

Agrupamentos em Redes IIoT Baseados em Q-Learning Resilientes à Falhas de Quebra de Enlace

Carlos Pedroso¹, Vitor Fagundes², Aldri Santos^{1,2}

¹Núcleo de Redes Sem-Fio e Redes Avançadas (NR2) – Dinf – UFPR – Brasil

²Center for Computational Security sCience (CCSC) – DCC – UFMG – Brasil

capjunior@inf.ufpr.br, {vitorfagundes, aldri}@dcc.ufmg.br

Resumo. A adoção da Internet das Coisas Industrial (IIoT) é fundamental aos ambientes produtivos, exigindo redes de comunicação robustas e resilientes que garantam as suas operações. À medida que as falhas de enlace ameaçam a continuidade das operações industriais, elas comprometem a qualidade do serviço e paralisam processos críticos. Este trabalho propõe o SECTIONAL, uma abordagem baseada em Q-Learning para promover a resiliência das organizações topológicas de agrupamentos em redes IIoT. Um agente é treinado online em ambiente simulado para detectar falhas e tomar decisões corretivas de forma autônoma, reorganizando a estrutura da rede diante de alterações na conectividade. O SECTIONAL visa reduzir o impacto de falhas de comunicação, melhorar a continuidade dos serviços e demonstrar a viabilidade do uso de inteligência artificial para o gerenciamento adaptativo de redes industriais. O SECTIONAL foi avaliado no NS-3 gerenciando um ambiente IIoT e obteve uma eficiência na mitigação de falhas, alcançando taxas de realocação superiores a 96% em alguns casos e mantendo a rede operacional nos cenários avaliados.

Abstract. The adoption of the Industrial Internet of Things (IIoT) is fundamental to productive environments, requiring robust, resilient communication networks to ensure their operations. As link failures threaten the continuity of industrial operations, they compromise service quality and paralyze critical processes. This work proposes SECTIONAL, a Q-Learning-based approach to promote the resilience of topological clustering in IIoT networks. An agent is trained online in a simulated environment to detect failures and autonomously make corrective decisions, reorganizing the network structure in response to changes in connectivity. SECTIONAL aims to reduce the impact of communication failures, improve service continuity, and demonstrate the feasibility of using artificial intelligence for the adaptive management of industrial networks. SECTIONAL was evaluated on NS-3 in an IIoT environment and demonstrated efficiency in mitigating failures, achieving relocation rates exceeding 96% in some cases, and maintaining network operation across the evaluated scenarios.

1. Introdução

A Internet das Coisas Industrial (Industrial Internet of Things – IIoT) tem se consolidado como um dos principais pilares da automação e modernização dos ambientes produtivos, ao viabilizar a integração de sensores, atuadores e sistemas de controle em infraestruturas industriais complexas [Hu et al. 2024]. Essa integração tem impulsionado setores

industriais críticos, como fábricas de automóveis, plantas químicas e siderúrgicas, que demandam alta disponibilidade, reconfiguração dinâmica e resposta rápida a falhas. A IIoT favorece o monitoramento contínuo dos processos, a automação de tarefas e a tomada de decisão orientada por dados, contribuindo para o aumento da eficiência operacional e da confiabilidade dos sistemas. [Masum 2023]. Entretanto, a interconexão massiva de dispositivos heterogêneos impõe requisitos rigorosos à organização da rede, especialmente quanto à manutenção da conectividade, à disseminação contínua de dados e à resiliência da infraestrutura de comunicação [Sun et al. 2023].

Os ambientes IIoT se caracterizam por topologias densas e dinâmicas, compostas por dispositivos com diferentes capacidades de processamento, comunicação e armazenamento [Qiu et al. 2025]. Nesse contexto, a organização da rede por meio de agrupamentos possibilita lidar com a complexidade topológica, estruturando a comunicação de forma hierárquica, reduzindo a sobrecarga de mensagens e facilitando o gerenciamento distribuído dos dispositivos [Pedroso and Santos 2022]. Todavia, fatores como interferências eletromagnéticas, limitações físicas da infraestrutura e ataques cibernéticos tornam esses ambientes particularmente suscetíveis a falhas de enlace, afetando a comunicação entre os dispositivos agrupados [Kaur and Chanak 2022]. Tais falhas comprometem a troca de informações entre os dispositivos, degradam a qualidade do serviço e, em cenários críticos, paralisam operações industriais inteiras, resultando em prejuízos significativos [Gaba et al. 2025]. Diante desse cenário, mecanismos de resiliência tornam-se essenciais [Pedroso and Santos 2025]. Diferente da tolerância a falhas baseada em redundância, a resiliência permite detectar e recuperar-se adaptativamente de anomalias imprevistas, sendo essencial para sistemas ciber-físicos e sistemas críticos [Förster et al. 2026].

Diversas abordagens têm sido propostas para lidar com falhas de enlace em redes IoT [Nandyala et al. 2023], redes de sensores sem fio (WSN) [Yaraziz and Hill 2025] e redes sensíveis ao tempo (TSN) [Feng et al. 2022]. Em geral, esses trabalhos apresentam bons resultados em cenários com baixa densidade de dispositivos, mas enfrentam limitações em ambientes industriais complexos e densos nos quais a topologia da rede varia frequentemente e a demanda por conectividade é elevada. A maior parte das soluções existentes baseia-se em mecanismos de redundância, reconfiguração offline ou otimização topológica centralizada, o que limita a escalabilidade e a autonomia das redes [Zhang et al. 2024]. Além disso, poucos trabalhos exploram o emprego de agrupamentos como estrutura para organizar a rede e distribuir a tomada de decisão. Assim, pesquisas recentes têm explorado soluções baseadas em aprendizado por reforço, em particular Q-learning, que permite aos agentes aprenderem políticas ótimas por meio da interação com o ambiente, sem exigir um modelo explícito do sistema [Gu et al. 2024]. Estes trabalhos evidenciam o potencial de métodos baseados em Q-learning na reorganização resiliente de topologias e rotas em redes sem fio frente a falhas de conectividade, possibilitando adaptações dinâmicas com baixo custo computacional [Nandyala et al. 2023, Cardel et al. 2024].

Este trabalho propõe um sistema de gestão de configuração topológica contra falhas de enlace em agrupamentos de redes IIoT, chamado *SECTIONAL (reSiliEnt ClusTering In iiOt Networks bAsed q-Learning)*. Ele busca detectar e mitigar falhas de enlace em dispositivos dentro dos agrupamentos IIoT. O *SECTIONAL* emprega agrupamentos por similaridade de capacidade sensoriais para lidar com a massividade de dispositivos e um módulo de Q-learning para tomada de decisão sobre falhas. A avaliação do *SECTIONAL* no simulador NS-3 gerenciando um ambiente IIoT mostrou que o sistema

manteve taxas de realocação superiores a 96% na maioria dos cenários avaliados, mesmo sob diferentes níveis de falhas de enlace. Além disso, o SECTIONAL reduziu significativamente o percentual de dispositivos órfãos e preservou a disponibilidade dos agrupamentos formados, assegurando a continuidade das operações. Esses resultados devem-se ao mecanismo de reconfiguração autônoma orientado à qualidade de conectividade, que permite detectar falhas e realocar dinamicamente os dispositivos afetados.

