COPPE UFRJ
Instituto Alberto Luiz Coimbra de
Pós-Graduação e Pesquisa de Engenharia

SELEÇÃO DE PARÂMETROS DE TRANSMISSÃO DE REDES LORA ATRAVÉS DE TÉCNICAS DE APRENDIZADO DE MÁQUINAS

Mário Nascimento Carvalho Filho

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Miguel Elias Mitre Campista

Rio de Janeiro
Junho de 2022

SELEÇÃO DE PARÂMETROS DE TRANSMISSÃO DE REDES LORA ATRAVÉS DE TÉCNICAS DE APRENDIZADO DE MÁQUINAS

Mário Nascimento Carvalho Filho

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientador: Miguel Elias Mitre Campista

Aprovada por: Prof. Miguel Elias Mitre Campista
Prof. Rodrigo de Souza Couto
Prof. Leobino Nascimento Sampaio

RIO DE JANEIRO, RJ – BRASIL
JUNHO DE 2022

Carvalho Filho, Mário Nascimento

Seleção de Parâmetros de Transmissão de Redes LoRa Através de Técnicas de Aprendizado de Máquinas/Mário Nascimento Carvalho Filho. – Rio de Janeiro: UFRJ/COPPE, 2022.

XII, 62 p.: il.; 29, 7cm.

Orientador: Miguel Elias Mitre Campista

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2022.

Referências Bibliográficas: p. 59 – 62.

1. LoRa. 2. Aprendizado de Máquina. 3. Seleção. 4. Parâmetros de Transmissão. 5. Aprendizado por Reforço. I. Campista, Miguel Elias Mitre. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

A Deus e à minha família.

Agradecimentos

Primeiramente, agradeço a Deus por conceder-me saúde para enfrentar os desafios da vida. Agradeço aos meus pais que sempre apoiaram o meu crescimento pessoal e profissional e que trabalharam muito para proporcionar-me sempre o melhor.

Expresso minha gratidão ao meu orientador Miguel Elias Mitre Campista pela orientação acadêmica e atenção que me foram dadas durante todo o mestrado. Agradeço também ao professor Rodrigo de Souza Couto pelo apoio nos assuntos acadêmicos e na fase de integração ao mestrado.

Agradeço a todo o Grupo de Sistemas de Armas do Instituto de Pesquisas da Marinha (IPqM) por auxiliar-me na escolha do tema e no fornecimento de informações que me auxiliaram no desenvolvimento da dissertação.

Gostaria de agradecer também aos meus amigos e a todos que, de alguma forma, me ajudaram na condução dos trabalhos.

Finally, I would like to acknowledge Pixel perfect, Freepik and netscript from https://www.flaticon.com/ for the icons used in this paper.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

SELEÇÃO DE PARÂMETROS DE TRANSMISSÃO DE REDES LORA ATRAVÉS DE TÉCNICAS DE APRENDIZADO DE MÁQUINAS

Mário Nascimento Carvalho Filho

Junho/2022

Orientador: Miguel Elias Mitre Campista

Programa: Engenharia Elétrica

O aumento da densidade das redes *Long Range* (LoRa) leva o mecanismo Adative Data Rate (ADR) do protocolo LoRaWAN a apresentar queda de desempenho em termos de taxa de entrega de pacotes e energia consumida, uma vez que as interferências internas devido às transmissões concorrentes aumentam. Tais interferências aumentam a dinamicidade do ambiente, exigindo um mecanismo proativo de ajuste de parâmetros de transmissão que garanta a capacidade dos enlaces. Além disso, o mecanismo ADR se mostra incapaz de ajustar os parâmetros quando o sinal de rádio é degradado, o que pode ocorrer, por exemplo, caso a rede seja submetida à interferência externa. Esta dissertação propõe a utilização de técnicas de aprendizado supervisionado e de aprendizado por reforço para ajustar os parâmetros de transmissão. As técnicas de aprendizado supervisionado Árvore de Decisão e Máquina de Vetor de Suporte são utilizadas para selecionar os parâmetros de transmissão e, ao mesmo tempo, reduzir o consumo de energia. A técnica de aprendizado por reforço *Q-learning* é utilizada para ajustar os parâmetros de transmissão, mapeando a probabilidade de ocorrência de colisão e otimizando a taxa de entrega de pacotes e a energia consumida. Dois estudos de caso são avaliados. O primeiro representa uma rede densa com o *gateway* localizado no centro da topologia, enquanto o segundo representa uma rede com dispositivos distantes entre 5 e 10 km do *gateway* submetida a uma fonte de interferência externa. No primeiro estudo de caso, os resultados mostram que, para uma quantidade de até 100 dispositivos, o desempenho da técnica *Q-learning* supera o mecanismo ADR. No segundo estudo de caso, o mecanismo ADR se mostra incapaz de manter a capacidade dos enlaces e o desempenho da técnica *Q-learning* indica que esta pode ser uma alternativa para ajustar os parâmetros.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

SELECTION OF LORA NETWORK TRANSMISSION PARAMETERS THROUGH MACHINE LEARNING TECHNIQUES.

Mário Nascimento Carvalho Filho

June/2022

Advisor: Miguel Elias Mitre Campista

Department: Electrical Engineering

The increase in LoRa networks density lead the LoRaWAN protocol's Adaptive Data Rate (ADR) mechanism to show a drop in performance, in terms of packet delivery rate and energy consumed, since internal interference due to concurrent transmissions increases. Such interference increases the dynamics of the environment, requiring a proactive mechanism to adjust transmission parameters to guarantee the capacity of the links. Furthermore, the ADR mechanism is unable to adjust the parameters when the radio signal is degraded, which can occur, for example, if the network is subjected to external interference. This dissertation proposes the use of supervised learning and reinforcement learning techniques to adjust the transmission parameters. The decision tree and support vector machine supervised learning techniques are used to select the transmission parameters and, at the same time, reduce the power consumption. The reinforcement learning technique Q-learning is used to adjust the transmission parameters online, mapping the probability of collision occurrence for each selected parameter and optimizing the packet delivery rate and the energy consumed in the transmission. Two case studies are evaluated. The first represents a dense network with the gateway located in the center of the topology, while the second represents a network with devices located among 5 and 10 km subjected to an external source of interference. In the first case, the results show that, considering a number up to 100 devices, the performance of the Q-learning technique outperforms the ADR mechanism. In the second case, the ADR mechanism is unable to maintain the capacity of the links and the performance of the Q-learning technique indicates that this can be an effective alternative to dynamically adjust the transmission parameters.

Sumário

Lista de Figuras **x**

Lista de Tabelas **xii**

1 Introdução **1**

1.1 LPWANs e a importância das Redes LoRa 1

1.2 Motivação 3

1.3 Contribuição 4

1.4 Submissões 5

1.5 Organização da Dissertação 5

2 Trabalhos Relacionados **7**

2.1 Seleção de parâmetros sem a utilização de aprendizado de máquinas . 7

2.2 Seleção de parâmetros com a utilização de aprendizado de máquinas . 9

3 Tecnologias LPWAN e LoRa **12**

3.1 Tecnologia LoRa 13

3.2 Protocolo LoRaWAN 18

4 Aprendizado Supervisionado e o Aprendizado por Reforço **22**

4.1 Aprendizado Supervisionado 23

4.1.1 Árvore de Decisão 23

4.1.2 Máquina de Vetor de Suporte 25

4.2 Aprendizado por Reforço 26

5 Propostas de Seleção de Parâmetros **31**

5.1 Proposta de Utilização do Aprendizado Supervisionado 31

5.2 Proposta de utilização do aprendizado por reforço 32

6 Experimentos e Resultados **38**

6.1 Estudo de caso 1: *Gateway* no centro da topologia e dispositivos com posições aleatórias 38

6.2 Estudo de caso 2: Topologia circular com *gateway* a 5-10 km do centro da topologia 40
6.3 Parâmetros de simulação 42
6.4 Resultados do estudo de caso 1: *Gateway* no centro da topologia e dispositivos com posições aleatórias 44
6.5 Resultados do estudo de caso 2: Topologia circular com *gateway* de 5 a 10 km do centro da topologia 52
6.5.1 Experimento sem a implementação das propostas 52
6.5.2 Experimento com a implementação das propostas 54

7 Conclusões e Trabalho Futuros **57**

Referências Bibliográficas **59**

Lista de Figuras

1.1 Relação entre parâmetros das tecnologias LPWAN, celular e *ZigBee*. . 2

3.1 Ilustração de uma rede LoRa e seus componentes. 13
3.2 Modulação de símbolos com aumento de frequência [1]. 15
3.3 Espectrograma dos fatores de espalhamento. Da esquerda para a direita são os fatores de espalhamento 7, 8, 9, 10, 11 e 12 respectivamente. 16
3.4 Espectrograma de um exemplo de transmissão LoRa com fator de espalhamento 9 com vários símbolos. 16
3.5 Efeito do ajuste dos parâmetros no alcance da comunicação dos dispositivos finais (círculos pretos) com seus respectivos *gateways* (triângulos vermelhos). 18
3.6 Diagrama temporal das classes de dispositivo. 19
3.7 O formato da unidade de dados da camada física e MAC [2]. 20

4.1 Exemplo de Árvore de Decisão a partir de uma base de dados composta pelas coordenadas x e y da posição dos dispositivos, o fator de espalhamento e a potência de transmissão selecionados. 24
4.2 Hiperplanos de separação sem os vetores de suporte. 25
4.3 Hiperplanos de separação com os vetores de suporte. 25
4.4 Ciclo do Aprendizado por Reforço. 27

5.1 Processo de aprendizado dos agentes. 32
5.2 Criação e treinamento dos agentes a nível de servidor. O treinamento do agente 1 é evidenciado com a captação do status da última transmissão e configuração do dispositivo 1 com os parâmetros selecionados através do sinal *downlink*. 33
5.3 Processo de aprendizado dos agentes. 37

6.1 Seleção do fator de espalhamento para os dispositivos próximos ao *gateway* para evitar a colisão. 39

6.2 Ilustração do estudo de caso 2 no qual um conjunto de 5 sensores acústicos submarinos, conectados cada um a um dispositivo LoRa, em uma determinada área marítima com formação circular. O *gateway* da rede está instalado em um navio distante. 40
6.3 Topologia da rede no estudo de caso 2 com a fonte de interferência. . 41
6.4 Taxa de entrega de pacotes obtida pelas propostas selecionando fator de espalhamento para uma quantidade de 40 a 100 dispositivos. . . . 45
6.5 Recompensa média obtida pelos agentes ao longo do treinamento com 25000 episódios no estudo de caso 1 com 100 dispositivos para seleção de fator de espalhamento. 46
6.6 Energia consumida na transmissão obtida pelas propostas para quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento. . 46
6.7 Eficiência energética obtida pelas técnicas utilizadas para uma quantidade de 40 a 100 dispositivos na seleção do fator de espalhamento. . 47
6.8 Taxa de entrega de pacotes obtida pelas propostas selecionando fator de espalhamento e potência de transmissão para quantidade de 40 a 100 dispositivos. 48
6.9 Recompensa média obtida pelos agentes ao longo do treinamento com 25000 episódios no estudo de caso 1 com 100 dispositivos para seleção de fator de espalhamento e potência de transmissão. 49
6.10 Energia consumida na transmissão obtida pelas propostas para uma quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento e potência de transmissão. 49
6.11 Eficiência energética obtida pelas propostas para uma quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento e potência de transmissão. 50
6.12 Seleção dos Fatores de espalhamento e potências de transmissão das técnicas de aprendizado para a quantidade de 100 dispositivos. 51
6.13 Taxa de entrega de pacotes mantendo fixos cada fator de espalhamento em um cenário com e sem interferência e com *gateway* a uma distância de 10 km. 53
6.14 Taxa de entrega de pacotes para cada fator de espalhamento em um cenário com e sem interferência e com *gateway* a uma distância de 5 km. 54
6.15 Fatores de espalhamento selecionados pelas propostas no estudo de caso 2. 55

Lista de Tabelas

3.1 Taxa de bits para cada fator de espalhamento. 17

6.1 Parâmetros de simulação do estudo de caso 1. 44
6.2 Parâmetros de simulação do estudo de caso 2. 44
6.3 Taxa de entrega de pacotes obtida pelas propostas no cenário cuja rede é submetida à interferência externa. 55

Capítulo 1

Introdução

As tecnologias associadas ao conceito de *Internet of Things* (IoT), que vão desde microeletrônica até computação em nuvem, têm se tornado a base do desenvolvimento de diversas aplicações e serviços modernos. Consequentemente, a quantidade de dispositivos IoT tem crescido nos últimos anos, exigindo pesquisas que atendam as demandas da sociedade. Nesse contexto, existem aplicações que, apesar da grande quantidade de dispositivos conectados, geram um montante relativamente pequeno de dados. Esse cenário favorece redes que ofereçam conectividade em larga escala, cobertura de grandes áreas e baixo consumo de energia, como é o caso das redes *Low Power Wide Area Networks* [3, 4] (LPWAN). A tecnologia *Long Range* (LoRa) é atualmente a principal representante, já sendo bastante utilizada em uma gama considerável de aplicações.

1.1 LPWANs e a importância das Redes LoRa

As LPWANs permitem a comunicação entre centenas de dispositivos sem fio e a estação base a grandes distâncias com baixo consumo de energia. Isso ocorre em virtude da alta sensibilidade da estação base, o que não é uma característica de outras tecnologias concorrentes, como o WiFi, as redes celulares e o *zigbee*. A redução do consumo de energia é alcançada graças ao compromisso entre a taxa de transmissão de dados e o alcance das comunicações [4]. Por um lado, as LPWANs, as redes WiFi e as redes *Bluetooth* possuem consumo de energia reduzido, sendo as LPWANs as redes que possuem maior alcance de rádio. Por outro lado, as redes celulares e as LPWANs possuem raios de cobertura extensos, sendo as LPWANs menos custosas energeticamente. Essas características fazem com que as LPWANs sejam mais atrativas para aplicações de larga escala que exigem baixo consumo de energia [3]. O compromisso vem na taxa de transmissão que, nas redes LPWAN, é baixa dada a reduzida largura de banda. A Figura 1.1 representa a relação entre os parâmetros de cada uma das tecnologias.

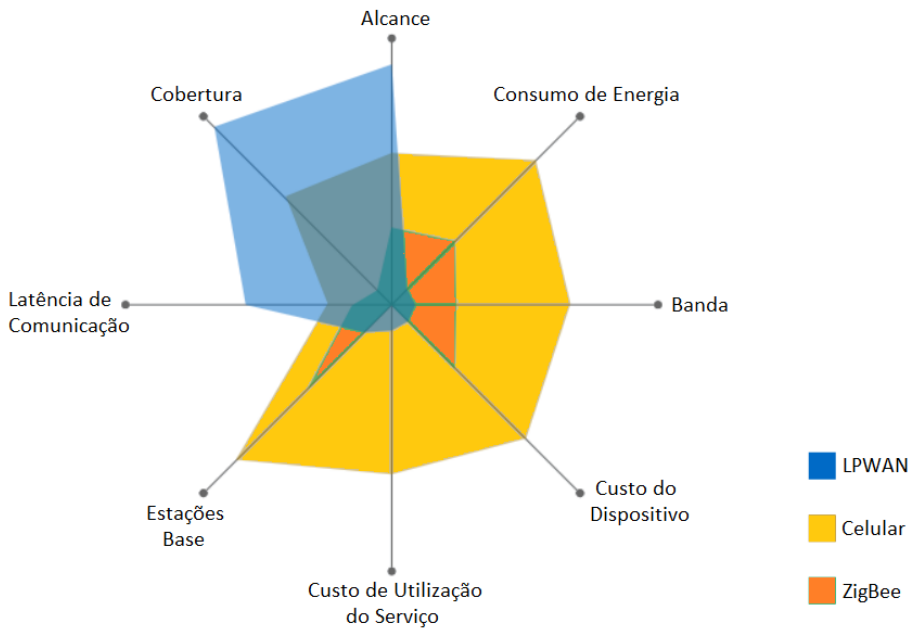


Figura 1.1: Relação entre parâmetros das tecnologias LPWAN, celular e *ZigBee*.

Nas LPWANs, os dispositivos finais tipicamente utilizam baterias e, portanto, o consumo de energia não pode ser alto [3]. Além disso, os dispositivos finais se comunicam com a estação base a distâncias de dezenas de quilômetros, como consequência da sua resiliência à interferência e à facilidade na decodificação. Essas redes suportam redes densas de forma que seu desempenho não seja consideravelmente afetado com o aumento do número de dispositivos finais implementados. O *chip* de rádio implementado nos dispositivos LPWANs é de baixo custo, uma vez que não necessitam de grandes unidades de processamento.

As tecnologias *Long Range* (LoRa), Sigfox, *Long Term Evolution - Machine Type Communication* (LTE-M) e *Narrow Band Internet of Things* (NB-IoT) são exemplos de LPWANs [3]. A NB-IoT e Sigfox utilizam faixas de frequência licenciadas, o que as tornam mais onerosas apesar de serem bastante conhecidas e utilizadas. As tecnologias LoRa e LTE-M utilizam faixas de frequência não licenciadas e consomem menos energia em relação ao NB-IoT e Sigfox, além de ter maior resiliência à interferência. Já a NB-IoT e LTE-M oferecem uma maior taxa de transmissão de dados. O LoRa possui um protocolo de acesso ao meio chamado de *Long Range Wide Area* (LoRaWAN) de padrão aberto. Tanto o LoRaWAN quanto o protocolo de acesso ao meio do Sigfox são baseados no protocolo de acesso ao meio ALOHA puro [5]. As redes NB-IoT e Sigfox podem ser disponibilizadas a partir de um contrato de operação firmado, enquanto as redes LoRaWAN podem ser implementadas como redes privadas, assim como ocorre com as redes WiFi. O Sigfox suporta uma taxa de transmissão máxima de 100 bps, enquanto NB-IoT e o LoRa suportam 200 e 50 kbps, respectivamente [3].

Como consequência das características destacadas, dentre as LPWANs, a tecnologia LoRa se destaca. O LoRa é uma tecnologia de camada física que modula os sinais nas faixas de frequência sub-GHz Industrial, Científica e Médica não licenciadas (variando entre 6 MHz e 244 GHz). Dessa forma, o LoRa permite a criação e implementação de redes privadas nas localizações dos clientes, sem restrições, sendo essa uma grande vantagem em comparação ao Sigfox e ao NB-IoT, já que ambas são controladas por operadores de rede.

A especificação LoRa inclui duas camadas, sendo uma delas a camada física (LoRa RF) e, a segunda, a camada de enlace, chamada de *Long Range Wide Area* (LoRaWAN). O baixo consumo de energia e a comunicação através de longas distâncias são características alcançadas pelo LoRa graças ao esquema de modulação *Chirp Spread Spectrum* (CSS), utilizado na camada física, e da utilização do LoRaWAN, baseada no ALOHA puro e que permite a utilização do mecanismo *Adaptive Data Rate* (ADR) para ajustar o fator de espalhamento e a potência de transmissão. Porém, o aumento de colisões devido ao maior *Time on Air* (ToA) das transmissões de pacotes tem se tornado uma desvantagem da tecnologia. Esse problema está relacionado à seleção inadequada de parâmetros da camada física, especificamente valores de potência de transmissão e de fator de espalhamento dos dispositivos finais. De forma simplificada, enquanto maiores potências de transmissão diminuem as oportunidades de reaproveitamento espacial, maiores fatores de espalhamento tornam o tempo no ar dos sinais transmitidos mais longos. Fatores de espalhamento selecionados com valores iguais por diferentes dispositivos com instantes de transmissão sobrepostos também contribuem para o aumento das colisões [6]. Além disso, mesmo com diferentes fatores de espalhamento, a colisão entre sinais no mesmo canal pode ocorrer devido à ortogonalidade imperfeita dos fatores de espalhamento [7, 8]. Portanto, para minimizar a probabilidade de colisões e, consequentemente, aprimorar o acesso ao meio, deve-se buscar a configuração apropriada de fatores de espalhamento e potências de transmissão.

1.2 Motivação

Em aplicações de LPWANs, as redes LoRa podem se tornar densas devido ao aumento na demanda por sensoriamento. Em tais casos, o fenômeno da interferência entre dispositivos muito próximos do *gateway* é maior, o que exige uma seleção adequada dos parâmetros de rede, em especial os parâmetros da camada física. O desempenho do mecanismo de seleção de parâmetros utilizado pelo protocolo LoRaWAN tem sua eficácia comprometida quando a quantidade de dispositivos finais é escalado. Isso ocorre pois, antes do mecanismo de ajuste dos parâmetros ser invocado, muitos pacotes precisam ser perdidos ou retransmitidos. Essa é uma ca-

racterística do mecanismo de ajuste de parâmetros do LoRaWAN que é reativo. Além disso, o tempo de reação desse mecanismo se torna um problema para aplicações nas quais a tomada de decisão de ajuste de parâmetros precisa ser realizada em um intervalo de tempo muito curto. Por isso, algoritmos inteligentes e proativos de ajuste de parâmetros da camada física devem ser utilizados para dirimir a queda de desempenho do mecanismo utilizado pelo LoRaWAN.

Além das aplicações de LPWANs, as redes LoRa também têm sido utilizadas em aplicações, como monitoramento ambiental, redes de sensores subaquáticos, ou até mesmo em *smartgrids*. Um exemplo destas aplicações é a utilização de redes LoRa para testes de sistemas de armas. A Marinha do Brasil (MB), por exemplo, utiliza redes LoRa para realizar testes de precisão de seus canhões. Os dispositivos finais da rede são conectados a hidrofones e dispostos em uma formação circular de raio entre 250 e 500 metros, enquanto o *gateway* fica localizado no navio artilheiro a uma distância entre 5 e 20 km do centro da formação dos dispositivos. O navio realiza disparos de canhão e, para avaliar a precisão dos disparos, cada hidrofone da rede subaquática capta a intensidade do sinal acústico gerado pela queda do projétil na água, transmitindo essa informação para o navio artilheiro através de um módulo LoRa (dispositivo final da rede LoRa).

