

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**DOUGLAS ALISSON TUBIANA**

**ACESSO ALEATÓRIO NÃO-ORTOGONAL BASEADO EM *Q-LEARNING* PARA  
REDES IOT-SATÉLITE AUXILIADAS POR *RELAYS***

**CURITIBA**

**2022**

**DOUGLAS ALISSON TUBIANA**

**ACESSO ALEATÓRIO NÃO-ORTOGONAL BASEADO EM  $Q$ -LEARNING PARA  
REDES IOT-SATÉLITE AUXILIADAS POR *RELAYS***

**$Q$ -Learning NOMA Random Access for IoT-Satellite Terrestrial Relay  
Networks**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre em Engenharia Elétrica e Informática Industrial do Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná.

Orientador: Prof. Dr. Glauber Gomes de Oliveira Brante

Coorientador: Dr. Jamil de Araújo Farhat

**CURITIBA**

**2022**



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



**Ministério da Educação  
Universidade Tecnológica Federal do Paraná  
Campus Curitiba**



DOUGLAS ALISSON TUBIANA

**ACESSO ALEATÓRIO NÃO-ORTOGONAL BASEADO EM Q-LEARNING PARA REDES IOT-SATÉLITE  
AUXILIADAS POR RELAYS**

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Ciências da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Telecomunicações E Redes.

Data de aprovação: 03 de Junho de 2022

Dr. Glauber Gomes De Oliveira Brante, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Joao Luiz Rebelatto, Doutorado - Universidade Tecnológica Federal do Paraná

Dr. Marcelo Eduardo Pellenz, Doutorado - Pontifícia Universidade Católica do Paraná (Pucpr)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 03/06/2022.

## **AGRADECIMENTOS**

Agradeço ao meu orientador, Glauber Brante, pelo apoio e incentivo durante todos esses anos. Obrigado pela paciência e pela confiança depositada.

Agradeço ao meu coorientador, Jamil Farhat, por toda a ajuda no desenvolvimento do projeto e escrita do artigo.

Agradeço a todos os meus familiares e amigos, que sempre deram apoio e suporte, por proporcionar momentos inesquecíveis.

Agradeço a Universidade Tecnológica Federal do Paraná, por ser sempre esse local acolhedor e sempre proporcionar o suporte necessário a todos aqueles buscando aumentar seu conhecimento em todos os momentos.

Agradeço a todos os colegas e colaboradores da Consilux, em especial ao Paulo Shio-kawa, pela amizade e incentivo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

São as perguntas que não podemos responder que mais nos ensinam. Elas nos ensinam a pensar. Se você der uma resposta ao homem, tudo o que ele ganha é um fato qualquer. Mas lhe dê uma pergunta e ele procurará suas próprias respostas (ROTHFUSS, Patrick, 2011).

## RESUMO

Em um cenário de múltiplos dispositivos IoT, espalhados por uma área muito extensa, a maior dificuldade é prover acesso, tendo em vista a grande competição por recursos disponíveis em uma estação central que provê conectividade. Aqui, o uso de satélites LEO (do inglês *Low Earth Orbit*) é uma alternativa. Entretanto, os dispositivos IoT não possuem potência de transmissão necessária para alcançar ao satélite, por isso o uso de *relays* intermediários que captam essas informações é crucial. Pensando nesse cenário, este trabalho propõe um protocolo de acesso aleatório, baseado em múltiplo acesso não ortogonal (NOMA) com o controle de acesso através de aprendizado de máquina por *Q-Learning*. O algoritmo QL-NOMA proposto decide, a partir de uma *Q-Table*, o canal de comunicação e o *timeslot* em que cada dispositivo irá transmitir na conexão com um *relay* que retransmite para um satélite, tendo como objetivo melhorar a taxa de transferência no enlace dispositivo-*relay*. Ainda assim, mesmo que hajam colisões, os *relays* aplicam cancelamento sucessivo de interferência (SIC), buscando decodificar sinais de diferentes dispositivos transmitidos no mesmo *timeslot*-canal. A partir do sucesso da transmissão de cada dispositivo, a *Q-Table* de cada um é atualizada. Como comparação, foi escolhido o protocolo de acesso aleatório *Slotted-ALOHA*, que escolhe o canal e o *timeslot* de forma aleatória. Os resultados numéricos mostram que o método proposto melhora significativamente o desempenho em comparação com o *Slotted-ALOHA* com NOMA, apesar de ser uma estratégia descentralizada. Além disso, é possível concluir que existe uma quantidade ótima de *relays* e canais dependendo do número de dispositivos IoT tentando acessar a rede.

**Palavras-chave:** *q-learning*; acesso múltiplo não-ortogonal; rede terrestre de satélite-iot; múltiplos *relays*.

## ABSTRACT

In a scenario with multiple IoT devices, spread over a very large area, the biggest difficulty is providing access, in view of the great competition for resources available in a central station that provides connectivity. Here, the use of LEO satellites is an alternative. However, IoT devices do not have the transmission power needed to establish connection with the satellite, so the use of intermediate relays that capture this information is crucial. With this scenario in mind, this work proposes a random access protocol, based on non-orthogonal multiple access (NOMA) with access control through machine learning with  $Q$ -Learning. The proposed QL-NOMA algorithm decides, from a  $Q$ -Table, the communication channel and the timeslot in which each device will transmit in connection with a relay that retransmits to a satellite, aiming to improve the throughput in the device-relay link. Still, even if there are collisions, *relays* apply successive interference cancellation (SIC), seeking to decode signals from different devices transmitted on the same timeslot-channel. From the successful transmission of each device, the  $Q$ -Table of each one is updated. As a comparison, the Slotted-ALOHA random access protocol was chosen, which chooses the channel and timeslot at random. Numerical results show that the proposed method significantly improves performance compared to Slotted-ALOHA with NOMA, despite being a decentralized strategy. Furthermore, it is possible to conclude that there is an optimal number of relays and channels depending on the number of IoT devices trying to access the network.

**Keywords:**  $q$ -learning; non-orthogonal multiple access; iot-satellite terrestrial networks; multiple relays.

## LISTA DE ALGORITMOS

Algoritmo 1 – Método de RA com <i>Q-Learning</i> baseado em NOMA. . . . .	33
---	----



## LISTA DE FIGURAS

Figura 1 – Cenários de uso de acordo com o International Mobile Telecommunications (IMT). . . . .	18
Figura 2 – Estimativa global de usuários móveis e seu respectivo tráfego de 2020 a 2030.	19
Figura 3 – Usos dos novos cenários previstos para o 6G. . . . .	21
Figura 4 – Modelo de aprendizado por reforço. . . . .	25
Figura 5 – Exemplo de <i>Slotted</i> -ALOHA. . . . .	27
Figura 6 – Esquema SIC. . . . .	28
Figura 7 – Modelo de sistema com $D$ dispositivos IoT distribuídos em uma área circular de raio $\mathcal{R}$ , auxiliados por $R$ relays que fornecem conectividade a um satélite LEO. . . . .	29

## LISTA DE GRÁFICOS

Gráfico 1 – Taxa de transferência em função de $D$ , com $R = 4$ relays e $C = 1$ canal disponível. . . . .	35
Gráfico 2 – Taxa de mensagens redundantes em função de $D$ , com $R = 4$ relays e $C = 1$ canal disponível. . . . .	35
Gráfico 3 – Taxa de transferência em função de $C$ , com $D = 400$ dispositivos IoT servidos por $R = 4$ relays. . . . .	36
Gráfico 4 – Taxa de transferência normalizada como uma função de $C = R$ , com $D = 1500$ dispositivos IoT. . . . .	37
Gráfico 5 – Taxa de transferência normalizada em função de $P$ com $D = 1500$ , $R = 3$ e $C \in \{1, 3\}$ . . . . .	38
Gráfico 6 – Taxa de transferência em função de $\alpha$ do algoritmo $Q$ -Learning com $\gamma = 0,5$ , $D = 400$ , $C = 3$ e $R = 4$ . . . . .	38
Gráfico 7 – Taxa de transferência em função de $\gamma$ do algoritmo $Q$ -Learning com $\alpha = 0,25$ , $D = 400$ , $C = 3$ e $R = 4$ . . . . .	39

## LISTA DE QUADROS

Quadro 1 – Comparativo dos requisitos do 5G e do 6G. . . . .	20
--	----

## LISTA DE ABREVIATURAS E SIGLAS

### Abreviaturas

ACK Reconhecimento, do inglês *Acknowledgement*

### Siglas

3G Terceira geração de comunicações sem fio

4G Quarta geração de comunicações sem fio

5G Quinta geração de comunicações sem fio

6G Sexta geração de comunicações sem fio

AWGN Ruído Branco Gaussiano Aditivo, do inglês *Additive White Gaussian Noise*

BLE Bluetooth de Baixa Energia, do inglês *Bluetooth Low Energy*

DQL-JRSAC Distribuição Conjunta de Seleção de *Relay* e Controle de Acesso baseado em *Q-Learning*, do inglês *Distributed Q-Learning Based Joint Relay Selection and Access Control*

eMBB Banda Larga Móvel melhorada, do inglês *enhanced Mobile BroadBand*

HSTN Rede Terrestre de Satélite Híbrida, do inglês *Hybrid Satellite-Terrestrial Network*

IoT Internet das Coisas, do inglês *Internet of Things*

LEO Órbita Terrestre Baixa, do inglês *Low Earth Orbit*

LPWA Grandes Áreas com Baixa Potência, do inglês *Low-Power Wide-Area*

LTE *Long Term Evolution*

LTE-A *Long Term Evolution Advanced*

LTE-M *Long Term Evolution MTC*

M2M Máquina-Máquina, do inglês *Machine to Machine*

MAP *Maximum a Posteriori Probability*

MBB	Banda Larga Móvel, do inglês <i>Mobile BroadBand</i>
ML	Aprendizado de Máquina, do inglês <i>Machine Learning</i>
mMTC	Comunicação do Tipo Máquina massiva, do inglês <i>massive Machine Type Communication</i>
MTC	Comunicação do Tipo Máquina, do inglês <i>Machine-Type Communications</i>
mULC	Comunicação massiva Ultra-Confíável e de Baixa Latência, do inglês <i>massive Ultra-Reliable Low-Latency Communication</i>
NB-IoT	<i>Narrow-Band IoT</i>
NOMA	Acesso Múltiplo Não-Ortogonal, do inglês <i>Non-Orthogonal Multiple Access</i>
PC	Computador Pessoal, do inglês <i>Personal Computer</i>
QL-NOMA	Método proposto de RA com <i>Q-Learning</i> baseado em NOMA
RA	Acesso Aleatório, do inglês <i>Random Access</i>
SA-NOMA	<i>Slotted-ALOHA</i> com NOMA
SIC	Cancelamento Sucessivo de Interferência, do inglês <i>Successive Interference Cancellation</i>
SINR	Relação Sinal-Interferência mais Ruído, do inglês <i>Signal to Interference plus Noise Ratio</i>
SNR	Relação Sinal-Ruído, do inglês <i>Signal to Noise Ratio</i>
STRN	Rede de Retransmissão Terrestre de Satélite, do inglês <i>Satellite Terrestrial Relay Network</i>
ULBC	Comunicação Banda Larga Ultra-Confíável e de Baixa Latência, do inglês <i>Ultra-Reliable Low-Latency BroadBand Communication</i>
uMBB	Banda Larga Móvel onipresente, do inglês <i>ubiquitous Mobile BroadBand</i>
URLLC	Comunicação Ultra-Confíável e de Baixa Latência, do inglês <i>Ultra-Reliable and Low Latency Communication</i>