O restante do artigo está organizado da seguinte forma: A Seção 2 revisa os trabalhos relacionados. A Seção 3 apresenta o sistema SECTIONAL. A Seção 4 descreve a avaliação e os resultados obtidos. A Seção 5 apresenta a conclusão e os trabalhos futuros.

2. Trabalhos Relacionados

Diversas abordagens têm sido propostas para lidar com falhas de enlace em redes IoT, redes de sensores sem fio (WSN) e redes sensíveis ao tempo (TSN). Em geral, esses trabalhos apresentam bons resultados em cenários controlados ou estáticos, mas enfrentam limitações em ambientes industriais dinâmicos, nos quais a topologia da rede varia frequentemente e a demanda por conectividade é elevada. Em [Abdulrab et al. 2022], um modelo de roteamento multipercorso tolerante a falhas para redes WirelessHART estabelece três caminhos distintos entre origem e destino, incluindo nós de backup para garantir a continuidade do serviço. Embora resultados indiquem melhorias em atraso, tempo de vida da rede e taxa de entrega de pacotes, o uso de múltiplos caminhos introduz sobrecarga de gerenciamento, especialmente em redes de grande escala, e a ausência de uma organização baseada em agrupamentos dificulta o gerenciamento distribuído. Em [Feng et al. 2022], uma abordagem de escalonamento tolerante a falhas é adaptada ao padrão IEEE 802.1Qbv TSN, reservando slots alternativos para fluxos Time-Triggered. Apesar de a solução melhorar a confiabilidade e a disponibilidade ao lidar com falhas transitórias e reduzir o desperdício de recursos, ela não considera reconfiguração em tempo de execução nem explora agrupamentos para organizar os dispositivos, o que resulta em retransmissões e aumento do tempo de resposta.

No contexto de aprendizado por reforço, em [Cardel et al. 2024] um modelo Q-learning controla o escalonamento em redes 6TiSCH/TSCH, adaptando dinamicamente o cronograma de comunicação às variações de tráfego. Embora a abordagem reduza a latência e mantenha a confiabilidade com baixo consumo energético, o desempenho depende fortemente da modelagem dos estados e do processo de aprendizagem, podendo introduzir sobrecarga, além de não considerar estruturas de agrupamento para distribuir o aprendizado. Em [Yaraziz and Hill 2025], uma abordagem baseada em Q-learning federado atua na alocação de recursos em ambientes IIoT, com foco em manutenção preditiva e eficiência energética. Apesar dos ganhos em consumo de energia e desempenho decisório, a solução não aborda reconfiguração topológica nem tolerância a falhas de comunicação. Em [Kaur and Chanak 2022], um esquema de roteamento tolerante a falhas para IIoT organiza a rede em agrupamentos e trata falhas de nós e enlaces. Embora apresente ganhos relevantes, a solução depende de decisões centralizadas no *sink*, limitando a autonomia local dos agrupamentos. De forma semelhante, em [Rawal and Chanak 2024], são combinados agrupamento, *sink* móvel e Q-learning para roteamento tolerante a falhas, obtendo melhorias expressivas em tempo de vida da rede e latência. Contudo, a complexidade do modelo e o foco no roteamento restringem a reorganização autônoma dos agrupamentos frente a falhas. Em [Nandyala et al. 2023], é proposto o protocolo QTAR, baseado em

Q-learning e sensível à topologia, que identifica implicitamente nós de corte e evita a formação de sub-redes isoladas. Apesar dos ganhos em consumo energético e latência, a abordagem não utiliza agrupamentos como abstração para simplificar a gestão topológica e a tomada de decisão distribuída.

3. SECTIONAL: Um sistema de Controle de Falhas de Enlace em Redes IIoT

Esta seção apresenta o sistema SECTIONAL (*reSiliEnt ClusTering In iiOt Networks bAsed q-Learning*) para o provimento da resiliência topológica em redes IIoT organizadas em agrupamentos. Ele atua na detecção e mitigação de falhas de enlace de maneira autônoma, reorganizando dinamicamente a estrutura dos agrupamentos para preservar a conectividade e a continuidade dos serviços da rede. O sistema organiza a rede em agrupamentos por similaridade e proximidade espacial dos dispositivos conforme Figura 1. O SECTIONAL concentra-se na manutenção da conectividade entre os agrupamentos e o *Access Point* (AP), empregando Aprendizado por Reforço baseado em Q-Learning para selecionar estratégias de reconfiguração local diante de falhas de enlace.

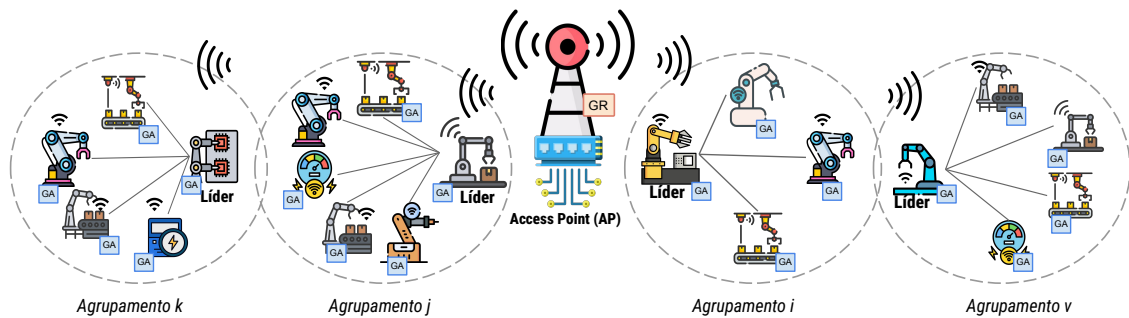


Figura 1. Exemplo de um ambiente de rede IIoT

Organização da Rede IIoT: consiste de uma rede implantada em uma área bidimensional (X_x, Y_y) , composta por um conjunto de dispositivos IIoT denotado por $D = \{d_1, d_2, \dots, d_n\}$. Cada dispositivo $d_i \in D$ possui um identificador único Id_i e diferencia-se pelos seus conjuntos de capacidades $C = \{c_1, c_2, \dots, c_m\}$, e por características de processamento e memória. Os dispositivos são estáticos e distribuídos aleatoriamente na área de cobertura da rede, com coordenadas $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Além disso, assume-se a ausência de restrições energéticas severas, permitindo a execução contínua de mecanismos de monitoramento e reconfiguração da conectividade.