O navio artilheiro, portanto, é capaz de saber a localização da queda de cada projétil disparado para calibrar o canhão. Porém, problemas de conectividade são relatados, o que pode indicar a ocorrência de interferências devido a fontes externas à rede LoRa utilizada. Nesse caso, nem a pré-seleção dos parâmetros de transmissão pelo operador da rede nem o mecanismo ADR do protocolo LoRaWAN se mostram eficazes na manutenção da capacidade da rede, indicando a necessidade de um mecanismo proativo de ajuste de parâmetros de transmissão.

Essa aplicação precisa coletar informações sobre a localização da queda do projétil para a calibração do canhão, exigindo a maior quantidade de informações possível a ser enviada para o centro de controle. Logo, é fundamental que sejam utilizados mecanismos que garantam a capacidade dos enlaces durante toda a sua operação, uma vez que qualquer fonte de interferência externa pode degradar os enlaces e, consequentemente, comprometer a aplicação.

1.3 Contribuição

Abordagens de aprendizado de máquinas, como o aprendizado supervisionado e por reforço, se mostram alternativas adequadas na seleção inteligente de parâmetros da camada física para dirimir as interferências internas e manter a economia de energia. Além disso, técnicas de aprendizado por reforço permitem não só dirimir as interferências internas, mas também as interferências externas que podem degradar

a capacidade da rede previamente estabelecida, garantindo a conectividade entre os dispositivos com o ajuste de seus parâmetros de forma *online*.

Esta dissertação propõe dois estudos de caso para avaliar o desempenho de técnicas de aprendizado de máquinas diversas no ajuste de parâmetros de transmissão. São utilizadas técnicas de aprendizado supervisionado Árvore de Decisão e Máquina de Vetor de Suporte e uma técnica de aprendizado por reforço para seleção dos parâmetros fator de espalhamento e potência de transmissão.

Em relação à utilização de técnicas de aprendizado supervisionado, não há na literatura trabalhos que busquem ajustar tanto o fator de espalhamento quanto a potência de transmissão. Já em relação à utilização da técnica de aprendizado por reforço, o algoritmo utilizado para otimizar a taxa de entrega de pacotes e a energia consumida é diferente dos algoritmos já utilizados na literatura.

Sabe-se que o desempenho do ADR é comprometido com o aumento do número de dispositivos finais, aumentando as colisões e impedindo a escalabilidade da rede. Além disso, o ADR tem seu desempenho comprometido quando a dinamicidade da rede é grande. Portanto, esta dissertação propõe utilizar técnicas de aprendizado supervisionado e por reforço para reduzir as colisões devido à interferências internas e externas. São realizadas comparações entre o desempenho das técnicas supramencionadas, em termos de taxa de entrega de pacotes e energia consumida, com o desempenho do mecanismo ADR.

1.4 Submissões

Este trabalho já resultou nas seguintes submissões de trabalhos científicos:

- Filho, C. N. M., Campista, M. E. M. - "Economia de Energia Através da Configuração Inteligente de Parâmetros da Camada Física em Redes LoRa", in XL Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuído (SBRC 2022), Fortaleza, Brazil, May 2022 (Aceito para publicação).

1.5 Organização da Dissertação

A partir deste ponto, este trabalho é organizado conforme se segue:

- Capítulo 2 - Trabalhos Relacionados

  Este capítulo descreve trabalhos da literatura que investigam o desempenho do protocolo LoRaWAN em termos de escalabilidade e que propõem melhorias do protocolo LoRaWAN.

- Capítulo 3 - Tecnologias LPWAN e LoRa

  Este capítulo aborda as características das redes LoRa, assim como discorre sobre a importância da utilização do aprendizado de máquinas para a seleção de parâmetros. O protocolo LoRaWAN também é apresentado neste capítulo.

- Capítulo 4 - Aprendizado Supervisionado e o Aprendizado por Reforço

  Este capítulo discorre sobre o aprendizado supervisionado e o aprendizado por reforço com o objetivo de ajudar no entendimento das propostas e dos experimentos realizados.

- Capítulo 5 - Propostas de Seleção de Parâmetros

  Neste capítulo, são apresentadas as propostas de utilização do aprendizado supervisionado e do aprendizado por reforço para selecionar os parâmetros fator de espalhamento e potência de transmissão.

- Capítulo 6 - Experimentos e Resultados

  Este capítulo apresenta os experimentos realizados e os resultados obtidos. As propostas são implementadas em dois estudos de caso diversos. O primeiro estudo de caso considera uma topologia na qual o *gateway* ocupa uma posição central da topologia e os dispositivos finais ocupam posições diversas aleatórias na área da topologia que é circular. O segundo estudo de caso trata de uma rede cujos dispositivos finais formam um círculo e cujo *gateway* está a uma distância entre 5 km e 20 km de distância do centro da formação circular dos dispositivos. É realizada uma comparação entre o desempenho das propostas de aprendizado supervisionado, aprendizado por reforço e do desempenho do ADR.

- Capítulo 7 - Conclusão e trabalhos futuros

  A partir dos resultados dos experimentos, as conclusões a respeito do desempenho das propostas são apresentadas neste capítulo. Algumas propostas também são sugeridas como futuros trabalhos para posterior comparação com os resultados desta dissertação.

Capítulo 2

Trabalhos Relacionados

Diversos algoritmos desenvolvidos para otimizar transmissões em redes LoRa já foram propostos até o momento. Em geral, existem duas abordagens principais para lidar com redes LoRa de maior escala. Uma das abordagens busca a seleção de parâmetros sem a utilização de aprendizado de máquinas, enquanto a outra abordagem foca na seleção de parâmetros com a utilização de aprendizado de máquinas.

2.1 Seleção de parâmetros sem a utilização de aprendizado de máquinas

Cuomo, F. et al. [9] propuseram dois algoritmos chamados EXPLoRa-SF e EXPLoRa-AT para selecionar o fator de espalhamento de rede LoRaWAN.

O algoritmo EXPLoRa-SF seleciona fatores de espalhamento baseado no número total de dispositivos conectados, enquanto o EXPLoRa-AT realiza a seleção dos parâmetros baseado no ToA para cada fator de espalhamento. Os resultados mostraram que o desempenho do protocolo LoRaWAN foi superado pelo desempenho dos algoritmos propostos. Apesar dos resultados favoráveis obtendo desempenho superior ao do ADR, as propostas apresentadas preconizaram a seleção apenas do fator de espalhamento.

Bor et. al. [6] propuseram diferentes combinações de parâmetros da camada física para comparar o desempenho de tais combinações com o desempenho do protocolo LoRaWAN em redes densas. O desempenho foi avaliado em termos de taxa de extração de dados e de consumo de energia. São definidas diferentes combinações de parâmetros e três cenários diferentes. As combinações se diferenciam entre si pelos valores de fator de espalhamento, largura de banda e taxa de código. No primeiro cenário, os dispositivos são configurados com mesma combinação e os valores dos parâmetros são fixos. No segundo cenário, os dispositivos são configurados com a mesma combinação, porém o fator de espalhamento e potência de transmissão

são reduzidos em sequência. O terceiro cenário é semelhante ao segundo, porém são considerados mais de um *gateway*. Objetivo era avaliar o impacto da seleção dinâmica dos parâmetros e da utilização de mais de um *gateway*. Para realizar os experimentos, os autores desenvolveram e utilizaram o simulador chamado LoRaSIM. Os autores concluíram que a escalabilidade da rede aumenta quando o fator de espalhamento e a potência de transmissão são selecionados de forma dinâmica e quando a quantidade de *gateway* aumenta. A técnica utiliza, porém, um mecanismo semelhante ao ADR, reduzindo gradativamente o fator de espalhamento e a potência de transmissão a partir de seus valores máximos, tornando a proposta limitada em relação à seleção dinâmica de parâmetros para otimizar a taxa de extração de dados e a energia consumida. Além disso, como a proposta se baseia em regras e não considera a possível dinamicidade da rede, não se mostra capaz de se adaptar às mudanças do ambiente.

Slabicki et al. [10] propuseram uma versão aprimorada do mecanismo ADR, alterando o indicador de qualidade dos enlaces como também a política de incremento da potência de transmissão nos dispositivos finais. Os resultados mostraram que o ADR convencional é efetivo apenas quando a variância do canal é nula ou muito baixa e que a proposta obteve um desempenho superior em termos de na taxa de entrega de pacotes, mantendo um baixo consumo de energia em canais cuja variância é alta. Para realizar os experimentos, os autores desenvolveram o simulador chamado Flora em OMNET++. Porém, segundo os próprios autores, a proposta não obtém o mesmo desempenho em redes densas, necessitando de mecanismos que levem em conta a probabilidade de colisões e a distribuição dos parâmetros de transmissão entre os dispositivos finais.

Uma proposta de seleção de diferentes valores para o fator de espalhamento para os dispositivos finais foi apresentada por Aggarwal e Nassipuri [11]. A seleção dos fatores de espalhamento foi feita para dispositivos finais localizados próximo ao *gateway*, promovendo uma melhoria na taxa de entrega de pacotes. Porém, a atribuição de fatores de espalhamento é feita de forma explícita, não de forma inteligente e automática como a proposta desta dissertação.

O algoritmo proposto por Chinchilla et al. [12] para reduzir as colisões em redes LoRa divide o meio sem fio em blocos de recursos, no qual cada bloco utiliza um valor de fator de espalhamento e uma frequência para o canal. Embora o algoritmo aumente a capacidade da rede, a otimização da energia consumida na transmissão não foi considerada.

Adbelfadeel et. al. [13] propuseram um mecanismo de agendamento de transmissão de dados mais eficiente energicamente chamado FREE. Nesse mecanismo, em vez de transmitir os dados assim que são gerados, são programadas janelas de tempo fixas definidas através do algoritmo proposto. Embora a proposta elimine o

problema das colisões no LoRaWAN, essa solução de agendamento não é escalável para redes densas, uma vez que cada dispositivo final deve aguardar sua janela de tempo alocada. A proposta desta dissertação difere de [13] uma vez que permite reduzir as colisões de pacotes sem a necessidade de agendar janelas de transmissão. Além disso, vale salientar que o agendamento de janelas de transmissão provoca atrasos que são prejudiciais em aplicações nas quais o tempo seja um fator crítico.

Uma estratégia para seleção de fatores de espalhamento chamada *"Sequential Waterfilling"* foi apresentada por Bianchi et al. [14]. O mecanismo utilizado equaliza o ToA dos pacotes transmitidos pelos dispositivos finais em cada grupo de fator de espalhamento. O algoritmo utilizado para realizar a equalização considera uma configuração de rede com vários *gateways* e leva em consideração o efeito de captura do canal, uma vez que este efeito possui um papel mais significativo em redes LoRa em relação à outras tecnologias. Os resultados mostraram que a proposta supera o desempenho do ADR em termos de taxa de extração de dados em redes densas com múltiplos *gateways*. Porém, a proposta ajusta apenas o fator de espalhamento e não foram apresentadas informações sobre seu desempenho caso a capacidade dos enlaces seja degradada devido a uma fonte de interferência externa.

Portanto, as propostas supramencionadas que tratam da seleção eficiente dos parâmetros da camada física ora ajustam apenas o fator de espalhamento, ora ajustam fator de espalhamento e potência de transmissão, mas não são capazes de prover um ganho significativo em termos de taxa de entrega de pacotes e energia consumida por se tratarem de mecanismos reativos e incapazes de ajustar os parâmetros quando os enlaces sofrem alta variação na sua qualidade. Esta dissertação propõe mecanismos proativos de ajuste de parâmetros em redes densas e em redes submetidas a uma alta variação dos enlaces.

2.2 Seleção de parâmetros com a utilização de aprendizado de máquinas

Abordagens baseadas em aprendizado de máquinas têm sido utilizadas em muitos problemas de redes, como tarefas de classificação de tráfego, previsão de falhas, roteamento, gerenciamento de recursos, controle de congestionamento e segurança de redes [15, 16]. Para tarefas que envolvem ambientes dinâmicos, o aprendizado por reforço tem ganhado maior atenção devido à facilidade na implementação dos algoritmos e por ajudar em tarefas que exijam tomadas de decisão em aplicações que envolvem redes, como agendamento de transmissão, adaptação de parâmetros e alocação de recursos.

Yatagan e Oktug, por outro lado, exploraram técnicas de aprendizado supervisi-

onado para seleção do fator de espalhamento que fosse capaz de otimizar a taxa de entrega de pacotes em redes LoRa [17]. Os autores utilizaram as técnicas de Árvore de Decisão e Máquina de Vetor de Suporte. Os autores avaliaram apenas o efeito da ortogonalidade imperfeita dos fatores de espalhamento, já que ponto havia sido implementado no simulador desenvolvido e utilizado pelos próprios autores. Todos os demais parâmetros da camada física do LoRa foram mantidos fixos.

Park et al. propuseram a utilização da técnica de aprendizado por reforço profundo para ajustar o fator de espalhamento e a potência de transmissão [18]. A área na qual os dispositivos finais foram dispostos era de apenas 1,5 quilômetros quadrados e a quantidade de dispositivos foi de apenas 30. Os parâmetros de simulação não permitem avaliar o desempenho da proposta em redes densas, enquanto esta dissertação realiza testes com uma quantidade de dispositivos de até 1000 sensores.

Algoritmos como *Multi Armed Bandit* (MAB) e *Q-learning* têm sido propostos para melhorar o desempenho de protocolos de acesso ao meio em redes de sensores sem fio (WSN) [19]. Alguns desses algoritmos são o ALOHA-Q [20] e RL-MAC [21], que utilizam algoritmos baseados em *Q-learning* para ajustar de forma adaptativa o acesso ao meio. Tais protocolos agendam as transmissões dos dispositivos considerando alguns parâmetros como a largura de banda e a carga de tráfego.

Outras propostas focam na redução da latência e da energia consumida em redes de sensores sem fio, como em [22], que promove um algoritmo baseado em *Q-learning* para ajustar os parâmetros de transmissão. Algoritmos de aprendizado por reforço também são utilizados para ajustar os parâmetros de transmissão de dispositivos IoT para maximizar a eficiência energética e a confiabilidade na transmissão de dados [23].

Uma proposta de utilização da técnica Multi Armed Bandit (MAB), chamada LoRa-MAB, foi proposta por Ta et al. [24] para ajustar o fator de espalhamento, a potência de transmissão e a frequência de uma rede LoRaWAN de forma distribuída (técnica implementada nos dispositivos finais) através do uso do algoritmo *Exponential Weights for Exploration and Exploitation* (EXP3). Apesar do aumento da taxa de entrega de pacotes e da redução no consumo de energia, a complexidade computacional da proposta é alta, a ponto das simulações considerarem apenas 100 dispositivos finais em uma grande área. Tal problema impediu que o desempenho da proposta fosse avaliado em redes com uma quantidade acima de 100 dispositivos finais de forma a tornar a rede mais escalável. A proposta desta dissertação considera uma quantidade de dispositivos acima de 100 com o intuito de avaliar o desempenho em redes densas e com um algoritmo que promove uma convergência de treinamento mais rápida do que o algoritmo proposto pela referência. Além disso, não é realizada a otimização da energia, uma vez que os dispositivos finais têm a opção de escolher qualquer dos níveis de potência disponíveis, sem buscar a redução

do consumo de energia. A proposta desta dissertação, contudo, busca otimizar o consumo de energia.

Portanto, o protocolo LoRaWAN apresenta oportunidades para sua melhoria, o que exige uma melhor compreensão da camada física e de acesso ao meio deste protocolo. As abordagens de aprendizado por reforço têm se mostrado soluções promissoras para selecionar os parâmetros de transmissão dos dispositivos finais. Esta dissertação propõe a utilização de um algoritmo de aprendizado por reforço que permite superar o desempenho do mecanismo ADR em termos de taxa de entrega de pacotes e energia consumida em redes densas e em redes cuja capacidade dos enlaces possui alta variação.

Capítulo 3

Tecnologias LPWAN e LoRa

As LPWANs são redes sem fio que oferecem conectividade entre os dispositivos finais e os *gateways* a um baixo consumo de energia para uma grande quantidade de dispositivos finais distribuídos em uma grande área. Tais atributos existem graças aos esquemas de modulação utilizados por essas redes, como o *narrow-band* e o *chirp spread spectrum* [3], como também à operação na faixa de frequências sub-GHz. Estas características favorecem a propagação, uma vez que existe menos atenuação, menor congestionamento e menor desvanecimento em frequências mais baixas quando comparado à banda de 2,4 GHz na faixa de frequências não-licenciada *Industrial, Scientific and Medical* (ISM).

As tecnologias LPWANs adotam diferentes ciclos de trabalho para reduzir a potência de operação e o consumo de energia das baterias dos dispositivos finais. Tais tecnologias definem o ciclo de trabalho a partir do padrão do tráfego, tipo de fonte de energia e dos requisitos da aplicação. As LPWANs possuem, em geral, uma topologia estrela e protocolos que permitem o acesso ao meio de forma aleatória. A topologia estrela permite que os dispositivos finais se comuniquem diretamente com o *gateway* sem a necessidade de encaminhar pacotes de outros dispositivos. Além disso, protocolos de acesso ao meio mais simples permitem a economia de energia uma vez que não introduzem sobrecarga excessiva na rede [3]. Além da economia de energia, a utilização de faixas de frequência não-licenciadas, como a banda ISM [3] em 169 MHz, 433 MHz, 868 MHz, 915 MHz e 2,4 GHz proporcionam uma redução no custo para as operadoras de rede.

Diversas técnicas têm sido utilizadas para ajustar o esquema de modulação e selecionar os parâmetros de transmissão com o intuito de permitir o aumento do número de dispositivos implementados. Isso requer um monitoramento eficiente da qualidade do enlace, e a sincronização entre os *gateways* e os dispositivos finais. Em geral, a utilização de dispositivos finais mais complexos e o aumento na sobrecarga de comunicação tanto no *uplink* quanto no *downlink*, resultam no aumento do consumo de energia e na utilização de dispositivos mais custosos.

3.1 Tecnologia LoRa

A Figura 3.1 representa uma rede LoRa em sua topologia estrela e seus componentes.

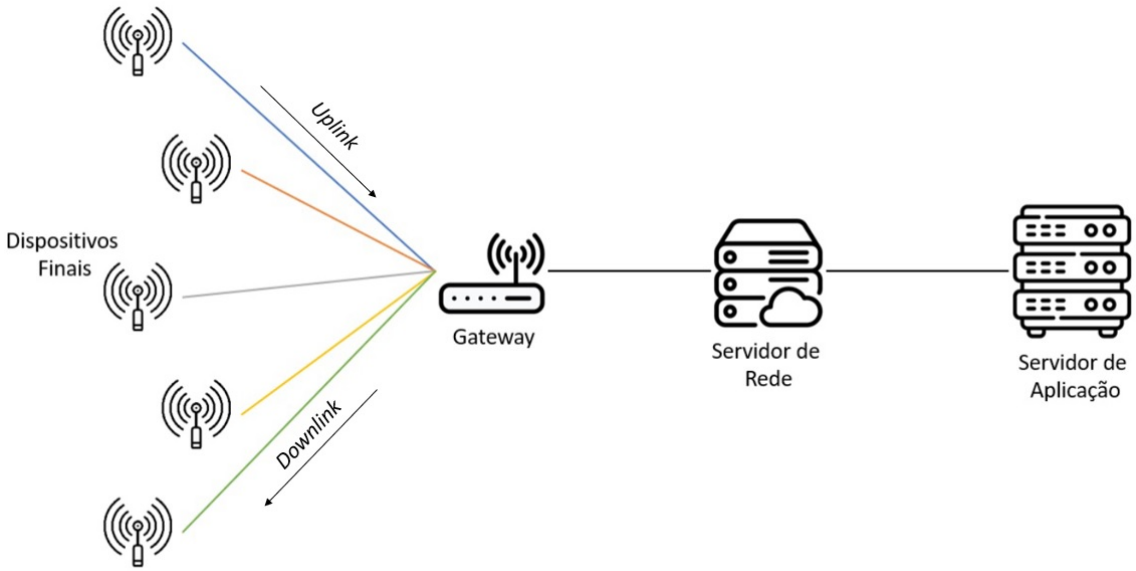


Figura 3.1: Ilustração de uma rede LoRa e seus componentes.

A topologia típica de uma rede LoRa é do tipo estrela e consiste de três componentes, os quais são os dispositivos finais, o *gateway*, o servidor de rede e o servidor de aplicação.

- Dispositivo final: é um dispositivo de baixo consumo de energia que se comunica apenas com o *gateway*, conforme ilustrado na Figura 3.1. Um dispositivo final pode se comunicar com vários *gateways*.

- *Gateway*: Dispositivo que transmite e recebe pacotes dos dispositivos finais através de múltiplos canais ao mesmo tempo. Este dispositivo costuma ser conectado à rede elétrica.

- Servidor de rede: Fornece processamento na camada de controle de acesso ao meio. Realiza o encaminhamento dos pacotes do servidor de aplicação para os dispositivos finais e vice-versa. O ajuste de parâmetros de transmissão dos dispositivos finais, tais como fator de espalhamento, potência de transmissão e canal, pode ser realizado no servidor de rede.

- Servidor de aplicação: Instala e executa aplicações, centralizando-as e dispensando a instalação destas em computadores clientes.

A modulação do LoRa é baseada no método de espalhamento espectral *Chirp Spread Spectrum* (CSS), utilizada em aplicações espaciais e militares, sendo desenvolvida originalmente para aplicações de radar. A codificação da informação é

realizada através da tecnologia de espalhamento do sinal a ser transmitido no espectro de frequência, gerando um sinal de portadora chamado de *chirp*. Nesse tipo de modulação, a taxa de transmissão varia em função da sensibilidade dos receptores. Assim como nos outros tipos de modulação por espalhamento espectral, como o *Direct Sequence Spread Spectrum* (DSSS) e o *Frequency Hoping Spread Spectrum* (FHSS), a tecnologia utiliza toda a largura de banda alocada para realizar a transmissão em difusão. Como consequência, o sinal torna-se resiliente ao efeito *Doppler*, ao desvanecimento por múltiplos caminhos e ao ruído do canal.

