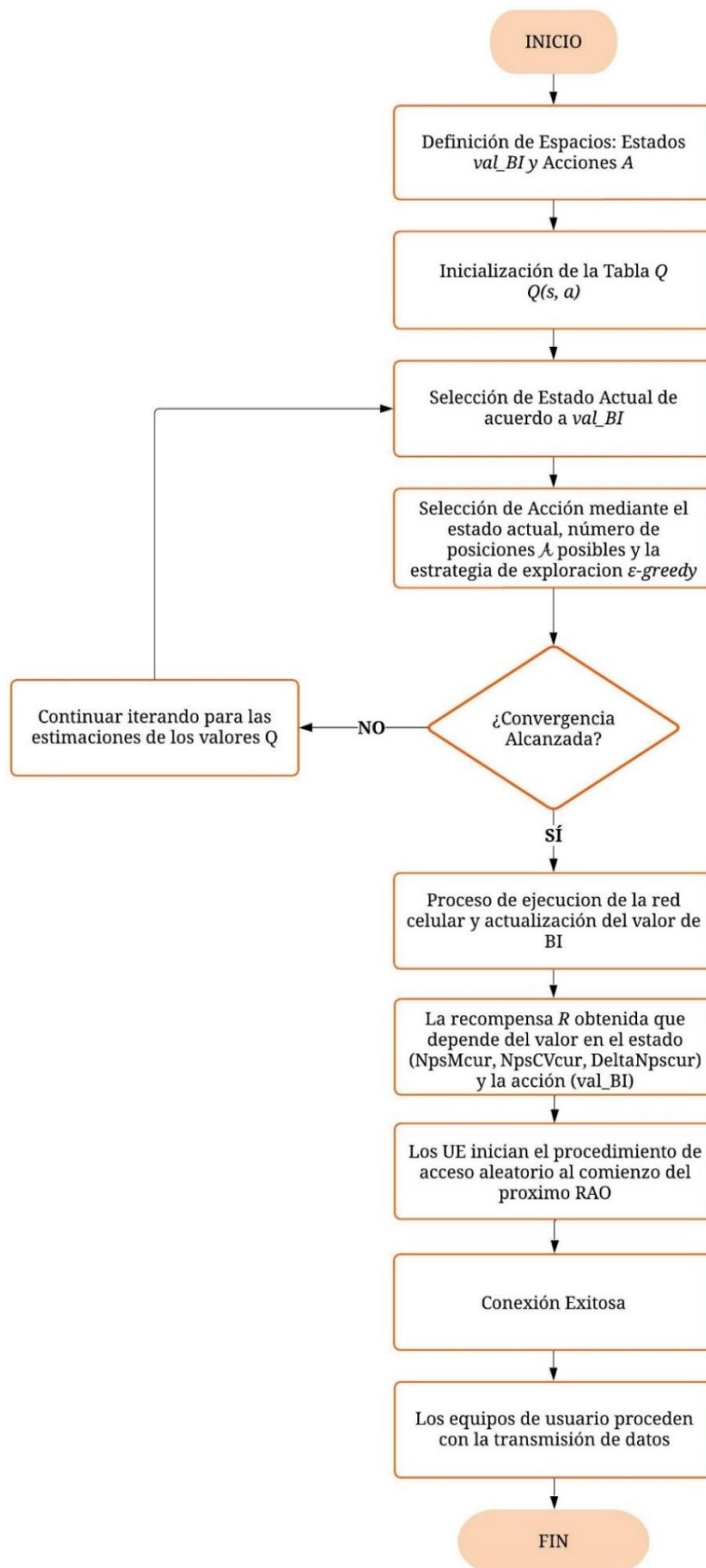


Figura 7 Diagrama de Flujo de la implementación del algoritmo de QL

La Figura 7 presenta el enfoque para evaluar el rendimiento del sistema de acceso exitoso en la red celular con comunicaciones masivas. La simulación se desarrolla en varios pasos, en primer lugar, se genera el tráfico que simula la llegada de datos de dispositivos MTC, lo que implica crear patrones de tráfico que reflejen la actividad de estos dispositivos en la red.

A continuación, se configuran los parámetros del sistema, esto incluye la configuración del RACH como la disponibilidad de preámbulos, la periodicidad RAO y los parámetros de retardo. Luego, se ejecuta el bucle de simulación donde se modela el acceso aleatorio de los dispositivos en cada ciclo RAO. Aquí, se siguen las reglas y procedimientos definidos para gestionar el acceso de dispositivos MTC.

Durante la simulación, se recopilan los datos que proporcionan información sobre el rendimiento del sistema. Estos datos incluyen la probabilidad de acceso exitoso ( $P_s$ ), la distribución de preámbulos transmitidos ( $k$ ) y el retraso de acceso ( $D$ ) para dispositivos MTC. Estos datos son esenciales para evaluar la eficiencia y confiabilidad del sistema en condiciones variadas de tráfico. Además, para mejorar el rendimiento del acceso exitoso, se implementa el algoritmo de aprendizaje  $Q$  que optimiza el valor de BI en función de las condiciones de la red, aprendiendo una política óptima para seleccionar el mejor valor de BI, que determina cuánto tiempo un dispositivo espera antes de intentar acceder a la red nuevamente después de un intento fallido.

Se inicia con una tabla  $Q$  vacía y se actualiza durante múltiples ciclos de simulación. Durante cada ciclo, el dispositivo elige un estado (valor del BI) basada en una política  $\epsilon$ -greedy, que equilibra la explotación de acciones conocidas con la exploración de nuevas. Después de cada intento de acceso aleatorio, se calcula una recompensa y se actualiza la tabla  $Q$  para reflejar la utilidad esperada de cada acción en cada estado.

El algoritmo continúa aprendiendo y actualizando la tabla  $Q$  hasta que converja, lo que significa que los valores  $Q$  ya no cambian significativamente. Una vez que ha convergido, se sigue la política óptima aprendida para seleccionar el valor de BI, mejorando así la eficiencia del acceso.

### 3.6.4 Función de la asignación de Recompensas a las posibles acciones tomadas

El algoritmo adaptativo de *backoff* utiliza técnicas de aprendizaje por refuerzo para ajustar dinámicamente los parámetros en función de la congestión de la red y mejora el rendimiento de la misma. El agente aprende a tomar decisiones óptimas sobre cuánto tiempo debe esperar una estación base antes de intentar acceder al medio de transmisión, considerando las recompensas o castigos recibidos en función del rendimiento de la red.

*Tabla 6* Asignación de recompensas a diferentes acciones en diversas condiciones

NpsMcur	NpsCVcur	DeltaNpscur	val_BI	R	val_BI	R	val_BI	R	val_BI	R
$\leq 3$	$< 0,4$	$= 1$	960	<b>50</b>	$\geq 240$	<b>40</b>	$\geq 80$	<b>10</b>	$\geq 20$	<b>-10</b>
$\leq 3$	$< 0,4$	$= 3$	960	<b>100</b>	$\geq 240$	<b>80</b>	$\geq 80$	<b>5</b>	$\geq 20$	<b>-20</b>
$\leq 3$	$\geq 0,4$	$= 3$	960	<b>60</b>	$\geq 240$	<b>50</b>	$\geq 80$	<b>3</b>	$\geq 20$	<b>30</b>
$< 7$	$< 0,4$	$= 1$	960	<b>30</b>	$\geq 240$	<b>40</b>	$\geq 80$	<b>50</b>	$\geq 20$	<b>50</b>
$< 7$	$< 0,4$	$= 3$	960	<b>60</b>	$\geq 240$	<b>70</b>	$\geq 80$	<b>60</b>	$\geq 20$	<b>60</b>
$< 7$	$\geq 0,4$	$= 1$	960	<b>1</b>	$\geq 240$	<b>5</b>	$\geq 80$	<b>40</b>	$\geq 20$	<b>40</b>
$\leq 10$	$\geq 0,2$	$= 1$	960	<b>0.5</b>	$\geq 240$	<b>1</b>	$\geq 80$	<b>5</b>	$\geq 20$	<b>6</b>
$\leq 10$	$\geq 0,2$	$= 3$	960	<b>5</b>	$\geq 240$	<b>8</b>	$\geq 80$	<b>15</b>	$\geq 20$	<b>16</b>
$\leq 10$	$< 0,2$	$= 1$	960	<b>-100</b>	$\geq 240$	<b>-90</b>	$\geq 80$	<b>-60</b>	$\geq 20$	<b>-50</b>
$> 10$	$< 0,2$	$= 3$	960	<b>-90</b>	$\geq 240$	<b>-85</b>	$\geq 80$	<b>-45</b>	$\geq 20$	<b>-35</b>
$> 10$	$\geq 0,2$	$= 3$	960	<b>-90</b>	$\geq 240$	<b>-85</b>	$\geq 80$	<b>-45</b>	$\geq 20$	<b>-35</b>

La recompensa R se define como una función dependiente de las variables NpsMcur, NpsCVcur, DeltaNpscur y val\_BI. Dadas las numerosas posibilidades, se representa algunas de las posibles combinaciones estado/recompensa. La Tabla 6 describe cómo se asignan las recompensas a diversas acciones en función de condiciones específicas, con el objetivo de optimizar los valores de BI mediante el algoritmo QL. Cada configuración de condiciones y recompensas permite al algoritmo ajustar estratégicamente los valores, contribuyendo a la mejora del rendimiento del sistema.

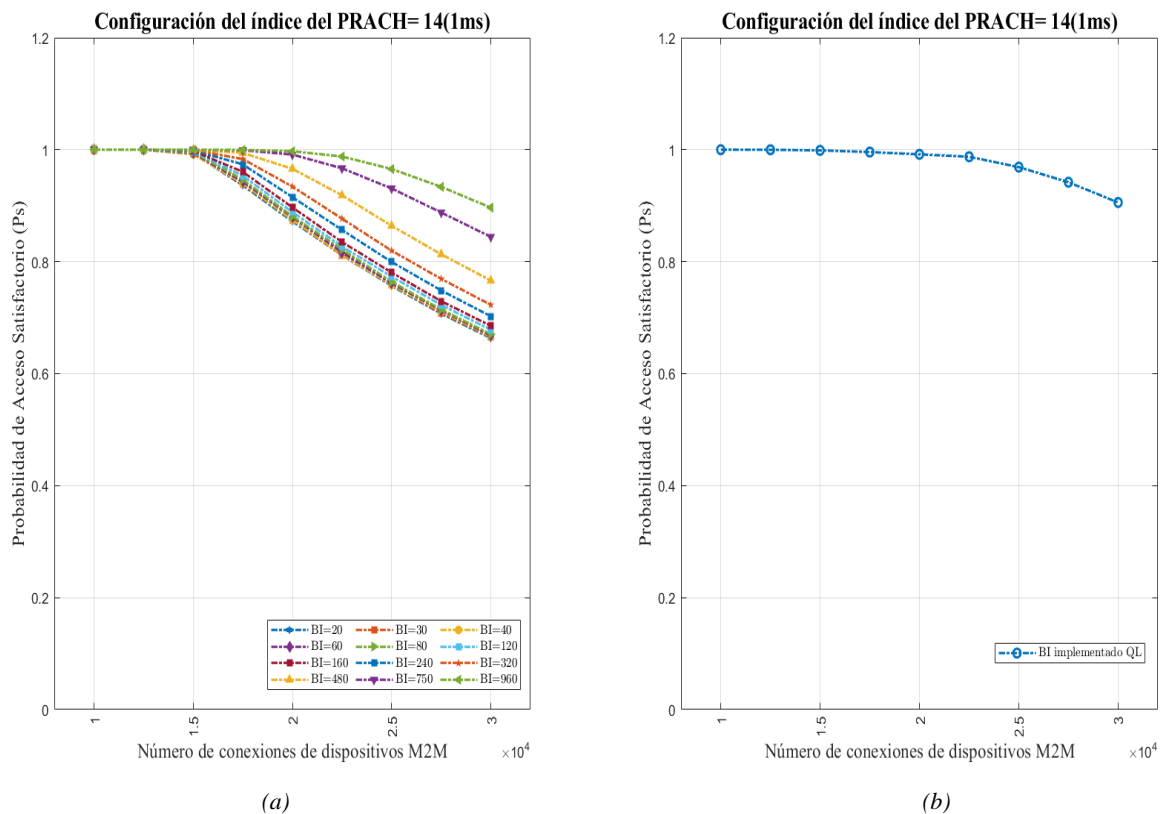


## CAPÍTULO IV RESULTADOS Y DISCUSIÓN

### 4.1 Análisis de los resultados de las métricas de rendimiento

Los resultados obtenidos de las métricas de rendimiento se basan en una configuración en el RACH que utiliza una periodicidad de RAOs de 1 ms (configuración del índice del PRACH=14). Estos hallazgos se representan en 3 figuras que muestran la probabilidad de acceso exitoso ( $P_s$ ), el promedio preámbulos transmitidos ( $k$ ) y el retardo de acceso con un percentil 95 ( $D_{95}$ ). Se lleva a cabo una comparación entre el comportamiento de un valor de BI estático y un valor de BI dinámico mediante el uso del algoritmo de QL.