Organização dos Agrupamentos: Para possibilitar o gerenciamento da conectividade e permitir intervenções localizadas, o SECTIONAL organiza a rede em agrupamentos. A formação dos agrupamentos baseia-se na similaridade de capacidade dos dispositivos e na proximidade espacial. Assim, cada agrupamento elege um dispositivo líder, responsável por coordenar a comunicação interna e manter o enlace direto com o AP. Essa organização hierárquica permite reconfigurações localizadas, reduzindo a complexidade do gerenciamento do desempenho e evitando impactos na topologia global.

Comunicação e Mensagens na Rede IIoT: A comunicação na rede IIoT ocorre por meio de um canal sem fio assíncrono e baseia-se na troca de mensagens de controle ao longo do ciclo de operação da rede, composta por cinco tipos de mensagens. As mensagens

de *CapabilityDissemination*, utilizadas na fase inicial para a formação dos agrupamentos; as mensagens *LeaderRegister* permitem aos líderes anunciar sua eleição ao AP; as mensagens *HeartbeatMsg* monitoram a integridade dos enlaces entre líderes e o AP; as mensagens *FailureNotification*, que acionam o processo de reconfiguração; as mensagens *Reconfigurationcontrol* conforme a ação selecionada pelo agente de aprendizado.

Modelo de Falha: consiste em falhas de enlace lógico nos enlaces de comunicação entre os dispositivos da rede IIoT, exceto no dispositivo AP. Elas comprometem a disponibilidade dos agrupamentos ao degradar ou interromper a rota principal de disseminação de dados e controle. Elas podem ser causadas por interferências no meio sem fio, degradação do canal, limitações de processamento ou sobrecarga de tráfego levando à indisponibilidade do líder.

3.1. Arquitetura

A estrutura do SECTIONAL compreende os módulos de **Gestão de Agrupamentos (GA)** e o **Gestão de Resiliência (GR)**, conforme ilustrado na Figura 2, que atuam de forma integrada para organizar estruturalmente a rede IIoT e assim permitir sua adaptação dinâmica diante de falhas de enlace. O módulo **GA** atua na formação e manutenção dos agrupamentos, assegurando uma organização hierárquica coerente e funcional da rede IIoT. Já o módulo **GR** realiza a estimativa do nível de resiliência e executa, de forma autônoma, ações corretivas sempre que falhas são identificadas na comunicação entre os dispositivos e o AP. Essa separação modular melhora a escalabilidade do sistema, reduz o acoplamento entre funções e facilita a evolução independente dos componentes. Essa abordagem localizada garante que apenas os agrupamentos particionados sejam reconfigurados, evitando perturbações desnecessárias na estrutura topológica global da rede.

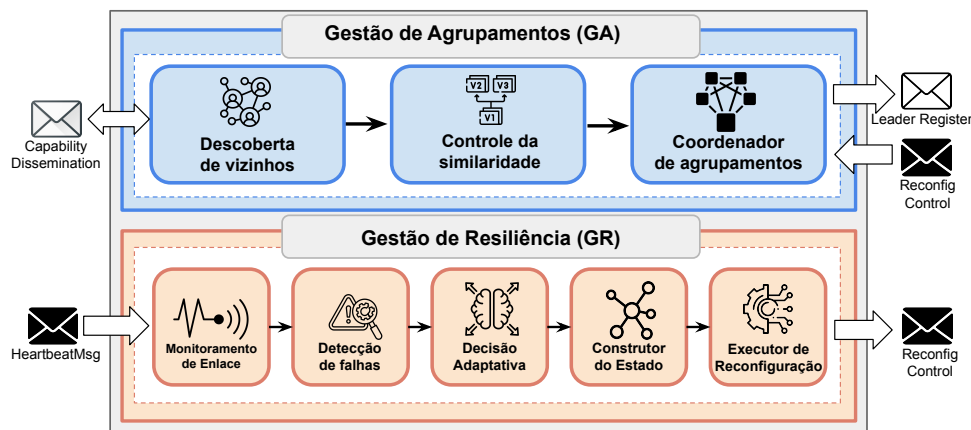


Figura 2. Arquitetura do SECTIONAL

O **Módulo de Gestão de Agrupamentos (GA)** baseia-se na similaridade de capacidades entre dispositivos próximos, definindo aqueles aptos a participarem do mesmo agrupamento. Para isso, o GA atua por meio de três componentes: *Descoberta de Vizinhos (DV)*, *Controle de Similaridade (CS)* e *Coordenador de Agrupamentos (CA)*. O DV dissemina e interpreta mensagens *CapabilityDissemination*, que contêm o identificador, as capacidades de sensoriamento e a posição do dispositivo. Assim, com base nessas informações, cada dispositivo constrói sua vizinhança local e mantém atualizadas as informações dos vizinhos detectados. O CS calcula a similaridade funcional entre dispositivos

vizinhos, com base na sobreposição de suas capacidades de sensoriamento. O CA utiliza essas informações para formar os agrupamentos, considerando a vizinhança e um limiar mínimo de similaridade, além de eleger líderes. O líder selecionado coordena a comunicação interna do agrupamento e mantém o enlace direto com o AP. Após a formação dos agrupamentos, os líderes registram-se no AP, garantindo a consistência da estrutura hierárquica e possibilitando correções localizadas sem afetar a rede IIoT globalmente.

O **Módulo de Gestão de Resiliência (GR)** é responsável por manter a conectividade dos agrupamentos diante de falhas de enlace. Ele atua sobre a estrutura definida pelo GA, usando informações de agrupamentos, líderes e topologia para monitorar a rede e coordenar ações corretivas de forma autônoma. O GR possui cinco componentes: *Monitoramento de Enlace (ME)*, *Detecção de Falhas (DF)*, *Construção do Estado (CE)*, *Decisão Adaptativa (DA)* e *Executor de Reconfiguração (ER)*. O ME monitora os enlaces líder-AP por meio da troca periódica de mensagens *HeartbeatMsg*. O DF identifica falhas a partir da ausência dessas mensagens em um intervalo predefinido e classifica os dispositivos do agrupamento afetado como órfãos. O CE constrói o estado atual da rede, incorporando a topologia vigente, os dispositivos órfãos e métricas locais relevantes. Com base nesse estado, o DA utiliza um agente baseado em Q-Learning para selecionar a estratégia de reconfiguração mais adequada. Por fim, o ER executa a ação selecionada, dissemina os comandos de reconfiguração e atualiza as estruturas de agrupamentos e liderança, viabilizando recuperação localizada com impacto mínimo sobre os demais agrupamentos.

3.1.1. Formação dos Agrupamentos

A heterogeneidade de capacidades em redes IIoT impõe desafios para o gerenciamento das interações e para a manutenção de conectividade estável. Desta maneira, o GA organiza a rede em agrupamentos com base em proximidade espacial e similaridade funcional, criando uma estrutura hierárquica que reduz a complexidade do gerenciamento e facilita intervenções localizadas. Cada dispositivo dissemina uma mensagem *CapabilityDissemination* contendo seu identificador, conjunto de capacidades e posição. A partir dessas mensagens, a vizinhança é definida pelo raio de comunicação r , isto é, apenas dispositivos dentro de alcance físico são considerados candidatos a formar o agrupamento.