## LISTA DE SÍMBOLOS

$\alpha$	Taxa de aprendizado
$\bar{n}_{\text{dist}}^{(\text{sch})}$	Média do total de mensagens distintas dos dispositivos IoT recebidos com sucesso pelo satélite, em cada quadro
$\bar{n}_{\text{total}}^{(\text{sch})}$	Número total de mensagens recebidas pelo satélite
$\bar{n}_D^{(\text{sch})}$	Número médio de dispositivos IoT que transmitiram com sucesso
$\beta$	Eficiência espectral
$\epsilon$	Recompensa
$\gamma$	Fator de desconto
$\mathbf{n}_r$	Ruído Térmico
$\mathbf{x}_i$	Sinal transmitido
$\mathcal{R}$	Raio
$\rho^{(\text{sch})}$	Taxa de mensagens redundantes
$\tau^{(\text{sch})}$	Número médio de dispositivos IoT
$A_i$	Ação no modelo de Aprendizado por Reforço
$B$	Largura de Banda
$C$	Número de Canais
$D$	Número de Dispositivos
$d_{r,i}$	Distância entre o <i>relay</i> e o dispositivo IoT
$F$	Figura de Ruído
$h_{r,i}$	Coefficiente de desvanecimento do canal
$N_0$	Densidade espectral de potência de ruído
$P$	Potência de transmissão
$R$	Número de <i>Relays</i>
$S_i$	Estado no modelo de Aprendizado por Reforço
$T$	Número de <i>timeslots</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
1.1	Sistemas 5G	16
1.2	Sistemas 6G	18
1.3	Requisitos de Internet das Coisas	21
1.4	Delimitação do Tema	22
1.5	Objetivos	24
1.5.1	Objetivo Geral	24
1.5.2	Objetivos Específicos	24
1.6	Estrutura do Documento	24
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>25</b>
2.1	Aprendizado por reforço	25
2.2	Acesso Aleatório para IoT	26
2.3	Acesso Múltiplo Não-Ortogonal	27
<b>3</b>	<b>MÉTODO PROPOSTO</b>	<b>29</b>
3.1	Modelo do sistema	29
3.1.1	Modelo de Transmissão	30
3.1.2	Métricas do Sistema	31
3.2	Protocolo Proposto	32
3.2.1	<i>Q-Learning</i>	32
3.3	Método proposto de RA com <i>Q-Learning</i> baseado em NOMA	32
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>34</b>
4.1	Desempenho Geral do QL-NOMA	34
4.2	Efeito do Número de <i>Relays</i>	36
4.3	Efeito da Potência de Transmissão	36
4.4	Avaliação dos Parâmetros do <i>Q-Learning</i>	37
<b>5</b>	<b>CONCLUSÕES E PERSPECTIVAS DE TRABALHOS FUTUROS</b>	<b>40</b>
	REFERÊNCIAS	42

## 1 INTRODUÇÃO

Os últimos anos viram um crescimento exponencial na quantidade de tráfego de dados via conexões móveis, as quais ocupam hoje papel fundamental na transformação tecnológica global via digitalização de inúmeros serviços, processos e aplicações. É muito interessante notar que, dentro deste contexto, a quantidade de máquinas conectadas tem crescido mais rapidamente do que a população mundial (CISCO, 2019). Com esse crescimento, é esperado que o número de aparelhos de Internet das Coisas, do inglês *Internet of Things* (IoT), conectados até 2025 supere a marca de 27 bilhões (SINHA, 2021). Além disso, em 2020 o número de conexões entre máquinas já superou o de outras conexões (tais como *smartphones, tablets, PCs, etc.*), de 11,7 contra 9,9 bilhões (IOT-ANALYTICS, 2020).

Comunicação do Tipo Máquina, do inglês *Machine-Type Communications* (MTC), é uma das tecnologias habilitadoras da IoT (IEEE Innovate, 2021), cujos potenciais ganhos econômicos advindos de sua exploração e massificação são enormes. Mesmo com a crise dos chips e com o impacto da COVID-19 na cadeia de fornecimento, o mercado da IoT continua a crescer, movimentado por novas tecnologias como os sistemas móveis de quinta geração (5G) e por investimentos em aplicações que requerem uma integração de equipamentos. Por exemplo, os investimentos em IoT entre 2017 a 2025 são estimados em 15 trilhões de dólares (IEEE Innovate, 2021).

Com a evolução da tecnologia, novas implementações de IoT demandam por conexões de larga cobertura de área e longo alcance, com isso, novas tecnologias de comunicação de longo alcance são desenvolvidas. Exemplo disso é o próprio 5G, que não só é capaz de melhorar a qualidade das comunicações móveis, mas também é capaz de suprir a demanda por baixa latência e alta confiabilidade que as novas tecnologias IoT requerem.

### 1.1 Sistemas 5G

As comunicações sem fio passam por uma evolução de tecnologia aproximadamente a cada 10 anos, caracterizando uma mudança de geração, trazendo melhorias de desempenho nas mais diversas características. De olho neste mercado, o 5G (IEEE Spectrum, 2020) oferece ou promete suporte não apenas à altas taxas de dados para aplicações multimídia em banda larga, mas, de forma inovadora, também à aplicações relacionadas com IoT. Para o 5G, o principal foco foi atender a crescente utilização de dispositivos IoT, que cada vez mais demandam taxas de dados incompatíveis com as gerações anteriores. Isso representa uma evolução em relação às tecnologias 3G e 4G, que não são capazes de atender as demandas requeridas pelo 5G e não podem ser utilizadas para os casos de Grandes Áreas com Baixa Potência, do inglês *Low-Power Wide-Area* (LPWA) e comunicações de longa distância (CHETTRI; BERA, 2020), pois essas tecnologias operam com frequências de portadoras na faixa de GHz, o que limita o alcance, além de demandarem um alto consumo energético, um limitante para IoT.



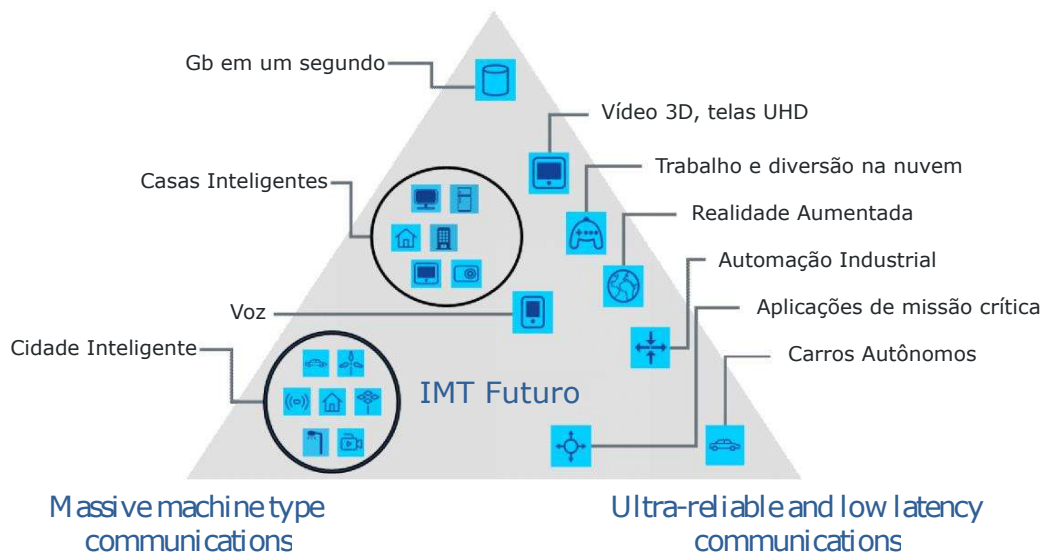
O 5G foi desenvolvido para atender uma variedade de cenários de três principais categorias (POPOVSKI *et al.*, 2018):

- Banda Larga Móvel melhorada, do inglês *enhanced Mobile BroadBand* (eMBB), pode ser considerado uma extensão direta do serviço de cobertura 4G. Caracterizado pelas altas quantidades de dados transmitidos e pela longa estabilidade nos padrões de ativação dos dispositivos. Com isso, a rede permite que diferentes dispositivos acessem os mesmos recursos simultaneamente. O objetivo do serviço eMBB é maximizar a taxa de dados garantindo uma certa confiabilidade;
- Comunicação do Tipo Máquina massiva, do inglês *massive Machine Type Communication* (mMTC), é caracterizado por um número elevado de dispositivos ativos intermitentemente e geralmente com uma taxa de transmissão baixa que podem estar conectados em uma base central. Porém, somente uma parte destes dispositivos se torna ativo em um certo período de tempo e transmite, assim não há a necessidade de se alocar recursos para cada dispositivo individualmente, apenas priorizando recursos que podem ser compartilhados por acesso aleatório. O objetivo deste serviço é maximizar a taxa de chegada de tráfego, permitindo uma taxa de erro um pouco maior;
- Comunicação Ultra-Confiável e de Baixa Latência, do inglês *Ultra-Reliable and Low Latency Communication* (URLLC), também é caracterizado por dispositivos ativos intermitentemente, porém, com um número reduzido comparado com o mMTC. Para garantir a alta confiabilidade, o URLLC precisa ter um controle avançado da distribuição de recursos, além de acesso aleatório para que os recursos não sejam desperdiçados durante períodos de inatividade. Para atender os requisitos de baixa latência, uma transmissão URLLC precisa ser localizada no tempo. O objetivo é garantir baixa latência, com alta confiabilidade e disponibilidade.

Esta divisão encaixa os mais diversos usos existentes e os que ainda estão por vir. Para eMBB, os principais cenários de uso incluem alta cobertura de área e *hotspots*, em mMTC, os principais cenários se encontram em situações com um montante elevado de dispositivos como cidades inteligente e em URLLC se encontram os cenários que requerem baixa latência e alta confiabilidade como certos processos de produção, cirurgias remotas e segurança dos transportes (ITU-R, 2015a). A Figura 1 ilustra os principais cenários de uso de cada categoria. Nesse contexto, vemos aplicações como a manufatura inteligente dentro do conceito da Indústria 4.0, realidades virtual e aumentada como ferramentas de projeto, produção e manutenção industrial, bem como o desenvolvimento de carros autônomos como realidades tecnológicas em plena expansão.

É evidente que a demanda por aplicações de IoT só tende a crescer, levando a indústria a criar novas classes de serviço, atendendo a requisitos cada vez mais exigentes e especializados (ERICSSON, 2020; MAHMOOD *et al.*, 2021; 6G Flagship, 2019; TARIQ *et al.*, 2020).

**Figura 1 – Cenários de uso de acordo com o International Mobile Telecommunications (IMT).  
Enhanced mobile broadband**



Vale ainda ressaltar que com a expansão das aplicações de IoT, em geral teremos um número enorme de dispositivos sem fio, substituindo soluções cabeadas ou criando novas. Em alguns cenários, como na indústria, substituir dispositivos cabeados por soluções sem fio operadas a bateria é muito interessante, pela flexibilidade, facilidade de instalação e custo.

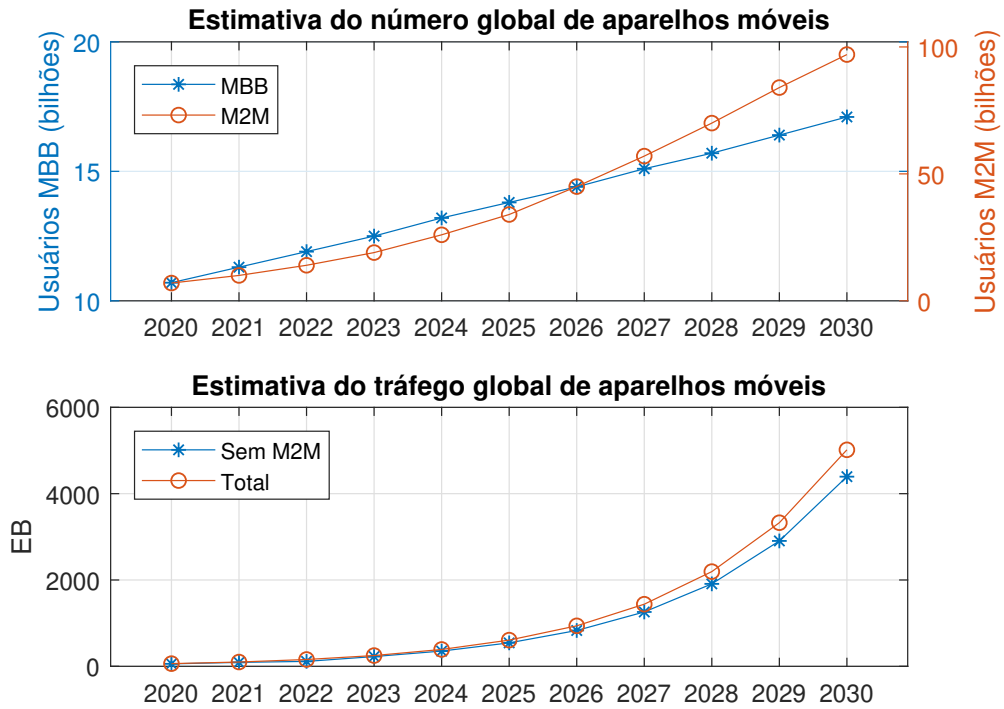
## 1.2 Sistemas 6G

A evolução dos sistemas 5G rumo aos sistemas 6G já vem sendo discutida na academia e na indústria. Tal evolução é ainda bastante conceitual, carregando a visão de novos indicadores-chave de desempenho, mais ligados a tendências sociais, metas de sustentabilidade, redução das emissões de dióxido de carbono, bem como demandas de produtividade cada vez maiores (6G Flagship, 2019). Sua implantação ainda é prevista para por volta de 2030, de modo que vários países já iniciaram suas pesquisas em 6G, tanto em universidades quanto em empresas de tecnologia, buscando prever e desenvolver os novos recursos que serão requeridos no futuro (LIU *et al.*, 2020).