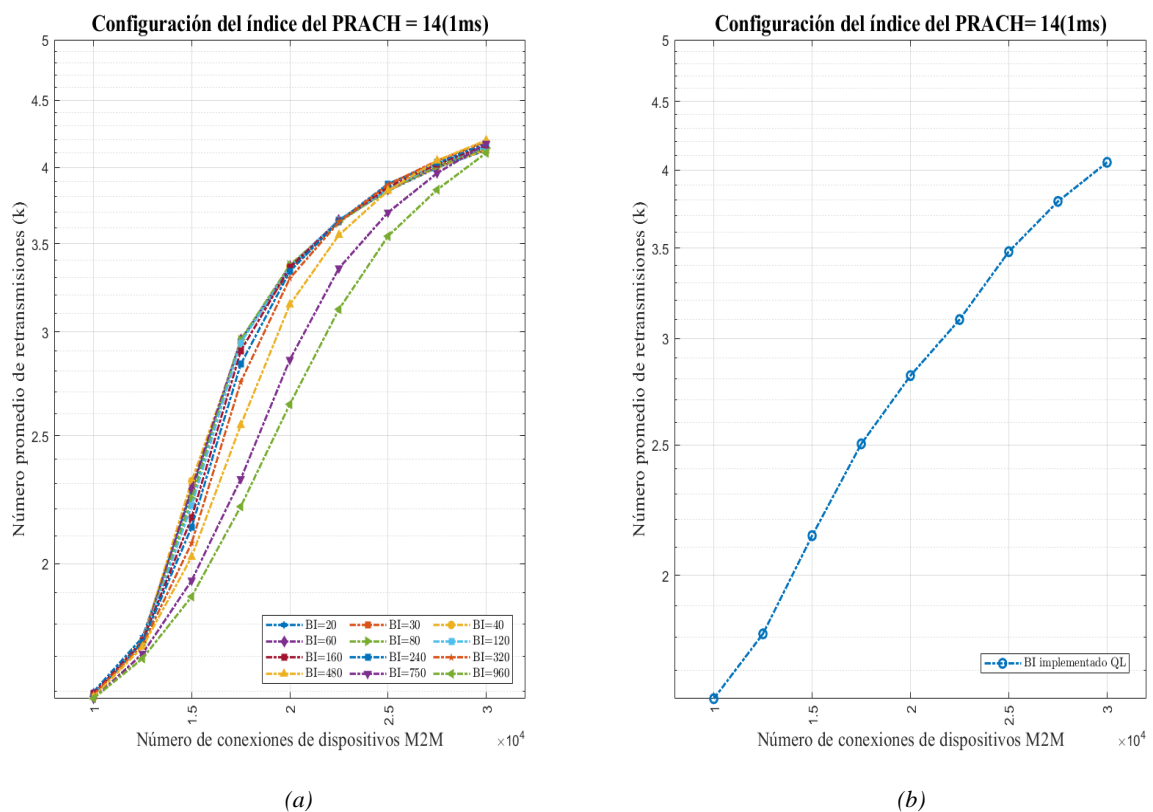
#### 4.1.1 Resultados de la probabilidad de acceso exitoso, número promedio de preámbulos transmitidos y retardo en el acceso con un percentil 95 en un entorno tanto de BI estático como con BI dinámico



**Figura 8** Probabilidad de acceso exitoso ( $P_s$ ) varía con diferentes cargas de tráfico; (a) Utilizando distintos valores de BI estáticos y (b) Al obtener valores óptimos de BI mediante la implementación del algoritmo QL.

Los valores obtenidos de la probabilidad de acceso exitoso ( $P_s$ ) en distintos contextos se muestra en la Figura 8, teniendo en cuenta varios escenarios de tráfico (MTC) y el índice de configuración PRACH=14, para el análisis de esta métrica se utilizan 12 valores de BI descritos en la Tabla 1.

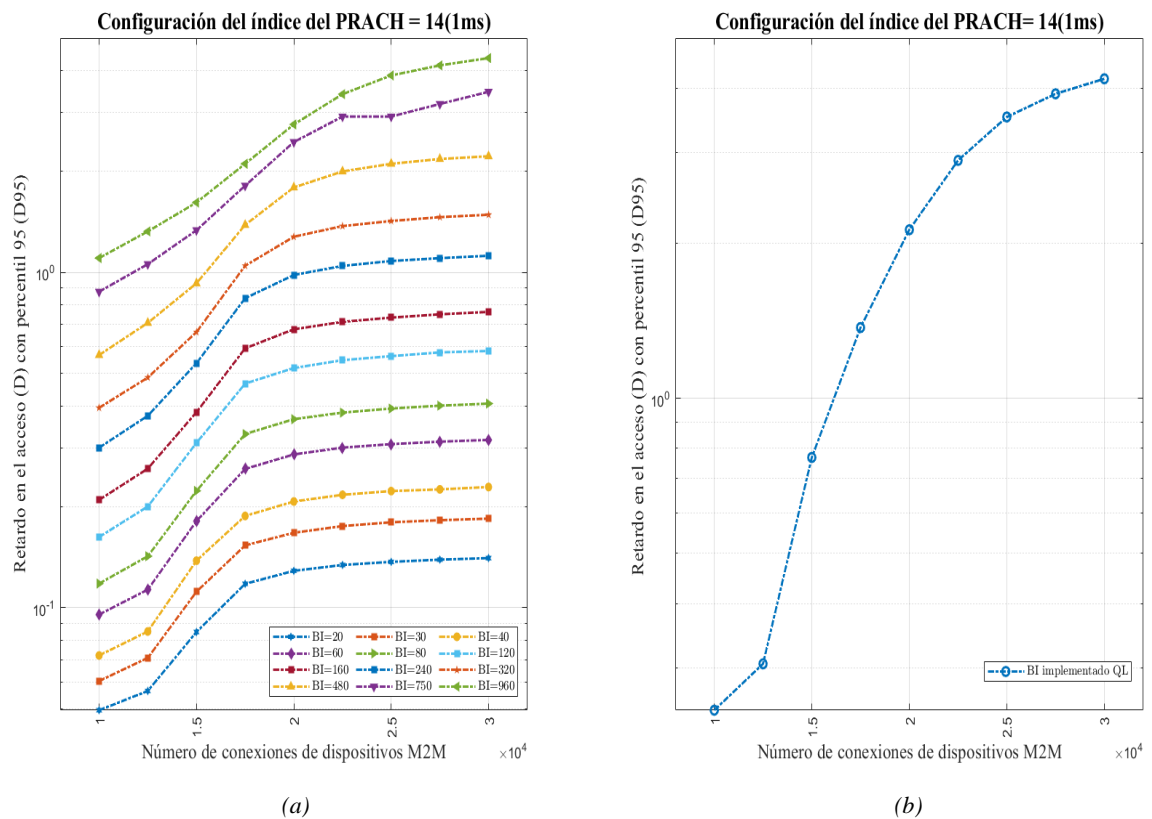
Para evaluar cómo varía la probabilidad de acceso exitoso, se determinó el valor de  $P_s$  en (a) la relación con distintos valores de BI estáticos. En (b), se observa el valor de  $P_s$  a través de la implementación del algoritmo QL para obtener el mejor valor de BI. Los resultados más relevantes se detallan en las Tablas 7, 8, 9, 10 y 11 teniendo en cuenta diversas cargas de tráfico (10000, 20000 y 30000) y una periodicidad de RAO de 1 ms.



**Figura 9** Número promedio de retrasmisiones ( $k$ ) cambia en función de las diversas cargas de tráfico; (a) Utilizando distintos valores de BI estáticos y (b) Al obtener valores óptimos de BI con la implementación del algoritmo QL.

La Figura 9 se ilustra el comportamiento en las retrasmisiones que el sistema experimenta donde (a) muestra el comportamiento bajo valores de BI estáticos mientras que en (b) se observa un comportamiento donde el valor de BI se ajusta en función del estado del canal.

Este análisis visual muestra la manera que el valor de BI afecta directamente las retransmisiones experimentadas por los dispositivos MTC durante su proceso de acceso, resaltando la capacidad de optimización que el algoritmo de QL proporciona al ajustar de forma dinámica el valor de BI de acuerdo con las recompensas. En las tablas 7, 8, 9, 10 y 11 se detallan los resultados más relevantes.



**Figura 10** Retardo en el Acceso con percentil 95 (D95) varía en función de las diversas cargas de tráfico; (a) Utilizando diferentes valores de BI estáticos y (b) Al obtener valores óptimos de BI mediante la implementación de algoritmo QL.

En la Figura 10 se muestra las variaciones en los resultados del percentil D95 en relación para distintos escenarios de tráfico y valores de BI con una configuración de 1 ms de RAOs.

En (a) se observa que aumentar la cantidad de tráfico, así como el valor de BI, también incrementa el retardo en el acceso. Estos hallazgos concuerdan con la idea de que mayores cargas de tráfico y valores de BI más altos pueden tener un impacto negativo en el retardo. En (b), la implementación del algoritmo de QL al obtener un valor de BI óptimo intenta mitigar este efecto negativo. Los valores más relevantes obtenidos se detallan en las Tablas 7, 8, 9, 10 y 11.

#### 4.1.1.1 Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 20 ms y un valor de BI aplicado el algoritmo de QL

*Tabla 7 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 20 y valor de BI óptimo aplicando el algoritmo QL.*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5977	49,5000	20	10000	1	1,6160	239,4000	20
12500	1	1,7577	56,5000	20	12500	0,9999	1,8086	310,1000	20
15000	0,9931	2,2696	84,7000	20	15000	0,9988	2,1225	742,0000	40
17500	0,9363	2,9534	117,9000	20	17500	0,9953	2,5180	1373,1000	20
20000	0,8720	3,3555	128,9000	20	20000	0,9917	2,8261	2131,4000	20
22500	0,8118	3,6417	134,2000	20	22500	0,9862	3,1081	2900,4000	30
25000	0,7568	3,8399	137,1000	20	25000	0,9702	3,4761	3505,2000	240
27500	0,7072	3,9981	139,2000	20	27500	0,9404	3,8054	3901,9000	120
30000	0,6635	4,1228	140,7000	20	30000	0,9049	4,0497	4160,6000	480

Se presenta una comparación detallada de las métricas de rendimiento en diversas condiciones de tráfico para un sistema de comunicación respaldado por el algoritmo de QL como se muestra en la Tabla 7. En un escenario con un valor de BI estático en 20 ms, se observa que la probabilidad de acceso exitoso (Ps) disminuye al 66% con un tráfico alto de 30000 UEs. Al emplear un escenario de tráfico medio de 20000 UEs, se tiene un Ps superior al 70%. Finalmente, para una carga de tráfico baja de 10000 UEs, la tasa se establece en el 100%, lo que indica que la mayoría de los dispositivos MTC logran acceder a la red.

El número promedio de preámbulos transmitidos (k) para cargas de tráfico bajas y medianas se demuestra eficiencia en el acceso al canal, mientras que para cargas de tráfico alto 30000 UEs, se necesita un mayor número de retransmisiones antes de lograr un acceso exitoso al canal de comunicación. El retardo de acceso con percentil 95 (D95) muestra valores relativamente más altos, en cargas de tráfico medio y alto lo que implica que algunos dispositivos experimentan retrasos en el acceso debido a retransmisiones adicionales. Por el contrario, los valores bajos indican que la mayoría de los 10000 UEs experimentan tiempos de acceso más cortos.

En cambio, al aplicar el algoritmo QL para optimizar el valor de BI resulta en una mejora significativa en las métricas de rendimiento. La Ps se mantiene alta en el orden del 90% y 100% para las distintas cargas de tráfico. Esto sugiere que el valor de BI optimizado por el algoritmo QL permite una mayor eficiencia conjuntamente con los términos de retransmisiones y retardo de acceso, mejorando así la calidad y rapidez de la comunicación para dispositivos MTC.

#### 4.1.1.2 Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 160 ms y un valor de BI aplicado el algoritmo de QL

*Tabla 8 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 160 y valor de BI óptimo aplicando el algoritmo QL.*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5912	209,8000	160	10000	1	1,6064	249,2000	480
12500	1	1,7463	260,0000	160	12500	0,9999	1,8120	308,9000	240
15000	0,9978	2,1675	383,0000	160	15000	0,9987	2,1476	763,5000	40
17500	0,9604	2,9007	594,0000	160	17500	0,9956	2,5081	1352,4000	160
20000	0,8969	3,3580	676,9000	160	20000	0,9920	2,8138	2156,7000	40
22500	0,8351	3,6406	712,3000	160	22500	0,9865	3,1030	2920,4000	20
25000	0,7806	3,8629	733,5000	160	25000	0,9699	3,4757	3519,0000	120
27500	0,7294	4,0242	750,2000	160	27500	0,9408	3,7950	3909,8000	80
30000	0,6854	4,1575	762,8000	160	30000	0,9050	4,0522	4165,9000	480

En la Tabla 8, es posible observar cómo el ajuste del valor de BI afecta directamente el desempeño del sistema de comunicación. Al se utiliza un valor estático de BI de 160 ms, se muestra que la probabilidad de acceso exitoso (Ps) disminuye al 68% con un alto volumen de tráfico de 30000 UEs. Al emplear un escenario de tráfico medio de 20000 UEs, se obtiene una Ps que es superior al 89%. El resultado final indica que la mayoría de los dispositivos MTC podrán acceder a la red en una carga de tráfico baja de 10000 UEs, que se establece en el 100%.