A variação do fator de espalhamento permite adaptar a taxa de transferência a depender da aplicação, alterando também a taxa de dados transmitidos. Outros parâmetros também influenciam na taxa de bits efetiva da modulação, na resiliência ao ruído e na decodificação do sinal, como abordado nas seções subsequentes.

A técnica de espalhamento espectral CSS codifica o sinal de uma determinada faixa de frequências espalhando-o no mesmo domínio da frequência. A largura de banda do sinal passa a ser maior do que a do original, mas com uma baixa densidade de potência. Como consequência, o sinal apresenta uma relação sinal ruído baixa o suficiente para se tornar imperceptível para alguns receptores [25]. Os bits são codificados em símbolos pelo espalhamento espectral. Cada símbolo é representado por um conjunto de *chirps*, sendo os pulsos de *chirps* modulados pelo espalhamento espectral [26], permitindo a variação da frequência mantendo a fase do sinal [27].

A frequência do sinal senoidal varia com o tempo, sendo através dessa variação que ocorre a transmissão dos símbolos. O aumento da frequência é denominada *up-chirp*, enquanto a diminuição da frequência é denominada *down-chirp*. A codificação do sinal ocorre durante o aumento da frequência, no caso do movimento *up-chirp*, partindo da frequência base f_b até a frequência máxima f_{max}. A frequência, então, salta para o valor mínimo f_{min} quando a frequência máxima é atingida e cresce até retornar ao valor base f_b, sendo esta a frequência da portadora. A frequência, portanto, aumenta linearmente e varre toda a largura de banda. A modulação de símbolos com aumento de frequência (*up-chirp*) de uma largura de banda de 125 kHz é representada pela Figura 3.2.

A largura de banda na tecnologia LoRa pode ser programada em três valores, 125 kHz, 250 kHz e 500 kHz, sendo definida pela diferença entre a frequência máxima f_{max}, e a frequência mínima f_{min}.

O parâmetro que define a quantidade de bits que são acrescidos como dados de redundância por pacote é denominado taxa de codificação. Esses dados são usados para recuperação de erros, assumindo quatro possíveis valores. A taxa de código ou *Code Rate* (CR) pode ser 1, 2, 3 e 4 para uma razão de 4/5, 4/6, 4/7 e 4/8 entre bits e bits mais redundância, conforme a Equação 3.1.

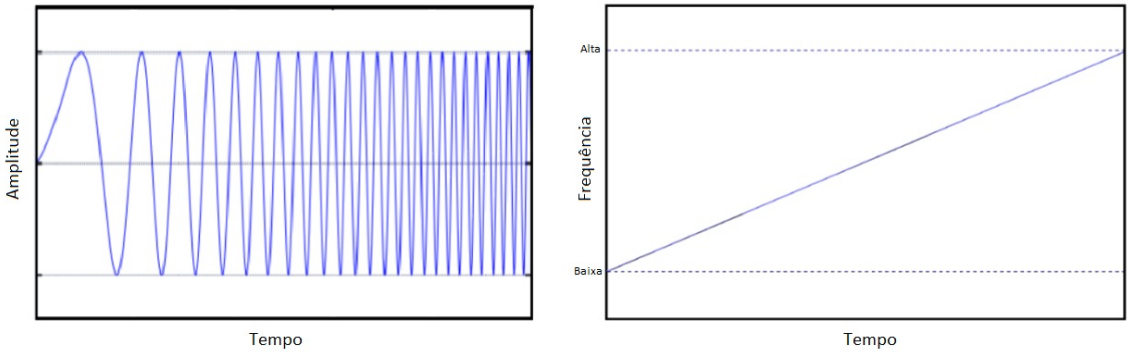


Figura 3.2: Modulação de símbolos com aumento de frequência [1].

$$T_c = \frac{4}{4 + CR}. \tag{3.1}$$

A quantidade de *chirps* enviados a cada segundo é definida pelo parâmetro de fator de espalhamento ou *Spreading Factor* (SF). O SF representa a quantidade de bits de informação codificados em um único símbolo (taxa de *chirp*), o que permite reduzir a relação sinal ruído [28]. Seis valores são definidos para esse parâmetro, de acordo com as especificações do LoRa, sendo estes SF7, SF8, SF9, SF10, SF11 e SF12. Tais valores permitem que enlaces de comunicação com fatores de espalhamento diferentes tenham interferência mínima entre si. Os diferentes fatores de espalhamento mantêm entre si enlaces cuja codificação é quase ortogonal (mas não totalmente ortogonais). A Equação 3.2 representa a quantidade de *chirps* por símbolo.

$$M = 2^{SF}. \tag{3.2}$$

A sensibilidade do limiar de recepção, tempo de duração de símbolos e o alcance do sinal são diretamente proporcionais ao valor do fator de espalhamento, enquanto a taxa de transmissão é inversamente proporcional a este parâmetro. O tempo de transmissão de um símbolo (em segundos) e a taxa de transmissão teórica (R_b) ou taxa de bits (em bits por segundo) são definidos pelas Equações 3.3 e 3.4.

$$T_s = \frac{2^{SF}}{BW}. \tag{3.3}$$

$$R_b = SF \times \frac{\frac{4}{4+CR}}{\frac{2^{SF}}{BW}} * 1000. \tag{3.4}$$

A taxa entre símbolo é definida por 2^{SF}, considerando a largura de banda fixa. O espectrograma de cada fator de espalhamento pode ser visto na Figura 3.3.

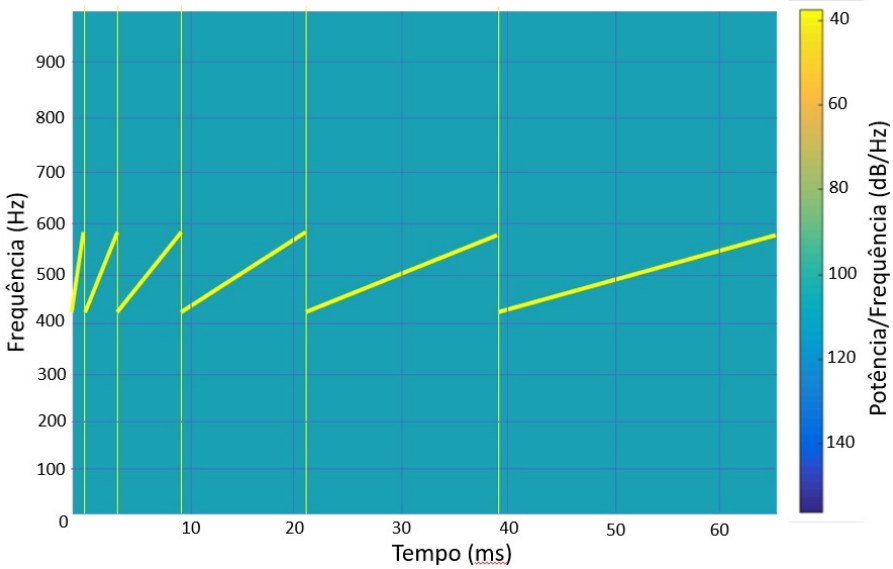


Figura 3.3: Espectrograma dos fatores de espalhamento. Da esquerda para a direita são os fatores de espalhamento 7, 8, 9, 10, 11 e 12 respectivamente.

A Figura 3.4 representa o espectrograma de um exemplo de transmissão LoRa com vários símbolos.

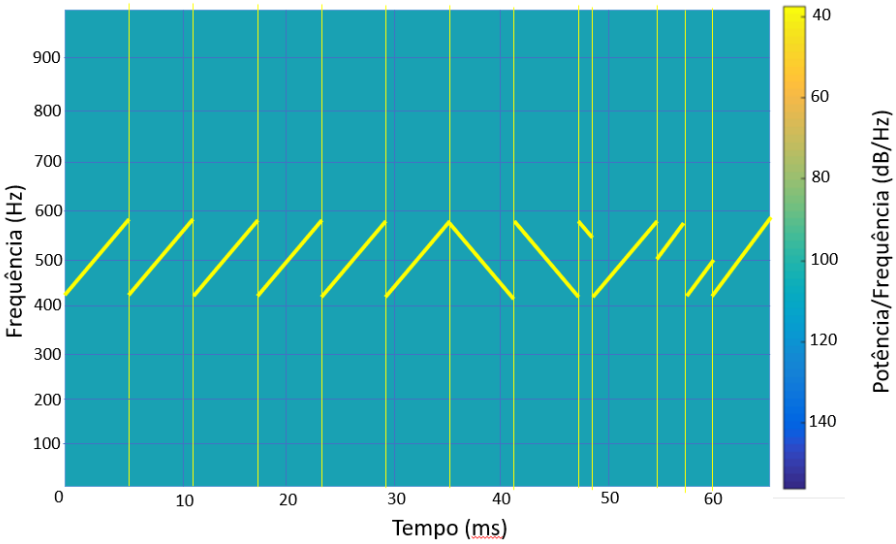


Figura 3.4: Espectrograma de um exemplo de transmissão LoRa com fator de espalhamento 9 com vários símbolos.

A taxa de bits para cada fator de espalhamento pode ser encontrada na Tabela 3.1. De acordo com a Equação 3.4, quando a largura de banda (BW) e o *Coding Rate* (CR) são constantes, a taxa de transmissão aumenta quando o fator de espalhamento aumenta. O alcance na transmissão se torna maior quando o fator de espalhamento aumenta em virtude do aumento da resiliência à erros. Porém, aumentar o fator de espalhamento acarreta no aumento na duração da transmissão e, consequentemente, no aumento do consumo de energia. Fatores de espalhamento

mais baixos reduzem o alcance das transmissões, uma vez que reduzem o ganho de processamento e aumentam a taxa de transmissão de dados. Fatores de espalhamento maiores promovem ganhos de processamento maiores, fazendo com que o sinal seja recebido com menos erros em relação a um sinal com fator de espalhamento menor e. portanto, pode percorrer uma distância maior. O tempo no ar (ToA) precisa ser maior para um sinal com fator de espalhamento alto quando comparado com um sinal com fator de espalhamento menor, considerando um montante de dados e largura de banda fixos. Logo, existe um compromisso entre consumo de energia e alcance da transmissão definido pelo fator de espalhamento.

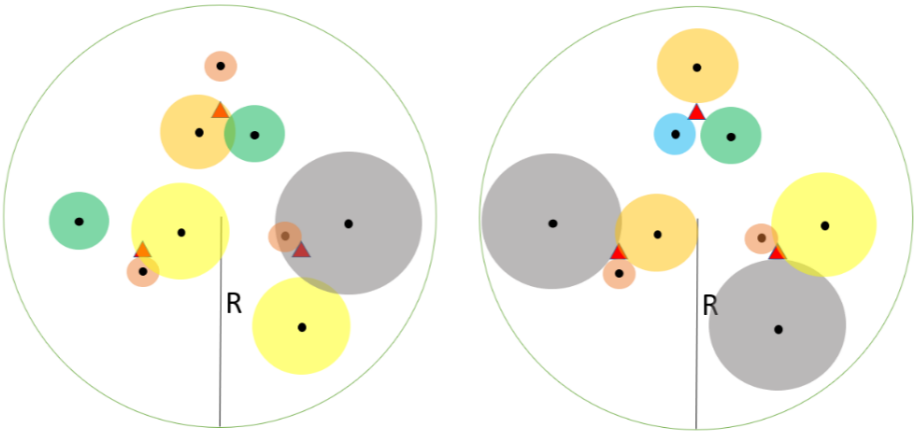
Tabela 3.1: Taxa de bits para cada fator de espalhamento.

BW (kHz)	`SF7`	`SF8`	`SF9`	`SF10`	`SF11`	`SF12`
125	5469	3125	1758	977	537	293
250	10938	6250	3516	1953	1074	586
500	21875	12500	7031	3906	2148	1172

A ortogonalidade não é perfeita, mas permite que o *gateway* receba múltiplas transmissões simultâneas com diferentes fatores de espalhamento. Contudo, transmissões simultâneas com o mesmo fator de espalhamento podem não ser recebidas pelo *gateway* devido às colisões. Dessa forma, é essencial que os fatores de espalhamento dos dispositivos sejam selecionados de forma a aumentar o desempenho da rede [29].

Inicialmente, para verificar a distância entre a sua posição e o *gateway*, o dispositivo observa a potência do sinal recebido das vinte últimas transmissões em *downlink*. Caso a potência do sinal recebido na transmissão *downlink* for muito baixa, o dispositivo aumenta o valor da potência na transmissão subsequente. Caso atinja a máxima potência de transmissão e a potência do sinal recebido permaneça baixa, o dispositivo aumenta o fator de espalhamento. De forma similar, caso a potência do sinal recebido na transmissão *downlink* for muito alta, o dispositivo reduz o valor do fator de espalhamento. Caso atinja o valor mínimo de fator de espalhamento e a potência do sinal recebido permaneça alta, o dispositivo reduz, então, a potência de transmissão. Esse esquema de seleção dos menores valores de potência de transmissão e fator de espalhamento é comumente utilizado em implementações LoRaWAN e é chamado de *Adaptive Data Rate* (ADR).

A Figura 3.5 representa os possíveis efeitos do alcance das transmissões na seleção dos parâmetros de transmissão, ora com alcance abaixo do necessário (Figura 3.5(a)), ora com o alcance além do necessário (Figura 3.5(b)).



(a) Sem ajuste de parâmetros. (b) Com ajuste de parâmetros.

Figura 3.5: Efeito do ajuste dos parâmetros no alcance da comunicação dos dispositivos finais (círculos pretos) com seus respectivos *gateways* (triângulos vermelhos).

Os dispositivos finais próximos ao *gateway* selecionam o menor valor possível de fator de espalhamento e potência de transmissão na maior parte do tempo de operação. A probabilidade de colisões entre transmissões com o mesmo fator de espalhamento aumenta proporcionalmente à densidade de dispositivos finais implementados na rede.

3.2 Protocolo LoRaWAN

O protocolo utilizado na camada de controle de acesso ao meio na tecnologia LoRa é chamado de LoRaWAN. Desenvolvido e mantido pela *LoRa Alliance* [30], organização sem fins lucrativos dedicada à padronização do LoRaWAN. Este protocolo é um requisito necessário para garantir a interoperabilidade entre diferentes redes LoRa. O LoRaWAN é baseado no método de acesso ao meio ALOHA puro, no qual os dispositivos finais não verificam se o canal está livre ou não antes de realizar a transmissão, assumindo o risco de colisões.

Três classes de dispositivos finais são definidos pelo padrão LoRaWAN, sendo essas as classes A, B e C, fornecendo diferentes soluções para diferentes aplicações. A comunicação de classe A é sempre iniciada pelo dispositivo final. Um dispositivo pode enviar uma mensagem de *uplink* a qualquer momento. Assim que a transmissão de *uplink* estiver concluída, o dispositivo abre duas janelas curtas de recepção (*downlink*). Há um atraso entre o final da transmissão do *uplink* e o início das janelas de recepção. Se o servidor de rede não responder durante essas duas janelas de recepção, o próximo *downlink* será após a próxima transmissão do *uplink*. A Figura 3.6 apresenta o diagrama temporal das janelas de recepção e de transmissão. Os dispositivos classe A consomem pouca energia, já que possuem apenas duas

janelas de recepção, sendo a classe de dispositivos utilizada no modo de operação padrão do LoRaWAN. O servidor pode responder durante a primeira janela de recepção (RX1) ou durante a segunda janela de recepção (RX2), mas não usa ambas as janelas.

Os dispositivos classe B, além das janelas de recepção iniciadas pela classe A, também abrem janelas de recepção programadas para receber mensagens de *downlink* do servidor de rede. Usando *beacons* sincronizados no tempo transmitidos pelo *gateway*, os dispositivos abrem *slots* de ping de *downlink* de forma programada para receber mensagens de *downlink* do servidor de rede. Os dispositivos classe B também abrem janelas de recepção após enviar um *uplink*, como pode ser visto na Figura 3.6. Os dispositivos finais classe B têm latência menor do que os dispositivos finais classe A, porque são alcançáveis em horários pré-configurados e não precisam enviar um uplink para receber um downlink. A duração da bateria é menor na classe B do que na classe A, pois o dispositivo passa mais tempo no modo ativo, durante *beacons* e *slots* de ping.

Os dispositivos de classe C estendem a classe A mantendo as janelas de recepção abertas, a menos que estejam transmitindo, conforme mostrado na Figura 3.6. Isso permite comunicação de baixa latência, mas consome muitas vezes mais energia do que os dispositivos classe A. Neste trabalho, os dispositivos finais são considerados como classe A, já que além de menos custoso energeticamente, é o modo de operação padrão do LoRaWAN.

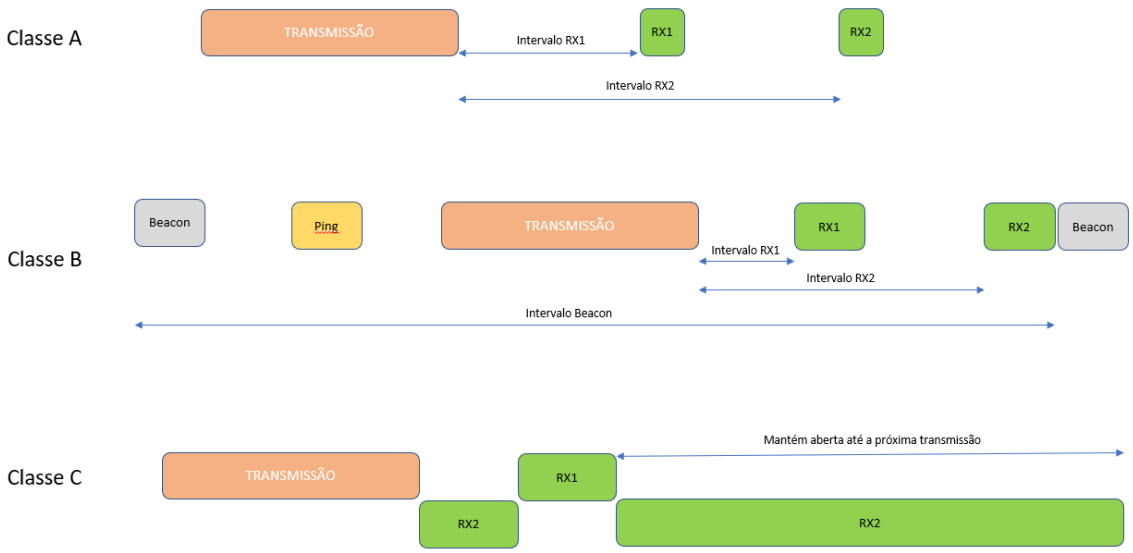


Figura 3.6: Diagrama temporal das classes de dispositivo.

O formato da unidade de dados e o protocolo de comunicação são descritos na especificação do LoRaWAN [2]. Os formatos das unidades de dados das camadas física e MAC do LoRaWAN podem ser vistos na Figura 3.7. A unidade de dados

da camada física é composta pelo preâmbulo, um cabeçalho, um cabeçalho verificação cíclica de redundância ou *cyclic redundance check* (CRC), presente apenas nas transmissões do dispositivo final para o *gateway* (*uplink*), e um *payload*. Já o pacote da camada MAC consiste de um cabeçalho, um campo referente à porta e um *payload*, sendo o cabeçalho composto pelo endereço do dispositivo, um campo de controle, um campo de contagem e um campo de opções. O campo de controle do cabeçalho do pacote MAC é utilizado para o ADR, através do qual o servidor de rede é capaz de acionar o dispositivo final para ajustar o fator de espalhamento e potência de transmissão. Caso o bit referente ao ADR esteja habilitado, então o fator de espalhamento e potência de transmissão das futuras transmissões passam a ser controlados pelo servidor de rede, fornecendo os novos valores dos parâmetros no campo de opções. A implementação do mecanismo ADR deve ser realizada pelo operador de rede, já que este não é um mecanismo padrão do LoRaWAN. Caso o mecanismo ADR seja habilitado, então o fator de espalhamento selecionado deve lidar com o compromisso entre a confiabilidade do enlace e tempo de transmissão para evitar colisões, além de economizar energia.

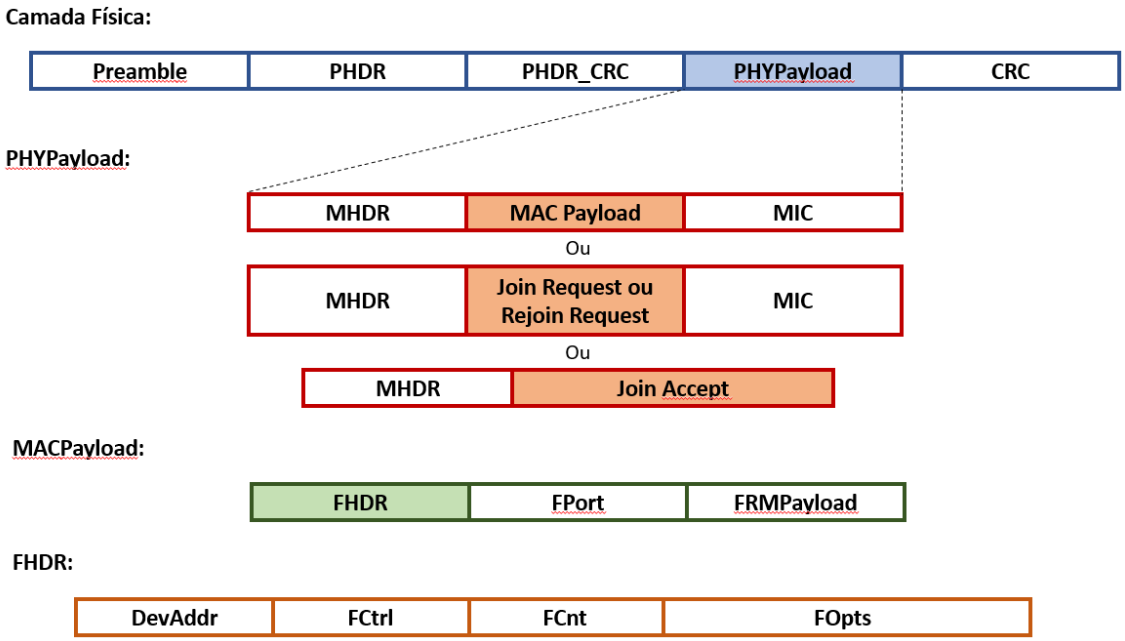


Figura 3.7: O formato da unidade de dados da camada física e MAC [2].