O desenvolvimento de uma nova geração de sistemas não depende apenas do crescimento exponencial do tráfego de comunicações móveis, mas também dos novos serviços e aplicações que virão a surgir no futuro próximo. Além disso, o sistema 5G não será capaz de acomodar o volume de tráfego de comunicações móveis esperado para 2030, justificando assim a implantação de uma nova geração de sistemas de comunicação (JIANG *et al.*, 2021). É esperado que a quantidade de *smartphones* e *tablets*, assim como comunicações Máquina-Máquina, do inglês *Machine to Machine* (M2M), cresça exponencialmente até 2030, com o número de usuários de Banda Larga Móvel, do inglês *Mobile BroadBand* (MBB), podendo chegar

a 17 bilhões e o número de usuários M2M podendo chegar a 97 bilhões, o que é cerca de 14 vezes mais do que em 2020 (ITU-R, 2015b). A Figura 2 demonstra esse crescimento e também mostra o crescimento no volume de tráfego de dados esperados para esta década. É possível notar que até o final da década o tráfego móvel chegará a 5016 EB<sup>1</sup> por mês no ano de 2030 comparado com 62 EB por mês no ano de 2020.

**Figura 2 – Estimativa global de usuários móveis e seu respectivo tráfego de 2020 a 2030.**



Fonte: Adaptado de ITU-R (2015b).

Segundo o relatório da Cisco (2020), é esperado que mais de 70% da população mundial possua algum tipo de conectividade móvel até 2023, alcançando cerca de 5,7 bilhões de usuários móveis. Com isso, o número de dispositivos conectados a uma rede será 3,6 vezes maior que a população mundial, o que representa cerca de 29,3 bilhões de dispositivos, onde mais da metade serão de M2M.

Devido ao crescimento extensivo do número de dispositivos IoT, a nova geração de sistemas de comunicação deve atender casos de alta eficiência espectral e energética, baixa latência e conectividade massiva (AKHTAR *et al.*, 2020). Vários cenários são esperados com aplicações em diferentes categorias, como vida inteligente que moldará as interações das pessoas com aprendizado, compras, trabalho e saúde; produção inteligente com tecnologias emergentes no campo agrícola e industrial que levará ao desenvolvimento de uma economia digital; e sociedade inteligente que trará inovações no campo da mobilidade com novos meios de transporte e implantação de inteligências artificiais na condução de veículos, na integração social

<sup>1</sup> 1 exabyte (EB) = 1 000 000 terabytes (TB), 1 TB = 1000 gigabytes (GB)

com novos meios de conectar populações isoladas a serviços públicos de saúde e educação, e melhor governança, trazendo novas ferramentas que permitirão que governos detectem possíveis desastres e monitorem o desenvolvimento de cidades com a implementação de cidades inteligentes (LIU *et al.*, 2020).

Com essas novas aplicações, é esperado que o 6G atinja certos patamares, em diferentes aspectos, como demonstrado no Quadro 1.

**Quadro 1 – Comparativo dos requisitos do 5G e do 6G.**

<b>Requerimentos</b>	<b>5G</b>	<b>6G</b>
Pico de taxa de dados	DL: 20Gbps UL: 10Gbps	Níveis de Tbps
Taxa de dados para a experiência de usuários	DL: 100 Mbps UL: 50 Mbps	10-100 Gbps
Latência	10 ms	<1 ms
Confiabilidade	99,999 %	>99,99999 %
Mobilidade	500 km/h	>1000 km/h
Capacidade de tráfego	10 Mbps/m <sup>2</sup>	0,1-10Gbps/m <sup>3</sup>
Densidade de conexões	1 milhão de dispositivos/km <sup>2</sup>	10-100 milhões de dispositivos/km <sup>3</sup>
Eficiência espectral	DL: 30 bps/Hz UL: 15 bps/Hz	2-3 vezes comparado com o 5G
Eficiência energética das redes	-	10-1000 vezes comparado com o 5G
Precisão de posicionamento	Ao ar livre: 10 m Interno: 3 m	Ao ar livre: nível abaixo de metro Interno: nível de centímetros

**Fonte: Adaptado de Liu *et al.* (2020).**

Para atender todos os requisitos do 6G, os cenários definidos para o 5G não serão suficientes, necessitando mesclar os aspectos de cada um, formando novos cenários assim definidos como: Banda Larga Móvel onipresente, do inglês *ubiquitous Mobile BroadBand* (uMBB), que junta comunicações de alta qualidade com conectabilidade global; Comunicação Banda Larga Ultra-Confiável e de Baixa Latência, do inglês *Ultra-Reliable Low-Latency BroadBand Communication* (ULBC), que garante confiabilidade e baixa latência com altas taxas de transferência; e Comunicação massiva Ultra-Confiável e de Baixa Latência, do inglês *massive Ultra-Reliable Low-Latency Communication* (mULC), que junta as características do mMTC e URLLC, facilitando a implementação de um grande número de sensores e atuadores em indústrias verticais (JIANG *et al.*, 2021).

Na Figura 3 é ilustrado essa intersecção dos cenários 5G, com os principais usos dos novos cenários do 6G.



- Cobertura: o nível de cobertura das aplicações IoT pode variar desde alguns metros até dezenas de quilômetros. Assim, é possível dividir em aplicações de baixo alcance chegando a dezenas de metros, como casas inteligentes, e aplicações de longo alcance chegando a dezenas de quilômetros, como agricultura inteligente e aeronaves não tripuladas;
- Potência: a potência afeta diretamente o custo de aparelhos IoT, além de poder impactar o meio ambiente, por isso se busca desenvolver formas de se economizar o máximo possível de energia. O maior desafio está em sensores de baixo consumo e longo alcance, pois não é possível recarregar estes aparelhos regularmente e com facilidade;
- Confiabilidade: podem ser divididos em dois grupos de aplicações de missão crítica e missão não crítica. Algumas aplicações como *smart grids*, robôs de manufatura e cuidados de saúde precisam ser confiáveis e por isso se encaixam no grupo de aplicações de missão crítica. Já a maioria das aplicações IoT podem ser consideradas do grupo de aplicações de missão não crítica, como sensores de umidade, estufas inteligentes e medidores de água e energia;
- Mobilidade: As aplicações de mobilidade podem ser divididas em baixa mobilidade e alta mobilidade. O maior desafio está em aplicações de alta mobilidade, como veículos, trens e aeronaves, onde a velocidade pode chegar a centenas de km/h, necessitando redirecionamento e re-seleção de células, além de aumentar o efeito Doppler diminuindo a confiabilidade.

#### 1.4 Delimitação do Tema

A evolução das redes sem fio em direção ao 6G deve intensificar a implantação da IoT em uma ampla gama de aplicações, com diferentes requisitos (MAHMOOD *et al.*, 2021; TARIQ *et al.*, 2020). Em particular, a IoT massiva é caracterizada por um grande número de dispositivos, baixo consumo de energia, baixa taxa de dados, baixo custo e alguma tolerância a atrasos (BACCO *et al.*, 2019). Os casos de uso típicos neste contexto estão relacionados ao monitoramento, rastreamento e navegação, que sofrem com a baixa conectividade em locais remotos. Uma solução interessante para esses casos é empregar satélites de Órbita Terrestre Baixa, do inglês *Low Earth Orbit* (LEO), para fornecer cobertura. No entanto, a escassa potência dos dispositivos IoT pode ser um fator limitante severo no *uplink*, exigindo o uso de *relays* dedicados para encaminhar dados IoT para o satélite. Esta arquitetura é comumente referida como uma Rede de Retransmissão Terrestre de Satélite, do inglês *Satellite Terrestrial Relay Network* (STRN), orientada para IoT (ZHAO *et al.*, 2021) ou como uma Rede Terrestre de Satélite Híbrida, do inglês *Hybrid Satellite-Terrestrial Network* (HSTN) modelo L (FANG *et al.*, 2021).

Dentro do contexto de STRNs, esse trabalho foca em garantir acesso à rede por parte de um número bastante grande de dispositivos IoT. Como será melhor detalhado no Capítulo 2, abordagens para Acesso Aleatório, do inglês *Random Access* (RA) baseados em algoritmos de Aprendizado de Máquina, do inglês *Machine Learning* (ML) têm sido bastante exploradas na literatura. Dentre as várias técnicas existentes, os métodos baseados em *Q-Learning* têm atraído bastante atenção nos últimos anos (SUTTON; BARTO, 2018). Por exemplo, um esquema de Distribuição Conjunta de Seleção de *Relay* e Controle de Acesso baseado em *Q-Learning*, do inglês *Distributed Q-Learning Based Joint Relay Selection and Access Control* (DQL-JRSAC) é proposto em (ZHAO *et al.*, 2021), que emprega vários *relays* para coletar informações dos dispositivos IoT e encaminhá-las para o satélite. O esquema proposto é capaz de melhorar a taxa de transferência, o atraso de acesso total e a taxa somada em comparação com os esquemas de RA, convencionais. Além disso, os autores em (SILVA *et al.*, 2020) propuseram um algoritmo *Q-Learning* baseado em Acesso Múltiplo Não-Ortogonal, do inglês *Non-Orthogonal Multiple Access* (NOMA), que aloca *slots* RA para que os dispositivos IoT se comuniquem diretamente com um *gateway*. Os resultados em (SILVA *et al.*, 2020) mostram que o método melhora a taxa de transferência da rede quando comparado ao *Slotted-ALOHA*, mesmo que este empregue o NOMA.

Neste trabalho, diferentemente de (SILVA *et al.*, 2020; ZHAO *et al.*, 2021), realizamos uma análise considerando as métricas de taxa de transferência normalizada e taxa de mensagens redundantes em um cenário prático de uma rede de retransmissão terrestre IoT-satélite. Vários dispositivos IoT tentam acessar a rede conectando-se a vários *relays* disponíveis, que por sua vez possuem vários canais ortogonais para fornecer conectividade e encaminhar as informações recebidas para o satélite. Para realizar a alocação de tais recursos, um algoritmo baseado no *Q-Learning* NOMA, chamado QL-NOMA, é projetado para melhorar o desempenho em uma arquitetura IoT-STRN. O NOMA é considerado para permitir que vários dispositivos IoT transmitam simultaneamente no mesmo *timeslot* e canal, enquanto a interferência entre usuários é mitigada usando Cancelamento Sucessivo de Interferência, do inglês *Successive Interference Cancellation* (SIC) nos *relays*. A técnica *Q-Learning* é empregada para alocar os dispositivos IoT em pares de *timeslot*-canal apropriados, exigindo um único bit de realimentação por *timeslot* de cada *relay*. Observamos que a otimização do número de *relays* e canais disponíveis é crucial para maximizar o desempenho do sistema, pois mais *relays* aumentam a chance de uma mensagem de um dispositivo IoT ser recebida com sucesso, mas também aumenta a quantidade de transmissões redundantes dos *relays* para os satélites, comprometendo o uso eficiente dos recursos da rede. Parte dos resultados apresentados nessa dissertação foram publicados em (TUBIANA *et al.*, 2022).