No obstante, este enfoque estático también proporciona un número promedio de preámbulos transmitidos (k), para cargas de tráfico baja y media se tiene una eficiencia en el acceso al canal, mientras que para cargas de tráfico alto 30000 UEs se necesita un mayor número de

retransmisiones antes de lograr un acceso exitoso al canal de comunicación. El retardo de acceso con percentil 95 (D95) indica valores que son relativamente altos, en cargas de tráfico medio y alto lo que sugiere que algunos dispositivos experimentan retardos en el acceso debido a retransmisiones adicionales. Por el contrario, valores bajos indican que la mayoría de las 10000 UEs experimentan tiempos de acceso más cortos.

Por otro lado, al utilizar el algoritmo QL para optimizar el valor de BI, se observa un cambio notorio en las métricas de rendimiento. La Ps se mantiene en niveles altos, indicando que el algoritmo simultáneamente con el número promedio de preámbulos transmitidos (k) y el retardo de acceso (D) logran mantener una alta tasa de acceso exitoso. Esto implica que el algoritmo QL es capaz de ajustar el valor de BI de manera eficaz, reduciendo la necesidad de retransmisiones y acortando los retrasos en el tiempo de acceso.

#### 4.1.1.3 Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 320 ms y un valor de BI aplicado el algoritmo de QL

*Tabla 9 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 320 y valor de BI óptimo aplicando el algoritmo QL.*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5896	394,8000	320	10000	1	1,6071	246,2000	60
12500	1	1,7350	486,0000	320	12500	0,9999	1,8030	307,3000	160
15000	0,9994	2,0728	664,1000	320	15000	0,9987	2,1305	761,4000	20
17500	0,9833	2,7477	1047,2000	320	17500	0,9952	2,5113	1372,8000	240
20000	0,9341	3,2943	1278,1000	320	20000	0,9919	2,8141	2139,4000	20
22500	0,8769	3,6349	1374,9000	320	22500	0,9868	3,0916	2907,0000	20
25000	0,8197	3,8802	1423,9000	320	25000	0,9697	3,4767	3510,8000	80
27500	0,7693	4,0459	1462,0000	320	27500	0,9497	3,7994	3914,8000	40
30000	0,7229	4,1851	1486,3000	320	30000	0,9056	4,0417	4165,3000	40

La Tabla 9, presenta una comparación detallada de las métricas de rendimiento en diversas situaciones de tráfico, considerando dos enfoques diferentes para el BI en un sistema de comunicación. En el primer escenario, donde se utiliza un valor fijo de BI de 320 ms, al encontrarse con un tráfico alto de 30000 UEs resulta en una disminución de la probabilidad

de acceso exitoso (Ps) al 72%. Al emplear un escenario de tráfico medio de 20000 UEs, se logra una Ps superior al 93%. Finalmente, para una carga de tráfico baja de 10000 UEs se estabiliza en 100%, lo que indica que la mayoría de los dispositivos MTC logran acceder a la red.

Sin embargo, el número promedio de preámbulos transmitidos (k) para cargas de tráfico baja y media se mantiene una eficiencia en el acceso al canal. Por otro lado, se requiere un mayor número de retransmisiones para cargas de tráfico alto 30000 UEs antes de lograr un acceso éxito al canal de comunicación. En cargas de tráfico alta y medio el retardo de acceso con percentil 95 (D95) son relativamente elevados, lo que sugiere que algunos dispositivos experimentan demoras en el acceso debido a retransmisiones adicionales. Por el contrario, los valores bajos indican que la mayoría de las 10000 UEs experimentan tiempos de acceso más cortos.

En contraste, los resultados muestran una mejora sustancial en las métricas clave cuando se implementa el algoritmo QL para determinar un valor óptimo de BI. La Ps sigue manteniendo su nivel alto, lo que denota una continua efectividad en el acceso. Pero lo más significativo se observa la disminución en los valores promedio de transmisiones (k) y los valores de retardo de acceso (D95). Este descenso resalta la capacidad del algoritmo QL para adaptar dinámicamente el valor de BI según las condiciones cambiantes de tráfico.

#### 4.1.1.4 Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 480 ms y un valor de BI aplicado el algoritmo de QL

*Tabla 10 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 480 y valor de BI óptimo aplicando el algoritmo QL.*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5835	567,2000	480	10000	1	1,6124	247,0000	120
12500	1	1,7285	707,4000	480	12500	0,9998	1,8203	327,7000	120
15000	0,9997	2,0234	927,9000	480	15000	0,9988	2,1525	771,7000	320
17500	0,9942	2,5479	1385,9000	480	17500	0,9958	2,5071	1360,7000	20
20000	0,9659	3,1463	1792,6000	480	20000	0,9920	2,8185	2117,4000	480
22500	0,9186	3,5534	2001,0000	480	22500	0,9870	3,0959	2898,9000	30

25000	0,8641	3,8395	2108,9000	480	25000	0,9697	3,4761	3510,0000	320
27500	0,8132	4,0475	2179,3000	480	27500	0,9402	3,8011	3907,6000	240
30000	0,7663	4,1926	2220,9000	480	30000	0,9061	4,0451	4162,3000	20

El impacto del ajuste dinámico del valor de BI en un sistema de comunicación bajo diversas condiciones de tráfico se detalla en la Tabla 10. Se obtiene una comprensión valiosa del rendimiento del sistema al examinar las métricas clave en ambos enfoques. Cuando se utiliza un valor de BI estático de 480 ms, se contrasta que la probabilidad de acceso exitoso (Ps) disminuye a un 76% con un tráfico alto de 30000 UEs. Al emplear un escenario de tráfico medio de 20000 UEs, se tiene una Ps superior al 96%. Por último, para una carga de tráfico baja de 10000 UEs, se establece en 100%, lo que indica que la mayoría de los dispositivos MTC logran acceder a la red.

El número promedio de preámbulos transmitidos (k) demuestra eficiencia en el acceso al canal para cargas de tráfico baja y media, mientras que se requiere un mayor número de transmisiones para cargas de tráfico alto 30000 UEs antes de lograr un acceso exitoso al canal de comunicación. El retardo de acceso con percentil 95 (D95) resulta relativamente alto, lo que indican ciertas deficiencias en la utilización del canal y la demora en el acceso.

Al contrario, el uso del algoritmo QL resulta en mejoras significativas en todas las métricas cuando se ajusta el valor de BI. La Ps permanece alta, aunque el número promedio de transmisiones (K) y el retardo en el acceso (D) disminuyen considerablemente. Esto permite una comunicación más efectiva entre los UEs. Los datos presentados en la tabla respaldan la conclusión de que la implementación del valor de BI dinámico a través del algoritmo QL tiene un efecto positivo en el rendimiento del sistema de comunicación, al permitir un acceso más eficiente y optimizado al canal en diversas situaciones de tráfico.

#### 4.1.1.5 Resultados numéricos de las métricas de rendimiento con un valor de BI estático de 960 ms y un valor de BI aplicado el algoritmo de QL

*Tabla 11 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 960 y valor de BI óptimo aplicando el algoritmo QL.*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
---------	----	---	---	----	---------	----	---	---	----



10000	1	1,5797	1103,7000	960	10000	1	1,6199	248,2000	40
12500	1	1,6939	1325,5000	960	12500	0,9999	1,8099	305,4000	480
15000	1	1,8875	1616,5000	960	15000	0,9988	2,1409	767,8000	30
17500	0,9997	2,2097	2109,0000	960	17500	0,9958	2,5049	1371,8000	120
20000	0,9973	2,6425	2763,1000	960	20000	0,9917	2,8150	2123,6000	20
22500	0,9878	3,1199	3401,5000	960	22500	0,9874	3,0983	2894,9000	80
25000	0,9655	3,5461	3868,6000	960	25000	0,9689	3,4798	3514,6000	160
27500	0,9337	3,8461	4145,1000	960	27500	0,9419	3,7919	3897,5000	40
30000	0,8966	4,1018	4357,3000	960	30000	0,9057	4,0544	4168,1000	40

En la Tabla 11, se destaca cómo la adaptación dinámica del valor de BI mediante el algoritmo QL puede influir en el rendimiento del sistema de comunicación en diversos escenarios de tráfico, representados por las distintas cantidades de UEs.

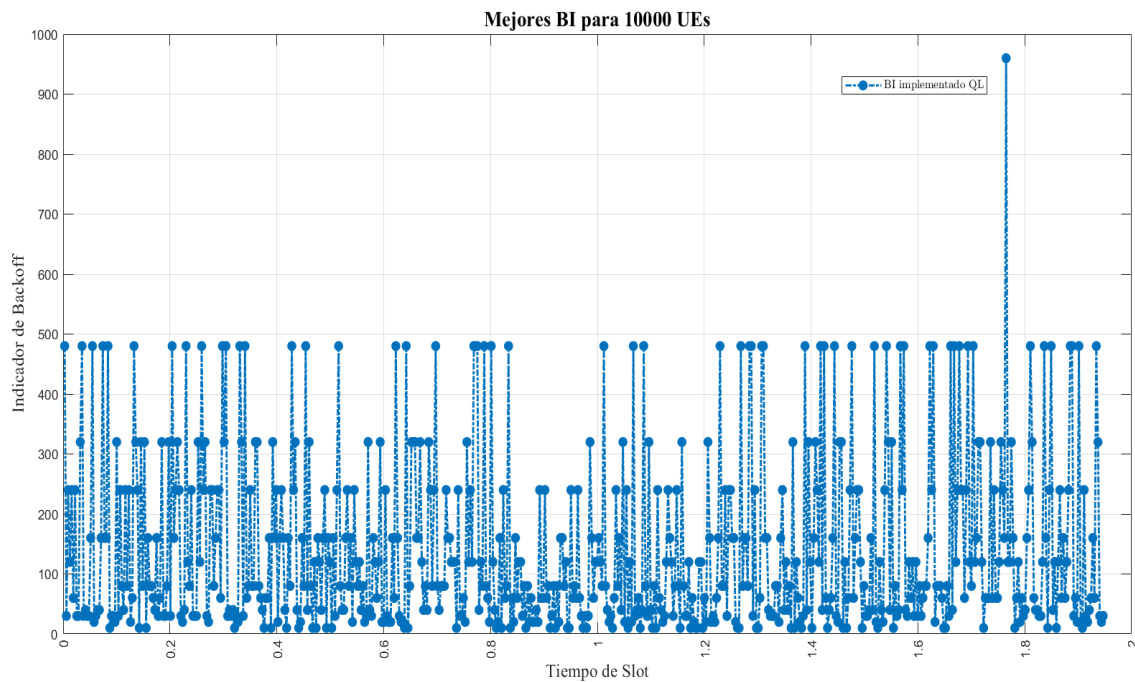
Un análisis detenido evidencia que, para un valor de BI estático de 960 ms, la probabilidad de acceso exitoso ( $P_s$ ) disminuye al 89% con un tráfico alto de 30000 UEs. Se logra un  $P_s$  de 99% al utilizar un escenario de tráfico medio de 20000 UEs. Al final, para una carga de tráfico baja de 10000 UEs alcanza el 100%, lo que indica que la mayoría de los dispositivos MTC logran acceder a la red.

En relación para cargas de tráfico baja, media y alta, al tener un valor de BI estático de 960 ms (siendo este el mejor valor de BI para el estudio), se observa que el número promedio de transmisiones ( $k$ ) y retardo de acceso con percentil 95 ( $D_{95}$ ) disminuye en relación a los valores de BI anteriores. Esta observación sugiere que el sistema mantiene eficiencia y tiempos de acceso adecuados incluso en situaciones de tráfico más intensa.

Sin embargo, con la introducción del algoritmo QL a medida que el tráfico aumenta, las UEs continúan manteniendo una alta probabilidad de acceso exitoso ( $P_s$ ), lo cual refleja la capacidad constante del sistema para permitir el acceso al canal de comunicación. Además, el número promedio de transmisiones ( $k$ ) y el retardo de acceso ( $D$ ) experimentan una disminución significativa a medida que aumenta el tráfico, lo que indica una adaptación efectiva del valor de BI a las condiciones cambiantes.

## 4.2 Resultados de los mejores valores de BI para diferentes cargas de tráfico

Los mejores valores de BI obtenidos después de implementar el algoritmo de QL a una carga de tráfico baja de 10000 UEs que se muestra en la Figura 11. Estos valores reflejan las decisiones óptimas para maximizar el rendimiento del sistema en términos de acceso al canal compartido y eficiencia en la comunicación.