O desempenho do ADR na seleção do fator de espalhamento e da potência de transmissão tem sua eficácia comprometida quando a quantidade de dispositivos finais sem fio escala. Isso ocorre pois, antes do mecanismo de ajuste dos parâmetros ser invocado, muitos pacotes precisam ser perdidos ou retransmitidos, uma vez que o ADR é um mecanismo reativo [6]. Além disso, o tempo de reação desse mecanismo se torna um problema para aplicações nas quais a tomada de decisão de ajuste de parâmetros precisa ser realizada em um intervalo de tempo muito curto. Por isso,

técnicas inteligentes e proativas de ajuste de parâmetros da camada física devem ser utilizadas para dirimir a queda de desempenho do mecanismo utilizado pelo LoRaWAN.

Diante do exposto acima, abordagens de aprendizado de máquinas são utilizadas nesta dissertação para ajustar o fator de espalhamento e a potência de transmissão. O objetivo é fornecer um ajuste mais adequado destes parâmetros de forma a reduzir a quantidade de pacotes perdidos e a quantidade de retransmissões.

No capítulo seguinte, uma revisão sobre o aprendizado por reforço será realizada, introduzindo conceitos, fundamentos e algoritmos do referido aprendizado. Embora muitos algoritmos de aprendizado por reforço pudessem ser apresentados, somente o *Q-learning* é descrito, uma vez que é o algoritmo de aprendizado por reforço utilizado neste trabalho, cujo desempenho é comparado com as técnicas de Árvore de Decisão e Máquina de Vetor de Suporte de aprendizado supervisionado.

Capítulo 4

Aprendizado Supervisionado e o Aprendizado por Reforço

Conforme comentado anteriormente, as abordagens de utilização de técnicas de aprendizado de máquinas para a seleção de parâmetros de transmissão em redes LoRa se mostram uma alternativa para dirimir as limitações do mecanismo ADR em redes densas. Uma vez que esta dissertação propõe a utilização de técnicas de aprendizado supervisionado e de reforço para a seleção dos parâmetros de transmissão, este capítulo discorre sobre as técnicas utilizadas para ajudar no entendimento das propostas e dos experimentos realizados.

As técnicas de aprendizado de máquina têm sido utilizadas para a resolução de problemas como reconhecimento facial, tarefas de detecção e classificação, serviços de recomendação, reconhecimento de padrões etc. Os algoritmos implementados, em geral, são utilizados para fazer melhores decisões e ajustar suas ações sem a necessidade de um operador humano. Existem basicamente três tipos de aprendizado: aprendizado supervisionado, aprendizado não-supervisionado e o aprendizado por reforço, assim como quatro categorias de tarefas que podem ser solucionadas, sendo elas a de classificação, regressão, agrupamento e extração de regras [16]. As tarefas de classificação e regressão são realizadas através do aprendizado supervisionado, no qual um conjunto de dados rotulados é utilizado para realizar o treinamento do modelo de aprendizado. Já o aprendizado não-supervisionado utiliza conjuntos de dados não rotulados para criar um modelo capaz de classificar as amostras do conjunto de dados em diferentes grupos, sendo utilizado em problemas de agrupamento [31]. O aprendizado por reforço utiliza algoritmos que aprendem com o ambiente (exploração) e tira proveiro do aprendizado (explotação). Nesse tipo de aprendizado, os agentes tomam ações que são penalizadas ou recompensadas, utilizando o resultado das ações para saber quais as melhores ações que permitam otimizar a recompensa. Considerando que esse tipo de aprendizado pode ignorar recompensas imediatas para explorar o ambiente e descobrir recompensas ainda maiores, tem se mostrado

o melhor tipo de aprendizado para tarefas de agendamento, planejamento e tomada de decisão [16].

4.1 Aprendizado Supervisionado

O aprendizado supervisionado é um ramo de aprendizado de máquinas que utiliza um conjunto de dados rotulados para treinar o modelo de aprendizado com o intuito de realizar tarefas de classificação e regressão. Esta dissertação propõe utilizar duas técnicas de aprendizado supervisionado, sendo elas a de Árvore de Decisão e Máquina de Vetor de Suporte, para prever o resultado de uma transmissão e, a partir do resultado da previsão, selecionar os parâmetros de transmissão dos dispositivos finais.

4.1.1 Árvore de Decisão

Esta técnica é capaz de realizar tarefas de classificação e regressão e é indicada para prever categorias discretas (tarefa de classificação) ou para prever valores numéricos (gasto no cartão de crédito em reais, por exemplo). Nesta dissertação, esta técnica prevê o resultado da transmissão, resultado este que pode ser representado pelos valores 1, 2 ou 3 indicando "transmitido com sucesso", "interferido"e "abaixo da sensibilidade do *gateway*"respectivamente.

A árvore pode ser definida como uma coleção de elementos chamados nós, dentre os quais existe um "raíz"que exerce uma posição hierárquica superior aos demais nós. Uma árvore de decisão é composta pelos seguintes componentes:

- Um nó folha (ou nó resposta) que contém uma nome de uma classe ou um símbolo nulo indicando que não é possível atribuir nenhuma classe ao nó por não haver nenhum exemplo que corresponda a esse nó.

- Um nó interno (ou nó de decisão) que contém o nome de um atributo. Para cada possível valor do atributo, corresponde um ramo para outra árvore de decisão.

- Ramos são condições que permitem passar de um nó interno para um nó folha. São representados por valores (atributos categóricos) ou intervalos, por exemplo "menor que 1 km"(>1 km) ou "fator de espalhamento maior que 7"(SF>7), sendo este um atributo numérico.

A Figura 4.1 representa um exemplo de Árvore de Decisão composta pelos elementos supracitados.

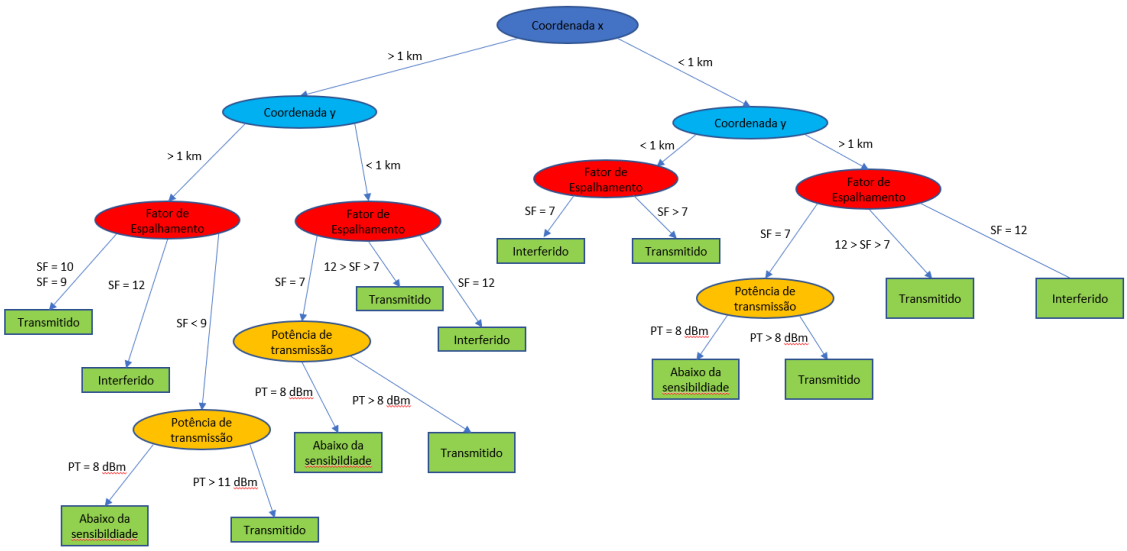


Figura 4.1: Exemplo de Árvore de Decisão a partir de uma base de dados composta pelas coordenadas x e y da posição dos dispositivos, o fator de espalhamento e a potência de transmissão selecionados.

Na técnica de Árvore de Decisão, os valores dos atributos de um determinado registro de base de dados determinam o caminho percorrido na árvore e, consequentemente, na classe (nó folha) obtida. A seleção de um atributo como nó raiz ou como nós internos depende do valor de cada atributo, sendo este valor determinado através do cálculo da entropia e do ganho de cada atributo, representados pelas equações 4.1 e 4.3 respectivamente. A entropia é o somatório da frequência de cada valor de um atributo (pi) multiplicada pelo logarítmo da frequência de cada valor de um atributo ($\log_2 pi$), conforme a Equação 4.1. Uma vez calculada a entropia, calcula-se o ganho relativo ao atributo. O ganho de informação é igual à diferença entre a entropia e o somatório da frequência de cada atributo multiplicada pela entropia, conforme a Equação 4.3. O valor obtido do ganho determina a posição do atributo na árvore e na formação da árvore de decisão. O atributo com maior ganho é o nó raiz da Árvore de Decisão [32].

$$Entropia(S) = \sum_{i}^{c} -p_i \log_2 pi\cdot \tag{4.1}$$

$$E(A) = \sum_{v \in V} \frac{|s_v|}{|s|} * Entropia\cdot \tag{4.2}$$

$$Ganho = Entropia(S) - E(A)\cdot \tag{4.3}$$

4.1.2 Máquina de Vetor de Suporte

A técnica de Máquina de Vetor de Suporte busca obter dois hiperplanos de separação dos registros de uma base de dados de forma que a margem de separação dos dados seja máxima. De forma simplificada, uma margem de separação máxima entre os hiperplanos é obtida a partir de vetores perpendiculares a estes hiperplanos, chamados de vetores de suporte.

As Figuras 4.2 e 4.3 representam o efeito dos vetores de suporte na determinação da margem máxima de separação entre os registros "transmitido", "interferido"e "abaixo da sensibilidade".

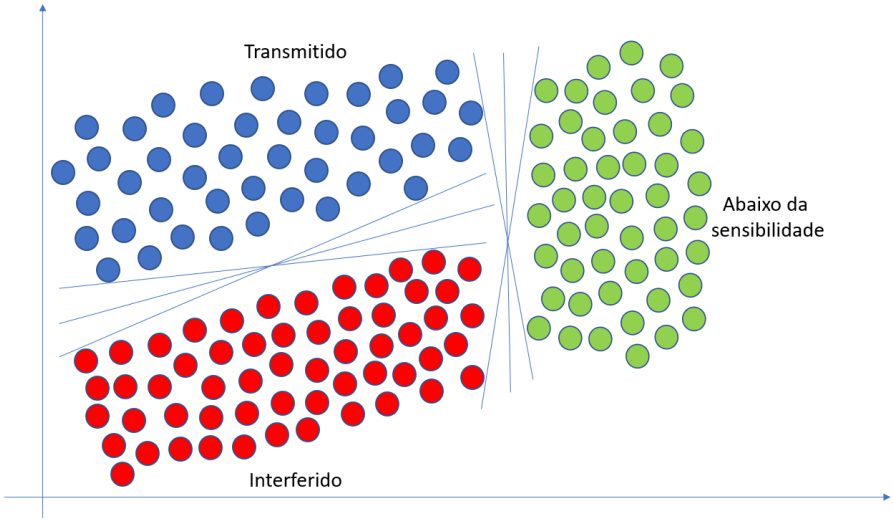


Figura 4.2: Hiperplanos de separação sem os vetores de suporte.

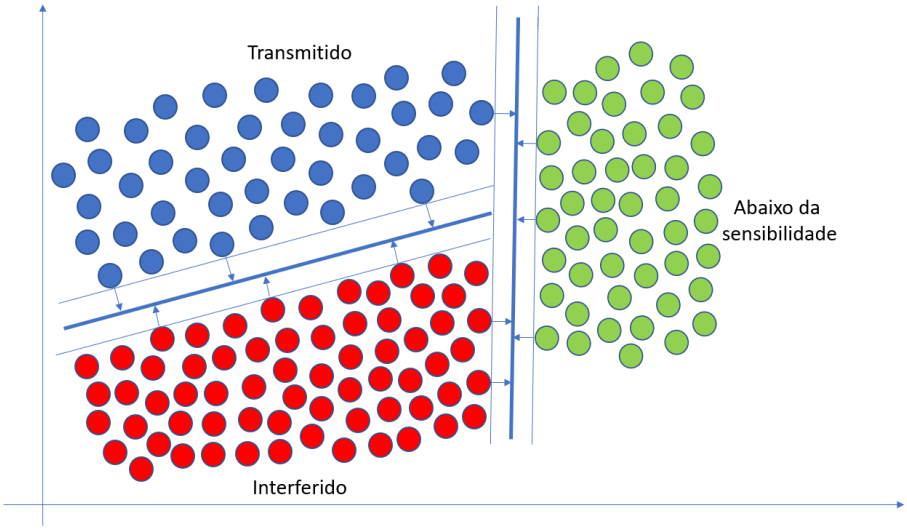


Figura 4.3: Hiperplanos de separação com os vetores de suporte.

O *Maximum-margin hyperplane* é um hiperplano central, equidistante e paralelo a esses hiperplanos que é responsável pela separação das classes. O SVM aplica uma

função kernel quando não é possível separar os dados linearmente para transformar os dados de treinamento em um espaço de maior dimensão. Então, a divisão linear é feita entre os pontos das duas classes, de forma a encontrar o *Maximum-margin hyperplane* no novo espaço. A determinação da margem máxima de separação pelos vetores de suporte permite a utilização desta técnica em tarefas complexas como classificação de dados lineares, não-lineares e reconhecimento de imagens, voz, caracteres. O SVM evita o problema de *overfitting*, ou seja, quando o modelo representa bem apenas os dados de treinamento, mas não consegue representar os demais dados. O tempo de treinamento do SVM é muito alto em relação às demais técnicas, sendo esta a sua desvantagem.

4.2 Aprendizado por Reforço

Um sistema de aprendizado por reforço é composto pelo ambiente, pela função recompensa, pelos agentes de aprendizado e pela política seguida pelos agentes. Estes tomam ações que os levam a diferentes estados em busca de alcançar um objetivo.

A cada ação tomada, os agentes recebem do ambiente um valor de recompensa de acordo com o benefício obtido. Tais recompensas fazem com que o agente aprenda como alcançar o objetivo, ou seja, aprenda a política. Em outras palavras, a política permite que os agentes mapeiem os estados e as ações a serem tomadas em cada estado para maximizar a recompensa obtida. O processo seguido pelo agente para encontrar a melhor política é um processo estocástico, no qual probabilidades estão associadas à seleção de cada possível ação [33].

O aprendizado ocorre através da iteração dos agentes com o ambiente. O agente de aprendizado descobre as melhores experiências a cada estado explorado no ambiente, sem ter inicialmente nenhum conhecimento sobre as melhores ações a serem tomadas.

O par estado-ação que determina a importância de cada estado para cada agente é chamado de função valor. O valor de cada estado corresponde à recompensa total que um agente espera receber no futuro a partir deste estado. O agente busca tomar as ações que o leva para estados com maior valor, em vez da maior recompensa, pois tais ações proporcionam o maior montante de recompensas ao longo do tempo.

A modelagem de aplicações que exigem a tomada de ações com o intuito de buscar a recompensa máxima (aprendizado por reforço) pode ser realizada através do Processo de Decisão de Markov (MDP). Esse processo preconiza que um problema de aprendizado por reforço é composto pelo agente, pelo ambiente e pelas informações trocadas entre estes, sendo essas informações as ações escolhidas pelo agentes, os estados nos quais as ações são tomadas e as recompensas. A Figura 4.4 representa

o típico ciclo de aprendizado por reforço contendo a relação entre o ambiente e o agente, ambos trocando as referidas informações entre si.

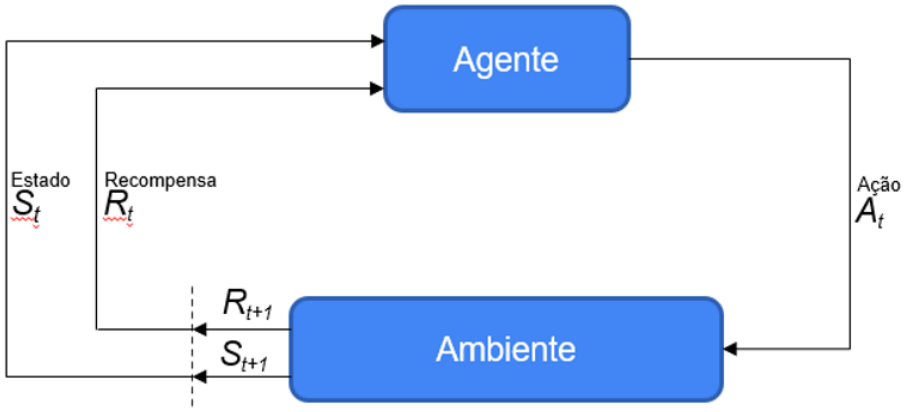


Figura 4.4: Ciclo do Aprendizado por Reforço.

Conforme mencionado anteriormente, o objetivo do aprendizado é maximizar o montante de recompensas (retorno), sendo este montante igual à soma das recompensas obtidas em cada instante de tempo, representado na Equação 4.4. Nessa equação, t representa o instante, N a quantidade de passos da tarefa em um episódio, R_t representa a recompensa obtida no instante de tempo t e G_t o retorno obtido em um episódio.

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + \ldots + R_{t+N} = \sum_{t=1}^{N} R_{t+N}. \tag{4.4}$$

A Equação 4.4 representa o retorno obtido em tarefas compostas por episódio, ou seja, tarefas nas quais existe um estado que entrega a maior entre todas as recompensas, representando o estado final.

Em tarefas contínuas, ou seja, tarefas nas quais não existe um estado final e, portanto, nunca acabam, o retorno que o agente busca maximizar pode ser infinito, já que a quantidade de passos é infinita. Nesse caso, o retorno esperado é definido em termos de um parâmetro chamado fator de desconto γ, o qual assume um valor entre 0 e 1 (Equação 4.5).

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + ... = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}. \tag{4.5}$$

De acordo com a Equação 4.5, quando o valor do fator de desconto está próximo de 0 (zero), o retorno esperado leva em conta recompensas imediatas, valorizando menos recompensas futuras, Já quando o fator de desconto é próximo de 1, o retorno esperado valoriza mais recompensas futuras. Este retorno esperado também pode ser expresso através da Equação 4.6. O agente deve escolher as ações que levam à

melhor solução global possível, não apenas a melhor solução imediata. O fator de desconto, portanto, determina essencialmente o quanto os agentes de aprendizado por reforço se preocupam com as recompensas no futuro distante em relação àquelas no futuro imediato. No limite, caso o fator de desconto seja 0 (zero), o agente será completamente míope e aprenderá apenas sobre ações que produzem uma recompensa imediata. Se o fator de desconto for 1, o agente avaliará cada uma de suas ações com base na soma total de todas as suas recompensas futuras.

$$G_t = R_{t+1} + \gamma G_{t+1}. \tag{4.6}$$

Em um MDP, a política é definida como um mapeamento dos estados em probabilidades de selecionar cada possível ação, sendo a política seguida por um agente definida como $\pi(a, s)$, na qual s representa o estado atual e a representa a ação. A função-valor em um MDP é definido como o valor esperado do retorno, dado que o estado atual é s, conforme representado na Equação 4.7 abaixo:

$$V_\pi(s) = \mathbb{E}[G_t|S_t = s] = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}|S_t = s\right]. \tag{4.7}$$

Na Equação 4.7, $\mathbb{E}[.]$ é o valor esperado da variável aleatória, dado que o agente segue a política π, e t é qualquer instante de tempo. A função-valor V_π é também chamada de função valor-estado para a política π. O retorno esperado, dado que o estado atual é s e que foi tomada a ação a, seguindo a política π, é definida pela Equação 4.8. Esse retorno esperado também é chamado de função valor-ação para a política π, como visto a seguir:

$$Q_\pi(s, a) = \mathbb{E}[G_t|S_t = s, A_t = a] = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}|S_t = s, A_t = a\right]. \tag{4.8}$$

A função valor-estado, que representa a importância de cada estado na obtenção das recompensas para o agente, pode ser representada como uma função da probabilidade $p(s', r|s, a)$, como visto na Equação 4.9. A probabilidade $p(s', r|s, a)$ reflete a probabilidade do agente se mover para o estado s' e receber a recompensa r dado que esteve no estado s e tomou a ação a.

$$V_\pi(s) = \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a) \left[r + \gamma V_\pi(s')\right]. \tag{4.9}$$

A Equação 4.9 é recursiva e expressa a relação entre a importância (valor) de um estado e a importância de estados subsequentes, sendo esta chamada de equação de *Bellman* para v_π. Portanto, para cada política π seguida pelo agente, existe

um valor ou importância associada a cada estado. Em um MDP finito, ou seja, cuja quantidade de estados, ações e recompensas são finitas, existe pelo menos uma política ótima denotada por π^* e uma função valor-estado associada a essa política ótima, expressa na Equação 4.10.

$$V_*(s) = \max(V_\pi(s))\cdot \tag{4.10}$$

De forma análoga, as funções valor-ação associadas às políticas são definidas como na Equação 4.11:

$$Q_*(s,a) = \max(q_\pi(s,a))\cdot \tag{4.11}$$

A Equação de *Bellman* para a v_* (Equação 4.9), conhecida como equação de otimalidade de *Bellman*, pode ser obtida a partir da Equação 4.6.

$$V_*(s) = max\left(\sum_{s',r} p(s',r|s,a)\left[r + \gamma V_\pi(s')\right]\right)\cdot \tag{4.12}$$

A partir da Equação 4.9, obtém-se o valor de V_* para cada estado, sendo V_* o valor de V_π, quando tomada uma ação que maximiza seu valor.