## 1.5 Objetivos

### 1.5.1 Objetivo Geral

O objetivo geral deste trabalho é propor um algoritmo baseado em *Q-Learning* e NOMA denominado QL-NOMA, que é projetado para melhorar o desempenho em uma arquitetura STRN, considerando NOMA para permitir que vários dispositivos IoT transmitam simultaneamente no mesmo *timeslot* e canal, com a interferência entre usuários sendo mitigado usando SIC, nos *relays*.

### 1.5.2 Objetivos Específicos

- Considerar um cenário prático com vários dispositivos IoT tentando acessar a rede, conectando-se a vários *relays* disponíveis, que por sua vez possuem vários canais ortogonais para fornecer conectividade e encaminhar as informações recebidas para o satélite;
- Buscar uma configuração otimizada de parâmetros para o algoritmo proposto, com o objetivo de garantir o melhor desempenho em relação a taxa de transmissão e atraso, exigidos pelos sistemas 5G/6G;
- Avaliar o algoritmo proposto em relação a taxa de transmissão e atraso, em cenários com diferentes configurações, comparando os resultados com outros algoritmos como o *Slotted-ALOHA* com NOMA.

## 1.6 Estrutura do Documento

Este documento está dividido da seguinte forma. No Capítulo 2 são introduzidos os conceitos de *Q-Learning*, dentro de ML; junto com uma explicação e demonstração do protocolo *Slotted-ALOHA* e dos conceitos de RA. Além disso, também são explicados os conceitos de NOMA e SIC. No Capítulo 3 é demonstrado o modelo do sistema utilizado com o modelo de transmissão; também é explicado como será o protocolo e o método proposto, com suas características e parâmetros. No Capítulo 4 são apresentados e discutidos os principais resultados encontrados, avaliando o desempenho geral, o efeito do número de *relays* e da potência de transmissão, e dos parâmetros do *Q-Learning*, comparando com outros algoritmos conhecidos. Por fim, o Capítulo 5 apresenta as conclusões gerais da dissertação junto com possíveis trabalhos futuros.



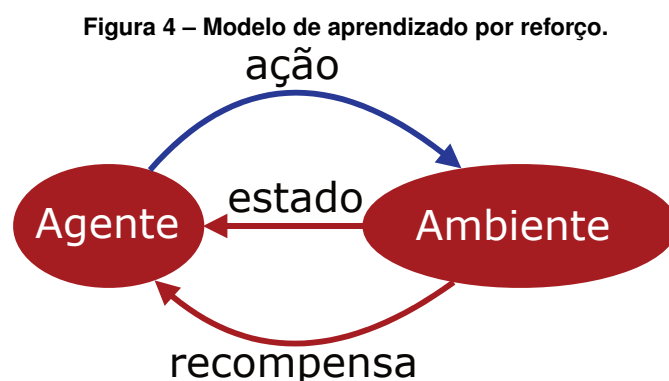
## 2 FUNDAMENTAÇÃO TEÓRICA

Para que o modelo desenvolvido seja entendido, alguns conceitos que foram empregados precisam ser apresentados. Na seção 2.1 será apresentada uma explicação geral do aprendizado por reforço, um tipo de ML, com foco no algoritmo de *Q-Learning*. A seção 2.2 explica e apresenta a motivação do uso de acesso aleatório para IoT, explicando o funcionamento do protocolo *Slotted-ALOHA*, utilizado neste trabalho. Por fim, os conceitos de NOMA e SIC, como estratégias para lidar com colisões, são apresentados na seção 2.3.

### 2.1 Aprendizado por reforço

Nos últimos anos, as técnicas de ML tem sido cada vez mais utilizadas nos processos de automação dos sistemas 5G, devido a alta demanda de taxas de dados e informação, com perspectivas semelhantes previstas para o 6G. Muitas abordagens para melhorar o RA em redes IoT massivas, recentemente, também recorreram a algoritmos de ML. Dentre as muitas técnicas de aprendizado de máquina existentes, o método *Q-Learning* tem atraído enorme atenção nos últimos anos devido à sua capacidade de ser implementado de forma distribuída e livre de modelos (SUTTON; BARTO, 2018).

*Q-Learning* é um algoritmo de aprendizado por reforço, onde um agente coleta informações através de ações interagindo com o ambiente, como resposta o agente recebe dois tipos de informações: o seu estado atual e um valor real de recompensa. Este modelo é ilustrado pela Figura 4. O objetivo do agente é maximizar a recompensa e determinar a melhor sequência de ações para alcançar o objetivo (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018).



Fonte: Adaptado de Mohri, Rostamizadeh e Talwalkar (2018).

*Q-Learning* é baseado em uma relação agente-ambiente com uma função de valor de ação (recompensa) representada pela *Q-Table*. Então, em cada *timeslot*, um agente  $i$  pertencente ao estado  $S_i$  realiza uma ação  $A_i$ , tentando maximizar sua recompensa. As atualizações

na  $Q$ -Table são definidas por

$$Q(S_i, A_i) \leftarrow (1 - \alpha)Q(S_i, A_i) + \alpha \left( \epsilon_{i+1} + \gamma \max_a Q(S_{i+1}, a) \right), \quad (1)$$

onde  $\alpha$  é a taxa de aprendizagem,  $\gamma$  é o fator de desconto, quantificando a importância das recompensas futuras e  $\epsilon$  é a recompensa.

É importante ressaltar que o  $Q$ -Learning surge como uma solução potencial para melhorar o acesso de rádio em redes mMTC (SILVA *et al.*, 2020; SHARMA; WANG, 2019). O método  $Q$ -Learning está altamente alinhado com os requisitos de uma rede IoT, uma vez que é livre de modelos, pode ser implementado de forma distribuída (SHARMA; WANG, 2019; SUTTON; BARTO, 2018) e pouca memória é necessária para armazenar uma única  $Q$ -Table. Além disso, são necessários baixos recursos computacionais em cada dispositivo para calcular (1) e a decisão da próxima ação associada à técnica de aprendizado de máquina é puramente algébrica.

Em relação a robustez, a taxa de aprendizado no método  $Q$ -Learning está relacionada a  $\alpha$ , que delimita a convergência do valor  $Q$  para o valor de recompensa. Um valor mais alto de taxa de aprendizado leva a uma convergência mais rápida, no entanto, um valor pequeno de taxa de aprendizado é geralmente empregado para aumentar a robustez por colisões infrequentes causadas por variações de canal (CHU *et al.*, 2015).

## 2.2 Acesso Aleatório para IoT

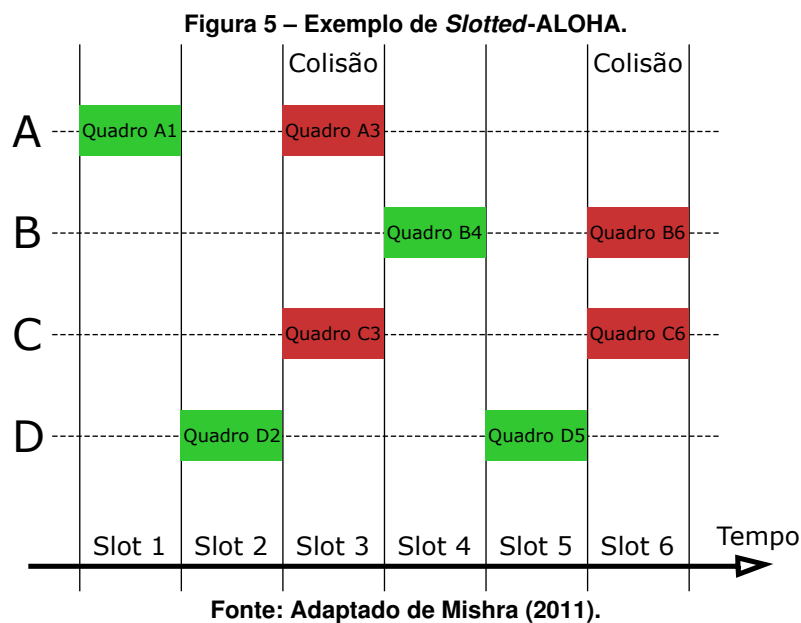
Para atender as demandas de comunicação das mais diversas aplicações de IoT, principalmente em MTC com um número elevado de nós, um tipo de protocolo de acesso múltiplo é requerido. Esse tipo de protocolo permite que vários usuários se comuniquem entre si através dos canais disponíveis, porém, se diferentes usuários utilizarem o mesmo canal para enviar dados ao mesmo tempo, uma colisão pode ocorrer (BERIOLI *et al.*, 2016).

À medida que o número de dispositivos IoT aumenta exponencialmente, o congestionamento para acesso à rede se torna um problema crítico. Algoritmos de RA típicos sofrem de alocação de recursos ineficiente quando o número de dispositivos aumenta, de modo que algumas técnicas de RA foram propostas recentemente para melhorar o desempenho em STRNs (MOON; LEE; LEE, 2018; KIM *et al.*, 2019; SU; REN; ZHANG, 2020; SU *et al.*, 2020). Os autores em (MOON; LEE; LEE, 2018) desenvolveram uma técnica de RA empregando acesso múltiplo de código esparsa, que permite reduzir o atraso de acesso e o número de transmissões de preâmbulos, bem como aumentar a probabilidade de sucesso de RA e a taxa de transmissão. Em uma direção diferente, o trabalho em (KIM *et al.*, 2019) propõe um esquema de RA de duas vias, *i.e.*, requerendo apenas duas mensagens para completar o procedimento de RA, o que reduz a latência ao custo de um número maior de preâmbulos.

Devido ao modo como esse tipo de comunicação é realizado, onde pequenas quantidades de dados são, muitas vezes, transmitidos de forma imprevisível, se torna muito custoso

e ineficiente aplicar um modelo fixo de escolha de canal. Em vista disso, uma alternativa pode ser utilizar protocolos de acesso aleatório, que por escolher o canal de forma sem coordenação, não depende de processos custosos (MUNARI; CLAZZER, 2019). Este tipo de protocolo garante que as estações sejam independentes entre si, onde não há uma prioridade entre os usuários, que podem transmitir dependendo apenas do estado do canal, entre ocupado e disponível.

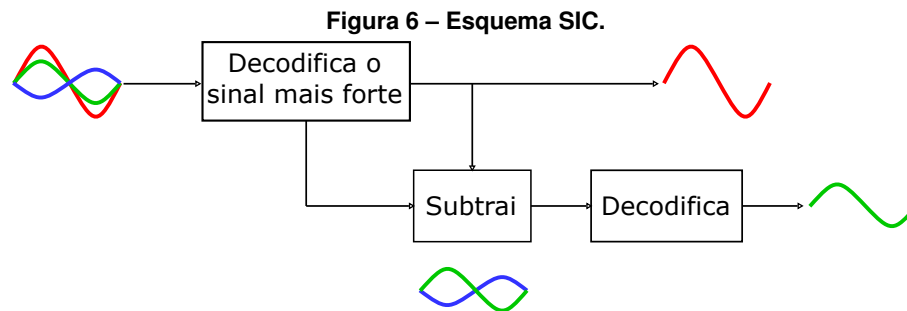
Dentre os protocolos de RA, talvez o mais relevante seja o *Slotted-ALOHA* para o projeto de esquemas de múltiplo acesso, já que através de técnicas de retransmissão é capaz de resolver colisões e otimizar métricas como taxa de transmissão ou atrasos (JADOON *et al.*, 2022). Este protocolo organiza o tempo em intervalos, chamados *timeslots*, assim cada aparelho aloca e transmite no início de um *slot*, depois da geração do pacote. No caso de colisões, os dados são retransmitidos depois de um intervalo de atraso aleatório (RECH; TOMASIN, 2021). A Figura 5 mostra um exemplo de caso do *slotted-ALOHA*, onde nos *slots* 1 e 2 não houve colisões, portanto os usuários A e D puderam transmitir suas informações com sucesso. Porém, no *slot* 3 houve uma colisão pois os usuários A e C transmitiram no mesmo *timeslot*. O usuário C tentou transmitir novamente no *slot* 6, colidindo novamente, dessa vez com o usuário B.