Algorithm 1 GA: Formação de Agrupamentos e Eleição de Líderes

Require: $D; C(d); (x_d, y_d)$; raio r ; AP; λ ; θ
Ensure: \mathcal{C}, \mathcal{L}

- 1: $\mathcal{C} \leftarrow \emptyset; \mathcal{L} \leftarrow \emptyset$ ▷ inicialização
- 2: **for all** $d \in D$ **do**
- 3: BROADCAST($\text{Id}(d), C(d), (x_d, y_d)$) ▷ disseminação de capacidades
- 4: **end for**
- 5: WAITINTERVAL ▷ sincronização
- 6: **for all** $d \in D$ **do**
- 7: $Nbr(d) \leftarrow \{u \in D \mid \text{DIST}(d, u) \leq r\}$ ▷ descoberta de vizinhos
- 8: **end for**
- 9: $\mathcal{C} \leftarrow \text{CLUSTERBUILD}(D, Nbr(\cdot), C(\cdot), \theta)$ ▷ formação dos agrupamentos
- 10: **for all** cluster $c \in \mathcal{C}$ **do**
- 11: $L_c \leftarrow \text{SELECTLEADER}(c, \text{AP}, \lambda)$ ▷ eleição de líder
- 12: $\mathcal{L} \leftarrow \mathcal{L} \cup \{L_c\}$ ▷ registro local do líder
- 13: $M \leftarrow \text{MEMBERS}(c)$ ▷ membros do agrupamento
- 14: SENDMESSAGE TO AP($\text{Id}(L_c), \text{Id}(c), M$) ▷ LeaderRegister
- 15: **end for**
- 16: **return** \mathcal{C}, \mathcal{L} ▷ estrutura inicial da rede

O Algoritmo 1 descreve o processo de formação de agrupamentos e eleição de lí-

deres conduzido pelo GA. Cada dispositivo $d \in D$ dissemina suas informações locais por meio de mensagens *CapabilityDissemination* (l.2 – l.4) e aguarda um intervalo de sincronização (l.5). A partir das informações recebidas, cada dispositivo constrói sua vizinhança $Nbr(d)$ com base no raio de comunicação r (l.6 – l.8), restringindo o agrupamento a enlaces fisicamente viáveis. Com a vizinhança definida, a função *CLUSTERBUILD* constrói os agrupamentos a partir da similaridade de capacidades e proximidade, resultando no conjunto \mathcal{C} (l.9). Para cada agrupamento $c \in \mathcal{C}$, o GA seleciona o líder L_c por meio da função *SELECTLEADER*, que utiliza métricas locais e a distância ao AP ponderadas por λ (l.10). O líder eleito é inserido em \mathcal{L} e os membros do agrupamento são obtidos para compor a mensagem de registro (l.11 – l.12). Por fim, o líder registra-se no AP por meio de *LeaderRegister*, informando seu identificador, o identificador do agrupamento e a lista de membros (l.13 – l.14). Ao final, o GA mantém a estrutura hierárquica consolidada (l.15), permitindo uma organização eficiente para a disseminação de dados.

$$sim(d_1, d_2) = \frac{|C_{d1} \cap C_{d2}|}{\sqrt{|C_{d1}| \times |C_{d2}|}} \quad (1)$$

A similaridade entre dois dispositivos d_1 e d_2 é calculada a partir da sobreposição de suas capacidades de sensoriamento. Considera-se que o conjunto de capacidades pode ser representado como um vetor binário (presença/ausência), permitindo o uso de uma medida normalizada de similaridade. A Equação 1 define essa relação em $[0, 1]$, onde valores próximos de 1 indicam dispositivos funcionalmente similares, aptos a participarem do mesmo agrupamento. A restrição espacial é garantida separadamente pelo critério de vizinhança $Nbr(d)$ limitado pelo raio r .

3.1.2. Coordenação da Reconfiguração Resiliente

A detecção e correção de falhas de enlace e as ações de reconfiguração são conduzidas pelo Módulo de Gestão de Resiliência (GR), que opera sobre a infraestrutura hierárquica definida pelo GA. O GR mantém informações atualizadas sobre os agrupamentos ativos e seus respectivos líderes, monitorando periodicamente os enlaces entre os líderes e o AP por meio de mensagens *Heartbeat*. Esse monitoramento ocorre em intervalos regulares, definidos pelo período T_{hb} , cujo valor pode ser ajustado de acordo com as características da rede. A partir dessas informações, o GR busca preservar a conectividade dos agrupamentos e reduzir o tempo de recuperação diante de falhas, atuando de forma localizada.

O Algoritmo 2 descreve o ciclo de monitoramento, detecção de falhas e recuperação pelo GR. Enquanto a rede permanece ativa (l.1), o GR percorre cada líder $L_c \in \mathcal{L}$ (l.2) e realiza o monitoramento do enlace crítico líder–AP por meio de mensagens *Heartbeat* (l.3). A função *MISSEDHEARTBEATS* verifica a ausência dessas mensagens de acordo com os parâmetros k e Δ_{hb} (l.4), caracterizando uma falha no enlace (L_c, AP). Quando a falha é detectada, o procedimento de recuperação *RECOVER* é acionado (l.5). Após a verificação de todos os líderes, o GR aguarda o período de monitoramento T_{hb} antes de reiniciar o ciclo (l.8 – l.9). No procedimento *RECOVER*, o GR inicia o tratamento da falha ao enviar uma notificação *FailureNotification* associada ao enlace comprometido (l.10 – l.11). O agrupamento afetado c e o conjunto de nós órfãos O_t são identificados a partir das estruturas \mathcal{C} e do líder L_c (l.12). Com base nessas informações, o GR constrói o estado atual da rede e seleciona a ação de reconfiguração correspondente por meio da

função DECIDE (l.13). A ação selecionada é aplicada à topologia pela função RECONFIGURE, atualizando G de acordo com o contexto observado (l.14). Por fim, a função UPDATEQ (l.15) atualiza o mecanismo de aprendizado, encerrando o procedimento.