Para resolver a equação de otimalidade de *Bellman*, em tarefas de aprendizado por reforço, são utilizados métodos de aproximação, como Programação Dinâmica e Diferença Temporal a partir das transições de estado já experimentadas. Este trabalho utiliza o algoritmo *Q-learning*, sendo esse um método de diferença temporal, como algoritmo de aprendizado por reforço para ajustar os parâmetros de transmissão de redes LoRa.

Para atualizar o valor do estado, os métodos de Diferença Temporal atuam de acordo com a Equação 4.10, na qual $\alpha \in (0,1]$ é um valor pequeno que controla a taxa de aprendizado.

$$V(s_t) \leftarrow V(s_t) + \alpha\left[R_{t+1} + \gamma V(s_{t+1}) - V(s_t)\right]\cdot \tag{4.13}$$

Nota-se através da Equação 4.13 que o novo valor estimado do estado S_{t+1} é obtido pela soma do valor estimado do estado S_t somado à diferença entre o valor estimado do estado S_t ($V(S_t)$) e a melhor estimativa ($R_{t+1} + \gamma V(S_t)$).

O algoritmo *Q-learning* é um dos métodos de Diferença Temporal [34], no qual os valores de ação ($Q(s,a)$) são atualizados de acordo com a Equação 4.11.

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[R_{t+1} + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)\right]\cdot \tag{4.14}$$

Existem dois tipos de algoritmos, sendo um deles é o algoritmo *On-policy*, que estima $Q_\pi(s,a)$ para a política π e para todos os estados s e ações a, enquanto

o algoritmo *Off-policy* realiza uma aproximação para a função valor-ação ótima $Q_*(s, a)$ independentemente da política a ser seguida, como é o caso do algoritmo *Q-learning*.

O algoritmo *Q-learning* pode ser utilizado em tempo real, exigindo baixa complexidade computacional, sendo expresso por equações simples e de fácil implementação. Esse algoritmo é utilizado nesta dissertação para selecionar os parâmetros de transmissão dos dispositivos finais em redes *LoRaWAN*. O desempenho desse algoritmo é comparado com as técnicas de aprendizado supervisionado Árvore de Decisão e Máquina de Vetor de Suporte, além do mecanismo *Adaptive Data Rate* (ADR) do *LoRaWAN*.

Capítulo 5

Propostas de Seleção de Parâmetros

Neste capítulo são apresentadas as propostas de seleção de parâmetros. Inicialmente será abordada a proposta de seleção dos parâmetros através da utilização das técnicas de Árvore de Decisão e Máquina de Vetor de Suporte. Em seguida, será abordada a proposta de utilização do algoritmo baseado em *Q-learning* para ajustar os parâmetros.

5.1 Proposta de Utilização do Aprendizado Supervisionado

Em termos práticos, a proposta de utilização do Aprendizado Supervisionado pode ser explicada da seguinte forma. O servidor de rede inicialmente configura os parâmetros dos dispositivos finais com valores aleatórios, dentro obviamente dos intervalos de valores possíveis, e inicializa uma rotina de recepção de dados transmitidos pelos dispositivos, durante um tempo definido pelo administrador de rede, para compor uma base se dados. A configuração com valores aleatórios é interessante, pois permite avaliar o desempenho da comunicação entre cada um dos dispositivos finais e o *gateway* sob diferentes condições. Durante a rotina de obtenção do conjunto de dados, deve ser prevista a implementação de uma aplicação do tipo cliente-servidor em que mensagens de controle possam ser enviadas pelo servidor aos dispositivos finais. Essas mensagens devem ser usadas para o ajuste de parâmetros do LoRa e monitoramento do sucesso das transmissões das mensagens. O conjunto de dados obtido representa o comportamento da transmissão dos dispositivos configurados com todas as combinações de parâmetros possíveis. Após o término do referido tempo, o conjunto de dados obtido é utilizado para realizar o treinamento e teste dos classificadores de forma *offline*. Uma vez realizados o treinamento e teste, o servidor de rede configura os dispositivos com os menores valores de cada combinação de parâmetros cujo resultado da transmissão foi bem sucedido.

A proposta de utilização das técnicas como classificadores no servidor de rede se mostra adequada para a obtenção de taxa de entrega de pacotes maior ou igual a obtida pelo mecanismo ADR. Na prática, o servidor de rede pode criar uma base de dados para classificação quando ocorrer modificações na rede, como inclusão de novos dispositivos ou degradação de enlaces devido a interferências.

O conjunto de dados obtido é composto por até quatro atributos. Nos casos em que a potência de transmissão não é configurada, o servidor requer amostras que contenham como atributos as coordenadas x e y da posição do dispositivo final na topologia da rede e o fator de espalhamento. Nos casos em que a potência de transmissão também é configurada, é necessária a adição deste parâmetro às coordenadas topológicas e ao fator de espalhamento. São três classes, sendo elas "transmitido", "interferido" e "abaixo da sensibilidade".

Esta proposta é utilizada em ambos os estudos de caso 1 e 2, de forma a obter um conjunto de dados com o comportamento da rede para cada combinação de parâmetros e selecionar o parâmetro cujo resultado da previsão foi "transmitido".

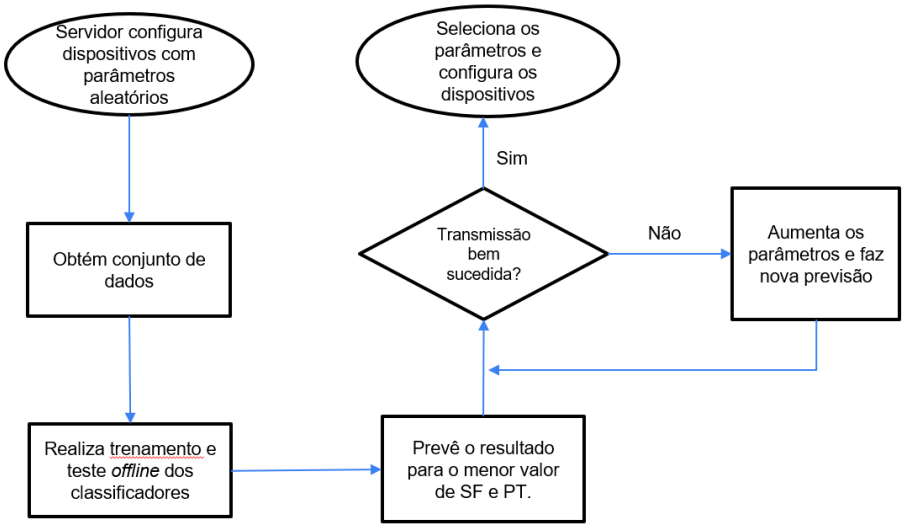


Figura 5.1: Processo de aprendizado dos agentes.

A Figura 5.1 representa o processo de seleção de parâmetros utilizando as técnicas de aprendizado supervisionado.

5.2 Proposta de utilização do aprendizado por reforço

Para esse algoritmo, os parâmetros utilizados e a função objetivo considerada para realizar a otimização da taxa de entrega de pacotes e da energia consumida são apresentados a seguir.

- Ambiente: A redes descritas nos estudos de caso 1 e 2 apresentados no capítulo de experimentos.

- Agentes: São representados pelas estruturas de dados criadas para cada dispositivo a nível de servidor de rede. Cada dispositivo final é um agente independente representado pela estrutura de dados própria gerada e treinada no servidor de rede. Enquanto no aprendizado supervisionado o treinamento é realizado remotamente no servidor a partir de uma base de dados, no por reforço, uma estrutura de dados é criada para cada dispositivo para realizar o treinamento no servidor. A Figura 5.2 ilustra a criação da estrutura de dados para treinamento de cada dispositivo (agente).

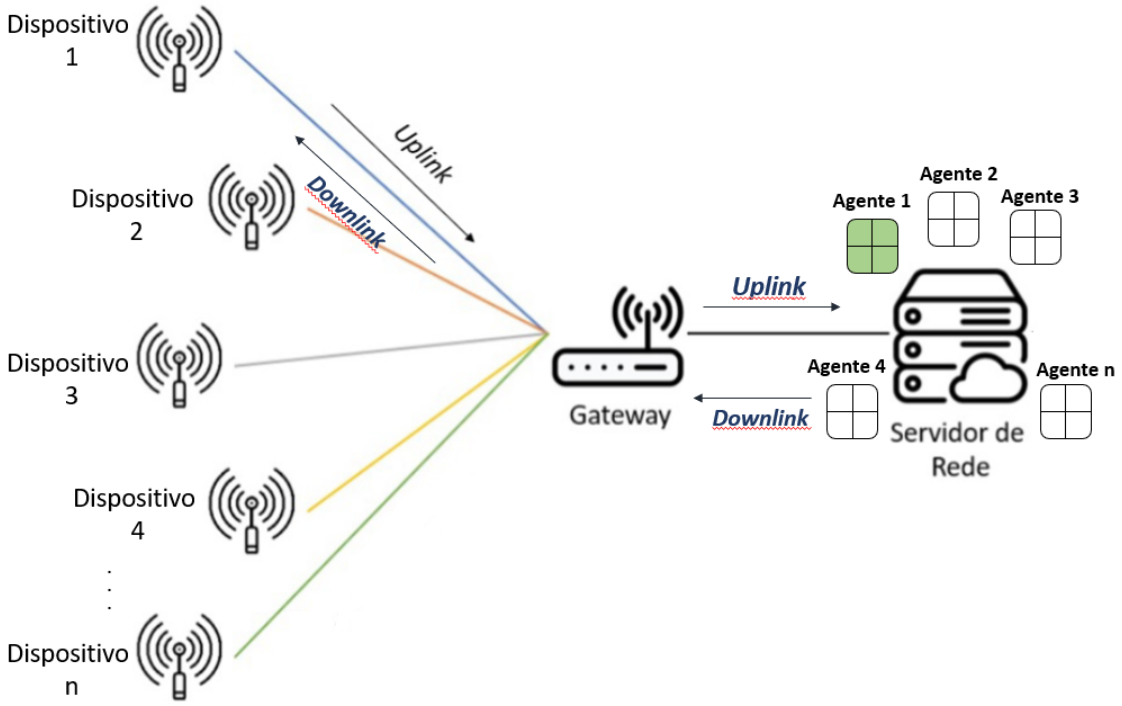


Figura 5.2: Criação e treinamento dos agentes a nível de servidor. O treinamento do agente 1 é evidenciado com a captação do status da última transmissão e configuração do dispositivo 1 com os parâmetros selecionados através do sinal *downlink*.

- Estados: O estado observável de um agente é definido pelo status da última transmissão, podendo ser uma "transmissão bem sucedida", "transmissão interferida" ou "transmissão abaixo da sensibilidade". Nota-se que, neste caso específico, o espaço de estados observáveis ("transmissão bem sucedida", "transmissão interferida" ou "transmissão abaixo da sensibilidade") não coincide com o espaço de estados dos agentes (combinações de fatores de espalhamento e potência de transmissão representadas por uma tupla), sendo o espaço de estados maior que o espaço de observações. Este é um caso que pode ser modelado como um *Partially Observable Markov Decision Process* (POMDP).

- Ações: No caso da seleção apenas do fator de espalhamento, cada agente pode executar 6 ações possíveis, sendo esta quantidade igual ao número de fatores de espalhamento (7 a 12). Já no caso da seleção do fator de espalhamento e potência de transmissão, o dispositivo pode executar 18 ações correspondentes às combinações entre fator de espalhamento (7 a 12) e potência de transmissão (8, 11 e 14 dBm). Isto permite que cada agente tenha a oportunidade de explorar todas as ações em um único episódio.

- Recompensa: a recompensa dada ao agente depende do status da última transmissão (estado observável). No estudo de caso 1, caso a última transmissão seja uma "transmissão bem sucedida", a recompensa r é dada pela razão entre a quantidade de pacotes transmitidos com sucesso e a energia total consumida na transmissão, conforme a Equação 5.1, na qual α é a taxa de aprendizado, P_i é a quantidade de pacotes recebidos com sucesso e E_i a energia total consumida. Essa razão é multiplicada por um número muito grande de forma a dar maior importância ao aumento da quantidade de transmissões bem sucedidas. Caso a última transmissão não seja uma transmissão bem sucedida, a recompensa r é de uma unidade negativa. No estudo de caso 2, caso a última transmissão seja uma "transmissão bem sucedida", a recompensa r tem valor 100. Caso a última transmissão não seja uma transmissão bem sucedida, a recompensa r é de uma unidade negativa.

- Política: o servidor de rede escolhe, para cada dispositivo, a ação que tem o maior valor de Q com probabilidade 1-ϵ, porém com uma pequena probabilidade ϵ de escolher uma ação diversa aleatoriamente de forma a explorar mais o ambiente [35].

- Atualização: cada dispositivo atualiza sua tabela-Q conforme a Equação 4.14. O fator de desconto tem valor de 0.95, sendo o valor típico utilizado no algoritmo *Q-learning*. O fator de desconto faz com que as recompensas futuras valham menos que as recompensas imediatas, aplicando um tipo de hedonismo de curto prazo no agente, mas sem deixar de valorizar as recompensas futuras. Um fator de desconto de 1 faria as recompensas futuras valerem tanto quanto as recompensas imediatas, o que traria um efeito indesejável de imediatismo.

- Inicialização: No caso de seleção apenas do fator de espalhamento, todos os dispositivos são inicializados com fator de espalhamento 7. No caso da seleção de fator de espalhamento e potência de transmissão, todos os dispositivos são inicializados com fator de espalhamento 7 e potência de transmissão de 8 dBm. O fator de espalhamento 7 e potência de transmissão 8 dBm foram escolhidos para inicializar os dispositivos pois sabe-se que os dispositivos mais distantes

do *gateway* não proporcionarão uma transmissão com relação sinal ruído acima do limiar de recepção do *gateway*, o que demandará o aumento dos valores de ambos os parâmetros durante o processo de aprendizado.

- Episódio: Representa uma sequência de iteração entre os agentes (dispositivos finais) e o ambiente no qual os dispositivos finais realizam ações com base no seu estado atual e avaliam a sua recompensa. O processo é repetido para um determinado número de iterações até que a quantidade máxima de episódios seja atingida. No caso da seleção apenas do fator de espalhamento, cada agente realiza 6 iterações por episódio, permitindo ao agente explorar até a quantidade máxima de estados possíveis (fatores de espalhamento). Já no caso da seleção de fator de espalhamento e potência de transmissão, cada agente realiza 18 iterações por episódio, permitindo ao agente explorar até a quantidade máxima de estados possíveis (tupla composta pelo fator de espalhamento e potência de transmissão). A quantidade total de episódios é de 25.000 em ambos os casos.

- Critério de parada: No estudo de caso 1, os critérios de parada utilizando o método proposto são: *i.*) caso o número máximo de iterações foi atingido, ou *ii.*) caso a energia consumida pelo agente foi menor que um valor limite, obtendo uma recompensa proporcional à energia consumida. Este valor limite é definido de acordo com a quantidade de dispositivos e da energia consumida pelo ADR para a mesma quantidade de dispositivos, com o intuito de obter uma energia consumida total menor que o mecanismo ADR, proporcionando uma taxa de entrega de pacotes maior ou igual a este mecanismo. No estudo de caso 2, os critérios de parada utilizando o método proposto são: *i.*) caso o número máximo de iterações foi atingido, ou *ii.*) caso a taxa de entrega de pacotes seja de 100%.

$$Recompensa = \frac{\alpha \sum_{i=0}^{N} P_i}{\sum_{i=0}^{N} E_i}. \tag{5.1}$$

O algoritmo proposto é executado de forma que uma estrutura de dados (*Q-table*) é criada para cada dispositivo final a nível de servidor de rede, representado no Algoritmo 1 para ajustar do fator de espalhamento e potência de transmissão. O processo de aprendizado está descrito no diagrama de blocos 5.3. A quantidade de agentes é n, que é exatamente a quantidade de dispositivos (n>0). M (M>0) é a quantidade de episódios. n_s corresponde à quantidade de possíveis estados do agente (n_s=3), ou seja, o status da transmissão do agente para os parâmetros selecionados ("transmitido", "interferido", ou "abaixo da sensibilidade"). n_a corresponde a quantidade de ações que o agente pode tomar. α representa a taxa de aprendizado (α=0.1). O fator de desconto γ tem valor 0.95. Caso a seleção seja apenas do

fator de espalhamento, a ação corresponde ao valor do fator de espalhamento (SF). Caso a seleção seja do fator de espalhamento e da potência de transmissão, a ação corresponde a uma tupla com os valores destes parâmetros (SF, P_t). z é a quantidade de passos por episódio, sendo z=6 quando o algoritmo ajusta apenas o fator de espalhamento e z=18 quando ajusta ambos fator de espalhamento e potência de transmissão.

Algoritmo: Multi-agente Q-learning para alocação de fator de espalhamento e potência de transmissão em redes LoRa.

Require: $n>0$, $M>0$, $n_s = 18$, $n_o = 2$ $n_a = 18$

Entrada: Número de identificação de cada dispositivo.

Saída: Fator de espalhamento e potência de transmissão para cada dispositivo.

Estado $s \in \mathcal{S} = \{1, \ldots, n_s\}$

Observações $o \in \mathcal{O} = \{1, \ldots, n_o\}$

Ação $a \in \mathcal{A} = \{1, \ldots, n_a\}$

Função Recompensa $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$

Função transição de probabilidade $T : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$

procedure QLEARNING($\mathcal{S}$, A, R, T, α, γ)
  Gera as n funções Q-learning $Q_i(S, a), i \in [1, n]$;
  Inicializa a função Q-learning $Q_i(S, a), i \in [1, n]$; para cada dispositivo;
  for episódio = 1,M **do**
    for dispositivo = 1,n **do**
      Seleciona o estado inicial s
      for t =1 : z ou até que s seja o estado final **do**
        Para cada agente:
        com probabilidade ϵ, seleciona uma ação aleatória
        $a_i = (SF, P_t)$, senão seleciona $a_i = \max_a Q_i(s, a_i)$;
        Calcula o parâmetro crítico de atualização
        $\Delta Q_i = r(s, a_i) + \gamma \cdot \max_{a \in A} Q_i(s', a'_i) - Q_i(s, a_i)$
        Atualiza o fator crítico utilizando a taxa de aprendizado α:
        $Q_i(s', a_i) = Q_i(s, a_i) + \alpha \cdot \Delta Q_i$;
        Seleciona o estado s'
  return $Q_i(s', a_i)$

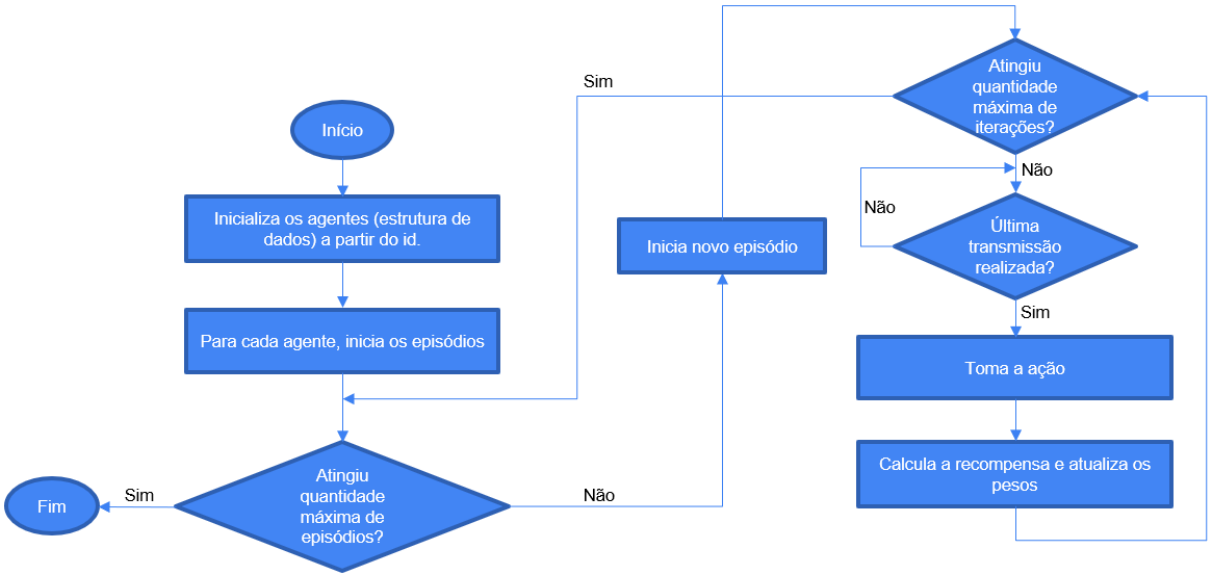


Figura 5.3: Processo de aprendizado dos agentes.

Esta proposta é utilizada no estudo de caso 1, com o intuito de otimizar a taxa de entrega de pacotes e a energia consumida por cada dispositivo a partir das experiências obtidas a medida que os agentes selecionam os parâmetros. No estudo de caso 2, o algoritmo é utilizado de forma análoga, porém buscando otimizar apenas a taxa de entrega de pacotes.

Capítulo 6

Experimentos e Resultados

A seguir, são apresentados os estudos de caso nos quais são utilizadas as propostas apresentadas.