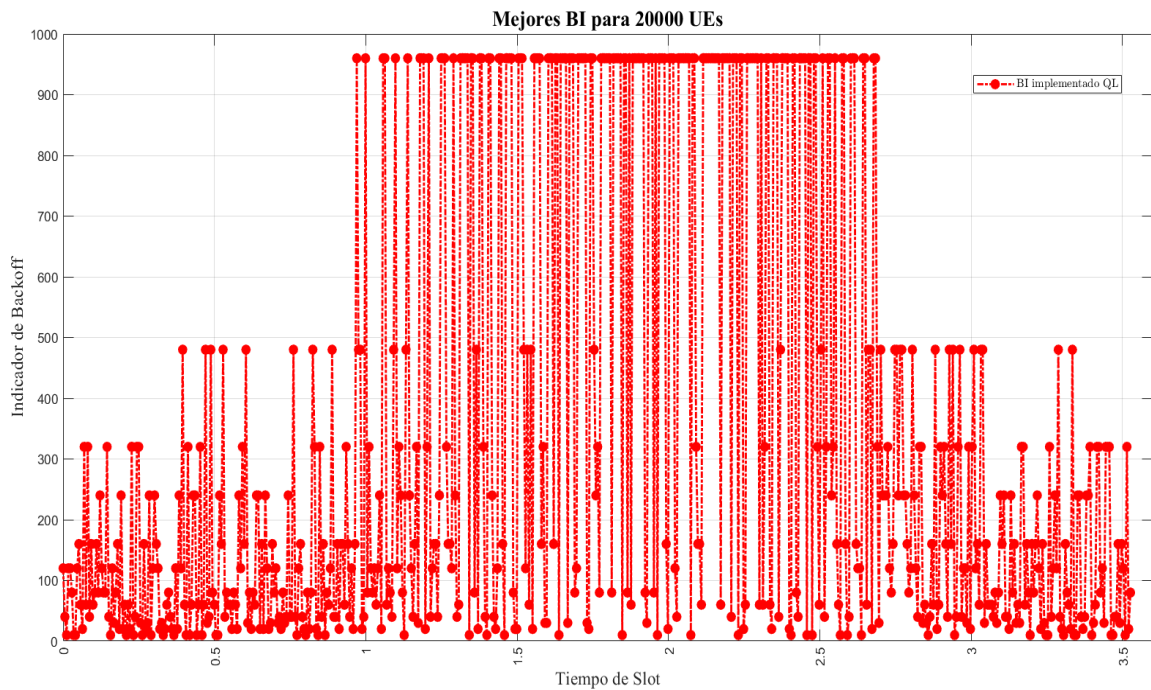
*Figura 11* Implementación del algoritmo de QL, obtención de los mejores BI para una baja carga de tráfico.

La adaptación del algoritmo QL permite mejorar el rendimiento del RACH al ajustar correctamente el valor de BI en función del número de dispositivos. En este sentido, al tener valores constantes de BI, indica que todos los dispositivos tienen acceso a la red, lo que sugiere que para el escenario de tráfico cualquier valor de BI resulta efectivo. Sin embargo, la presencia de un valor alto de BI, como 960 ms, demuestra que el algoritmo QL prioriza la confiabilidad en el acceso del canal de comunicación en situaciones de riesgo de colisión, lo que es esencial para prevenir la pérdida de datos y el retraso en la comunicación.

En la Figura 12 se ilustra los resultados de los diferentes valores de BI cuando la demanda de acceso al canal es media (20000 UEs). Los valores de BI representan estrategias para gestionar el acceso al canal compartido y mejorar la eficiencia, los valores de 960 ms señalan configuraciones óptimas que logra un rendimiento excepcional en el acceso, mientras que

los valores inferiores hasta el mínimo valor constante de 20ms. indican intentos de explorar y ajustar configuraciones para equilibrar el rendimiento y la estabilidad en el entorno.

Este análisis, basado en un algoritmo de aprendizaje QL, busca optimizar la calidad de la comunicación al ajustar los valores de BI de manera dinámica. Además, se observa que para esta carga de tráfico todos los dispositivos tienen un acceso exitoso lo que cumple y supera la condición de fiabilidad por lo que se podría considerar a cualquier valor de BI.



**Figura 12** Implementación del algoritmo de QL, obtención de los mejores BI para una carga de tráfico media.

El valor de BI puede aumentar durante periodos de alta congestión o de mayor demanda como se muestra en la Figura 13, llegando a su valor máximo constante de 960 ms. Esta medida asegura un acceso más confiable al canal y reduce las colisiones de transmisión.

Por tal razón, el valor óptimo de BI que cumple con la condición de fiabilidad para una alta carga de tráfico es de 960 ms, lo que se considera como un punto equilibrio óptimo entre la eficiencia y la confiabilidad. Aunque valores de BI más bajos podrían acelerar el acceso al canal y comprometer la confiabilidad en un entorno de alta congestión. Por otro lado, valores de BI más altos, como 960 ms mencionados, aseguran la confiabilidad a costa de una ligera demora en el acceso.

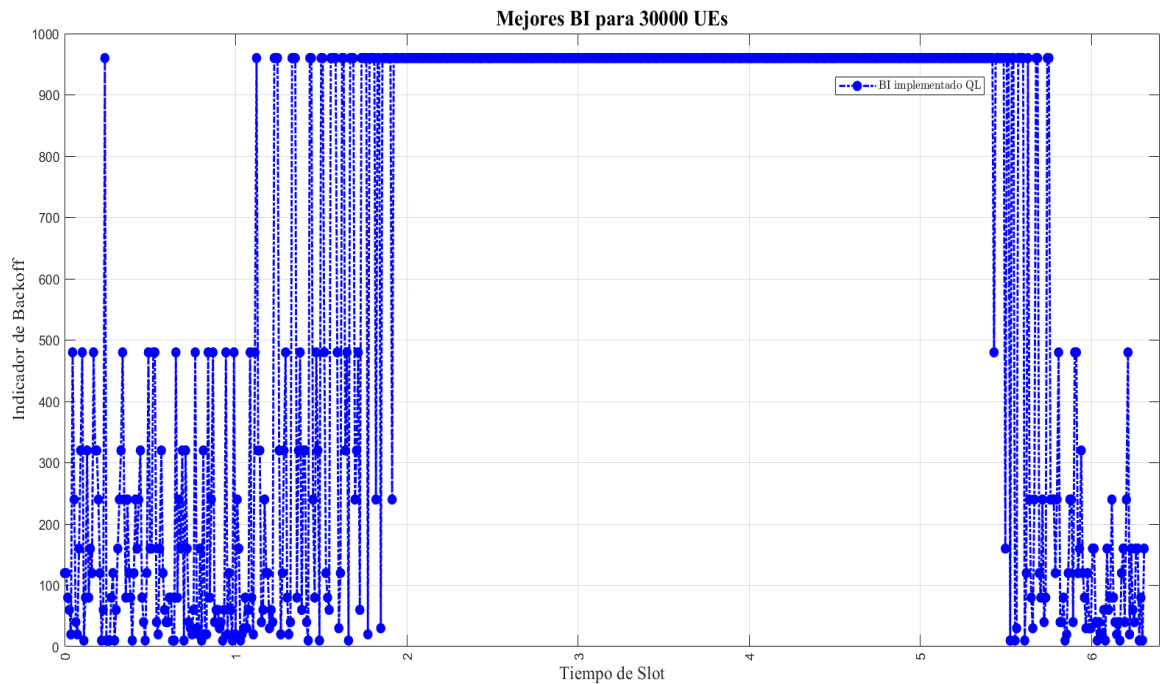


Figura 13 Implementación del algoritmo de QL, obtención de los mejores BI para una alta carga de tráfico.

### 4.3 Análisis estadístico de las métricas de rendimiento

Para el análisis comparativo de las métricas de rendimiento en diferentes condiciones de tráfico masivo, se llevará a cabo en función del parámetro del indicador de *backoff*. La primera configuración se basa en un valor de BI estático, mientras que la segunda configuración se apoya en un valor de BI óptimo aplicando el algoritmo QL. Se realizaron varias simulaciones para determinar las medianas de las diferentes métricas de rendimiento y se llevaron a cabo pruebas de hipótesis en cada una de ellas.

Tabla 12 Resumen del procesamiento de los casos

Tipo de Algoritmo	Caso	Casos					
		Válido		Perdidos		Total	
		N	Porcentaje	N	Porcentaje	N	Porcentaje
Métricas de rendimiento (Ps-k-D)	SQL	100	100,0%	0	0,0%	100	100,0%
	CQL	100	100,0%	0	0,0%	100	100,0%

La Tabla 12 muestra los resultados obtenidos para los grupos: sin aplicar el algoritmo de aprendizaje Q (SQL) y aplicando el algoritmo de aprendizaje Q (CQL), lo que indica un conjunto de datos completos y válidos para ambos grupos. La igualdad en el tamaño de la muestra y la validez de los datos facilita el análisis estadístico.

### 4.3.1 Prueba de hipótesis para la probabilidad de acceso exitoso (Ps).

#### 4.3.1.1 Diagrama de cajas

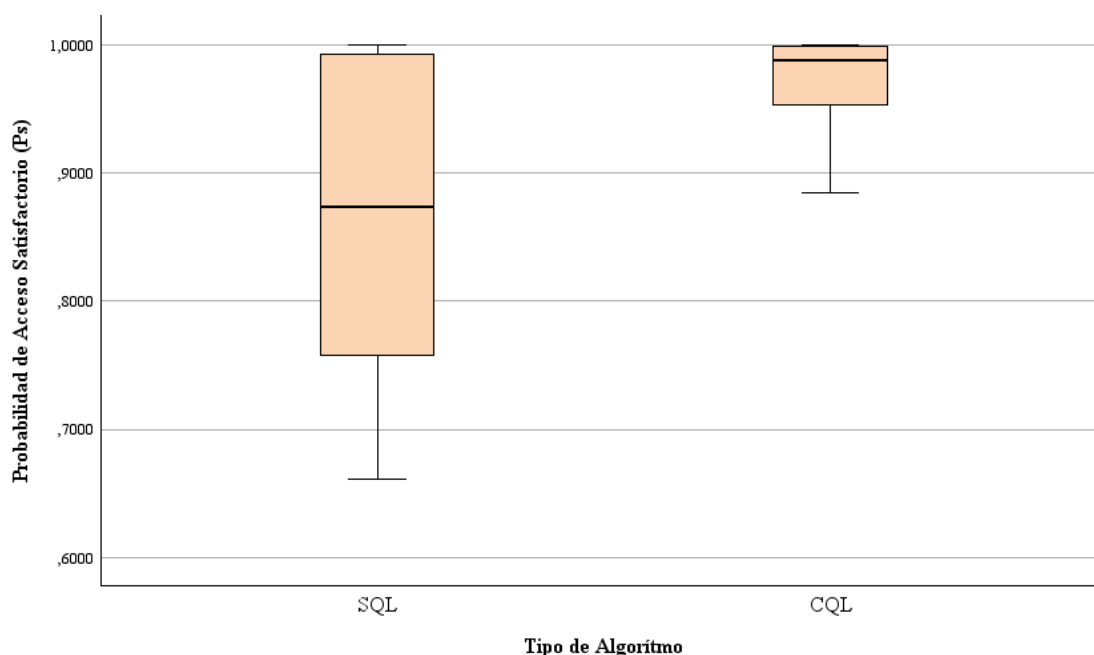


Figura 14 Diagrama de Cajas de la probabilidad de acceso exitoso en función de las dos configuraciones.

En la Figura 14 se puede observar el diagrama de cajas correspondiente al análisis de la Ps en diferentes situaciones de tráfico. Para el grupo CQL la mediana es de 0,9882, mientras que para el grupo SQL, la mediana es de 0,8736. Esto sugiere que, en promedio CQL presenta un mejor rendimiento en términos de probabilidad de acceso exitoso en condiciones de tráfico masivo.

En el grupo SQL, la probabilidad de acceso exitoso varía desde un mínimo de 0,6614 hasta un máximo de 1,00, lo que implica una mayor variabilidad en los datos. Esto significa que en condiciones de altas cargas de tráfico (30000 UEs), SQL tiene un rendimiento más bajo. En el caso de CQL, la variabilidad en los datos es menor, ya que la probabilidad varía desde un mínimo de 0,8849 y un máximo de 1,00. Esto sugiere que CQL tiene una probabilidad de acceso exitoso más consistente en comparación con SQL.