Algorithm 2 GR – Monitoramento, Detecção de Falhas e Reconfiguração

Require: $G, \mathcal{C}, \mathcal{L}, AP, T_{hb}, k, \Delta_{hb}$

Ensure: G updated

```

1: while active network do
2:   for all  $L_c \in \mathcal{L}$  do
3:      $Heartbeat(L_c, AP)$  ▷ monitoramento do enlace líder-AP
4:     if MISSEDHEARTBEATS( $L_c, k, \Delta_{hb}$ ) then
5:       RECOVER( $G, \mathcal{C}, \mathcal{L}, L_c, AP$ ) ▷ início da recuperação
6:     end if
7:   end for
8:   WAIT( $T_{hb}$ ) ▷ intervalo de monitoramento
9: end while
10: procedure RECOVER( $G, \mathcal{C}, \mathcal{L}, L_c, AP$ )
11:    $FailureNotification$  para ( $L_c, AP$ ) ▷ notificação de falha
12:    $(c, O_t) \leftarrow GETAFFECTED(\mathcal{C}, L_c)$  ▷ cluster afetado e nós órfãos
13:    $(S_t, A_t) \leftarrow DECIDE(G, \mathcal{C}, L_c, O_t)$  ▷ estado e ação
14:    $G \leftarrow RECONFIGURE(G, \mathcal{C}, \mathcal{L}, c, O_t, A_t)$  ▷ aplica reconfiguração
15:   UPDATEQ( $S_t, A_t, G$ ) ▷ aprendizado
16: end procedure

```

3.1.2.1 Modelagem do Agente de Aprendizado por Reforço

A tomada de decisão no SECTIONAL ocorre por um agente de Aprendizado por Reforço baseado em Q-Learning. O processo de aprendizagem ocorre por meio da tupla $\langle S, A, R \rangle$, que define o estado observado da rede, o conjunto de ações disponíveis e a função de recompensa utilizada para guiar o processo de adaptação. O estado da rede no instante t , denotado por S_t , é definido como: $S_t = (G_t, O_t, \text{deg}(\cdot), \text{dist}(\cdot, AP), C_{move}(\cdot))$. aem que G_t representa a topologia atual da rede, O_t corresponde ao conjunto de nós órfãos resultantes de uma falha de enlace, e as métricas $\text{deg}(\cdot)$, $\text{dist}(\cdot, AP)$ e $C_{move}(\cdot)$ são computadas para os nós candidatos envolvidos na reconfiguração (por exemplo, nós órfãos e potenciais nós de suporte), sendo agregadas no vetor de estado. Com base no estado observado, o agente seleciona uma ação A_t dentre um conjunto finito de estratégias de recuperação. No escopo deste trabalho, avaliação, o agente considera exclusivamente a ação *ReassignCluster*, responsável pela realocação de nós órfãos para líderes ativos. As demais ações representam extensões conceituais do modelo e não são exploradas nos resultados apresentados.

A função de recompensa R_t incentiva soluções rápidas e eficientes e penaliza reconfigurações custosas. Ela é calculada como: $R_t = \alpha \cdot \Delta QI - \beta \cdot C_{move} - \gamma \cdot T_{rec}$. Onde ΔQI representa a variação de um índice de qualidade de conectividade (por exemplo, fração de nós reconectados e estabilidade do enlace) antes e após a ação, C_{move} corresponde ao custo total associado à reconfiguração e T_{rec} é o tempo necessário para restabelecer a comunicação com o AP. Os parâmetros α , β e γ são pesos ajustáveis que controlam a influência relativa de cada termo. O Algoritmo 3 descreve o mecanismo de decisão do agente de aprendizado do SECTIONAL baseado em política ϵ -greedy e atualização Q-Learning. A partir do estado atual S_t , o agente seleciona a ação A_t combinando exploração e aproveitamento: com probabilidade ϵ , escolhe uma ação aleatória no conjunto A (l.1 – l.2), caso contrário, seleciona a ação com maior valor estimado na tabela Q para o estado S_t (l.3 – l.4). A ação selecionada é então executada no ambiente, permitindo observar o próximo estado S_{t+1} e a recompensa R_t (l.5). Com base nessa transição, a

função UPDATEQ atualiza a tabela Q , incorporando a taxa de aprendizado α e o fator de desconto γ (l.6). Ao final, o algoritmo retorna a tabela Q atualizada, que representa a política aprendida pelo (l.7).

Algorithm 3 Agente RL do SECTIONAL

Require: Table Q , rate α , discount γ , exploitation ϵ

Require: Current state S_t , action set A

Ensure: Implicit policy stored in Q

```

1: if RAND([0, 1]) <  $\epsilon$  then
2:    $A_t \leftarrow \text{RANDOMACTION}(A)$  ▷ exploração
3: else
4:    $A_t \leftarrow \text{GREEDYACTION}(Q, S_t, A)$  ▷ exploração gulosa
5: end if
6: Executar  $A_t$  e observar  $(S_{t+1}, R_t)$  ▷ interação com o ambiente
7: UPDATEQ( $Q, S_t, A_t, R_t, S_{t+1}, \alpha, \gamma$ ) ▷ aprendizado
8: return  $Q$ 

```

3.1.3. Funcionamento

Considere uma rede IIoT com $N = 200$ nós distribuídos em uma área de $200 \times 200\text{m}^2$. Esses nós possuem capacidades distintas e transmitem dados periodicamente ao AP por meio de um líder de agrupamento. O GA organiza a rede em três agrupamentos (AG_1 , AG_2 e AG_3), considerando a proximidade espacial e similaridade funcional. Em cada agrupamento, um líder é eleito segundo critérios de conectividade e distância ao AP. Na **Figura 3(A)**, ilustra-se uma falha no enlace (L_2, AP) decorrente de interferência ou degradação do canal. Inicialmente, o GR monitora os enlaces via *heartbeats* ($T_{hb} = 2\text{s}$); após a ausência de duas mensagens consecutivas, a falha é detectada e os nós de AG_2 tornam-se órfãos. O tempo T_{hb} pode ser definido conforme o tipo de rede.

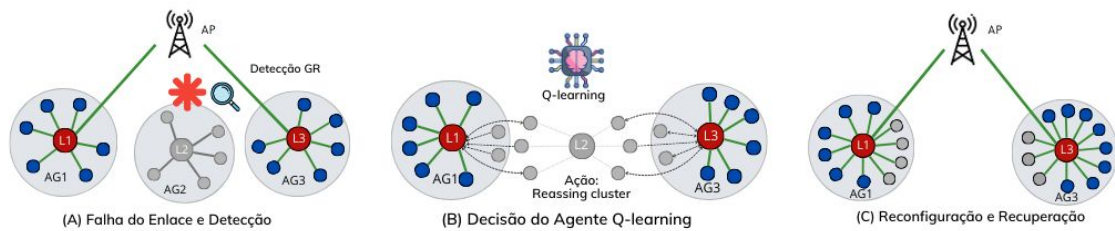


Figura 3. Exemplo de controle de falhas pelo SECTIONAL

Na **Figura 3(B)**, uma vez confirmada a falha, o agente constrói o estado S_t a partir da topologia vigente G_t , do conjunto de nós órfãos O_t e de métricas como conectividade, distância até líderes vizinhos e custo de realocação (C_{move}). Com base nesse estado, o agente seleciona a ação corretiva *ReassignCluster*, definindo líderes alternativos para os nós órfãos, priorizando a proximidade espacial e evitando sobrecarga estrutural. Na **Figura 3(C)**, os nós órfãos estabelecem novos enlaces com líderes e retomam o encaminhamento de dados ao AP. A reconfiguração ocorre de forma localizada, afetando apenas o subconjunto impactado pela falha, sem necessidade de reeleição de líderes ou reestruturação global da rede. Assim, o tempo de recuperação é $T_{rec} = 4\text{s}$, apresentando baixo custo operacional, impacto mínimo na comunicação e mantendo localizada a correção. Ao final o sistema prove a resiliência e estabilidade dos agrupamentos autonomamente.