6.1 Estudo de caso 1: *Gateway* no centro da topologia e dispositivos com posições aleatórias

A topologia da rede referente ao estudo de caso 1 está ilustrada na Figura 6.1. Os dispositivos são considerados classe A e estão dispostos de forma aleatória dentro da topologia circular com *gateway* central. O triângulo vermelho central representa o *gateway* e os círculos representam os dispositivos. Cada dispositivo possui uma cor representando o fator de espalhamento selecionado. Os círculos de cores representam o alcance relativo a cada fator de espalhamento. Nota-se que alguns dispositivos possuem cores diversas da cor do círculo que os limitam, indicando que o fator de espalhamento pode estar acima ou abaixo do valor ótimo, o que pode ocasionar queda na taxa de entrega de pacotes ou consumo elevado de energia na transmissão.

É possível verificar que existe uma alta probabilidade de ocorrer colisões entre dispositivos que estejam próximos do *gateway* devido às transmissões concorrentes e à seleção de fatores de espalhamento iguais. Este problema das colisões envolvendo dispositivos próximos ao *gateway* pode ser resolvido forçando alguns dispositivos que estejam próximos ao *gateway* a selecionar fatores de espalhamento maiores (destacados na Figura 6.1) nas cores diversas da cor azul). Porém, o aumento do fator de espalhamento desses dispositivos deve ser feito de forma cuidadosa para não desperdiçar energia. Além disso, os dispositivos mais distantes do *gateway* podem selecionar fatores de espalhamento e potências de transmissão ora acima do necessário (o que aumenta a energia consumida na transmissão), ora abaixo do necessário, ocasionando a perda de pacotes devido à impossibilidade de captura do sinal por parte do *gateway*. O mesmo cuidado na seleção dos fatores de espalhamento também

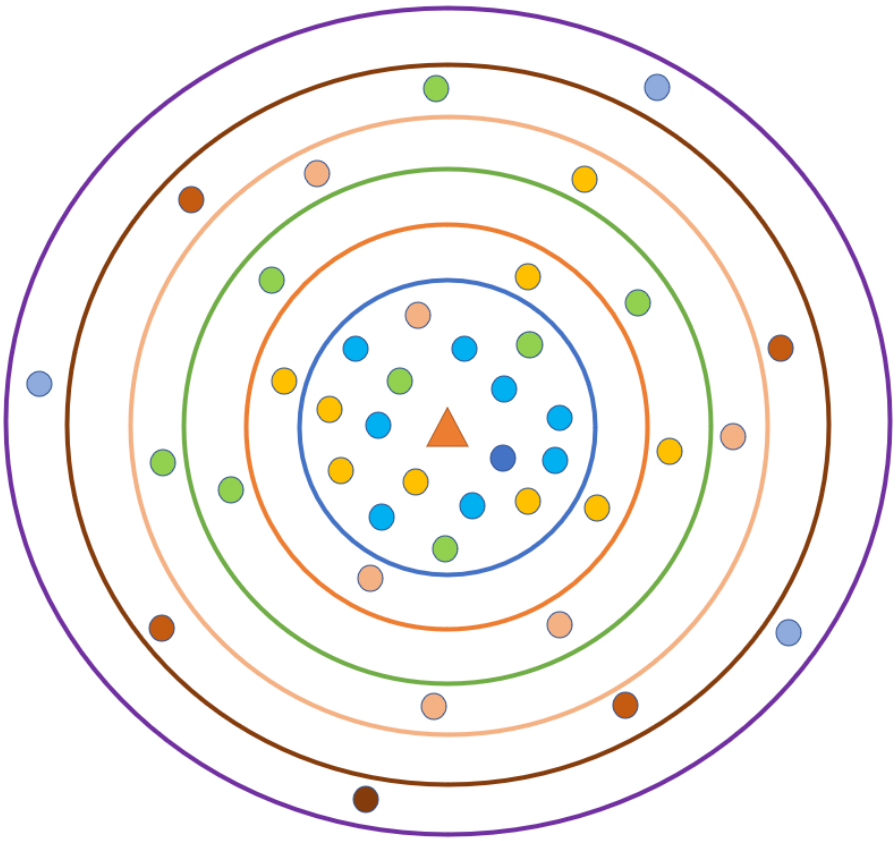


Figura 6.1: Seleção do fator de espalhamento para os dispositivos próximos ao *gateway* para evitar a colisão.

vale para dispositivos mais distantes do *gateway*, mas com menos importância, uma vez que os dispositivos estão menos concentrados. Para dispositivos mais distantes do *gateway*, uma seleção adequada da potência de transmissão se mostra relevante para a obtenção de uma elevada taxa de entrega de pacotes e consequentemente uma maior eficiência energética.

Para este estudo de caso, foram utilizadas técnicas de aprendizado supervisionado Árvore de Decisão e Máquina de Vetor de Suporte e do algoritmo de aprendizado por reforço *Q-learning* para selecionar o fator de espalhamento e da potência de transmissão nos dispositivos finais. O objetivo foi investigar a escalabilidade do mecanismo ADR e comparar o seu desempenho com o desempenho das referidas propostas em termos de taxa de entrega de pacotes e energia consumida na transmissão.

Para cada transmissão, o classificador prevê o resultado da transmissão para os menores fatores de espalhamento e potência de transmissão possíveis. Se o resultado da transmissão é classificado como interferido, então o simulador aumenta a potência de transmissão e, quando alcançada a potência de transmissão máxima, segue aumentando o fator de espalhamento e prevê o resultado da nova transmissão. Se o resultado da nova transmissão é classificada como bem sucedida, então as

transmissões do dispositivo seguem com os parâmetros selecionados.

Inicialmente, assume-se que os dispositivos selecionam a potência máxima de transmissão de 14 dBm para os canais da banda europeia. Com o intuito de avaliar o efeito da seleção dos fatores de espalhamento. Uma segunda análise é realizada considerando a variação da potência de transmissão e do fator de espalhamento para avaliar o efeito de ambos os parâmetros. É utilizado apenas um canal de 868 MHz com o intuito de avaliar apenas o efeito do fator de espalhamento e potência de transmissão.

6.2 Estudo de caso 2: Topologia circular com *gateway* a 5-10 km do centro da topologia

A Figura 6.2 ilustra o caso no qual um conjunto de 5 dispositivos finais formam um círculo e se comunicam com um *gateway* localizado a uma distância de 5 a 10 km de distância do centro da formação dos dispositivos.

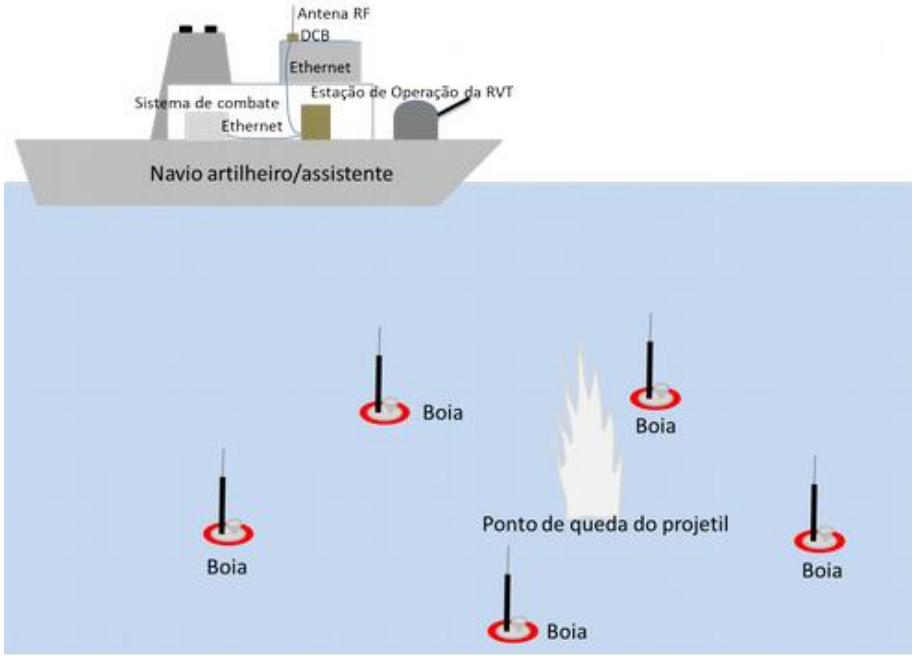


Figura 6.2: Ilustração do estudo de caso 2 no qual um conjunto de 5 sensores acústicos submarinos, conectados cada um a um dispositivo LoRa, em uma determinada área marítima com formação circular. O *gateway* da rede está instalado em um navio distante.

Os dispositivos representam as 5 sonobóias e o *gateway* está localizado no navio. Cada boia possui alimentação, central de processamento, sensor acústico, comunicação via rádio e receptor de sistema de posicionamento via satélite independentes. As sonobóias detectam o som da queda de um projétil que se propaga através da

água, sensibilizando cada hidrofone em instantes de tempo diferentes. A partir desta detecção, cada sonobóia transmite via rádio, para o *gateway*, representado na figura como DCB, sua posição georreferenciada, com a marcação temporal, obtida pela referência via satélite. O *gateway* então repassa essas informações à Estação de Operação (EO), onde é calculada e disponibilizada a estimativa da localização do ponto de queda do projétil.

O objetivo é avaliar o desempenho das propostas e do mecanismo ADR utilizados na seleção do fator de espalhamento de forma a impedir a degradação da capacidade da rede quando esta é submetida a uma fonte de interferência externa. Os dispositivos possuem o mesmo fator de espalhamento uma vez que todos eles estão a uma distância de 5 a 10 km. Uma fonte de interferência de 433 MHz e modulação *chirp spread spectrum* com fator de espalhamento de valores aleatórios entre 7, 8 e 9 foi introduzida com o intuito de degradar os enlaces entre os dispositivos e o *gateway* para forçar o algoritmo a ajustar o fator de espalhamento de forma a recuperar a capacidade dos enlaces. Esta fonte de interferência realiza uma transmissão a cada 3 s de pacotes cujo tamanho é de 200 bytes e está localizado nas coordenadas (2500, 1000) em metros, conforme a Figura 6.3. A fonte de interferência atuou durante todo o tempo de simulação.

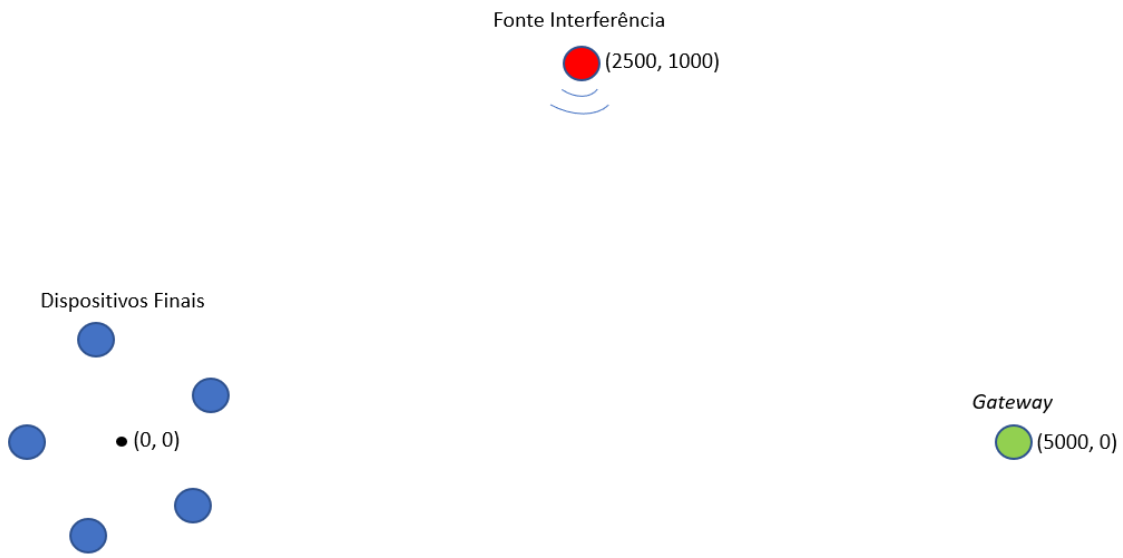


Figura 6.3: Topologia da rede no estudo de caso 2 com a fonte de interferência.

Busca-se selecionar o fator de espalhamento de forma a obter a taxa de entrega de pacotes em 100%. A potência de transmissão é mantida fixa em 22 dBm. As características da fonte de interferência externa são descritas na seção seguinte.

6.3 Parâmetros de simulação

O simulador utilizado foi o desenvolvido por Yatagan et. al. [17] escrito em linguagem Python, no qual o efeito da ortogonalidade imperfeita dos fatores de espalhamento e o efeito das colisões devido às transmissões concorrentes estão implementados, tornando o efeito das colisões mais próximo possível da realidade. Foram realizadas alterações no simulador para permitir a variação da potência de transmissão, originalmente fixada em 14 dBm. O modelo de propagação utilizado foi o de espaço livre. A soma dos ganhos e perdas de sinal considerada nos dispositivos transmissores e receptores foi de 7 dB. O CR utilizado tem valor 1.

Para a utilização dos modelos de aprendizado supervisionado, foi utilizada a biblioteca scikit-learn desenvolvida em linguagem Python [36]. O mecanismo de busca em grade (*GridSearch*) da biblioteca *scikit-learn* foi utilizado para realizar a busca exaustiva dos melhores hiper parâmetros das técnicas Árvore de Decisão e Máquina de Vetor de Suporte.Para a técnica de Árvore de Decisão, a função "criterion" selecionada (que mede a qualidade da divisão de um nó em outros nós hierarquicamente inferiores) foi a "gini". A estratégia selecionada para realizar a divisão em cada nó foi a "best". A máxima profundidade "max depth" foi selecionada como "None", fazendo com que os nós sejam expandidos até que todos os nós sejam do tipo folha. O número de atributos considerados para encontrar a melhor divisão do nó "max features" foi configurado como "None", sendo "max features" igual ao número de atributos da base de dados. O número máximo de nós folha "max leaf nodes" foi considerado ilimitado (igual a "None"). Os pesos atribuídos às classes foram consideradas os mesmos e iguais a 1, ou seja, "class weight" igual a "None". Para a técnica de Máquina de Vetor de Suporte, o parâmetro de regularização "C" foi selecionado em 1. O tipo de núcleo utilizado no algoritmo foi "rbf" e o coeficiente utilizado foi "scale".

O tamanho dos pacotes enviados pelos dispositivos é de 60 Bytes no estudo de caso 1 e de 200 bytes no estudo de caso 2. Foram considerados dispositivos classe A, já que estes possuem menor consumo de energia.

O tempo de simulação considerado no estudo de caso 1 foi de 3600s, enquanto no estudo de caso 2 foi de 600 segundos. Os intervalos de geração e transmissão de pacotes LoRa no estudo de caso 1 seguem uma distribuição de *Poisson* com média $0,01t$ [37, 38]. No estudo de caso 2, os dispositivos acessam o meio utilizando multiplexação temporal, na qual é realizada a primeira transmissão no instante de tempo $t = 2id$, sendo id o número de identificação do dispositivo. Considerando que $id \in \{0, 1, 2, 3, 4\}$, os dispositivos transmitem inicialmente nos instantes 0,2,4,6 e 8 s. Após a primeira transmissão, cada dispositivo segue uma transmissão periódica, passando a transmitir a cada 10 s. A fonte de interferência utilizada no estudo de

caso 2 utiliza uma frequência de 433 MHz e modulação *chirp spread spectrum* com fator de espalhamento de valores aleatórios entre 7, 8 e 9. Esta fonte de interferência realiza uma transmissão a cada 3 s de pacotes cujo tamanho é de 200 bytes e está localizado nas coordenadas (2.5 km, 1 km). A fonte de interferência foi gerada pelo simulador de forma semelhante à geração de um dispositivo final, mas com acesso ao meio a cada 3 segundos e a sua contabilização de pacotes transmitidos e perdidos foi desabilitada, uma vez que não é considerado um dispositivo da rede e que o objetivo é apenas gerar colisões entre pacotes desta fonte de interferência e dos dispositivos finais.

A qualidade do enlace sem fio pode ser calculada através da diferença entre os ganhos e as perdas do dispositivo transmissor para o dispositivo receptor, conforme a Equação 6.1 [39].

$$P_{rx}^{dbm} = P_{tx}^{dbm} + G_{SYS}^{db} - L_{SYS}^{db} - L_{PATH}^{db}. \tag{6.1}$$

Na Equação 6.1, P_{RX} é a potência recebida, P_{TX} é a potência transmitida. G_{SYS} é o ganho do sistema composto pelo ganho da antena transmissora e receptora. L_{SYS} é a perda no sistema, composta pela perda no circuito e na antena. L_{PATH}. No simulador, a soma dos ganhos e das perdas do sistema é considerada igual a +7 dB.

A propagação no espaço livre é calculada de acordo com a Equação 6.2 [40], na qual h é a altura do *gateway* e f é a frequência do sinal.

$$L_{PATH}^{db} = 40(1 - 4 * 10^{-3} * h)\log_{10} R - 18\log_{10} h + 21\log_{10} f + 80. \tag{6.2}$$

Caso a potência do sinal recebido for maior que a sensibilidade do *gateway*, o sinal pode ser decodificado pelo receptor com sucesso quando não há transmissão interferente.

Os parâmetros de simulação referentes aos estudos de caso 1 e 2 estão contidos nas Tabelas 6.1 e 6.2. Os valores do fator de desconto e taxa de aprendizagem foram 0.95 e 0.1 respectivamente, uma vez que estes são valores típicos utilizados no algoritmo *q-learning* e que permitem aos agentes um aprendizado eficiente.

Tabela 6.1: Parâmetros de simulação do estudo de caso 1.

Parâmetros	`Valor`
Frequência da portadora	868 MHz
Altura do *gateway*	15 m
Altura do dispositivo	1,5 m
Raio da topologia	5 km
Número total de episódios	25000
Número de iterações por episódio	6 (SF) e 18 (SF e TP)
Taxa de aprendizagem(α)	0.1
Fator de desconto(λ)	0.95
Tempo de simulação em s	3600 s

Tabela 6.2: Parâmetros de simulação do estudo de caso 2.

Parâmetros	`Valor`
Frequência da portadora	433 MHz
Altura do *gateway*	15 m
Altura do dispositivo	1,5 m
Distância entre *gateway* e centro da topologia	5-20 km
Raio da topologia circular de dispositivos	250 m
Número total de episódios	25000
Número de iterações por episódio	6
Taxa de aprendizagem(α)	0.1
Fator de desconto(λ)	0.95
Tempo de simulação em	600 s

6.4 Resultados do estudo de caso 1: *Gateway* no centro da topologia e dispositivos com posições aleatórias

A Figura 6.4 representa a taxa de entrega de pacotes obtida pelas técnicas de aprendizado utilizadas, assim como pelo mecanismo ADR, em função do número de dispositivos. Este resultado considera apenas o ajuste do fator de espalhamento, mantendo a potência de transmissão fixa em 14 dBm. Verifica-se que as técnicas de aprendizado supervisionado obtiveram uma taxa de entrega de pacotes superior ao ADR para todas as quantidades de dispositivos, enquanto o algoritmo *Q-learning* obteve aproximadamente o mesmo desempenho do ADR.

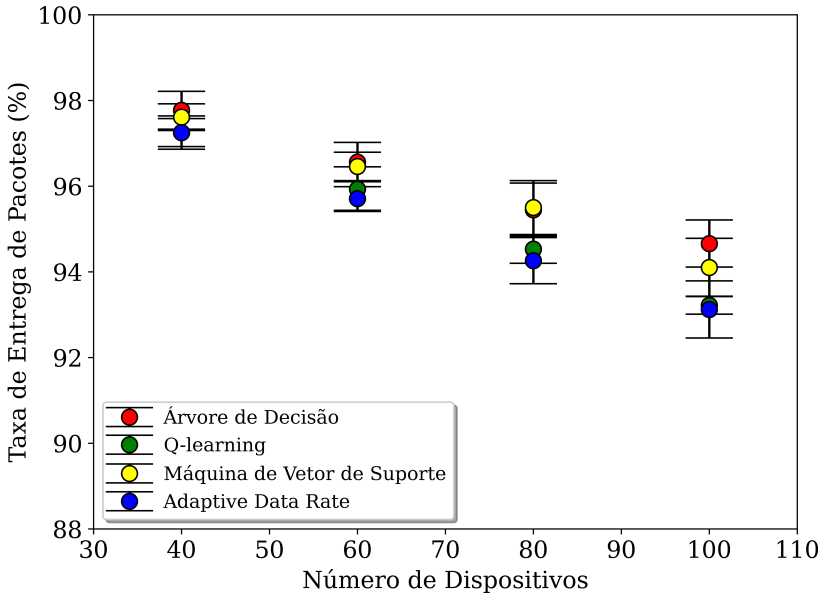
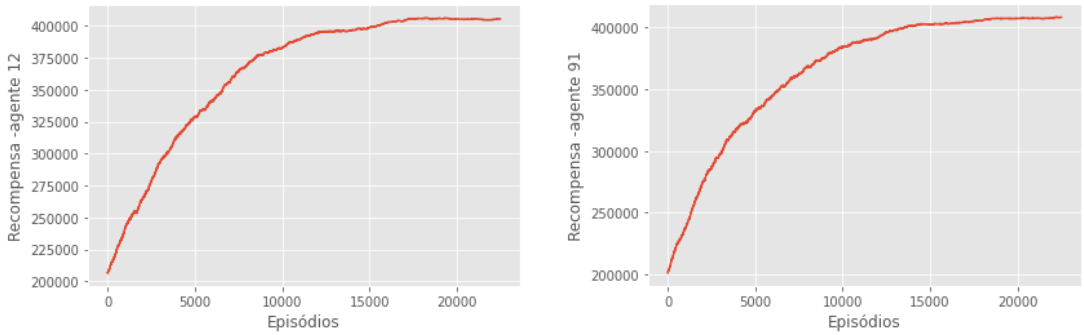


Figura 6.4: Taxa de entrega de pacotes obtida pelas propostas selecionando fator de espalhamento para uma quantidade de 40 a 100 dispositivos.