### 2.3 Acesso Múltiplo Não-Ortogonal

Algumas estratégias para lidar com colisões de pacotes em RA são propostas em (SU; REN; ZHANG, 2020; SU *et al.*, 2020). O trabalho em (SU; REN; ZHANG, 2020) emprega multiplexação de intercalação aleatória, de forma que diferentes intercaladores são usados para diferenciar os sinais dos usuários, enquanto um detector de *Maximum a Posteriori Probability* (MAP) é empregado para decodificar os pacotes sobrepostos. Além disso, o trabalho em (SU *et al.*, 2020) usa sequências complexas não ortogonais de vários níveis para pré-codificar os pa-

cotes transmitidos, enquanto os pacotes sobrepostos no receptor são decodificados usando um algoritmo SIC. Quando SIC é empregado no receptor, a interferência entre os usuários pode ser mitigada conforme ilustra a Figura 6, em que o sinal mais forte é decodificado e então subtraído do sinal original para que o próximo sinal seja decodificado.



Fonte: Adaptado de Abdelmoniem *et al.* (2019).

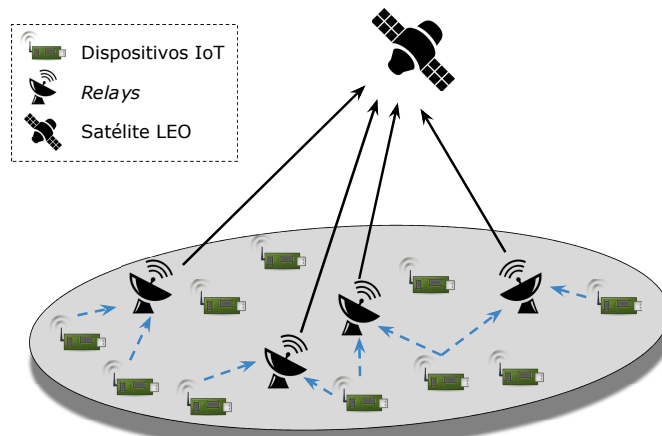
Portanto, o NOMA permite que vários usuários se comuniquem com o receptor usando o mesmo intervalo de tempo e frequência. Voltando ao exemplo da Figura 5, o emprego de NOMA e SIC poderia permitir (caso as iterações ilustradas pela Figura 6 tivessem sucesso) que no *slot* 3 os usuários A e C fossem decodificados simultaneamente, o que aumenta a taxa de transferência do sistema quando comparado ao acesso múltiplo ortogonal (YUAN *et al.*, 2021).

### 3 MÉTODO PROPOSTO

#### 3.1 Modelo do sistema

Consideramos o *uplink* de uma STRN, onde  $D$  dispositivos IoT e  $R$  relays são uniformemente distribuídos em uma área circular de raio  $\mathcal{R}$ , sob a cobertura de um satélite LEO, conforme ilustrado na Figura 7. Assumimos que o acesso direto dos dispositivos ao satélite não é viável devido à potência de transmissão limitada e/ou tipo da antena dos dispositivos (ZHAO *et al.*, 2021). Portanto, os dispositivos IoT transmitem informações a todos os *relays* disponíveis em seu alcance, enquanto a função dos *relays* é permitir a conectividade entre os dispositivos IoT e o satélite.

**Figura 7 – Modelo de sistema com  $D$  dispositivos IoT distribuídos em uma área circular de raio  $\mathcal{R}$ , auxiliados por  $R$  relays que fornecem conectividade a um satélite LEO.**



Fonte: Autoria própria (2022).

O acesso ao meio é baseado no *Slotted-ALOHA grant-free*, onde cada dispositivo IoT transmite em um dos  $T$  *timeslots* disponíveis, escolhendo um dos  $C$  canais disponíveis para se comunicar com os *relays*. Tal enlace de comunicação é ilustrado pelas linhas azuis tracejadas na Figura 7. Então, ao final de cada *frame*, composto por  $T$  *timeslots*, os *relays* enviam pelo canal de retorno um bit de Reconhecimento, do inglês *Acknowledgement* (ACK) por *timeslot*, de forma centralizada, informando o sucesso ou não do recebimento da mensagem. Além disso, usando enlaces dedicados e não interferentes, os *relays* encaminham as mensagens de informação decodificadas com sucesso para o satélite. Observe que a mesma mensagem pode ser recebida por mais de um *relay* e, portanto, ser encaminhada ao satélite mais de uma vez. Consideramos que as multiplicidades são tratadas pelas camadas superiores do protocolo, seja no satélite ou na estação terrestre.

### 3.1.1 Modelo de Transmissão

Dada a transmissão do  $i$ -ésimo dispositivo IoT,  $i \in [1, D]$ , em um *timeslot*  $t \in [1, T]$  e canal  $c \in [1, C]$ , o sinal recebido no  $r$ -ésimo *relay*,  $r \in [1, R]$ , é

$$\mathbf{y}_{r,i} = \sqrt{P \kappa_{r,i}} h_{r,i} \mathbf{x}_i + \mathbf{n}_r, \quad (2)$$

onde eliminamos os índices  $t$  e  $c$  para tornar a notação mais leve, já que os pares de canal de período de atividade são ortogonais entre si. Então,  $P$  é a potência de transmissão dos dispositivos IoT,  $\forall i$ ,  $\mathbf{x}_i$  é o sinal transmitido por um canal complexo de desvanecimento quase-estático, de modo que  $h_{r,i}$  é modelado de acordo com uma distribuição Nakagami- $m$ , e o ruído térmico  $\mathbf{n}_r$  é modelado como um Ruído Branco Gaussiano Aditivo, do inglês *Additive White Gaussian Noise* (AWGN), de média zero e potência  $N = N_0 B F$ , com  $N_0$  sendo a densidade espectral de potência unilateral de ruído,  $B$  a largura de banda e  $F$  a figura de ruído.

Além disso, a perda de percurso entre os dispositivos IoT e os *relays*, que estão em terra, segue (3rd Generation Partnership Project, 2015), dada em dB por

$$\kappa_{r,i} = -128,1 - 37,6 \log_{10}(d_{r,i}) \quad (3)$$

com  $d_{r,i}$  sendo a distância entre o  $r$ -ésimo *relay* e o  $i$ -ésimo dispositivo IoT.

Além disso, como mais de um dispositivo IoT pode escolher o mesmo par canal-*timeslot* para se comunicar com os *relays*, a comunicação está sujeita a interferências. Assim, a mensagem transmitida do  $i$ -ésimo dispositivo IoT, (em um *timeslot*  $t$  e canal  $c$ ), é decodificada com sucesso pelo *relay*  $r$  se a Relação Sinal-Interferência mais Ruído, do inglês *Signal to Interference plus Noise Ratio* (SINR), é maior do que um determinado limiar que permite decodificação com sucesso. Assumindo a capacidade de Shannon, podemos escrever

$$\log_2(1 + \text{SINR}_{r,i}) \geq \beta, \quad (4)$$

que é necessário para uma decodificação bem-sucedida, em que  $\beta$  é a eficiência espectral em bps/Hz.

Assumindo um decodificador SIC, quando o *relay*  $r$  recebe  $k$  transmissões simultâneas, o processo de decodificação começa com o sinal recebido com a maior SINR, enquanto as demais mensagens recebidas simultaneamente são tratadas como ruído. Então, uma vez que o usuário mais forte é decodificado, ele é removido do sinal geral recebido na iteração seguinte (XIA *et al.*, 2018) e a decodificação SIC continua com o próximo dispositivo com a SINR mais alta. Sem perda de generalidade, presumimos que os dispositivos IoT são ordenados em diminuição da potência recebida, de modo que a iteração  $s = 1$  do decodificador SIC começa com o dispositivo IoT mais forte, que é removido do sinal recebido para a próxima iteração (HI-

GUCHI; BENJEBBOUR, 2015), até  $s = k$ , que é a iteração que decodifica o dispositivo IoT mais fraco. Assim, a SINR na  $s$ -ésima iteração do decodificador SIC é

$$\text{SINR}_{r,s} = \frac{P \kappa_{r,s} |h_{r,s}|^2}{\sum_{l=s+1}^k P \kappa_{r,l} |h_{r,l}|^2 + N}, \quad (5)$$

onde  $P \kappa_{r,s} |h_{r,s}|^2$  é a potência instantânea recebida na  $s$ -ésima iteração, enquanto que  $\sum_{l=s+1}^k P \kappa_{r,l} |h_{r,l}|^2$  é tratado como ruído, na iteração seguinte esta potência é removida (HIGUCHI; BENJEBBOUR, 2015).

Por fim, como os *relays* são dispositivos dedicados com antenas adequadas e alta potência de transmissão, possuindo linha de visada para o satélite, consideramos que a probabilidade de falha nos enlaces entre os *relays* e o satélite é desprezível.

### 3.1.2 Métricas do Sistema

A fim de avaliar o desempenho do protocolo proposto, primeiro definimos a taxa de transferência normalizada dos esquemas como

$$\tau^{(\text{sch})} = \beta \frac{\bar{n}_D^{(\text{sch})}}{T \cdot C} \cdot \frac{\bar{n}_{\text{dist}}^{(\text{sch})}}{\bar{n}_{\text{total}}^{(\text{sch})}}, \quad (6)$$

onde  $\bar{n}_D^{(\text{sch})}$  é o número médio de dispositivos IoT que transmitiram com sucesso uma mensagem para, pelo menos, um dos *relays* disponíveis, em cada quadro. Além disso,  $\bar{n}_{\text{dist}}^{(\text{sch})}$  é a média do total de mensagens distintas dos dispositivos IoT recebidos com sucesso pelo satélite, em cada quadro, enquanto  $\bar{n}_{\text{total}}^{(\text{sch})}$  representa o número total de mensagens recebidas pelo satélite<sup>1</sup>. A taxa de transferência normalizada expressa a eficiência com que o quadro está sendo utilizado pelos *relays* e dispositivos IoT.

Além disso, também definimos a taxa de mensagens redundantes na rede como

$$\rho^{(\text{sch})} = 1 - \frac{\bar{n}_{\text{dist}}^{(\text{sch})}}{\bar{n}_{\text{total}}^{(\text{sch})}}, \quad (7)$$

que tende a aumentar quando a mesma mensagem de um dispositivo IoT é corretamente decodificada por vários *relays*.

<sup>1</sup> Notemos que  $\bar{n}_{\text{total}}^{(\text{sch})}$  inclui multiplicidades, pois diferentes *relays* podem ter recebido a mesma mensagem de um dispositivo IoT.

## 3.2 Protocolo Proposto

O *uplink* da STRN enfrenta grandes desafios, especialmente com um grande número de dispositivos, o que pode levar ao congestionamento da rede. Além disso, as estratégias de acesso centralizado muitas vezes não são práticas, seja devido à complexidade do protocolo, sincronização ou mesmo fornecimento limitado de energia dos dispositivos. Para enfrentar esses desafios, uma alternativa depende do uso de algoritmos de aprendizagem inteligentes sem modelo, que podem ser implementados de forma distribuída (SILVA *et al.*, 2020). Neste trabalho, estamos particularmente interessados nas múltiplas transmissões simultâneas entre os dispositivos IoT e os *relays*, que posteriormente encaminharão as informações recebidas para o satélite. Uma vez que muitos dispositivos IoT podem escolher o mesmo par de canais e *timeslot*, empregamos um método baseado em *Q-Learning* para otimizar as escolhas de *timeslot* e canais, para cada dispositivo IoT, de maneira descentralizada.

### 3.2.1 *Q-Learning*

Consideramos que os dispositivos IoT são os agentes do algoritmo *Q-Learning*, cada um manipulando sua própria *Q-Table*, com o par estado-ação representado pelo par de acesso *timeslot*-canal usado para se comunicar com os *relays*. Além disso, consideramos que cada dispositivo sempre escolhe o par com o maior valor em sua *Q-Table*, enquanto a recompensa é definida como

$$\epsilon = \begin{cases} -1, & \text{se a transmissão não ocorreu} \\ 1, & \text{se a transmissão ocorreu.} \end{cases} \quad (8)$$

Notemos que, diferentemente de (ZHAO *et al.*, 2021), propomos um método que requer apenas um único bit de retorno por *timeslot*, da mesma forma que (SILVA *et al.*, 2020), informando o sucesso ou não das transmissões, independentemente do número de dispositivos IoT e canais disponíveis.