4. Avaliação de Desempenho

Esta seção apresenta a avaliação de desempenho do sistema SECTIONAL¹. O sistema foi implementado no simulador NS-3, versão 3.36, em conjunto com a biblioteca ns3-gym, possibilitando a troca de informações entre o ambiente e o agente Q-learning e permitindo representar uma rede IIoT distribuída com características próximas ao real. O cenário avaliado corresponde a um ambiente de manufatura inteligente, com dispositivos IIoT integrados a robôs industriais, esteiras, prensas e sistemas de inspeção de qualidade para monitoramento de variáveis operacionais [Hu et al. 2024]. A escolha do conjunto de capacidades dos dispositivos baseia-se nas funções de sensoriamento que um dispositivo IIoT é habilitado a desempenhar, seguindo as definições de [Pedroso et al. 2020]. A rede IIoT consiste de $N = 200$ dispositivos (nós) distribuídos aleatoriamente em uma área de $200m \times 200m$, cuja densidade de dispositivos foi a partir baseada em [International Telecommunication Union 2017]. O tempo total de operação da rede compreende 900s, com (*Warm-up*) de 150s, para formação inicial dos agrupamentos. A comunicação entre os dispositivos emprega os protocolos IPv6 e UDP, estabelecendo uma rede *ad-hoc* no padrão IEEE 802.15.4. Além disso, na formação inicial dos agrupamentos, há um intervalo de 2ms entre as transmissões de mensagens para evitar a perda de pacotes e gargalos.

As falhas consideradas são do tipo enlace lógico, direcionadas aos enlaces entre os dispositivos líderes e o AP, localizado no centro da área. Essas falhas comprometem a disponibilidade do agrupamento ao interromper a rota principal de disseminação de dados e de controle, resultando de interferências no meio sem fio, degradação do canal ou sobrecarga de tráfego. A avaliação foi organizada em três cenários principais: **Cenário 1 (Baseline)**, em que o SECTIONAL opera sem falhas; **Cenário 2 (BASE+FALHA)**, que representa o sistema base submetido a falhas; e **Cenário 3 (SECTIONAL+FALHA)**, em que o SECTIONAL opera sob condições de falha. Nos cenários 2 e 3, foram consideradas variações na taxa de falhas de 10%, 20% e 30%, ocorrendo de forma controlada no instante $t = 310$ s. Essas taxas foram definidas a partir do trabalhos [Pedroso and Santos 2022] e buscam representar um comportamento mais próximo ao real. Além disso, avaliou-se um cenário dinâmico **SECTIONAL (FR)** em que as falhas ocorrem em instantes aleatórios, distribuídos entre 180 e 600s segundos, podendo comprometer entre 10% a 30% dos líderes. Essa configuração permite uma análise mais abrangente do comportamento do sistema sob condições de falha não determinísticas. A tomada de decisão é conduzida por um agente de Q-Learning, cujos hiperparâmetros foram definidos com base em [Kumar and Singh 2024], visando equilibrar aprendizado e estabilidade. Adotou-se uma taxa de aprendizado (α_{RL}) de 0,1 e um fator de desconto (γ_{RL}) de 0,9. A política de exploração ϵ -greedy foi configurada com $\epsilon = 0,1$, garantindo 10% de exploração e 90% de aproveitamento (*exploitation*) das políticas aprendidas.

Para avaliar a resiliência e a eficiência do SECTIONAL, foram empregadas dois tipos de métricas. As **métricas de disponibilidade** incluem o número de agrupamentos aptos, o número de agrupamentos ociosos, a porcentagem de nós órfãos e a porcentagem de nós realocados. As **métricas de resiliência** incluem: Tempo de Recuperação (T_{rec}), que mede o intervalo em segundos entre a detecção da falha e o restabelecimento completo da conectividade dos nós órfãos; Custo de Reconfiguração (C_{move}), calculado como a soma das distâncias euclidianas dos nós órfãos até seus novos líderes, $C_{move} = \sum_{d \in O_t} dist(d, novo_lder)$, onde valores menores indicam reconfigu-

¹Disponível em <https://github.com/vitor-fagundes/sectional>

rações mais eficientes; e Índice de Qualidade de Conectividade (ΔQI), que representa a variação da qualidade de conectividade após a reconfiguração, calculado como $\Delta QI = \frac{|O_{reconectados}|}{|O_t|} \times (1 - \frac{C_{move}}{C_{max}})$, onde C_{max} é a distância máxima possível na rede e valores próximos a 1 indicam alta qualidade. Os resultados apresentados correspondem à média de 35 simulações com intervalo de confiança de 95%.

4.1. Resultados

A Tabela 1 apresenta o desempenho do SECTIONAL em diferentes níveis de falhas de enlace, considerando o total de agrupamentos formados, aptos, ociosos, falhas, nós órfãos e nós realocados. Ressalta-se que a formação de agrupamentos considera que o número de membros é não determinístico. No cenário **BASE+FALHA**, o aumento da taxa de falhas resulta em um crescimento progressivo do percentual de nós órfãos, que evolui de 10% no cenário com 10% de falhas para 25% quando a taxa de falhas atinge 30%, evidenciando que a ausência de mecanismos eficientes de detecção e correção de falhas prolonga o isolamento dos nós afetados após falhas de enlace. Em contraste, no **SECTIONAL**, a detecção autônoma da quebra do enlace líder-AP possibilita a identificação imediata dos agrupamentos afetados e o acionamento rápido do processo de recuperação. Assim, o percentual de nós órfãos permanece controlado em todos os cenários avaliados, mantendo-se em 10% com 10% de falhas e reduzindo-se para 13% e 22% nos cenários com 20% e 30% de falhas, respectivamente, em função da realocação dinâmica dos nós órfãos.