As curvas da recompensa obtida pelos agentes 12 e 91 estão representadas na Figura 6.5 nas quais é possível verificar que o recompensa média convergiu para um valor ótimo, indicando que a quantidade de episódios definida para o treinamento foi suficiente para prover um aprendizado adequado, fazendo o agente encontrar a política ótima de seleção do fator de espalhamento. O mesmo comportamento foi obtido no treinamento dos demais agentes. As curvas dos agentes 12 e 91 foram escolhidas para fins de ilustração do treinamento dos agentes e, uma vez que todos os agentes obtiveram resultado de treinamento com comportamento de convergência semelhantes, quaisquer outras curvas de outros agentes poderiam ter sido escolhidos para ilustrar o treinamento. Para o caso específico, o mesmo valor ótimo de 400.000 foi obtido nas curvas dos agentes 12 e 91, indicando que os dispositivos selecionaram o mesmo fator de espalhamento, mas o valor ótimo de convergência pode variar a depender dos parâmetros selecionados por cada agente. O tempo mínimo requerido pelo algoritmo para atingir a convergência na seleção do fator de espalhamento foi de 25000 episódios, conforme pode ser visto na Figura 6.5.



(a) Recompensa do agente 12.

(b) Recompensa do agente 91.

Figura 6.5: Recompensa média obtida pelos agentes ao longo do treinamento com 25000 episódios no estudo de caso 1 com 100 dispositivos para seleção de fator de espalhamento.

Na Figura 6.6, verifica-se que as técnicas de aprendizado supervisionado consumiram mais energia em relação ao mecanismo ADR, enquanto a técnica de aprendizado por reforço consumiu a mesma quantidade de energia consumida pelo ADR para todas as quantidades de dispositivos.

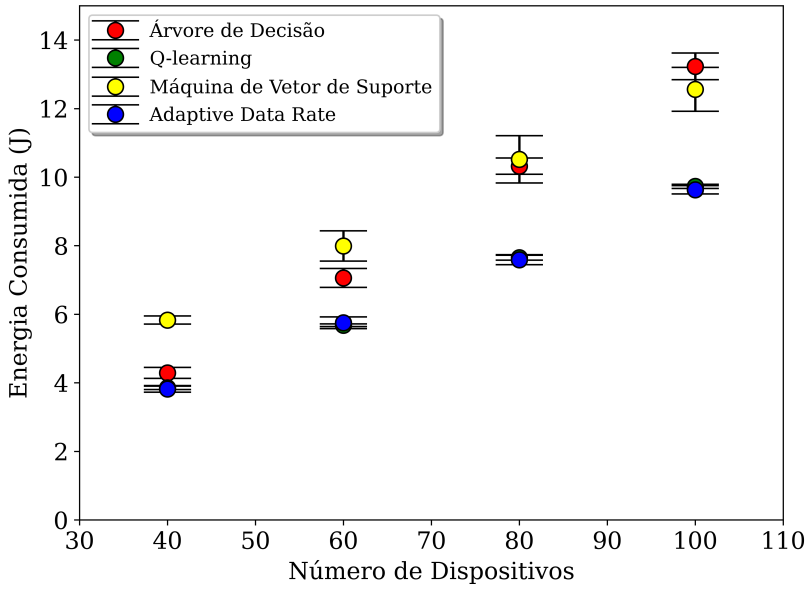


Figura 6.6: Energia consumida na transmissão obtida pelas propostas para quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento.

Uma vez que as técnicas de aprendizado supervisionado consumiram mais energia em relação ao mecanismo ADR, apesar de superar este em termos de taxa de entrega de pacotes, uma análise da eficiência energética de cada técnica utilizada, ou seja, a quantidade de pacotes transmitidos com sucesso a cada J (Joule) consumido, se fez

necessária. Para isso, dividiu-se a quantidade de pacotes transmitidos com sucesso pela quantidade total de energia consumida na simulação.

A Figura 6.7 representa a eficiência energética das técnicas de aprendizado, assim como a do mecanismo ADR. Uma eficiência energética maior indica que a proposta consegue transmitir mais pacotes com sucesso gastando menos energia, obtendo melhor resultado. Verifica-se que a técnica de aprendizado por reforço obteve uma eficiência energética igual a do mecanismo ADR, enquanto as técnicas de aprendizado supervisionado obtiveram uma eficiência inferior ao do ADR. Este resultado indica que, para a seleção apenas do fator de espalhamento de uma quantidade de até 100 dispositivos, o mecanismo ADR se mostra adequado para fins de economia de energia.

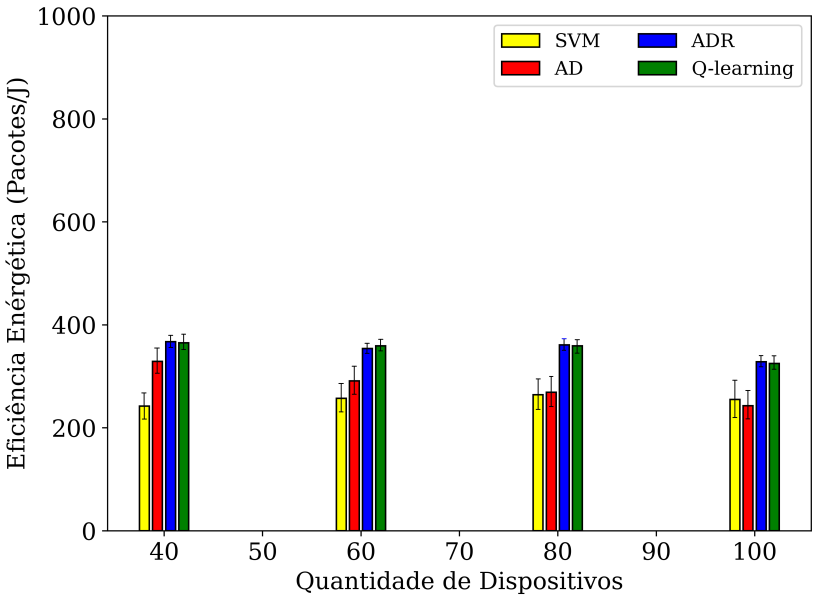


Figura 6.7: Eficiência energética obtida pelas técnicas utilizadas para uma quantidade de 40 a 100 dispositivos na seleção do fator de espalhamento.

A Figura 6.8 representa a taxa de entrega de pacotes obtida pelas técnicas de aprendizado utilizadas, assim como pelo mecanismo ADR, considerando o ajuste do fator de espalhamento e da potência de transmissão. Verifica-se que todas as técnicas obtiveram uma taxa de entrega de pacotes superior ao ADR para todas as quantidades de dispositivos.Foi considerado um intervalo de confiança de 95% no cálculo do erro associado à taxa de entrega obtida por cada técnica.

As curvas da recompensa obtida pelos agentes 25 e 77 estão representadas na Figura 6.9 nas quais é possível verificar que a recompensa média convergiu para um valor ótimo, indicando que a quantidade de episódios definida para o treinamento foi suficiente para prover um aprendizado adequado, fazendo o agente encontrar a

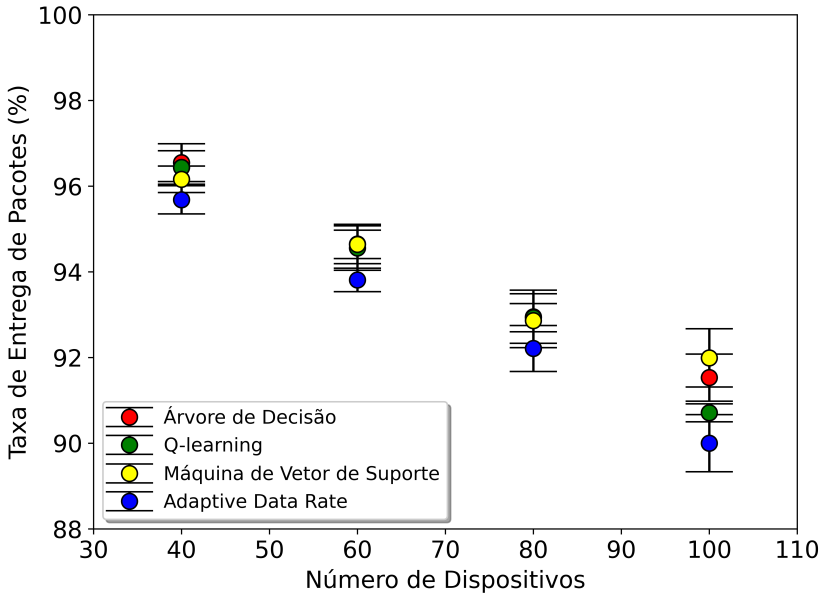
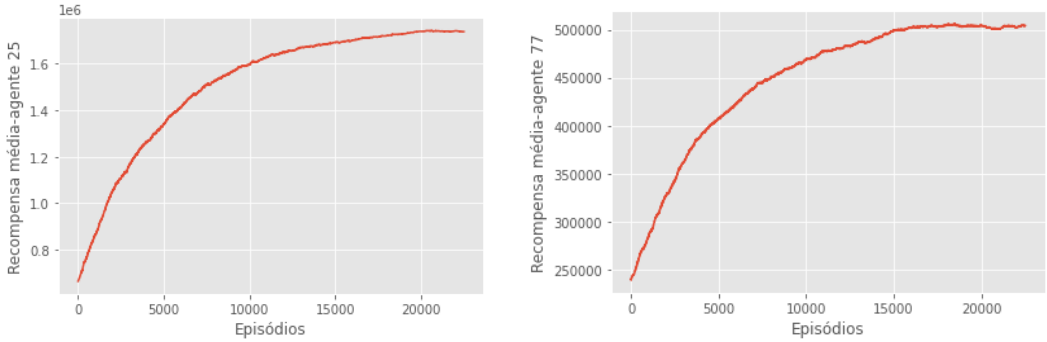


Figura 6.8: Taxa de entrega de pacotes obtida pelas propostas selecionando fator de espalhamento e potência de transmissão para quantidade de 40 a 100 dispositivos.

política ótima de seleção do fator de espalhamento e potência de transmissão. O mesmo comportamento foi obtido no treinamento dos demais agentes. As curvas dos agentes 25 e 77 foram escolhidas para fins de ilustração do treinamento dos agentes e, uma vez que todos os agentes obtiveram resultado de treinamento com comportamento da convergência semelhantes, quaisquer outras curvas de outros agentes poderiam ter sido escolhidos para ilustrar o treinamento. Para o caso da Figura 6.9, os agentes 25 e 77 convergiram para os valores ótimos diversos entre si de 1.000.000 e 500.000 respectivamente. Isto ocorreu devido à seleção de fatores de espalhamento e potências de transmissão diferentes entre si por parte dos dispositivos, já que a recompensa dada ao agente depende os valores selecionados de ambos os parâmetros. A quantidade de episódios requerida pelo algoritmo para atingir a convergência na seleção do fator de espalhamento e potência de transmissão também foi de 25000 episódios, conforme pode ser visto na Figura 6.9, o que representa aproximadamente 5 horas para a convergência do algoritmo (para a quantidade de 100 dispositivos).



(a) Recompensa do agente 25.

(b) Recompensa do agente 77.

Figura 6.9: Recompensa média obtida pelos agentes ao longo do treinamento com 25000 episódios no estudo de caso 1 com 100 dispositivos para seleção de fator de espalhamento e potência de transmissão.

Na Figura 6.10, verifica-se que as técnicas de aprendizado supervisionado consumiram mais energia em relação ao mecanismo ADR, enquanto a técnica de aprendizado por reforço consumiu menos energia para todas as quantidades de dispositivos.

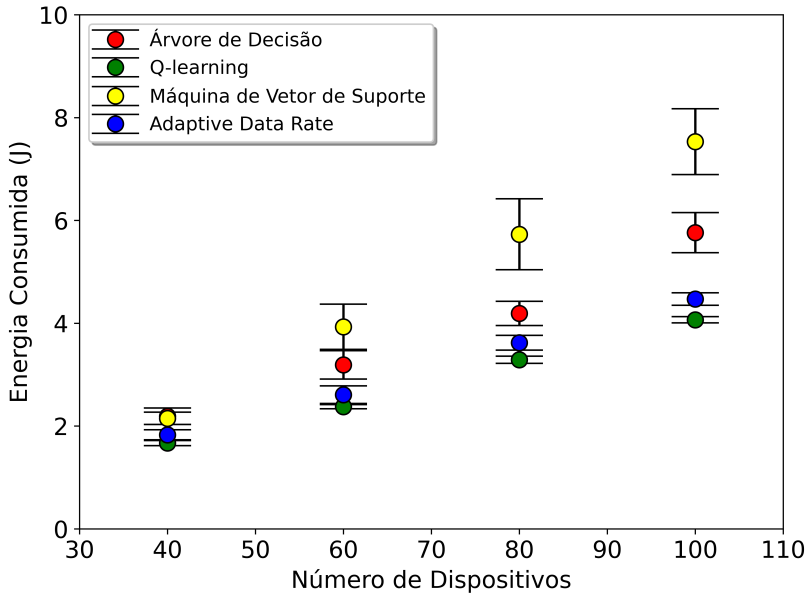


Figura 6.10: Energia consumida na transmissão obtida pelas propostas para uma quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento e potência de transmissão.

De forma análoga ao caso de seleção apenas do fator de espalhamento, a eficiência energética de cada técnica utilizada está representada na Figura 6.11, assim como do mecanismo ADR, através da qual verifica-se que a técnica de aprendizado por reforço foi a única proposta que de fato obteve uma eficiência energética maior que

a do mecanismo ADR. Este resultado indica a efetividade do algoritmo de aprendizado por reforço utilizado para dirimir as interferências internas em detrimento do mecanismo ADR.Foi considerado um intervalo de confiança de 95% no cálculo do erro associado à energia consumida e eficiência energética obtida por cada técnica.

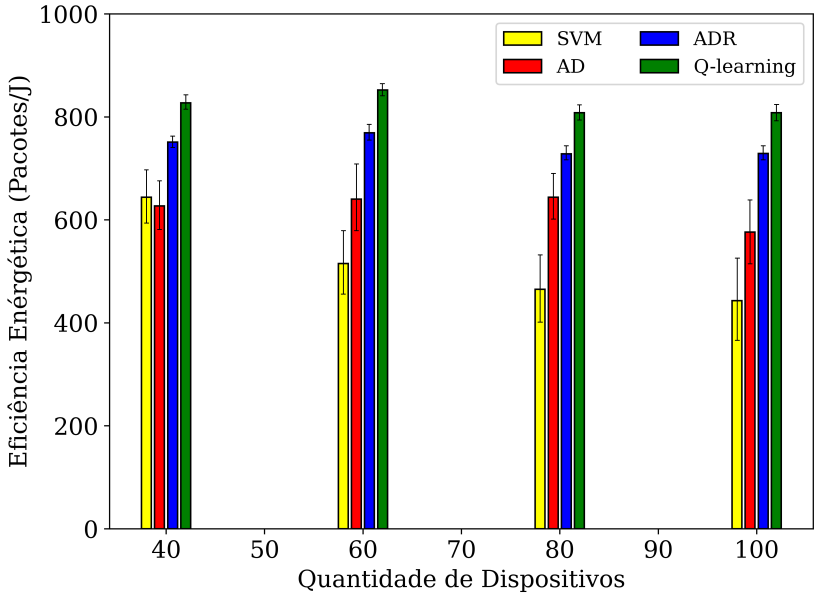
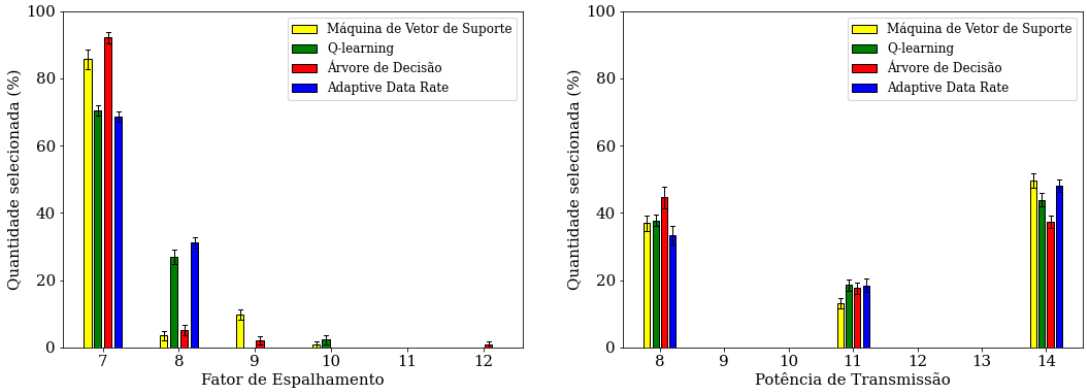


Figura 6.11: Eficiência energética obtida pelas propostas para uma quantidade de 40 a 100 dispositivos na seleção de fator de espalhamento e potência de transmissão.

A seleção dos fatores de espalhamento e potências de transmissão para quantidade de 100 dispositivos está abaixo representada na Figura 6.12. O intervalo de confiança considerado no cálculo do erro associado à quantidades selecionadas de cada parâmetro também foi de 95%.



(a) Fatores de espalhamento selecionados por cada técnica de aprendizado para 100 dispositivos.

(b) Potências de transmissão selecionadas por cada técnica de aprendizado para 100 dispositivos.

Figura 6.12: Seleção dos Fatores de espalhamento e potências de transmissão das técnicas de aprendizado para a quantidade de 100 dispositivos.

Através da Figura 6.12(a), verifica-se que o algoritmo *Q-learning* selecionou 2,5% do total dos fatores de espalhamento como SF10, 70,5% como SF7 e 27% como SF8, enquanto o mecanismo ADR selecionou 69% como SF7 e 31% como SF8, mostrando uma diversidade maior na seleção dos fatores de espalhamento por parte do algoritmo *Q-learning* em relação ao ADR. Este resultado indica que o algoritmo *Q-learning* tem melhor desempenho ao dirimir as interferências internas devido às transmissões concorrentes entre dispositivos próximos ao *gateway*, já que, seguindo o mecanismo ADR, os dispositivos próximos ao *gateway* tendem a selecionar os menores fatores de espalhamento. Análise semelhante pode ser feita com a seleção realizada pelas técnicas de aprendizado supervisionado dos fatores de espalhamento, as quais também obtiveram uma diversidade maior, dirimindo os efeitos das transmissões concorrentes próximas ao *gateway*. Esta diversidade maior de seleção de fatores de espalhamento, portanto, justifica os resultados superiores de taxa de entrega de pacotes das propostas em relação ao mecanismo ADR, uma vez que as propostas promoveram uma maior resiliência às interferências.

Conforme apresentado anteriormente, a seleção de fatores de espalhamento maiores acarreta no aumento da energia consumida. Na Figura 6.10, verificou-se que as técnicas de Árvore de Decisão e Máquina de Vetor de Suporte obtiveram maior consumo de energia. Conforme representado na Figura 6.12, este resultado pode ser justificado devido à seleção de maiores fatores de espalhamento por parte da técnica Árvore de Decisão (seleção de SF9 e SF12) e Máquina de Vetor de Suporte (seleção de SF9 e SF10), além da seleção de maiores potências de transmissão por parte da técnica Máquina de Vetor de Suporte (maior seleção de potências de 14 dBm).

Na Figura 6.12(b), verifica-se que o algoritmo *Q-learning* permitiu a seleção de

potências de transmissão menores em relação ao mecanismo ADR. Portanto, diante dos resultados apresentados, conclui-se que o algoritmo *Q-learning* se apresenta como a melhor proposta para aplicações cuja energia consumida pelos dispositivos seja um fator crítico, já que encontrou o melhor compromisso entre taxa de entrega e economia de energia. Porém, o tempo necessário para a convergência do algoritmo *q-learning* pode representar uma desvantagem do algoritmo *q-learning* em relação ao aprendizado supervisionado para aplicações cuja energia consumida pelos dispositivos finais não seja um fator crítico. As técnicas de aprendizado supervisionado, apesar de obter maior taxa de entrega de pacotes com o custo do aumento da energia consumida, o tempo necessário para o treinamento dos modelos utilizados foi de apenas 3 minutos, representando uma vantagem para aplicações cujo consumo de energia da rede não seja um fator crítico.

6.5 Resultados do estudo de caso 2: Topologia circular com *gateway* de 5 a 10 km do centro da topologia

O objetivo deste experimento foi avaliar o desempenho das propostas, em termos de taxa de entrega de pacotes, em ajustar o fator de espalhamento diante do surgimento da fonte de interferência externa. Neste estudo de caso, os dispositivos possuem o mesmo fator de espalhamento uma vez que todos eles estão a uma distância de 5 a 10 km. Conforme apresentado anteriormente, uma fonte de interferência de 433 MHz e modulação *chirp spread spectrum* com fator de espalhamento de valores aleatórios entre 7, 8 e 9 foi introduzida com o intuito de dar uma maior dinamicidade ao estudo de caso. Esta fonte de interferência realiza uma transmissão a cada 3 s de pacotes cujo tamanho é de 200 bytes e está localizado nas coordenadas (2.5 km, 1 km). A métrica avaliada foi apenas a taxa de entrega de pacotes, pois, para a aplicação em questão, a energia consumida não é um fator crítico, uma vez que os dispositivos são lançados ao mar para a realização dos testes do sistema de canhão e, após os testes, os dispositivos são recolhidos, permitindo o teste constante das baterias e sua substituição caso seja necessário.

6.5.1 Experimento sem a implementação das propostas

A Figura 6.13 representa a taxa de entrega de pacotes obtida para cada seleção de fator de espalhamento, na qual verifica-se que, no cenário sem interferência, apenas os fatores de espalhamento 9 e 10 pode ser selecionados, pois ao selecionar os fatores de espalhamento 7 ou 8, o *gateway* não consegue capturar as transmissões, uma vez

que estão abaixo da sensibilidade do *gateway*. Já quando os fatores de espalhamento 11 ou 12 são selecionados, ocorrem colisões devido ao aumento excessivo do ToA.