## 3.3 Método proposto de RA com *Q-Learning* baseado em NOMA

No protocolo QL-NOMA proposto, consideramos que uma mensagem no  $t$ -ésimo intervalo de tempo usando o canal  $c$ , na  $s$ -ésima iteração do decodificador SIC, é decodificada com sucesso quando a SINR em (5) satisfaz  $\text{SINR}_{r,s} \geq 2^\beta - 1$ .

Então, QL-NOMA começa com a *Q-Table* de cada dispositivo IoT inicializada com 0 em todas as posições. Em seguida, os dispositivos IoT escolhem o valor máximo em suas respectivas *Q-Table*, e transmitem suas mensagens com o par *timeslot*-canal representado pela respectiva posição na *Q-Table*. Em cada *timeslot*, os múltiplos *relays* disponíveis tentam deco-



dificar as mensagens recebidas nos mesmos canais empregando o SIC. No final do quadro, consideramos uma mensagem de retorno centralizada entre os *relays* e os dispositivos IoT, empregando um único bit por canal-*timeslot* para informar se as mensagens foram decodificadas com sucesso ou não naquele par de canal-*timeslot* específico. A confirmação só é positiva quando todos os usuários no mesmo canal-*timeslot* forem decodificados corretamente via SIC. Caso contrário, mesmo que só alguns sejam decodificados com sucesso e outros não, a confirmação do *relay* é negativa.

Com esse retorno, cada dispositivo IoT atualiza sua *Q-Table* com a recompensa definida em (8), antes de prosseguir com a próxima tentativa de transmissão.

O protocolo proposto é descrito no Algoritmo 1, com  $f \in [1, F]$  representando um quadro entre um total de  $F$  quadros. Notemos que o protocolo proposto está altamente alinhado com os requisitos de uma rede IoT com baixos recursos computacionais, uma vez que *Q-Learning* adiciona complexidade computacional mínima em cada dispositivo para calcular (1), e baixa necessidade de memória para armazenar uma única tabela *Q-Table*, com tamanho  $C \times T$ . Por outro lado, a complexidade nos *relays* é maior, principalmente devido ao processamento do SIC. No entanto, os *relays* são dispositivos dedicados, com provavelmente muito mais energia, processamento e capacidades de armazenamento do que os dispositivos IoT.

---

**Algoritmo 1 – Método de RA com *Q-Learning* baseado em NOMA.**

---

**requer** *Q-Table* inicializada com 0 em todas as posições

- 1: **para todos**  $f \in [1, F]$  **faça**
  - 2:   **para todos**  $t \in [1, T]$  **faça**
  - 3:     Dispositivos IoT com o maior *Q-value* correspondente ao *timeslot*  $t$  transmitem
  - 4:     Os dispositivos IoT selecionados transmitem no canal  $c$ , também definido pelo *Q-value* mais alto
  - 5:     Os *relays* empregam SIC para decodificar possíveis mensagens sobrepostas transmitidas no mesmo canal
  - 6:   **finaliza para**
  - 7: **finaliza para**
  - 8: Os *relays* transmitem um único bit de *feedback* por *timeslot*, seguindo (8)
  - 9: **para todos**  $i \in [1, D]$  **faça**
  - 10:   O dispositivo IoT  $i$  atualiza sua *Q-Table* usando (1)
  - 11: **finaliza para**
  - 12: Os *relays* encaminham as mensagens decodificadas para o satélite
- 

**Fonte: Autoria própria (2022).**

## 4 RESULTADOS E DISCUSSÃO

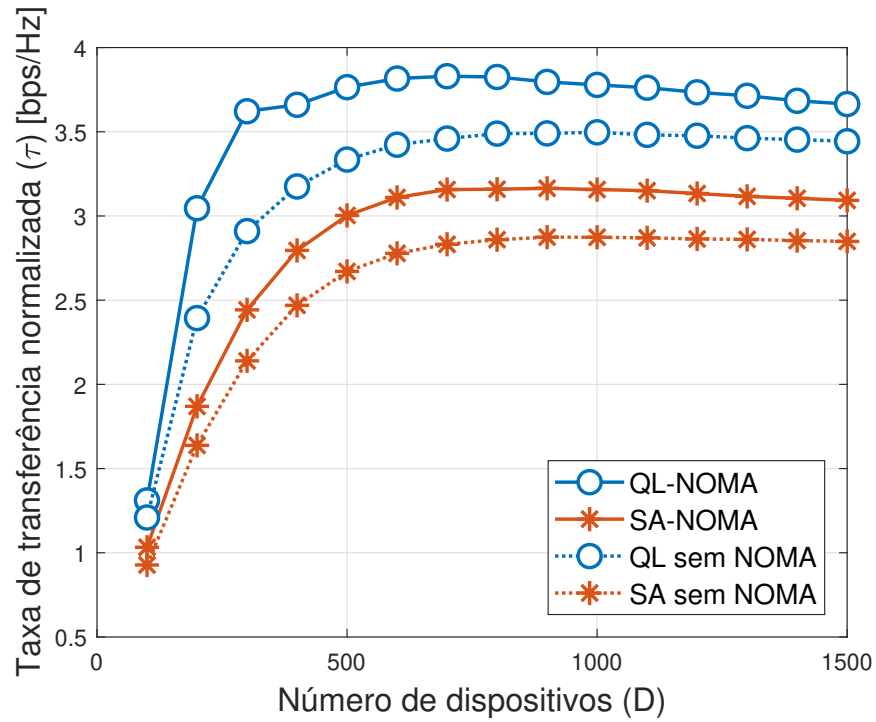
Nesta seção, comparamos numericamente o esquema QL-NOMA proposto com o protocolo SA-NOMA. Assumimos  $T = 100$  *timeslots* dentro de um quadro e um total de  $F = 50$  quadros. Os dispositivos IoT e os *relays* são considerados aleatoriamente distribuídos em uma célula com um raio  $\mathcal{R} = 5$  km, enquanto o satélite LEO está a uma altitude de 780 km. A potência de transmissão dos dispositivos IoT é  $P = 10$  dBm, distribuição Nakagami- $m$  com  $m = 2$  é assumida para o desvanecimento do canal,  $\beta = 3$  bps/Hz,  $N_0 = -174$  dBm/Hz,  $B = 100$  kHz e  $F = 6$  dB. Além disso, assumimos uma taxa de aprendizado de  $\alpha = 0,1$  e um fator de desconto de  $\gamma = 0,5$  para o algoritmo *Q-Learning*, que foram otimizados numericamente para melhorar o desempenho. Finalmente, as curvas apresentam o resultado médio sobre a execução de 100 simulações.

### 4.1 Desempenho Geral do QL-NOMA

Primeiro, o Gráfico 1 compara o desempenho da taxa de transferência de QL-NOMA e SA-NOMA como uma função do número de dispositivos IoT ( $D$ ), com  $R = 4$  *relays* e  $C = 1$  canal. Além disso, também consideramos ambos os protocolos sem empregar NOMA, usando um modelo de colisão suave, onde apenas o dispositivo mais forte em uma colisão pode ser potencialmente decodificado nos *relays*. Como observamos, o protocolo QL-NOMA proposto aumenta consideravelmente a taxa de transferência em comparação com SA-NOMA, com a taxa de transferência máxima de QL-NOMA sendo cerca de 20% maior do que a de SA-NOMA. O Gráfico 2 complementa a análise com a taxa de mensagens redundantes no mesmo cenário. Como podemos observar, a taxa de mensagens redundantes é muito semelhante para QL-NOMA e SA-NOMA, com um ligeiro aumento quando NOMA é utilizado. Note também que observamos no Gráfico 1 um efeito de saturação nos esquemas sem NOMA, uma vez que o modelo de colisão suave adotado assume que pelo menos o dispositivo com maior Relação Sinal-Ruído, do inglês *Signal to Noise Ratio* (SNR), consegue ser decodificado pelos *relays*. Tal efeito também se manifesta no Gráfico 2, uma vez que o número de mensagens redundantes cai com o aumento de dispositivos, tendo em vista que todos os pares canal-*timeslot* ficam ocupados pelos usuários mais próximos dos *relays*, isto é, de maior SNR.

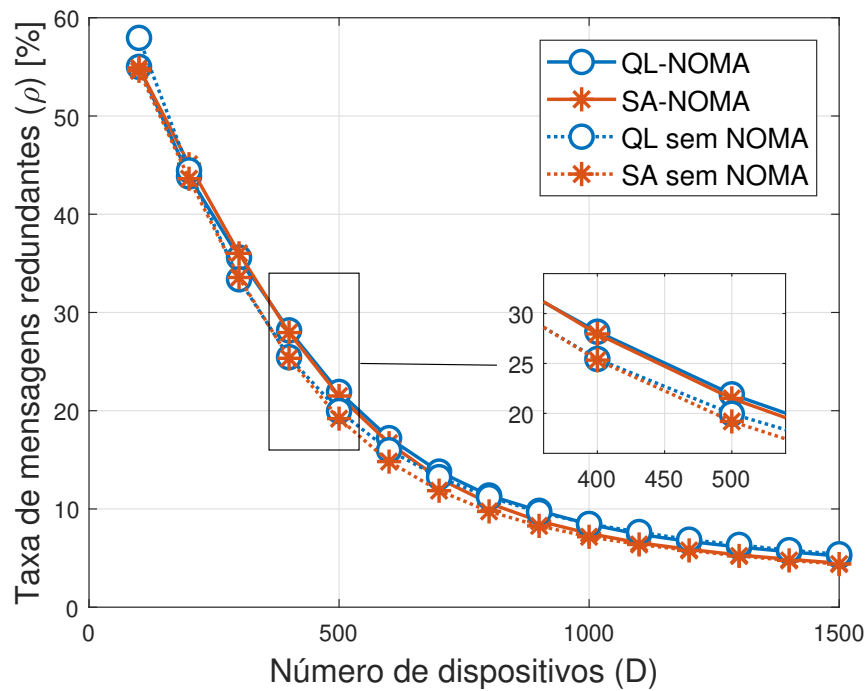
Para complementar a análise, o Gráfico 3 ilustra a importância da escolha adequada de  $C$  dependendo do número de dispositivos IoT ativos na rede. O Gráfico 3 plota a taxa de transferência da rede em função de  $C$  com  $D = 400$  dispositivos IoT servidos por  $R = 4$  *relays*. Como observamos, há uma melhora acentuada na taxa de transferência com o aumento de  $C$ , que diminui quando  $C > 3$ . Isso se deve ao fato de  $C$  estar no denominador da (6), de modo que há uma relação de troca entre um número maior de colisões quando  $C$  é pequeno e uma taxa de transferência diminuída quando  $C$  aumenta.

**Gráfico 1 – Taxa de transferência em função de  $D$ , com  $R = 4$  relays e  $C = 1$  canal disponível.**



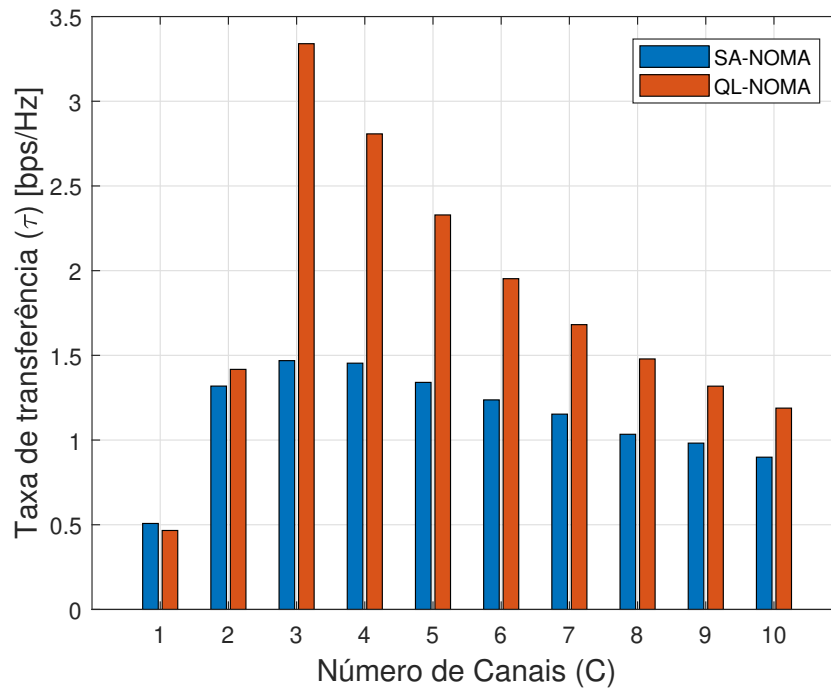
Fonte: Autoria própria (2022).