#### 4.3.1.2 Análisis Descriptivo

Tabla 13 Resultados estadísticos descriptivos de la métrica Ps

Tipo de Algoritmo	Estadístico	Error estándar
SQL Media	,865	,011

Probabilidad de Acceso exitoso (Ps)	95% de intervalo de confianza para la media	Límite inferior	,842	
		Límite superior	,888	
	Mediana	,873		
	Desv. Estándar	,116		
	Mínimo	,661		
	Máximo	1,000		
	Media	,972		,003
	95% de intervalo de confianza para la media	Límite inferior	,966	
		Límite superior	,979	
	CQL	Mediana	,988	
		Desv. Estándar	,033	
		Mínimo	,884	
		Máximo	1,000	

En el análisis de la probabilidad de acceso exitoso (Ps) para los algoritmos SQL y CQL, se observa en la Tabla 13 que el algoritmo CQL registra una media de 0,972 en comparación con SQL que es de 0,865, esto indica que se tiene un mejor rendimiento promedio por parte de CQL. Además, CQL presenta una menor variabilidad con una desviación estándar de aproximadamente 0,033, mientras que SQL tiene una desviación estándar de 0,116, lo que sugiere una mayor uniformidad en los resultados de CQL. En términos de Probabilidad de Acceso exitoso CQL muestra un rendimiento superior y una variabilidad de los datos menor en comparación con SQL.

#### 4.3.1.3 Pruebas de normalidad

Para llevar a cabo un estudio preciso, es esencial analizar los datos y determinar si los valores son paramétricos o no paramétricos, con el fin de garantizar que se utilicen las pruebas de estadística adecuadas y asegurar una interpretación precisa y confiable de los resultados. En este caso se busca incrementar la probabilidad de acceso exitoso con altas cargas de tráfico lo que implica la comparación de los resultados en dos escenarios SQL y CQL.

#### Hipótesis nula

$H_0$ : Los datos siguen una distribución normal.

#### Hipótesis Alternativa

$H_1$ : Los datos no siguen una distribución normal.

**Tabla 14** Resultados de las pruebas de Normalidad de la métrica Ps

Tipo de Algoritmo	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Probabilidad de Acceso exitoso (Ps) SQL	,144	100	<,001	,886	100	<,001
CQL	,223	100	<,001	,791	100	<,001

En el contexto de la Probabilidad de Acceso exitoso (Ps), la Tabla 14 revela que tanto el estadístico Kolmogorov-Smirnov como el estadístico Shapiro-Wilk indican un p-valor <0,001 para los algoritmos SQL y CQL. Lo que implica que la hipótesis nula ( $H_0$ ) se rechaza y en su lugar se acepta la hipótesis alternativa ( $H_1$ ), donde indica que los datos no siguen una distribución normal. Por consiguiente, en la investigación, se empleará la prueba de Mann-Whitney para comparar medianas.

#### 4.3.1.4 Prueba de Mann-Whitney U

La prueba de Mann-Whitney es un método no paramétrico utilizado para comparar las medianas de dos conjuntos de datos independientes. En relación con los datos de los rangos que se encuentra en la Tabla 15, esta prueba se utiliza para evaluar si existe una diferencia significativa entre las distribuciones de los datos de probabilidad de acceso exitoso para los dos algoritmos, ya sean SQL y CQL. Para abordar este propósito se plantea:

##### Hipótesis Nula

$$H_0: med_1=med_2$$

El promedio de la probabilidad de acceso exitoso SQL es significativamente igual al promedio de la probabilidad de acceso exitoso CQL.

##### Hipótesis alternativa

$$H_1: med_1 \neq med_2$$

El promedio de la probabilidad de acceso exitoso SQL es significativamente diferente al promedio de la probabilidad de acceso exitoso CQL.

*Tabla 15 Resultado comparativo de la prueba U para cada tipo de algoritmo*

	Tipo de Algoritmo	N	Rango promedio	Suma de rangos
Probabilidad de Acceso Exitoso (Ps)	SQL	100	77,740	7774,000
	CQL	100	123,260	12326,000
	Total	200		

Los rangos promedio y la suma de rangos son componentes que se utilizan para calcular estadísticas de prueba y determinar si las medianas de los dos grupos son diferentes.

*Tabla 16 Resultado comparativo de la prueba de Mann-Whitney*

	Probabilidad de acceso exitoso (Ps)
U de Mann-Whitney	2724,000
W de Wilcoxon	7774,000
Z	-5,568
Sig. asin. (bilateral)	<,001

Como se evidencia en las Tabla 16, la hipótesis nula ( $H_0$ ). Dado que los resultados de la prueba de Mann-Whitney indican una diferencia estadísticamente significativa ( $p$ -valor < 0,001) entre los dos grupos, se rechaza la hipótesis nula ( $H_0$ ). En consecuencia, se acepta la hipótesis alternativa ( $H_1$ ), lo que implica que el promedio de la probabilidad de acceso SQL es significativamente distinto al promedio de la probabilidad de acceso exitoso CQL. Con este resultado se puede corroborar que al implementar el algoritmo de QL se tiene un incremento en la probabilidad de acceso éxito con comunicaciones masivas.

#### **4.3.2 Prueba de hipótesis para el Número promedio de preámbulos transmitidos (k).**

En este caso se busca disminuir el número de preámbulos transmitidos con altas cargas de tráfico lo que implica la comparación de los resultados en dos escenarios SQL y CQL. Para abordar este propósito se plantea:

##### **Hipótesis nula**

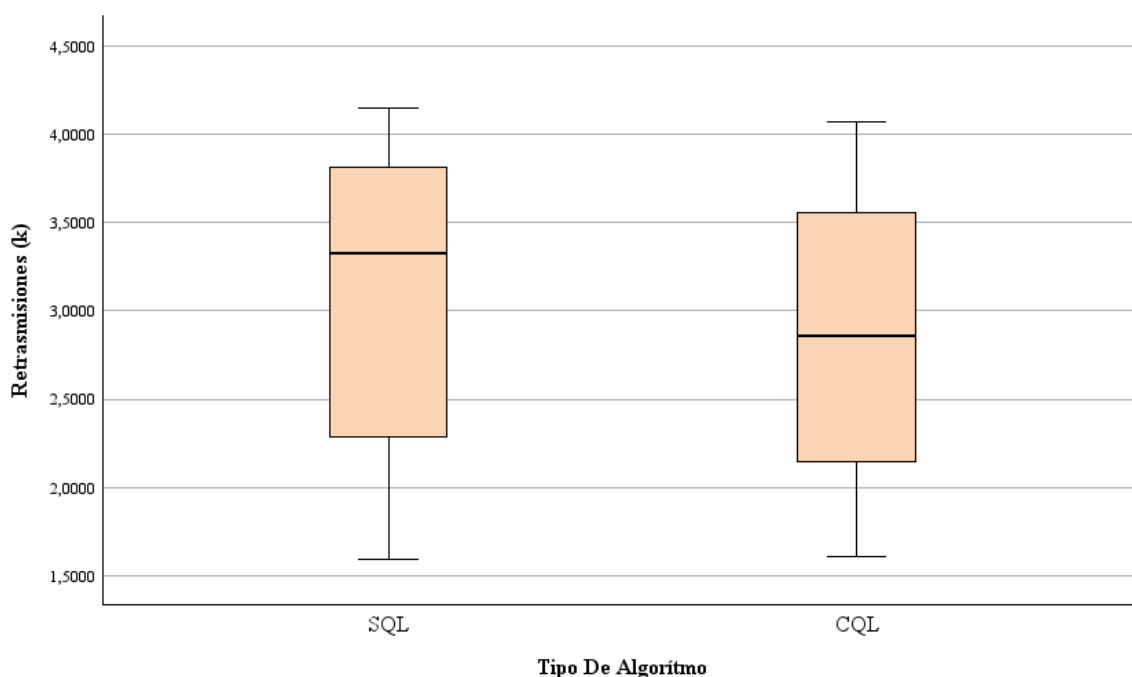
$H_0$ : Los datos siguen una distribución normal.

##### **Hipótesis Alternativa**

$H_1$ : Los datos no siguen una distribución normal.



### 4.3.2.1 Diagrama de cajas



*Figura 15* Diagrama de Cajas del número de retransmisiones en función de las dos configuraciones.

En la Figura 15 presenta un diagrama de cajas que ilustra el análisis del número promedio de preámbulos transmitidos (k) en diversas condiciones de tráfico masivo. Para el grupo CQL, se observa una mediana de 2,861, en contraste con el grupo SQL, que muestra una mediana de 3,329. Esto señala que, en promedio CQL presenta un menor número de preámbulos transmitidos en condiciones de tráfico masivo. Esta reducción en el número de transmisiones sugiere una mejora en el rendimiento del acceso, ya que se logra una eficiencia mayor al transmitir con éxito utilizando menos recursos de preámbulo.

En cuanto a los valores externos, en el grupo SQL, el número promedio de transmisiones varía desde un mínimo de 1,597 hasta un máximo de 4,150, mientras que en el caso de CQL, la variación abarca desde un mínimo de 1,611 y un máximo de 4,072. Esto respalda la idea de que CQL tiene un menor número de preámbulos transmitidos en comparación con SQL, lo que contribuye a mejorar la eficiencia y el rendimiento del acceso en situaciones de tráfico masivo.

### 4.3.2.2 Análisis Descriptivo

*Tabla 17* Resultados estadísticos descriptivos de la métrica k.

Tipo De Algoritmo	Estadístico	Error estándar
SQL Media	3,068	,086

Número promedio de transmisiones (k)	95% de intervalo de confianza para la media	Límite inferior	2,897		
		Límite superior	3,240		
	Mediana		3,329		
	Desv. Estándar		,865		
	Mínimo		1,597		
	Máximo		4,150		
	Media		2,840	,078	
	95% de intervalo de confianza para la media	Límite inferior	2,684		
		Límite superior	2,995		
	CQL	Mediana		2,861	
		Desv. Estándar		,783	
		Mínimo		1,611	
		Máximo		4,072	

Los datos ofrecen una visión detallada de la distribución del número promedio de transmisiones (k) en ambos algoritmos. Para SQL, la media es de 3,068 con un intervalo de confianza del 95% entre 2,897 y 3,240. La mediana es 3,329, y la varianza es 0,748, lo que resulta en una desviación estándar de 0,865. Los valores varían entre 1,597 y 4,150. En el caso de CQL, la media es de 2,840 con un intervalo de confianza entre 2,684 y 2,995. La mediana es 2,861, y la varianza es 0,614, con una desviación estándar de 0,783. Los valores oscilan entre 1,611 y 4,072. La Tabla 17 proporciona información detallada sobre el número promedio de transmisiones (k) en los algoritmos SQL y CQL, lo que es fundamental para comprender sus diferencias en situaciones de tráfico masivo.

#### 4.3.2.3 Pruebas de normalidad

*Tabla 18 Resultados de las pruebas de Normalidad de la métrica k*

	Tipo De Algoritmo	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Estadístico	Gl	Sig.	Estadístico	gl	Sig.
N transmisiones (k)	SQL	,138	100	<,001	,877	100	<,001
	CQL	,080	100	,114	,936	100	<,001

Los resultados de las pruebas de normalidad realizadas para k en los algoritmos SQL y CQL se muestran en la Tabla 18. En el caso SQL, el estadístico de Kolmogorov-Smirnov presentan valores bajos de significancia, inferior a 0,001, indicando que los datos en la muestra de SQL no siguen una distribución normal. Para el caso de CQL, el estadístico de

Kolmogorov-Smirnov es bajo 0,080 con un valor de significancia de 0,114. Esto sugiere que los datos en la muestra de CQL podría seguir una distribución normal. Estos hallazgos tienen implicaciones importantes para el análisis estadístico, ya que se puede requerir el uso de pruebas no paramétricas debido a la falta de normalidad en los datos. En este caso se utilizará la prueba de MANN-Whitney U.

#### 4.3.2.4 Prueba de Mann-Whitney U

Los datos resumen la distribución de las transmisiones promedio y permiten realizar comparaciones entre los algoritmos. Por lo que se plantea:

##### Hipótesis Nula

$$H_0: med_1 = med_2$$

El número promedio de preámbulos transmitidos SQL es significativamente igual al número promedio de preámbulos transmitidos CQL.

##### Hipótesis alternativa

$$H_1: med_1 \neq med_2$$

El número promedio de preámbulos transmitidos SQL es significativamente diferente al número promedio de preámbulos transmitidos CQL.

*Tabla 19 Resultado comparativo de la prueba U para cada tipo de algoritmo*

	Tipo De Algoritmo	N	Rango promedio	Suma de rangos
Número promedio de transmisiones (k)	SQL	100	108,980	10898,000
	CQL	100	92,020	9202,000
	Total	200		

La Tabla 19 presenta información sobre el número promedio de transmisiones (k) en los algoritmos SQL y CQL. En el caso de SQL, se tiene 100 datos con un rango promedio de 108.980 y una suma de rangos de 10,898. CQL, por su parte, también contó con 100 datos, registrando un rango promedio de 92,020 y una suma de rangos de 9,202. El total de datos entre ambos algoritmos es de 200.