Quando submetida à fonte de interferência, a capacidade da rede é degradada, comprometendo o desempenho da taxa de entrega de pacotes quando selecionado o fatore de espalhamento 9, indicando que apenas a seleção do fator de espalhamento 10 proporciona uma conectividade adequada.

Por ocasião da configuração da rede, o operador pode testar a conectividade para cada fator de espalhamento e selecionar aquele que fornece a melhor conectividade e que consome menos energia. Outra alternativa é utilizar o mecanismo ADR para a configuração da rede que, no cenário sem interferência, selecionaria o fator de espalhamento 9. Porém, uma vez configurado o fator de espalhamento para 9, com o surgimento da fonte de interferência externa, a capacidade dos enlaces é degradada. Neste caso, o ajuste do fator de espalhamento é um fator crítico e muito importante para continuar provendo conectividade entre os dispositivos e o *gateway*.

Diante do surgimento de uma fonte de interferência externa à rede, uma vez que é desconhecida do operador, obtém-se o resultado em vermelho da Figura 6.13, necessitando, então, uma nova configuração da rede que a conectividade entre os dispositivos.

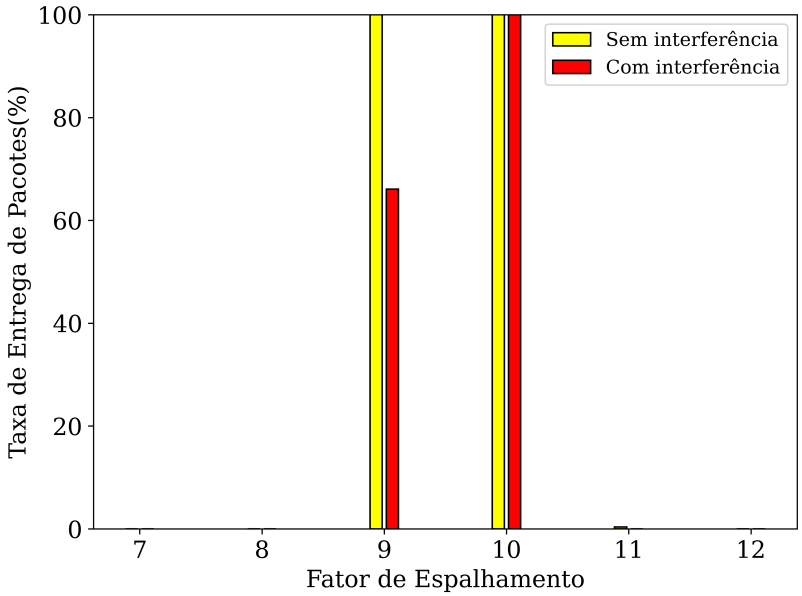


Figura 6.13: Taxa de entrega de pacotes mantendo fixos cada fator de espalhamento em um cenário com e sem interferência e com *gateway* a uma distância de 10 km.

De forma análoga ao caso com *gateway* à 10 km de distância, a Figura 6.14 ilustra o resultado da taxa de entrega de pacotes obtida para cada fator de espalhamento selecionado no caso de um *gateway* à 5 km do centro da topologia. Neste caso,

verifica-se que os fatores de espalhamento de valores 7, 8, 9 e 10 proporcionam uma taxa de entrega de 100% quando a rede não está degradada, enquanto apenas o fator de espalhamento 10 provê uma capacidade adequada aos enlaces quando a rede está degradada.

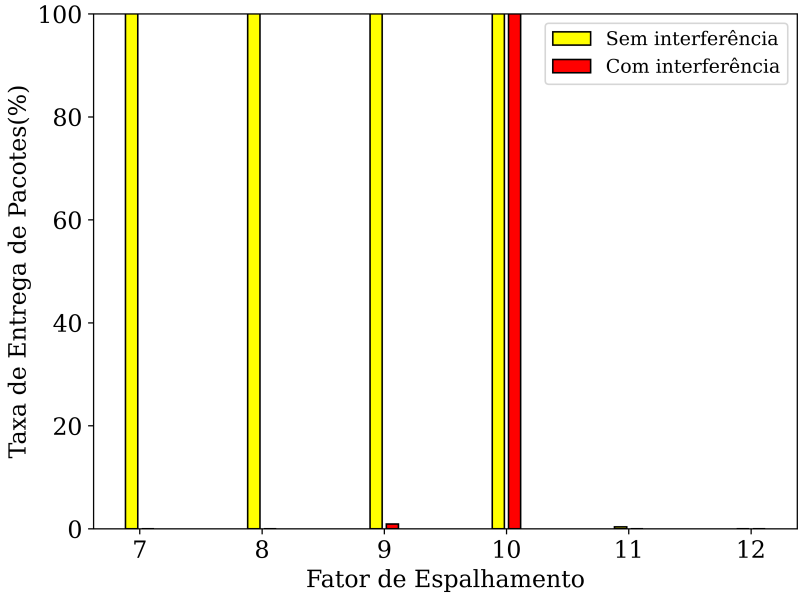


Figura 6.14: Taxa de entrega de pacotes para cada fator de espalhamento em um cenário com e sem interferência e com *gateway* a uma distância de 5 km.

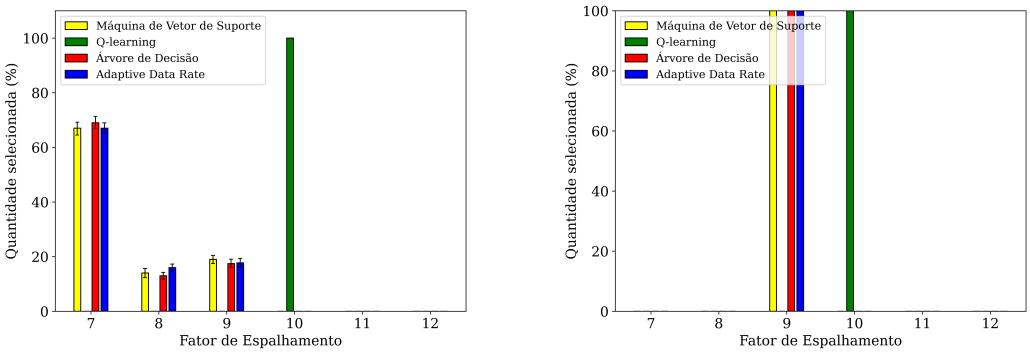
6.5.2 Experimento com a implementação das propostas

O objetivo desse experimento é verificar o desempenho do mecanismo ADR quando submetido a um ambiente de rede dinâmico. De acordo com SORNIN, N. et. al [41], o mecanismo ADR pode não ser eficaz quando a atenuação do canal de radio é rápida, frequente e constante, como ocorre neste experimento. A Tabela 6.3 contém o desempenho, em termos de taxa de entrega de pacotes, do mecanismo ADR e das propostas no ajuste dos fatores de espalhamento com a rede submetida à interferência. Verifica-se que, de fato, o desempenho do mecanismo ADR na seleção do fator de espalhamento é prejudicado quando em um ambiente submetido à interferência externa . O desempenho das propostas de aprendizado supervisionado obtiveram desempenho semelhante ao do ADR. O erro associado aos valores de taxa de entrega de pacotes obtidos pela técnica de Árvore de Decisão e Máquina de Vetor de Suporte são de 0,88% e 0.32% respectivamente, considerando um intervalo de confiança de 95%. Já o algoritmo *Q-learning* obteve desempenho superior ao ADR e às demais propostas com taxa de entrega de 100%, indicando ser a proposta adequada para dirimir o efeito da dinamicidade causada pela fonte de interferência.

Tabela 6.3: Taxa de entrega de pacotes obtida pelas propostas no cenário cuja rede é submetida à interferência externa.

Distância do Gateway	ADR	AD	SVM	Q-learning
5 km	96.5%	96.19%	96.18%	100%
10 km	87.4%	88.3%	85,14%	100%

A Figura 6.15 representa o percentual de seleção de cada fator de espalhamento, seleção esta realizada por cada proposta no experimento com interferência. As quantidades selecionadas foram obtidas através de simulação com os modelos de aprendizado já treinados. Para as distâncias de 5 km e 10 km do *gateway*, o mecanismo ADR, assim como as técnicas de aprendizado supervisionado, obtiveram desempenho inferior ao algoritmo *Q-learning*, devido à seleção apenas de fatores de espalhamento de 7 a 9, no caso da distância do gateway de 5 km, e à seleção apenas do fator de espalhamento 9, no caso da distância do gateway de 10 km, como pode ser visto na Figura 6.15, indicando a incapacidade destas propostas de encontrar a seleção ótima do parâmetro (fator de espalhamento 10 neste experimento) em um ambiente submetido a uma interferência externa. O fato de espalhamento 10 não é selecionado, pois ao identificar uma transmissão mal sucedida por interferência, o ADR e as propostas de aprendizado supervisionado realizam a retransmissão fazendo com que a transmissão se torne bem sucedida após a retransmissão, justificando a permanencia na seleção dos fatores de espalhamento entre 7 e 9. A proposta do algoritmo *Q-learning*, contudo, se mostra adequada para este fim, uma vez que consegue convergir para o parâmetro ótimo.



(a) Fatores de espalhamento selecionados com *gateway* a 5 km.

(b) Fatores de espalhamento selecionados com *gateway* a 10 km.

Figura 6.15: Fatores de espalhamento selecionados pelas propostas no estudo de caso 2.

Em relação ao algoritmo *Q-learning*, para a distância de 10 km do *gateway*, inicialmente foi selecionado o fator de espalhamento em 9 para todos os dispositivos,

uma vez que essa seleção seria realizada pelo mecanismo ADR caso não ocorresse a interferência externa. Verificou-se que o fator de espalhamento foi ajustado para 10 pelo algoritmo sem a necessidade de qualquer intervenção do operador, proporcionando uma taxa de entrega de pacotes de 100%. Já para a distância de 5 km do *gateway*, inicialmente foi selecionado o fator de espalhamento selecionado em 7 para todos os dispositivos, uma vez que essa seleção seria realizada pelo mecanismo ADR caso não ocorresse a interferência externa. Verificou-se também que o fator de espalhamento foi ajustado para 10 pelo algoritmo sem a necessidade de qualquer intervenção do operador, proporcionando uma taxa de entrega de pacotes de 100%.

Capítulo 7

Conclusões e Trabalho Futuros

Esta dissertação propôs a utilização de técnicas de aprendizado supervisionado e de aprendizado por reforço para selecionar o fator de espalhamento e potência de transmissão dos dispositivos finais de redes LoRa. O objetivo foi comparar o desempenho das propostas, com o desempenho do mecanismo ADR, em termos de taxa de entrega de pacotes e de energia consumida na transmissão.

Para avaliar o desempenho do mecanismo ADR e das propostas apresentadas, dois estudos de caso distintos foram identificados. O primeiro estudo de caso representou uma rede com o *gateway* localizado no centro da topologia e o segundo estudo de caso representou uma pequena rede submetida a uma fonte de interferência externa.

No primeiro estudo de caso, os resultados mostram que, para uma quantidade de até 100 dispositivos, o desempenho da técnica *Q-learning* supera o mecanismo ADR em termos de taxa de entrega de pacotes e de energia consumida na transmissão. Os resultados mostram também que as técnicas de aprendizado supervisionado, apesar de obterem uma taxa de entrega de pacotes superior ao do mecanismo ADR, estas técnicas consumiram mais energia do que o referido mecanismo. Porém, o tempo necessário para a convergência do algoritmo *Q-learning* pode representar uma desvantagem do algoritmo *Q-learning* em relação ao aprendizado supervisionado para aplicações cuja energia consumida pelos dispositivos finais não seja um fator crítico. As técnicas de aprendizado supervisionado, apesar de obter maior taxa de entrega de pacotes com o custo do aumento da energia consumida, o tempo necessário para o treinamento dos modelos é muito menor do que o treinamento do algoritmo *Q-learning*, representando uma vantagem para aplicações cujo consumo de energia não seja um fator crítico. No segundo estudo de caso, de acordo com os resultados, o desempenho do mecanismo ADR é comprometido, em termos de taxa de entrega de pacotes, quando a rede é submetida a uma fonte de interferência externa, indicando que o referido mecanismo pode não ser capaz de garantir a qualidade dos enlaces diante de um ambiente submetido a uma fonte de interferência. Os resultados mos-

tram também que as técnicas de aprendizado supervisionado obtiveram taxas de entrega de pacotes semelhantes ao ADR para a quantidade de até 100 dispositivos. Já o algoritmo *Q-learning* proposto mostrou-se capaz selecionar o fator de espalhamento ótimo para os dispositivos de forma a garantir uma taxa de entrega de 100%, indicando que pode ser uma alternativa para ajustar os parâmetros de transmissão em ambientes submetidos à inteferência externa. Para este estudo de caso, portanto, apenas a proposta de aprendizado por reforço mostrou ser uma alternativa ao mecanismo ADR para otimizar a taxa de entrega de pacotes.

Como trabalhos futuros, uma fonte de interferência externa pode ser incluída no primeiro estudo de caso com o intuito de avaliar o desempenho das propostas em uma rede densa, não só com interferências internas devido às transmissões concorrentes, mas também quando submetida a uma interferência externa. No segundo estudo de caso, o parâmetro de potência de transmissão pode ser considerado como parâmetro selecionável, assim como foi considerado no primeiro estudo de caso.

Referências Bibliográficas

[1] SAKSHAMA GHOSLYA. "LoRa: Symbol Generation". Disponível em: `https://www.sghoslya.com/p/lora-is-chirp-spread-spectrum.html`, 2015. Acessado em maio de 2022.

[2] SORNIN, N. AND YEGIN, A. "LoRaWAN 1.1 Specification, Technical Report". Disponível em: `https://lora-alliance.org/resource_hub/lorawan-specification-v1-1/`, 2019. Acessado em maio de 2022.

[3] RAZA, U., KULKARNI, P., SOORIYABANDARA, M. "Low power wide area networks: an overview", *IEEE Communications Surveys & Tutorials*, v. 19, n. 2, pp. 855–873, fev. 2017.

[4] GEORGIOU, O., RAZA, U. "Low power wide area network analysis: Can lora scale?" *IEEE Wireless Communications Letters*, v. 6, n. 2, pp. 162–165, fev. 2017.

[5] ABRAMSON, N. "THE ALOHA SYSTEM: Another Alternative for Computer Communications", *Fall Joint Computer Conference, AFIPS '70 (Fall), ACM*, v. 6, n. 2, pp. 281–285, nov. 1970.

[6] BOR, M. C., BOR, M. C., ROEDIG, U., et al. "Do LoRa low-power wide-area networks scale?" In: *19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, ser. MSWiM*, pp. 59–67, New York, USA, 2016. ACM.

[7] CROCE, D., ET AL. "Impact of spreading factor imperfect orthogonality in lora communications", *Communications in Computer and Information Science*, v. 6, n. 2, pp. 165–179, 2017.

[8] WARET, A., KANEKO, M., GUITTON, A., et al. "Lora throughput analysis with imperfect spreading factor orthogonality", *IEEE Wireless Comm. Letters*, pp. 1–1, 2018.

[9] CUOMO, F., CAMPO, M., CAPONI, A., et al. "Explora: Extending the performance of lora by suitable spreading factor allocations". In: *13th Inter-*

national Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Rome, Italy, 2017. IEEE.

[10] SLABICKI, M., PREMSANKAR, G., FRANCESCO, M. D. "Adaptive configuration of LoRa networks for dense IoT deployments". In: *Network Operations and Management Symposium*, pp. 1–9. IEEE/IFIP, 2018.

[11] AGGARWAL, S., NASIPURI, A. "Improving scalability of lorawan networks by spreading factor distribution". In: *SoutheastCon*, pp. 1–7. IEEE, 2021.

[12] CHINCHILLA-ROMERO, N., NAVARRO-ORTIZ, J., MUÑOZ, P., et al. "Collision avoidance resource allocation for LoRaWAN", *Sensors*, v. 21, n. 4, pp. 1218, 2021.

[13] ABDELFADEEL, K. Q., ZORBAS, D., CIONCA, V., et al. "FREE–Fine-Grained Scheduling for Reliable and Energy-Efficient Data Collection in LoRaWAN", *IEEE Internet of Things Journal*, v. 7, n. 1, pp. 669–683, 2019.

[14] BIANCHI, G., CUOMO, F., GARLISI, D., et al. "Capture Aware Sequential waterfilling for adaptive data rate allocation in LoRaWAN", *arXiv*, 2019.

[15] BOCHIE, K., GILBERT, M. S., GANTERT, L., et al. "A Survey on Deep Learning For Challenged Networks: Applications And Trends", *JOURNAL OF NETWORK AND COMPUTER APPLICATIONS*, v. 194, pp. 103213, 2021.

[16] BOUTABA, R., SALAHUDDIN, M. A., LIMAN, N., et al. "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities." *Journal of Internet Services and Applications*, v. 9, n. 1, pp. 16, dez. 2018.

[17] YATAGAN, T., OKTUG, S. "Smart spreading factor assignment for lorawans". In: *Symposium on Computers and Communications (ISCC)*, Barcelona, Spain, 2019. IEEE.

[18] PARK, G., LEE, W., JOE, I. "Network resource optimization with reinforcement learning for low power wide area networks." *EURASIP Journal onWireless Communications and Networking*, v. 176, 2020.

[19] ALSHEIKH, M. A., LIN, S., NIYATO, D., et al. "Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications." *IEEE Communications Surveys & Tutorials*, v. 16, n. 4, pp. 1996–2018, 2014.

[20] KOSUNALP, S., CHU, Y., MITCHELL, P. D., et al. "Use of Q-learning approaches for practical medium access control in wireless sensor networks." *Engineering Applications of Artificial Intelligence*, v. 55, pp. 146–154, 2016.

[21] LIU, Z., ELHANANY, I. "RL-MAC: A QoS-Aware Reinforcement Learning based MAC Protocol for Wireless Sensor Networks." In: *IEEE International Conference on Networking, Sensing and Control*, pp. 768–773. IEEE, 2006.

[22] NIU, J.; DENG, Z. "Distributed self-learning scheduling approach for wireless sensor network. Ad Hoc Networks." *Ad Hoc Networks*, v. 11, n. 4, pp. 1276–1286, 2013.

[23] AZARI, A.; CAVDAR, C. "Self-Organized Low-Power IoT Networks: A Distributed Learning Approach." In: *IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7. IEEE, 2018.

[24] TA, D.-T., KHAWAM, K., LAHOUD, S., et al. "Lora-mab: A flexible simulator for decentralized learning resource allocation in iot networks." In: *12th IFIP Wireless and Mobile Networking Conference (WMNC)*, Paris, France, 2019. IEEE.

[25] OLIVEIRA, F. T. X.; FILHO, H. B. "Spread Spectrum: O que é". Disponível em: `http://www.teleco.com.br/tutoriais/tutorialss/pagina_1.asp`, 2015. Acessado em maio de 2022.

[26] REYNDERS, B. E POLLIN, S. "Chirp spread spectrum as a modulation technique for long range communication." In: *IEEE Symposium on Communications and Vehicular Technologies (SCVT)*, pp. 1–5. IEEE, 2016.

[27] GOURSAUD, C. E GORCE, J.-M. "Dedicated networks for IoT: PHY/MAC state of the art and challenges." *EAI Endorsed Transactions on Internet of Things*, v. 1, n. 1, pp. 1–11, 2015.

[28] CENTENARO, M. E. A. "Long-range communications in unlicensed bands: the rising stars in the iot and smart city scenarios", *arXiv*, 2015.

[29] ADELANTADO, F., VILAJOSANA, X., TUSET-PEIRO, P., et al. "Understanding the Limits of LoRaWAN", *IEEE Communications Magazine*, v. 55, n. 9, pp. 34–40, 2017.

[30] "LoRa Alliance". Disponível em: `https://lora-alliance.org`, 2022. Acessado em maio de 2022.

[31] WANG, M., CUI, Y., WANG, X., et al. "Machine Learning for Networking: Workflow, Advances and Opportunities." *IEEE Network*, v. 32, n. 2, pp. 92–99, mar. 2018.

[32] HARTSHORN, S. *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners*. Seattle, USA, 2016.

[33] RAVICHANDIRAN, S. *Hands-On Reinforcement Learning with Python: Master reinforcement and deep reinforcement learning using OpenAI Gym and TensorFlow*. 1 ed. Birmingham, Reino Unido, Packt Publishing Ltd, 2018.

[34] WATKINS, C. J. C. H. *Learning from Delayed Rewards*. Cambridge, Reino Unido, Cambridge University, 1989.

[35] SUTTON, R. S., BARTO, A. G. *Reinforcement learning: An introduction*. Cambridge, Reino Unido, MIT press, 1998.

[36] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., et al. "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research*, v. 12, pp. 2825–2830, 2011.

[37] TIJMS, H. C. *A first course in stochastic models*. New York, USA, Wiley, 2003.

[38] ROSS, S. *Stochastic processes*. New York, USA, Wiley, 1996.

[39] SEMTECH CORPORATION WIRELESS SENSING AND TIMING PRODUCTS DIVISION. "LoRaTM Modulation Basics AN1200.22, Technical Report". Disponível em: `https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R0000001OJk/yDEcfAkD9qEz6oG3PJryoHKas3UMsMDa3TFqz1UQOkM`, 2015. Acessado em maio de 2022.

[40] EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE. "Radio Frequency (RF) system scenarios TR 136 942 V13.0.0, Technical Report". Disponível em: `https://www.etsi.org/deliver/etsi_tr/136900_136999/136942/13.00.00_60/tr_136942v130000p.pdf`, 2016. Acessado em maio de 2022.

[41] SORNIN, N. AND YEGIN, A. "LoRaWAN Specification v1.0.3 | LoRa Alliance". Disponível em: `https://lora-alliance.org/resource_hub/lorawan-specification-v1-0-3/`, 2018. Acessado em maio de 2022.