Tabela 1. Comparativo de Disponibilidade: Agrupamentos, Falhas e Realocação

Cenários	Total Agrup.	Agrup. Aptos	Agrup. Ociosos	Falhas	Órfãos (%)	Realoc. (%)
Baseline (0%)	20 ± 1	13 ± 1	7 ± 1	0 ± 0	0 ± 0%	-
BASE+FALHA (10%)	20 ± 1	12 ± 1	8 ± 1	1 ± 0	10 ± 2%	-
SECTIONAL (10%)	19 ± 1	12 ± 1	7 ± 1	1 ± 0	10 ± 1%	94 ± 2%
BASE+FALHA (20%)	19 ± 1	13 ± 1	7 ± 1	2 ± 0	17 ± 2,4%	-
SECTIONAL (20%)	19 ± 1	12 ± 1	8 ± 1	2 ± 0	13 ± 2%	95 ± 1,8%
BASE+FALHA (30%)	19 ± 1	11 ± 1	8 ± 1	3 ± 0	25 ± 3%	-
SECTIONAL (30%)	20 ± 1	12 ± 1	8 ± 1	3 ± 0	22 ± 3%	95 ± 2,1%
SECTIONAL (FR)	20 ± 1	12 ± 1	8 ± 1	2,3 ± 0	15 ± 2%	96 ± 1,4%

A taxa de realocação no sistema foi superior a 94% em todos os níveis de falha, alcançando 96% no cenário **SECTIONAL (FR)**, evidenciando a efetividade da política de decisão baseada em Q-Learning. O agente de aprendizado prioriza ações que maximizam o ΔQI , direcionando o processo decisório para reconfigurações que preservam a conectividade com menor impacto estrutural. Logo, o sistema mantém um maior número de agrupamentos aptos, tratando as falhas de enlace como eventos transitórios e assegurando a rápida reintegração dos nós afetados e a continuidade das operações na rede IIoT.

Os gráficos da Figura 4 evidenciam a eficácia do Q-Learning empregado pelo **SECTIONAL** na promoção da resiliência topológica diante de falhas de enlace. A Figura 4(a) apresenta a taxa média de realocação alcançada pelo sistema, a qual se mantém acima de 94% em todos os cenários avaliados, considerando taxas de falha de 10%, 20% e 30%, tanto para falhas com tempo fixo quanto aleatório. Esse comportamento indica que a maioria dos nós órfãos é reassociada a novos líderes, resultado direto da política de decisão orientada pela função de recompensa. A Figura 4(b) ilustra a convergência do

processo de aprendizado por meio da evolução do valor de $Q(\text{Alocar})$ ao longo das rodadas. Em todos os cenários, observa-se convergência progressiva para valores próximos de 100, refletindo a maximização acumulada das recompensas associadas ao aumento da qualidade de conectividade (ΔQI). O cenário com 30% de falhas apresenta convergência mais rápida nas rodadas iniciais, em função da maior frequência de eventos de falha, que intensifica a atualização das recompensas e acelera o aprendizado. Nos cenários com 10% e 20% de falhas, a convergência ocorre de forma gradual, atingindo estabilidade por volta de 25 rodadas. Observa-se que esse processo é influenciado pela frequência de falhas, evidenciando sua sensibilidade à dinâmica do ambiente. Ainda assim, o cenário dinâmico **SECTIONAL(FR)** mantém comportamento consistente, preservando a robustez do aprendizado mesmo diante de variações temporais e na intensidade das falhas, o que contribui para a manutenção da resiliência da topologia da rede IIoT.

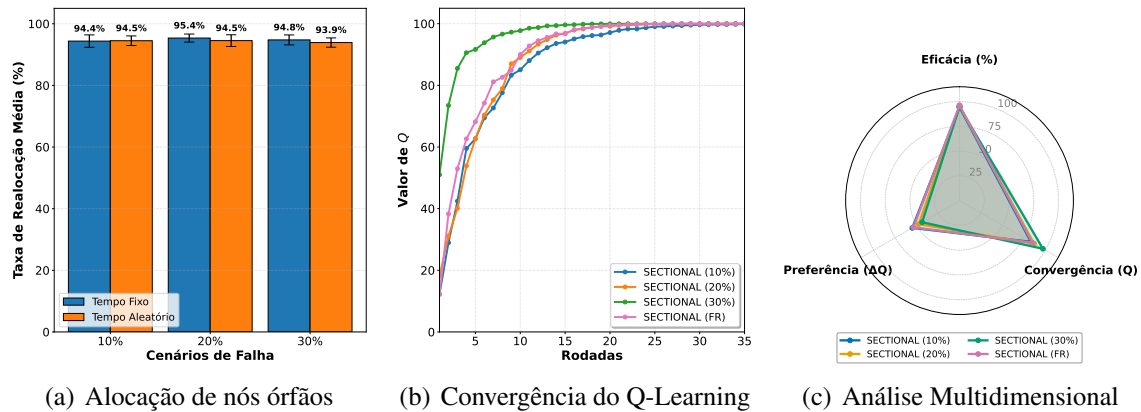


Figura 4. Eficácia na alocação de nós órfãos, Convergência do Q-learning e Análise da realocação com a convergência do aprendizado.

A Figura 4 (c) apresenta a análise multidimensional do sistema, ela correlaciona a eficácia da realocação com a convergência do aprendizado. Observa-se que cenários com maior convergência do valor de Q apresentam, simultaneamente, maiores taxas de realocação, o que evidencia que as decisões do agente estão diretamente associadas à preservação da conectividade funcional da rede. Esse comportamento confirma que o **SECTIONAL** adota ações que maximizam a qualidade de conectividade após falhas, reduzindo o impacto estrutural da reconfiguração ao mantê-la restrita ao contexto local. Os resultados destacam ainda a capacidade do sistema de lidar com falhas de enlace em eventos transitórios através do aprendizado contínuo, detecção autônoma e realocação eficiente dos nós órfãos, provendo a resiliência da topologia mesmo sob condições adversas.

A Figura 5(a) apresenta a velocidade de aprendizado do agente, medida pela variação absoluta do valor de $Q(\text{Alocar})$ entre rodadas. Nota-se que, nos estágios iniciais, todos os cenários apresentam variações elevadas de $|\Delta Q|$, o que reflete a fase exploratória do aprendizado. À medida que as rodadas avançam, há uma redução progressiva da variação do valor de Q , indicando a estabilização da política aprendida. Nos cenários com maiores taxas de falha, em especial com 30%, a redução de $|\Delta Q|$ ocorre mais rapidamente, devido à maior frequência de eventos de falha, que acelera o processo de ajuste da política. Após a convergência, as variações permanecem constantes caracterizando uma decisão estável, no qual o agente passa a reagir de forma mais previsível às falhas

de enlace. Esse comportamento contribui diretamente para a redução do tempo médio de recuperação (T_{rec}), uma vez que as decisões de realocação passam a ser tomadas com menor oscilação ao longo do tempo.