*Tabla 20 Resultado comparativo de la prueba de Mann-Whitney*

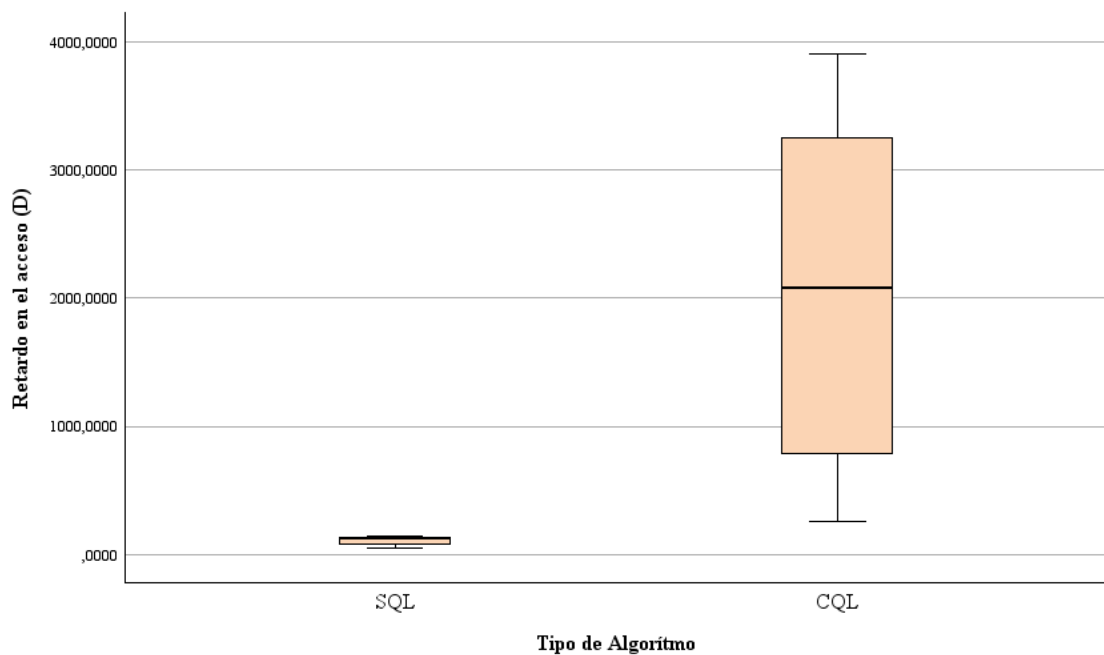
	Número promedio de transmisiones (k)
U de Mann-Whitney	4152,000

W de Wilcoxon	9202,000
Z	-2,072
Sig. asin. (bilateral)	,038

Los resultados de la prueba estadística revelan un p-valor de 0,038, que es menor al nivel de significancia comúnmente establecido de 0,050 en las pruebas estadísticas como se observa en la Tabla 20. Por lo tanto, se rechaza la hipótesis nula ( $H_0$ ) que afirmaba que los dos grupos tenían números promedio de transmisiones idénticos, en favor de la hipótesis alternativa ( $H_1$ ), que indica una diferencia significativa en k entre los dos grupos, SQL y CQL. Estos resultados respaldan la idea de que al implementar QL ofrece una mayor eficiencia en la transmisión exitosa, utilizando menos recursos de transmisión, lo cual es crucial en situaciones de alto tráfico.

### 4.3.3 Prueba de Hipótesis para el Retardo en el acceso (D).

#### 4.3.3.1 Diagrama de cajas



*Figura 16* Diagrama de Cajas del retardo en el acceso en función de dos configuraciones.

En la Figura 16 podemos observar el diagrama de cajas correspondiente al análisis del retardo en el acceso en diferentes condiciones de tráfico masivo. Para el grupo CQL la mediana es de 2082,600, mientras que para el grupo SQL, la mediana es de 129,250. Esto

sugiere que, en promedio CQL presenta un mayor retardo en el acceso en condiciones de tráfico masivo.

En el grupo SQL, el retardo en el acceso se encuentra en un rango que va desde un mínimo de 49,400 hasta 141,400 que representa el valor máximo. Por otro lado, en el grupo CQL, la variación del retardo es más significativa, abarcando desde un valor mínimo de 255,300 hasta 3,905,800 como un valor máximo. Estos datos indican claramente que, en términos de minimizar el retardo, el grupo SQL presenta una ventaja, ya que sus valores de retardo son considerablemente más bajos en comparación con el grupo CQL. No obstante, es importante destacar que la mejora de la eficiencia y el rendimiento en el acceso viene acompañada de un retardo significativo.

#### 4.3.3.2 Análisis Descriptivo

*Tabla 21 Resultados estadísticos descriptivos de la métrica D*

Tipo de Algoritmo		Estadístico	Error estándar	
Retardo en el acceso (D)	SQL	Media	111,374	
	95% de intervalo de confianza para la media	Límite inferior	104,843	
		Límite superior	117,904	
	Mediana	129,250		
	Desv. Estándar	32,911		
	Mínimo		49,400	
		Máximo	141,400	
	CQL	Media	2029,114	128,370
		95% de intervalo de confianza para la media	Límite inferior	1774,398
			Límite superior	2283,829
		Mediana	2082,600	
		Desv. Estándar	1283,707	
		Mínimo	255,300	
		Máximo	3905,800	

La Tabla 21 presenta datos relacionados con el retardo en el acceso para dos algoritmos, SQL y CQL. Para SQL, la media del retardo es de 111,374, con un intervalo de confianza del 95% entre 104,843 y 117,904. CQL, por otro lado, tiene una media de retardo de 2029,114, con un intervalo de confianza entre 1774,398 y 2283,829. Además, se proporciona

información adicional, como la mediana, varianza, desviación estándar, valores mínimo y máximo, entre otros, para ambas muestras, lo que ofrece una visión completa de la distribución de los datos.

#### 4.3.3.3 Pruebas de normalidad

Para abordar este propósito se plantea:

##### Hipótesis nula

H<sub>0</sub>: Los datos siguen una distribución normal.

##### Hipótesis Alternativa

H<sub>1</sub>: Los datos no siguen una distribución normal

*Tabla 22 Resultados de las pruebas de Normalidad de la métrica D*

	Tipo de Algoritmo	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Estadístico	gl	Sig.	Estadístico	gl	Sig.
Retardo en el acceso (D)	SQL	,239	100	<,001	,775	100	<,001
	CQL	,114	100	,003	,896	100	<,001

Los resultados de las pruebas de normalidad Kolmogorov-Smirnov realizadas en las muestras de los algoritmos SQL y CQL con respecto a la variable del Retardo en el acceso (D), tal como se detalla en la Tabla 22, indican que ninguna de las muestras sigue una distribución normal. En ambas muestras, los valores de significancia (Sig.) son bajos, lo que sugiere que se rechaza la H<sub>0</sub> ya que los datos no cumplen con el supuesto de normalidad. Esto tiene implicaciones importantes para el análisis estadístico, lo que podría requerir el uso de pruebas no paramétricas. En este sentido se utilizó la comparación de medianas aplicando la prueba de Mann-Whitney U.

#### 4.3.3.4 Prueba de Mann-Whitney U

Para realizar la prueba se toma en cuenta dos tipos de algoritmos SQL y CQL. El objetivo de la evaluación consiste en determinar si hay similitud o diferencia entre estos dos tipos de algoritmos. El proceso de comparar las medianas de algoritmos SQL y CQL, se define la hipótesis nula y la hipótesis alternativa de la siguiente forma:

##### Hipótesis Nula

$$H_0: med_1 = med_2$$

El promedio del retardo en el acceso SQL es significativamente igual al promedio del retardo en el acceso CQL.

### Hipótesis alternativa

$$H_1: med_1 \neq med_2$$

El promedio del retardo en el acceso SQL es significativamente diferente al promedio del retardo en el acceso CQL.

*Tabla 23 Resultado comparativo de la prueba U para cada tipo de algoritmo*

	Tipo de Algoritmo	N	Rango promedio	Suma de rangos
Retardo en el acceso (D)	SQL	100	50,500	5050,000
	CQL	100	150,500	15050,000
	Total	200		

La Tabla 23 proporciona datos relacionados con la variable del retardo en el acceso en dos tipos de algoritmos. Cada grupo consta de 100 datos en cada caso de SQL, el rango promedio es de 50,500 y la suma de los rangos es de 5050,000. En total, se tiene 200 datos en tabla, los valores indican diferencias en los retardos de acceso promedio entre los algoritmos con SQL y CQL

*Tabla 24 Resultado comparativo de la prueba de Mann-Whitney*

	Retardo en el acceso (D)
U de Mann-Whitney	,000
W de Wilcoxon	5050,000
Z	-12,217
Sig. asin. (bilateral)	<,001

Los resultados de la prueba de Mann-Whitney para la variable del Retardo en el acceso (D) indican una diferencia estadísticamente significativa entre los dos tipos de algoritmos (SQL y CQL). Lo que implica que existe diferencias estadísticas en las medianas. En base a las hipótesis planteadas, los resultados de la prueba de Mann-Whitney detallados en la Tabla 24, indica que se debe rechazar la hipótesis nula ( $H_0$ ) y aceptar la hipótesis alternativa ( $H_1$ ). Esto se debe a que el valor de significancia es menor que 0.001 ( $p$ -valor < 0,001), lo que respalda la evidencia estadística de que el promedio de retardo en el acceso SQL es distinto al promedio de retardo en el acceso CQL. Con estos resultados se puede confirmar que al implementar QL conlleva una notable mejora en la eficiencia y el rendimiento en el acceso.

Sin embargo, es importante tener en cuenta que esta mejora va de la mano con un retardo significativo.

#### **4.4 Discusión de los Resultados**

En la sección 4.1, se realizó la evaluación de las métricas de rendimiento en varios escenarios de tráfico (10000, 20000, 30000) utilizando una configuración de PRACH=14. En cuanto a la probabilidad de acceso exitoso ( $P_s$ ), se analizó su rendimiento con valores de BI estáticos. En un escenario con un valor de BI fijo de 960 ms y una alta carga de tráfico de 30000 UEs, se observó una disminución en la  $P_s$  al 89%. Sin embargo, en situaciones de tráfico medio (20000) y bajo (10000), la  $P_s$  alcanzó el 99% y el 100%, respectivamente. Esto sugiere que la mayoría de los dispositivos logran acceder a la red.

En relación al número promedio de preámbulos transmitidos ( $k$ ) y al retardo de acceso con percentil 95 ( $D_{95}$ ) para cargas de tráfico baja, media y alta, se observó una disminución en comparación con los valores de BI menores a 960ms. Esto indica que el sistema mantiene una eficiencia adecuada y tiempos de acceso razonables incluso en condiciones de tráfico masivo.

Sin embargo, la implementación del algoritmo de QL para minimizar la probabilidad de colisiones y lograr mejoras notables en la  $P_s$  se mantuvo al 90% en el escenario de tráfico alto y el 100% en los escenarios de tráfico medio y bajo. Esto sugiere que el valor de BI optimizado por el algoritmo de QL permite una mayor eficiencia en términos de retransmisiones y retardo de acceso, mejorando así la calidad y velocidad de la comunicación para los dispositivos. Esto se debe a la capacidad del algoritmo de QL al ajustar el valor del BI hasta encontrar un valor que le permita acceder al canal. Esto se basa en el uso de recompensas para actualizar el valor de BI, incrementando cada vez que el dispositivo recibe una recompensa negativa y disminuyendo cada vez que el dispositivo recibe una recompensa positiva.

En la Sección 4.2 se examina los resultados de los mejores valores de BI con la adaptación del algoritmo de QL. Para la carga de tráfico baja (10000), QL demostró mejorar el rendimiento, al priorizar el acceso a la red mediante la elección de un valor de BI pequeño para reducir el tiempo de espera para transmitir los datos, lo que sugiere que para este



escenario de tráfico cualquier valor de BI resulta efectivo. En un escenario de carga media (20000), se experimentan variaciones en los valores de BI encontrados, lo que refleja la exploración de diversas configuraciones para equilibrar el rendimiento y la estabilidad. No obstante, se logró un acceso exitoso para todos los dispositivos, superando la condición de fiabilidad. Y en el caso de altas cargas de tráfico (30000), el valor de BI aumenta a 960 ms para garantizar un acceso más confiable, la elección de este valor de BI permite la reducción de colisiones permitiéndole al agente mejorar la probabilidad de acceso exitoso, aunque esto implicara una ligera demora en el acceso.

## CAPÍTULO V CONCLUSIONES Y RECOMENDACIONES

### 5.1 Conclusiones

El diseño del algoritmo adaptativo de *backoff* basado en aprendizaje reforzado demostró su efectividad para incrementar la probabilidad de acceso exitoso. Su capacidad adaptativa permite aprender y ajustar de manera dinámica la política de *backoff* según las cambiantes condiciones de la red, lo que resulta en una eficiencia al utilizar un agente de aprendizaje reforzado. Este algoritmo ha mostrado un rendimiento destacado en diferentes niveles de carga de tráfico, priorizando la accesibilidad en situaciones de baja carga y equilibrando el rendimiento y la estabilidad en cargas medianas, además de garantizar un acceso confiable en condiciones de alta carga de tráfico.