**Gráfico 2 – Taxa de mensagens redundantes em função de  $D$ , com  $R = 4$  relays e  $C = 1$  canal disponível.**



Fonte: Autoria própria (2022).

**Gráfico 3 – Taxa de transferência em função de  $C$ , com  $D = 400$  dispositivos IoT servidos por  $R = 4$  relays.**



Fonte: Autoria própria (2022).

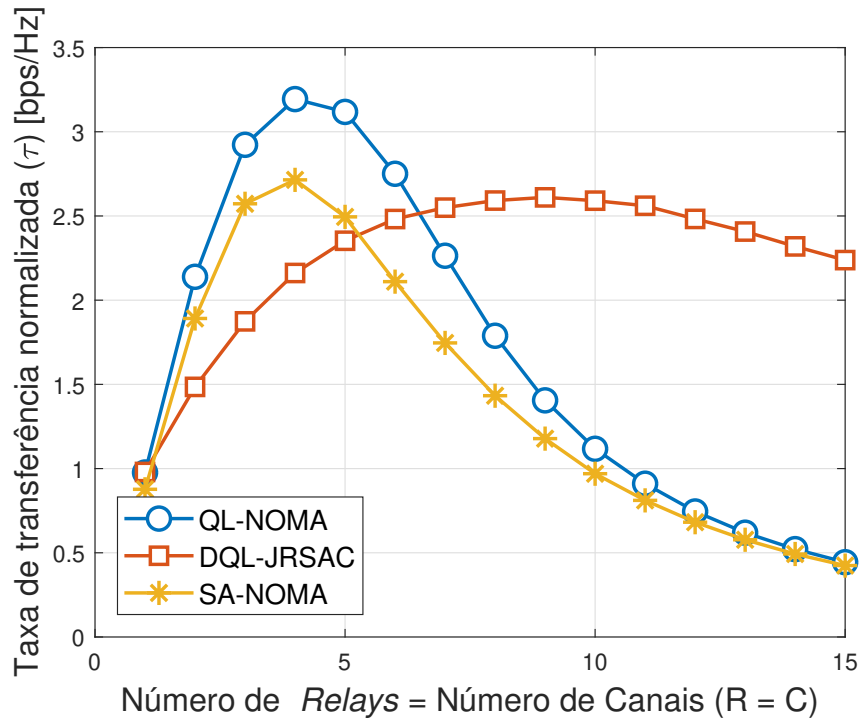
#### 4.2 Efeito do Número de Relays

O efeito do número de relays na taxa de transferência normalizada é avaliado no Gráfico 4, onde  $D = 1500$  e  $C = 3$ . No gráfico, plotamos  $\tau^{(\text{sch})}$  para um cenário em que o número de relays é igual ao número de canais ortogonais, ou seja,  $R = C$ . Além disso, também comparamos QL-NOMA e SA-NOMA com o protocolo DQL-JRSAC proposto em (ZHAO *et al.*, 2021). Nesse esquema, cada relay possui um único canal, ortogonal ao canal dos outros relays, enquanto a escolha do relay/canal e do timeslot é otimizada por uma abordagem *Q-Learning* nos dispositivos IoT. Como notamos, aumentar o número de relays (ou canais) nem sempre é benéfico para a taxa de transferência da rede, em nenhum dos esquemas considerados. No entanto, se o número de relays e canais for otimizado, a taxa de transferência normalizada do protocolo QL-NOMA proposto supera os outros esquemas, sendo 23% maior do que o de DQL-JRSAC e 18% maior do que o de SA-NOMA. Além disso, como observamos no Gráfico 4, um número menor de relays ( $R = 4$  relays) é exigido pelo QL-NOMA em comparação com o DQL-JRSAC, que requer  $R = 9$  relays para o desempenho otimizado.

#### 4.3 Efeito da Potência de Transmissão

A seguir, o Gráfico 5 mostra a taxa de transferência normalizada de QL-NOMA e SA-NOMA como uma função da potência de transmissão ( $P$ ) dos dispositivos IoT, considerando

**Gráfico 4 – Taxa de transferência normalizada como uma função de  $C = R$ , com  $D = 1500$  dispositivos IoT.**



Fonte: Autoria própria (2022).

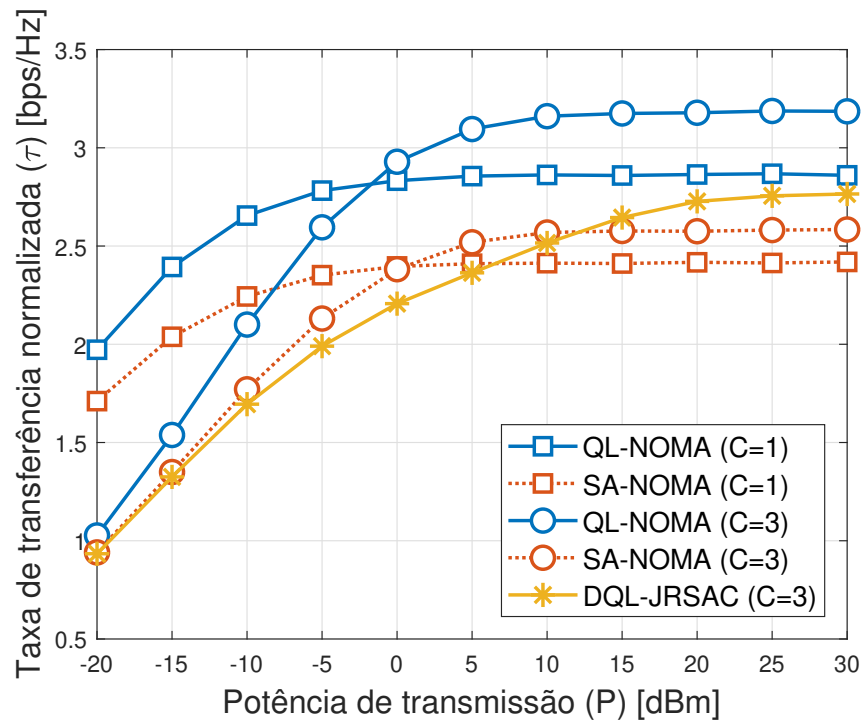
$D = 1500$ ,  $R = 3$  e  $C \in \{1, 3\}$ . Além disso, o DQL-JRSAC com  $R = 3$  também está incluído para comparação. Como podemos notar, a taxa de transferência normalizada aumenta com  $P$ , com QL-NOMA superando tanto SA-NOMA quanto DQL-JRSAC. Além disso, também observamos que DQL-JRSAC alcança a taxa de transferência de QL-NOMA apenas com alta potência de transmissão, indicando que uma abordagem mais eficiente em energia também é obtida com o esquema QL-NOMA proposto.

#### 4.4 Avaliação dos Parâmetros do $Q$ -Learning

Na sequência, os Gráficos 6 e 7 plotam o rendimento enquanto variam os atributos  $\alpha$  e  $\gamma$  do método  $Q$ -Learning. Como esperado, os valores de taxa de transferência permanecem constantes para SA-NOMA em ambos os gráficos. No Gráfico 6 investigamos o impacto da taxa de aprendizado em um cenário com um número fixo de dispositivos. Como observamos, existe um valor intermediário ótimo para  $\alpha$ , o que demonstra a importância de uma escolha adequada dessa variável.

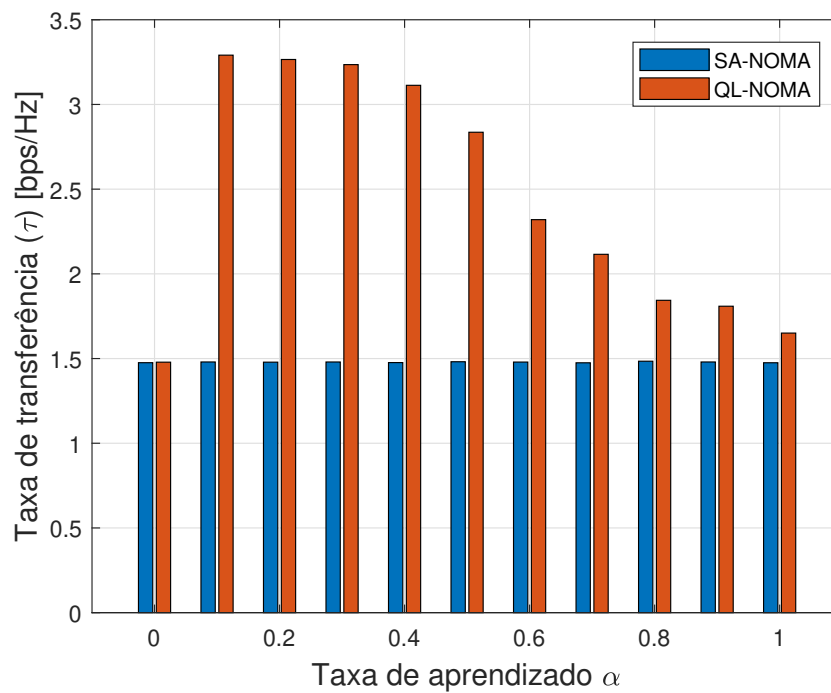
A seguir, a taxa de transferência para diferentes valores de  $\gamma$  é demonstrado no Gráfico 7. Como observamos, valores mais altos de  $\gamma$  permitem valores mais altos de taxa de transferência. Como  $\gamma$  representa a penalidade no método  $Q$ -Learning quando ocorre uma colisão, um  $\gamma$  menor pode descartar canais e intervalos de tempo em potencial muito rapidamente

**Gráfico 5 – Taxa de transferência normalizada em função de  $P$  com  $D = 1500$ ,  $R = 3$  e  $C \in \{1, 3\}$ .**



Fonte: Autoria própria (2022).

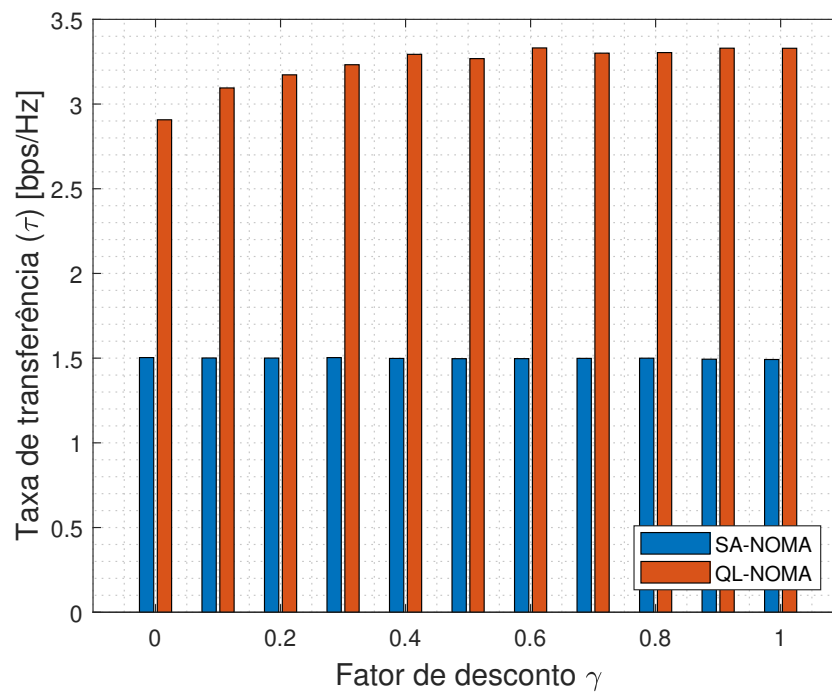
**Gráfico 6 – Taxa de transferência em função de  $\alpha$  do algoritmo  $Q$ -Learning com  $\gamma = 0,5$ ,  $D = 400$ ,  $C = 3$  e  $R = 4$ .**



Fonte: Autoria própria (2022).

após uma falha de transmissão ou decodificação SIC, enquanto um  $\gamma$  maior pode recuperar a rede mais rápido e melhor alocar os dispositivos IoT proporcionando maior rendimento.