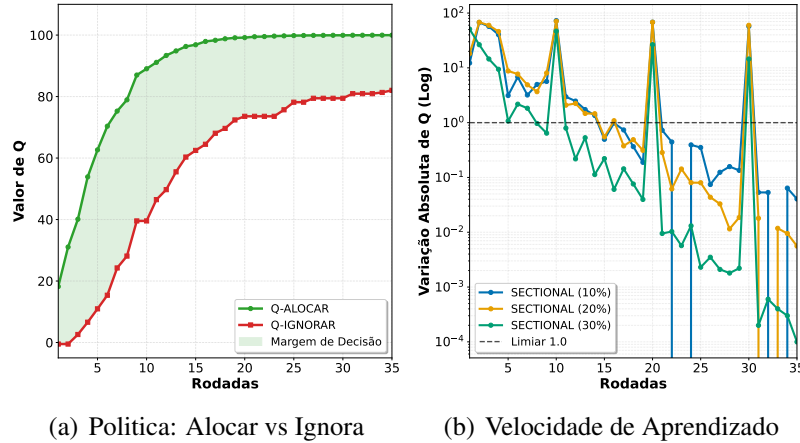


Figura 5. Velocidade do aprendizado e tomada de decisão do SECTIONAL

A Figura 5(b) mostra o aprendizado da política de decisão ao comparar os valores de $Q(Q(\text{Alocar}))$ e $Q(Q(\text{Ignorar}))$ ao longo das rodadas. Observa-se que o valor associado à ação de alocar cresce de forma consistente e se aproxima rapidamente do valor máximo, enquanto a ação de não alocar apresenta crescimento mais lento e permanece em um patamar inferior. A margem crescente entre as duas ações evidencia que o agente aprende a priorizar a realocação dos nós órfãos após a detecção de falhas, consolidando uma política orientada à preservação da conectividade. Esse comportamento explica as elevadas taxas de realocação observadas nos resultados e confirma que as decisões do agente estão alinhadas à maximização de (ΔQI) , evitando o isolamento progressivo dos dispositivos. Os resultados demonstram que a convergência do aprendizado não apenas estabiliza a política do agente, mas também sustenta decisões de realocação mais rápidas e consistentes, reforçando a capacidade do **SECTIONAL** de tratar falhas de enlace como eventos transitórios e de manter a resiliência topológica da rede IIoT.

5. Conclusão

Este trabalho apresentou o SECTIONAL, um sistema de controle resiliente de agrupamentos em redes IIoT baseado em Q-Learning, voltado à mitigação de falhas de enlace entre dispositivos. O sistema organiza a rede em agrupamentos hierárquicos baseado em similaridade de capacidades e aplica Q-Learning para tomada de decisão sobre reconfigurações topológicas de forma autônoma. Os resultados mostram que o SECTIONAL é capaz de reduzir o percentual de dispositivos órfãos e manter a conectividade mesmo em cenários críticos, realocando e preservando a disponibilidade da rede. Além disso, a convergência do aprendizado apoia decisões estáveis de realocação, contribuindo para a redução do tempo de recuperação e a continuidade das operações. Como trabalhos futuros, pretende-se ampliar a avaliação para outros tipos de falhas, considerar cenários com mobilidade e analisar o impacto em termos de consumo de recursos e escalabilidade.

Agradecimentos

Este trabalho foi financiado em parte pela CAPES - Código de Financiamento 001, Programa de Excelência Acadêmica (PROEX) e pelo CNPq nº 307752/2023-2.

Referências

- Abdulrab, H., Hussin, F. A., Abd Aziz, A., Awang, A., Ismail, I., and Devan, P. A. M. (2022). Reliable fault tolerant-based multipath routing model for industrial wireless control systems. *Applied Sciences*, 12(2):544.
- Cardel, V. S., Rettore, P. H., and Santos, B. P. (2024). Uma abordagem q-learning para escalonamento dinâmico de comunicação do tsch. In *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, páginas 113–126. SBC.
- Feng, Z., Deng, Q., Cai, M., and Li, J. (2022). Efficient reservation-based fault-tolerant scheduling for ieee 802.1 qbv time-sensitive networking. *Journal of Systems Architecture*, 123:102381.
- Förster, A., Hollick, M., Katzenbeisser, S., and Krauss, C. (2026). Resilience meets security in cyber-physical systems and critical infrastructures. *Commun. ACM*, 69(2):35–38.
- Gaba, G. S., Sari, A., Butun, I., Singh, P., Gurtov, A., and Liyanage, M. (2025). A survey on security and privacy of industry 4.0 and beyond: Technical aspects, use cases, challenges, and research directions. *IEEE Open Journal of the Communication Society*.
- Gu, S., Yang, L., Du, Y., Chen, G., Walter, F., Wang, J., and Knoll, A. (2024). A review of safe reinforcement learning: Methods, theories and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Hu, Y., Jia, Q., Yao, Y., Lee, Y., Lee, M., Wang, C., Zhou, X., Xie, R., and Yu, F. R. (2024). Industrial internet of things intelligence empowering smart manufacturing: A literature review. *IEEE Internet of Things Journal*, 11(11):19143–19167.
- International Telecommunication Union (2017). Minimum requirements related to technical performance for imt-2020 radio interface(s). Technical Report Report M.2410-0, ITU-R.
- Kaur, G. and Chanak, P. (2022). An intelligent fault tolerant data routing scheme for wireless sensor network-assisted industrial internet of things. *IEEE Transactions on Industrial Informatics*, 19(4):5543–5553.
- Kumar, A. and Singh, D. (2024). Adaptive epsilon greedy reinforcement learning method in securing iot devices in edge computing. *Discover Internet of Things*, 4(1):27.
- Masum, R. (2023). Cyber security in smart manufacturing (threats, landscapes challenges). *arXiv preprint arXiv:2304.10180*.
- Nandyala, C. S., Kim, H.-W., and Cho, H.-S. (2023). Qtar: A q-learning-based topology-aware routing protocol for underwater wireless sensor networks. *Computer Networks*, 222:109562.
- Pedroso, C., de Moraes, Y. U., Nogueira, M., and Santos, A. (2020). Managing consensus-based cooperative task allocation for iiot networks. In *2020 IEEE ISCC*, páginas 1–6. IEEE.
- Pedroso, C. and Santos, A. (2022). Dissemination control in dynamic data clustering for dense iiot against false data injection attack. *International Journal of Network Management*, 32(5):e2201.
- Pedroso, C. and Santos, A. (2025). A survey on faults and vulnerabilities in network topological connectivity: Logical and physical perspectives. *ACM Comput. Surv.*, 58(5).
- Qiu, F., Kumar, A., Hu, J., Sharma, P., Tang, Y. B., Xu Xiang, Y., and Hong, J. (2025). A review on integrating iot, iiot, and industry 4.0: A pathway to smart manufacturing and digital transformation. *IET Information Security*, 2025(1):9275962.
- Rawal, A. P. and Chanak, P. (2024). A q-learning-based fault-tolerance data routing scheme for iot-enabled wsns. *IEEE Internet of Things Journal*, 11(14):25283–25293.
- Sun, D., Hu, J., Wu, H., Wu, J., Yang, J., Sheng, Q. Z., and Dustdar, S. (2023). A comprehensive survey on collaborative data-access enablers in the iiot. *ACM Computing Surveys*, 56(2):1–37.
- Yaraziz, M. S. and Hill, R. (2025). Federated q-learning-based optimization for resource allocation in industrial iot networks. *IEEE Access*.
- Zhang, T., Xue, C., Wang, J., Yun, Z., Lin, N., and Han, S. (2024). A survey on industrial internet of things (iiot) testbeds for connectivity research.