Se estudió el procedimiento de *backoff* como una técnica eficaz para mitigar el problema de las colisiones en el RACH bajo condiciones de comunicación masiva generada por dispositivos IoT. El procedimiento de *backoff* funciona retrasando el siguiente intento de un dispositivo que no pudo acceder al canal. Esto reduce la probabilidad de que dos dispositivos intenten acceder al canal al mismo tiempo. En este estudio se observó la mejora del procedimiento de *backoff* modelado, permitiendo que los dispositivos retrasen su siguiente intento, lo que resulta en una reducción significativa en la probabilidad de colisiones.

Se implementó un algoritmo de *backoff* adaptativo basado en aprendizaje Q para aprender la política óptima de *backoff*. Mejorando el rendimiento del dispositivo usando las recompensas para aprender a esperar más tiempo en los estados en los que las solicitudes tienen más probabilidades de fallar. El dispositivo en este algoritmo actualiza sus estimaciones del valor de cada acción, en función de las recompensas que recibió, demostrando que aumenta la probabilidad de que tome la acción óptima en cada estado.

Se evaluó a la probabilidad de acceso exitoso  $P_s$ , considerando como el principal indicador de rendimiento (KPI), los resultados mostraron que la optimización del tiempo de *backoff*, facilitada por el algoritmo de aprendizaje por refuerzo QL, revelando una mejora en la calidad y la eficiencia de la comunicación. En particular, se observó que la implementación de QL para ajustar dinámicamente el tiempo de *backoff* permitió mejorar la  $P_s$  en un 90% en el escenario de tráfico alto y en un 100% en los escenarios de tráfico medio y bajo. Esto

sugiere que el valor de BI optimizado por el algoritmo de QL permite una mayor eficiencia, garantizando así un acceso exitoso con escenarios de tráfico masivo.

En conclusión, mantener un valor de tiempo de *backoff* estático puede resultar en una disminución significativa de la probabilidad de acceso exitoso (Ps) al 66% junto con un aumento en el número promedio de transmisiones y el retardo de acceso. Por lo que un valor de *backoff* estático no es suficiente para garantizar el acceso al canal en situaciones de alta carga de tráfico (30000). Sin embargo, al aplicar el algoritmo de aprendizaje por refuerzo (QL), se puede ajustar dinámicamente el tiempo de *backoff* en función de las condiciones del canal. Esto permite al dispositivo encontrar un valor de *backoff* que le proporcione la mayor probabilidad de acceso exitoso. En este estudio, se observó que la aplicación del algoritmo de aprendizaje Q resultó en una mejora notable de la Ps al 90%, así como una reducción en las retransmisiones y el retardo de acceso. Esto se debe a que el algoritmo QL es capaz de aprender de sus experiencias y adaptarse a las condiciones cambiantes del canal.

## **5.2 Recomendaciones**

Se recomienda que el algoritmo de *backoff* debe comenzar con una política inicial antes de comenzar a aprender. Los métodos de inicialización pueden incluir políticas aleatorias basadas en el conocimiento previo. Además, se debe implementar una estrategia de exploración que determine con qué frecuencia el algoritmo debe tomar decisiones exploratorias en lugar de seguir su política actual. Esto es crucial para garantizar que el algoritmo explore nuevas estrategias y evite quedarse atascado.

El algoritmo de *backoff* debe recibir información sobre el estado del entorno para tomar decisiones. Esto puede incluir datos como la carga actual de la red, el número de dispositivos en espera, la congestión actual y la probabilidad de colisión. Centrándose en el diseño de una representación adecuada de estas observaciones para el rendimiento del algoritmo de *backoff*.

Para un futuro trabajo de investigación se recomienda considerar un proceso de evaluación y validación continua del algoritmo adaptativo de *backoff* en un entorno de pruebas realistas. Esto implica la recolección de datos de rendimiento en condiciones de tráfico masivo y la comparación de los resultados con métricas de referencia. Además, sería importante la evaluación continua para garantizar que el algoritmo siga siendo eficaz a medida que las condiciones de la red y los patrones de tráfico evolucionan con el tiempo.

## BIBLIOGRAFÍA

- [1] L. Tello-Oquendo, J. R. Vidal, V. Pla y J. Martinez-Bauset, «Restricción de Acceso Extendida para Manejar Despliegues Masivos de Comunicaciones Tipo Máquina (mMTC),» *Novasinerгия*, ISSN 2631-2654, vol. 1, n° 2, pp. 38-44, 2018.
- [2] J.-R. Vidal, L. Tello-Oquendo, V. Pla y L. Guijarro, «Performance study and enhancement of access barring for massive machine-type communications,» *IEEE access: practical innovations, open solutions*, vol. 7, pp. 63745-63759, undefined 2019.
- [3] N. A. Shinkafi, L. M. Bello, D. S. Shu'aibu y P. Mitchell, «Learning automata based Q-learning RACH access scheme for cellular M2M communications,» de *2019 IEEE Global Conference on Internet of Things (GCIoT)*, 2019.
- [4] A. S. A. El-Hameed y K. M. F. Elsayed, «A Q-learning approach for machine-type communication random access in LTE-Advanced,» *Telecommunication systems*, vol. 71, n° 3, pp. 397-413, 2019.
- [5] D. A. Santander González, «Evaluación y caracterización del impacto del tiempo de backoff en la red de acceso radio de IoT celular con tráfico masivo,» Universidad Nacional de Chimborazo, Riobamba, 2021.
- [6] L. Tello-Oquendo, D. Pacheco-Paramo, V. Pla y J. Martinez-Bauset, «Reinforcement learning-based ACB in LTE-A networks for handling massive M2M and H2H communications,» de *IEEE International Conference on Communications (ICC)*, pp. 1-7, 2018.
- [7] J. Aranda, F. Astudillo-Salinas, D. Santander y K. Cicenía-Cárdenas, «Impacto del tiempo de backoff en el rendimiento de IoT celular en entornos de comunicación masiva,» *Novasinerгия, Revista Digital de ciencia, Ingenieria y Tecnología*, vol. 5, n° 1, pp. 17-30, 2022.
- [8] 3GPP TR 37.868, «RAN Improvements for Machine-type Communications,» 2022. [En línea]. Available: <https://itcspec.com/archive/3gpp-specification-tr-37-868/>.
- [9] A. H. Mohammed, A. S. Khwaja, A. Anpalagan y I. Woungang, «Base station selection in M2M communication using Q-learning algorithm in LTE-A networks,» de *IEEE 29th International Conference on Advanced Information Networking and Applications*, pp 17-22, 2015.

- [10] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare y J. Pineau, «An introduction to deep reinforcement learning,» *Foundations and Trends® in Machine Learning*, vol. 11, n° 3-4, pp. 219-354, 2018.
- [11] H. D. Althumali, M. Othman, N. K. Noordin y Z. M. Hanapi, «Dynamic Backoff collision resolution for massive M2M random access in cellular IoT networks,» *IEEE access: practical innovations, open solutions*, vol. 8, pp. 201345-201359, 2020.
- [12] D. Kim, W. Kim y S. An, «Adaptive random access preamble split in LTE,» *9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 814-819, 2013.
- [13] R.-H. Hwang, C.-F. Huang, H.-W. Lin y J.-J. Wu, «Uplink access control for machine-type communications in LTE-A networks,» *Personal and ubiquitous computing*, vol. 20, n° 6, pp. 851-862, 2016.
- [14] D. S. Team, «El aprendizaje de refuerzo y la importancia — Aprendizaje automático —,» DATA SCIENCE, 03 Mayo 2020. [En línea]. Available: <https://datascience.eu/es/aprendizaje-automatico/aprendizaje-automatico-para-humanos-parte-5-aprendizaje-de-refuerzo/>.
- [15] C. S. Moreno, «Evaluación de algoritmos de aprendizaje por refuerzo en un entorno Unity,» Universidad Politécnica de Cartagena, Cartagena, 2020.
- [16] J. I. Bagnato, «Aprendizaje por refuerzo,» Aprende Machine Learning, 2020. [En línea]. Available: <https://www.aprendemachinelearning.com/aprendizaje-por-refuerzo/>.
- [17] S. K. Sharma y X. Wang, «Collaborative distributed Q-learning for RACH congestion minimization in cellular IoT networks,» *IEEE communications letters: a publication of the IEEE Communications Society*, vol. 23, n° 4, pp. 600-603, 2019.
- [18] C. Rivera y H. Fernando, «Diseño de un mecanismo de generación de actualizaciones dinámico utilizando aprendizaje reforzado en comunicaciones IOT celular para optimizar la frescura de la información,» Riobamba, Universidad Nacional de Chimborazo, 2022. [En línea]. Available: <http://dspace.unach.edu.ec/handle/51000/9942>.
- [19] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik,

- I. Antonoglou, H. King y D. Kumaran, «Human-level control through deep reinforcement learning,» *Nature*, vol. 518, n° 7540, pp. 529-533, 2015.
- [20] 3GPP TS 36.331, «Radio Resource Control (RRC), Protocol specification,» 2017. [En línea]. Available: [https://www.3gpp.org/ftp/Specs/archive/36\\_series/36.331/](https://www.3gpp.org/ftp/Specs/archive/36_series/36.331/).
- [21] D. Pacheco-Paramo, L. Tello-Oquendo, V. Pla y J. Martinez-Bauset, «Deep reinforcement learning mechanism for dynamic access control in wireless networks handling mMTC,» *Ad hoc networks*, vol. 94, p. 101939, 2019.
- [22] D. Pacheco-Paramo y L. Tello-Oquendo, «Adjustable access control mechanism in cellular MTC networks: A double Q-learning approach,» *IEEE Fourth Ecuador Technical Chapters Meeting (ETCM)*, pp. 1-6, 2019.
- [23] N. A. Shinkafi, L. M. Bello, D. S. Shu'aibu y P. D. Mitchell, «Priority-based learning automata in Q-learning random access scheme for cellular M2M communications.,» *ETRI journal*, vol. 43, n° 5, pp. 4787-798, 2021.
- [24] L. Tello-Oquendo, I. Leyva-Mayorga, V. Pla, J. Martinez-Bauset, J.-R. Vidal, V. Casares-Giner y L. Guijarro, «Performance analysis and optimal access class barring parameter configuration in LTE-A networks with massive M2M traffic,» *IEEE transactions on vehicular technology*, vol. 67, n° 4, pp. 3505-3520, 2018.
- [25] 3GPP TS 36.211, «Physical Channels and Modulation,» vol. V15.7.0, 2019.
- [26] J. Moon y Y. Lim, «Access control of MTC devices using reinforcement learning approach,» *IEEE International Conference on Information Networking (ICOIN)*, pp. 641-643, 2017.

## ANEXOS

**Resultados correspondientes a las métricas de rendimiento con diferentes valores de BI estáticos y aplicando el algoritmo de QL en bajo condiciones de tráfico baja, media y alta.**

*Tabla 25 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 30 y valor de BI óptimo aplicando el algoritmo QL*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5882	60,5000	30	10000	1	1,6088	243,3000	160
12500	1	1,7484	70,9000	30	12500	0,9956	1,8077	321,4000	240
15000	0,9922	2,2941	111,7000	30	15000	0,9912	2,1575	775,2000	240
17500	0,9386	2,9532	153,5000	30	17500	0,9915	2,5023	1349,7000	40
20000	0,8745	3,3378	167,4000	30	20000	0,9958	2,8137	2116,5000	40
22500	0,8117	3,6447	175,3000	30	22500	0,9894	3,1077	2909,3000	20
25000	0,7585	3,8410	179,9000	30	25000	0,9752	3,4674	3518,0000	480
27500	0,7085	4,0040	182,5000	30	27500	0,9432	3,7928	3908,6000	20
30000	0,6653	4,1244	184,5000	30	30000	0,9594	4,0406	4165,0000	20

*Tabla 26 Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 40 y valor de BI óptimo aplicando el algoritmo QL*

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5884	72,1000	40	10000	1	1,6088	243,3000	160
12500	1	1,7483	85,0000	40	12500	0,9999	1,8002	304,5000	120
15000	0,9916	2,3104	138,2000	40	15000	0,9987	2,1429	774,5000	120
17500	0,9398	2,9438	187,9000	40	17500	0,9958	2,5180	1368,8000	240
20000	0,8753	3,3657	207,5000	40	20000	0,9918	2,8134	2106,9000	480
22500	0,8129	3,6446	217,3000	40	22500	0,9866	3,1014	2905,3000	60
25000	0,7583	3,8424	222,9000	40	25000	0,9700	3,4738	3510,9000	30
27500	0,7111	4,0015	225,5000	40	27500	0,9413	3,7917	3900,5000	80
30000	0,6659	4,1264	229,2000	40	30000	0,9059	4,0414	4168,5000	30