**Gráfico 7 – Taxa de transferência em função de  $\gamma$  do algoritmo *Q-Learning* com  $\alpha = 0,25$ ,  $D = 400$ ,  $C = 3$  e  $R = 4$ .**



Fonte: Autoria própria (2022).

## 5 CONCLUSÕES E PERSPECTIVAS DE TRABALHOS FUTUROS

Nesta dissertação foi apresentado um método distribuído de RA que utiliza *Q-Learning* para decidir o canal e o *timeslot* dentro de um sistema baseado em *Slotted-ALOHA*, com o objetivo de melhorar a taxa de transferência da rede. *Q-Learning* é um algoritmo de aprendizado por reforço, uma das técnicas de ML, que aprende a partir da relação entre um agente e o ambiente, onde uma ação do agente gera uma recompensa do ambiente, permitindo assim que o agente aprenda qual a melhor escolha de ação. A ação, neste trabalho, se caracteriza pela escolha do canal e do *timeslot* que cada dispositivo IoT utilizará para se comunicar dentro do protocolo *Slotted-ALOHA*. Como diferentes usuários ainda podem utilizar o mesmo *timeslot* para transmitir, uma colisão pode ocorrer, para mitigar esses casos foram empregados NOMA, que permite que vários usuários compartilhem do mesmo intervalo de tempo, frequência e recursos, e SIC, que é capaz de decodificar os sinais mais fortes recebidos.

O cenário utilizado consiste em uma área onde são distribuídos dispositivos IoT e *relays*, de forma que cada dispositivo transmite para um *relay*, que então retransmite para um satélite, para isso consideramos que o enlace entre o *relay* e o satélite é perfeito. Assim, o foco foi no enlace entre os dispositivos IoT e os *relays*, onde o sinal é transmitido por um canal complexo de desvanecimento quase-estático com AWGN e o canal e *timeslot* são escolhidos a partir de uma *Q-Table*, exigindo uma pequena quantidade de memória e processamento nos dispositivos. O sinal recebido é decodificado utilizando SIC e retorna um bit indicando o sucesso da transmissão e atualizando a *Q-Table*.

Observamos que o desempenho do sistema, em termos de taxa de transferência normalizada, depende da escolha do número de *relays* e canais disponíveis. Como  $C$  está no denominador de (6), há uma compensação clara. Quando  $C$  é pequeno, há um maior número de colisões, comprometendo a taxa de transferência. No entanto, aumentar  $C$  diminui a taxa de transferência devido ao uso ineficiente do espectro. Portanto, a escolha ótima de  $C$  é muito relevante. Além disso, implantar muitos *relays* nem sempre é benéfico para o rendimento da rede em nenhum dos esquemas considerados, uma vez que o número de mensagens redundantes também aumenta com  $R$ , ou seja, mais de um *relay* é capaz de receber com sucesso a mesma mensagem de um dispositivo IoT. Portanto, nossos resultados indicam que a otimização de  $C$  e  $R$  é importante para maximizar o uso eficiente dos recursos da rede, enquanto  $C$  e  $R$  ideais dependem do número de dispositivos IoT que se comunicam na rede. Os resultados numéricos mostram que o esquema QL-NOMA proposto supera o SA-NOMA e o esquema DQL-JRSAC (ZHAO *et al.*, 2021) em termos de taxa de transferência normalizada, enquanto uma otimização adequada do número de *relays* e canais ortogonais, que por sua vez depende da densidade da rede, é crucial.

Como sugestões de trabalhos futuros, formas de alocação de pilotos é um tema importante. Tendo em vista que não é prático alocar pilotos ortogonais para todos os dispositivos, o



impacto da alocação do piloto e da interferência do piloto na estimativa do canal e, portanto, no desempenho geral, é um aspecto importante a ser investigado em um futuro próximo.

Além disso, seria interessante considerar e investigar o fato de que o enlace entre os *relays* e o satélite não é perfeito, requerendo mais uma camada de otimização para garantir uma taxa de transferência adequada, minimizando a possível colisão de pacotes. Para isso, também é possível implementar alguma técnica de ML, até mesmo outro sistema de *Q-Learning*.

## REFERÊNCIAS

- 3rd Generation Partnership Project. *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (e-utra); Further Advancements for E-utra Physical Layer Aspects (release 9)*. [S.l.], 2015.
- 6G Flagship. Key drivers and research challenges for 6G ubiquitous wireless intelligence. *University of Oulu, Tech. Rep. 6G Research Visions 1*, v. 1, 2019.
- ABDELMONIEM, M. *et al.* Enhanced NOMA system using adaptive coding and modulation based on LSTM neural network channel estimation. **Applied Sciences**, v. 9, p. 3022, 07 2019.
- AKHTAR, M. W. *et al.* The shift to 6G communications: Vision and requirements. 2020.
- BACCO, M. *et al.* IoT applications and services in space information networks. **IEEE Wireless Communications**, v. 26, n. 2, p. 31–37, 2019.
- BERIOLI, M. *et al.* **Modern Random Access Protocols**. [S.l.]: NOW Publisher, 2016.
- CHETTRI, L.; BERA, R. A comprehensive survey on internet of things (IoT) toward 5G wireless systems. **IEEE Internet of Things Journal**, v. 7, n. 1, p. 16–32, 2020.
- CHU, Y. *et al.* Application of reinforcement learning to medium access control for wireless sensor networks. **Engineering Applications of Artificial Intelligence**, v. 46, p. 23–32, 2015. ISSN 0952-1976.
- CISCO. **Cisco visual networking index: Global mobile data traffic forecast update, 2017-2022**. [S.l.], 2019.
- CISCO. **Cisco Annual Internet Report (2018–2023)**. [S.l.], 2020.
- DING, J. *et al.* IoT connectivity technologies and applications: A survey. **IEEE Access**, Institute of Electrical and Electronics Engineers (IEEE), v. 8, p. 67646–67673, 2020.
- ERICSSON. **Cellular IoT Evolution for Industry Digitalization**. 2020. <https://www.ericsson.com/en/reports-and-papers/white-papers/cellular-iot-evolution-for-industry-digitalization>.
- FANG, X. *et al.* 5G embraces satellites for 6G ubiquitous IoT: Basic models for integrated satellite terrestrial networks. **IEEE Internet of Things Journal**, v. 8, n. 18, p. 14399–14417, 2021.
- HIGUCHI, K.; BENJEBBOUR, A. Non-orthogonal Multiple Access (NOMA) with Successive Interference Cancellation for Future Radio Access. **IEICE Transactions on Communications**, E98.B, n. 3, p. 403–414, 2015.
- IEEE Innovate. **IEEE is Fueling the Fourth Industrial Revolution**. 2021. <https://innovate.ieee.org/innovation-spotlight-ieee-fueling-fourth-industrial-revolution/>.
- IEEE Spectrum. **3GPP Release 15 Overview**. [S.l.], 2020. Accessed January 04, 2021. Disponível em: <https://spectrum.ieee.org/telecom/wireless/3gpp-release-15-overview>.
- IOT-ANALYTICS. **State of the IoT 2020: 12 billion IoT connections, surpassing non-IoT for the first time**. [S.l.], 2020. Disponível em: <https://iot-analytics.com/state-of-the-iot-2020-12-billion-iot-connections-surpassing-non-iot-for-the-first-time/>.

- ITU-R. **Framework and overall objectives of the future development of IMT for 2020 and beyond**. [S.l.], 2015. Disponível em: [https://www.itu.int/dms\\_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-!!!PDF-E.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-!!!PDF-E.pdf).
- ITU-R. **IMT Traffic Estimates for the Years 2020 to 2030**. [S.l.], 2015. Disponível em: [https://www.itu.int/dms\\_pub/itu-r/opb/rep/R-REP-M.2370-2015-PDF-E.pdf](https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2370-2015-PDF-E.pdf).
- JADOON, M. A. *et al.* Deep reinforcement learning for random access in machine-type communication. **CoRR**, abs/2201.09841, 2022. Disponível em: <https://arxiv.org/abs/2201.09841>.
- JIANG, W. *et al.* The road towards 6G: A comprehensive survey. **IEEE Open Journal of the Communications Society**, v. 2, p. 334–366, 2021.
- KIM, S. *et al.* Low latency random access for small cell toward future cellular networks. **IEEE Access**, v. 7, p. 178563–178576, 2019.
- LIU, G. *et al.* Vision, requirements and network architecture of 6G mobile network beyond 2030. **China Communications**, v. 17, n. 9, p. 92–104, 2020.
- MAHMOOD, N. H. *et al.* Machine type communications: key drivers and enablers towards the 6G era. **EURASIP Journal on Wireless Communications and Networking**, v. 2021, n. 1, p. 134, jun. 2021. ISSN 1687-1499.
- MISHRA, A. A. Enhanced slotted aloha. In: **2011 3rd International Conference on Electronics Computer Technology**. [S.l.: s.n.], 2011. v. 5, p. 278–282.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of Machine Learning**. Segunda edição. [S.l.]: The MIT Press, 2018.
- MOON, S.; LEE, H.-S.; LEE, J.-W. SARA: Sparse code multiple access-applied random access for IoT devices. **IEEE Internet of Things Journal**, v. 5, n. 4, p. 3160–3174, 2018.
- MUNARI, A.; CLAZZER, F. Modern random access for beyond-5g systems: a multiple-relay ALOHA perspective. **CoRR**, abs/1906.02054, 2019. Disponível em: <http://arxiv.org/abs/1906.02054>.
- POPOVSKI, P. *et al.* 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view. **IEEE Access**, v. 6, p. 55765–55779, 2018.
- RECH, A.; TOMASIN, S. Coordinated random access for industrial IoT with correlated traffic by reinforcement-learning. **CoRR**, abs/2109.08389, 2021. Disponível em: <https://arxiv.org/abs/2109.08389>.
- SHARMA, S. K.; WANG, X. Collaborative distributed Q-learning for RACH congestion minimization in cellular IoT networks. **IEEE Communications Letters**, v. 23, n. 4, p. 600–603, 2019.
- SILVA, M. V. da *et al.* A NOMA-based Q-learning random access method for machine type communications. **IEEE Wireless Communications Letters**, v. 9, n. 10, p. 1720–1724, 2020.
- SINHA, S. **State of IoT 2021: Number of connected IoT devices growing 9% to 12.3 billion globally, cellular IoT now surpassing 2 billion**. 2021. Disponível em: <https://iot-analytics.com/number-connected-iot-devices/>.
- SU, J. *et al.* Randomly pre-coded packets based random access scheme for IoT-oriented satellite networks. **IEEE Access**, v. 8, p. 221148–221161, 2020.

SU, J.; REN, G.; ZHANG, H. Random interleaving multiplexing based random access in IoT-oriented satellite networks. In: **IEEE 91st Vehicular Technology Conference (VTC2020-Spring)**. [S.l.: s.n.], 2020. p. 1–5.

SUTTON, R. S.; BARTO, A. G. **Reinforcement Learning: An Introduction**. Second. [S.l.]: The MIT Press, 2018.

TARIQ, F. *et al.* A speculative study on 6G. **IEEE Wireless Communications**, v. 27, n. 4, p. 118–125, 2020.

TUBIANA, D. A. *et al.* Q-Learning NOMA Random Access for IoT-Satellite Terrestrial Relay Networks. **IEEE Wireless Communications Letters**, p. 1–1, 2022.

XIA, B. *et al.* Outage Performance Analysis for the Advanced SIC Receiver in Wireless NOMA Systems. **IEEE Transactions on Vehicular Technology**, v. 67, n. 7, p. 6711–6715, 2018.

YUAN, Y. *et al.* NOMA for next-generation massive IoT: Performance potential and technology directions. **IEEE Communications Magazine**, v. 59, n. 7, p. 115–121, 2021.

ZHAO, B. *et al.* Distributed  $Q$ -learning based joint relay selection and access control scheme for IoT-oriented satellite terrestrial relay networks. **IEEE Commun. Letters**, v. 25, n. 6, p. 1901–1905, 2021.