**Tabla 27** Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 60 y valor de BI óptimo aplicando el algoritmo QL

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5929	95,5000	60	10000	1	1,6106	248,4000	40
12500	1	1,7440	113,3000	60	12500	0,9999	1,7974	309,6000	60
15000	0,9938	2,2803	181,4000	60	15000	0,9988	2,1364	773,5000	80
17500	0,9429	2,9576	259,5000	60	17500	0,9956	2,5094	1358,9000	30
20000	0,8796	3,3631	286,7000	60	20000	0,9920	2,8071	2122,4000	30
22500	0,8165	3,6468	300,4000	60	22500	0,9864	3,1013	2899,5000	320
25000	0,7614	3,8465	307,7000	60	25000	0,9691	3,4735	3511,4000	160
27500	0,7133	4,0076	312,9000	60	27500	0,9414	3,7956	3901,4000	80
30000	0,6680	4,1403	316,7000	60	30000	0,9069	4,0559	4176,7000	480

**Tabla 28** Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 80 y valor de BI óptimo aplicando el algoritmo QL

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5887	118,1000	80	10000	1	1,6107	239,5000	320
12500	1	1,7418	142,5000	80	12500	0,9999	1,8035	308,0000	480
15000	0,9952	2,2414	223,2000	80	15000	0,9988	2,1383	768,6000	240
17500	0,9464	2,9584	329,8000	80	17500	0,9956	2,5154	1351,6000	120
20000	0,8811	3,3712	365,2000	80	20000	0,9915	2,8281	2127,5000	30
22500	0,8218	3,6342	382,0000	80	22500	0,9870	3,0883	2894,7000	120
25000	0,7641	3,8468	392,9000	80	25000	0,9698	3,4786	3527,0000	160
27500	0,7146	4,0107	400,7000	80	27500	0,9410	3,8017	3919,6000	240
30000	0,6705	4,1372	406,4000	80	30000	0,9053	4,0493	4166,9000	80



**Tabla 29** Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 120 y valor de BI óptimo aplicando el algoritmo QL

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5881	162,6000	120	10000	1	1,6030	250,1000	20
12500	1	1,7427	200,5000	120	12500	0,9999	1,8089	309,2000	40
15000	0,9964	2,2129	310,3000	120	15000	0,9987	2,1524	785,1000	30
17500	0,9532	2,9431	466,3000	120	17500	0,9956	2,5124	1372,0000	160
20000	0,8885	3,3571	518,8000	120	20000	0,9920	2,8279	2132,1000	60
22500	0,8272	3,6496	547,6000	120	22500	0,9867	3,0918	2896,6000	120
25000	0,7733	3,8483	563,0000	120	25000	0,9704	3,4737	3510,5000	40
27500	0,7227	4,0252	577,6000	120	27500	0,9416	3,7935	3904,8000	30
30000	0,6783	4,1486	583,3000	120	30000	0,9060	4,0406	4168,7000	30

**Tabla 30** Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 240 y valor de BI óptimo aplicando el algoritmo QL

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5840	300,2000	240	10000	1	1,6127	248,3000	320
12500	1	1,7435	373,8000	240	12500	0,9999	1,8013	321,5000	60
15000	0,9987	2,1317	536,3000	240	15000	0,9990	2,1475	754,9000	120
17500	0,9738	2,8344	837,8000	240	17500	0,9955	2,5155	1368,1000	20
20000	0,9152	3,3345	982,8000	240	20000	0,9916	2,8108	2130,0000	240
22500	0,8567	3,6395	1046,0000	240	22500	0,9868	3,0951	2894,7000	20
25000	0,7996	3,8846	1081,9000	240	25000	0,9692	3,4768	3497,9000	160
27500	0,7486	4,0351	1103,3000	240	27500	0,9416	3,8049	3915,0000	80
30000	0,7025	4,1732	1121,5000	240	30000	0,9056	4,0583	4170,0000	60

**Tabla 31** Comparativa de las métricas de rendimiento en diferentes condiciones de tráfico con un valor de BI estático de 750 y valor de BI óptimo aplicando el algoritmo QL

UEs MTC	Ps	K	D	BI	UEs MTC	Ps	K	D	BI
10000	1	1,5789	876,2000	750	10000	1	1,6127	244,8000	30
12500	1	1,7081	1058,7000	750	12500	0,9999	1,8009	307,9000	320
15000	0,9999	1,9416	1336,3000	750	15000	0,9986	2,1442	804,0000	80
17500	0,9991	2,3186	1814,9000	750	17500	0,9956	2,5078	1348,2000	160
20000	0,9913	2,8565	2448,3000	750	20000	0,9919	2,8170	2124,4000	30
22500	0,9668	3,3526	2916,8000	750	22500	0,9863	3,1058	2929,5000	160
25000	0,9308	3,6959	2916,8000	750	25000	0,9693	3,4775	3512,0000	80
27500	0,8878	3,9591	3179,8000	750	27500	0,9413	3,7976	3901,9000	20
30000	0,8443	4,1685	3457,1000	750	30000	0,9060	4,0492	4171,1000	20

**Resultados de la asignación de las recompensas para diferentes valores de BI bajo diversas condiciones.**

**Tabla 32** Asignación de recompensas para un valor de BI de 960 en diversas condiciones

NpsMcur	NpsCVcur	DeltaNpscur	val_BI	R
$\leq 3$	$< 0,4$	$== 1$	960	50
$\leq 3$	$< 0,4$	$== 3$	960	100
$\leq 3$	$< 0,4$	$\neq 1, \neq 3$	960	100
$\leq 3$	$\geq 0,4$	$== 1$	960	30
$\leq 3$	$\geq 0,4$	$== 3$	960	60
$\leq 3$	$\geq 0,4$	$\neq 1, \neq 3$	960	60
$< 7$	$< 0,4$	$== 1$	960	30
$< 7$	$< 0,4$	$== 3$	960	60
$< 7$	$< 0,4$	$\neq 1, \neq 3$	960	60
$< 7$	$\geq 0,4$	$== 1$	960	1
$< 7$	$\geq 0,4$	$== 3$	960	15
$< 7$	$\geq 0,4$	$\neq 1, \neq 3$	960	15
$\leq 10$	$< 0,2$	$== 1$	960	1
$\leq 10$	$< 0,2$	$== 3$	960	10
$\leq 10$	$< 0,2$	$\neq 1, \neq 3$	960	15

$\leq 10$	$\geq 0,2$	$= 1$	960	0.5
$\leq 10$	$\geq 0,2$	$= 3$	960	5
$\leq 10$	$\geq 0,2$	$\neq 1, \neq 3$	960	8
$> 10$	$< 0,2$	$= 1$	960	-100
$> 10$	$< 0,2$	$= 3$	960	-90
$> 10$	$< 0,2$	$\neq 1, \neq 3$	960	-85
$> 10$	$\geq 0,2$	$= 1$	960	-100
$> 10$	$\geq 0,2$	$= 3$	960	-90
$> 10$	$\geq 0,2$	$\neq 1, \neq 3$	960	-85

*Tabla 33 Asignación de recompensas para un valor de BI de  $\geq 240$  en diversas condiciones*

<b>NpsMcur</b>	<b>NpsCVcur</b>	<b>DeltaNpscur</b>	<b>val_BI</b>	<b>R</b>
$\leq 3$	$< 0,4$	$= 1$	$\geq 240$	40
$\leq 3$	$< 0,4$	$= 3$	$\geq 240$	80
$\leq 3$	$< 0,4$	$\neq 1, \neq 3$	$\geq 240$	80
$\leq 3$	$\geq 0,4$	$= 1$	$\geq 240$	25
$\leq 3$	$\geq 0,4$	$= 3$	$\geq 240$	50
$\leq 3$	$\geq 0,4$	$\neq 1, \neq 3$	$\geq 240$	50
$< 7$	$< 0,4$	$= 1$	$\geq 240$	40
$< 7$	$< 0,4$	$= 3$	$\geq 240$	70
$< 7$	$< 0,4$	$\neq 1, \neq 3$	$\geq 240$	70
$< 7$	$\geq 0,4$	$= 1$	$\geq 240$	5
$< 7$	$\geq 0,4$	$= 3$	$\geq 240$	20
$< 7$	$\geq 0,4$	$\neq 1, \neq 3$	$\geq 240$	20
$\leq 10$	$< 0,2$	$= 1$	$\geq 240$	2
$\leq 10$	$< 0,2$	$= 3$	$\geq 240$	15
$\leq 10$	$< 0,2$	$\neq 1, \neq 3$	$\geq 240$	20
$\leq 10$	$\geq 0,2$	$= 1$	$\geq 240$	1
$\leq 10$	$\geq 0,2$	$= 3$	$\geq 240$	8
$\leq 10$	$\geq 0,2$	$\neq 1, \neq 3$	$\geq 240$	10
$> 10$	$< 0,2$	$= 1$	$\geq 240$	-90
$> 10$	$< 0,2$	$= 3$	$\geq 240$	-85
$> 10$	$< 0,2$	$\neq 1, \neq 3$	$\geq 240$	-80
$> 10$	$\geq 0,2$	$= 1$	$\geq 240$	-90
$> 10$	$\geq 0,2$	$= 3$	$\geq 240$	-85
$> 10$	$\geq 0,2$	$\neq 1, \neq 3$	$\geq 240$	-80

**Tabla 34** Asignación de recompensas para un valor de BI de  $\geq 80$  en diversas condiciones

NpsMcur	NpsCVcur	DeltaNpscur	val_BI	R
$\leq 3$	$< 0,4$	$== 1$	$\geq 80$	10
$\leq 3$	$< 0,4$	$== 3$	$\geq 80$	5
$\leq 3$	$< 0,4$	$\neq 1, \neq 3$	$\geq 80$	-5
$\leq 3$	$\geq 0,4$	$== 1$	$\geq 80$	8
$\leq 3$	$\geq 0,4$	$== 3$	$\geq 80$	3
$\leq 3$	$\geq 0,4$	$\neq 1, \neq 3$	$\geq 80$	-10
$< 7$	$< 0,4$	$== 1$	$\geq 80$	50
$< 7$	$< 0,4$	$== 3$	$\geq 80$	60
$< 7$	$< 0,4$	$\neq 1, \neq 3$	$\geq 80$	80
$< 7$	$\geq 0,4$	$== 1$	$\geq 80$	40
$< 7$	$\geq 0,4$	$== 3$	$\geq 80$	50
$< 7$	$\geq 0,4$	$\neq 1, \neq 3$	$\geq 80$	60
$\leq 10$	$< 0,2$	$== 1$	$\geq 80$	5
$\leq 10$	$< 0,2$	$== 3$	$\geq 80$	20
$\leq 10$	$< 0,2$	$\neq 1, \neq 3$	$\geq 80$	40
$\leq 10$	$\geq 0,2$	$== 1$	$\geq 80$	5
$\leq 10$	$\geq 0,2$	$== 3$	$\geq 80$	15
$\leq 10$	$\geq 0,2$	$\neq 1, \neq 3$	$\geq 80$	30
$> 10$	$< 0,2$	$== 1$	$\geq 80$	-60
$> 10$	$< 0,2$	$== 3$	$\geq 80$	-45
$> 10$	$< 0,2$	$\neq 1, \neq 3$	$\geq 80$	-40
$> 10$	$\geq 0,2$	$== 1$	$\geq 80$	-60
$> 10$	$\geq 0,2$	$== 3$	$\geq 80$	-45
$> 10$	$\geq 0,2$	$\neq 1, \neq 3$	$\geq 80$	-40

**Tabla 35** Asignación de recompensas para un valor de BI de  $\geq 20$  en diversas condiciones

NpsMcur	NpsCVcur	DeltaNpscur	val_BI	R
$\leq 3$	$< 0,4$	$== 1$	$\geq 20$	-10
$\leq 3$	$< 0,4$	$== 3$	$\geq 20$	-20
$\leq 3$	$< 0,4$	$\neq 1, \neq 3$	$\geq 20$	-30
$\leq 3$	$\geq 0,4$	$== 1$	$\geq 20$	-15
$\leq 3$	$\geq 0,4$	$== 3$	$\geq 20$	-30
$\leq 3$	$\geq 0,4$	$\neq 1, \neq 3$	$\geq 20$	-40

< 7	< 0,4	== 1	>= 20	50
< 7	< 0,4	== 3	>= 20	60
< 7	< 0,4	≠ 1, ≠3	>= 20	80
< 7	>= 0,4	== 1	>= 20	40
< 7	>= 0,4	== 3	>= 20	50
< 7	>= 0,4	≠ 1, ≠3	>= 20	60
<= 10	< 0,2	== 1	>= 20	6
<= 10	< 0,2	== 3	>= 20	21
<= 10	< 0,2	≠ 1, ≠3	>= 20	41
<= 10	>= 0,2	== 1	>= 20	6
<= 10	>= 0,2	== 3	>= 20	16
<= 10	>= 0,2	≠ 1, ≠3	>= 20	40
> 10	< 0,2	== 1	>= 20	-50
> 10	< 0,2	== 3	>= 20	-35
> 10	< 0,2	≠ 1, ≠3	>= 20	-30
> 10	>= 0,2	== 1	>= 20	-50
> 10	>= 0,2	== 3	>= 20	-35
> 10	>= 0,2	≠ 1, ≠3	>= 20	-30