



Um Novo Módulo Arquitetural de Controle Centralizado com DRL para Redes IoT Densas sem Bateria Baseadas em Retroespalhamento Ambiente

Edwardes A. Galhardo^{1,2}, Carlos B. Westphall², Antonio Oliveira Jr^{1,3}

¹ Instituto de Informática – Universidade Federal de Goiás (UFG)
Goiânia – Brasil

² Departamento de Informática e Estatística do CTC
– Universidade Federal de Santa Catarina (UFSC)
Florianópolis – Brasil

³ Fraunhofer Portugal AICOS
Porto – Portugal

edwardesamarogalhardo@inf.ufg.br,
antoniojr@ufg.br, carlosbwestphall@gmail.com

Abstract. *Dense IoT networks composed of batteryless devices based on AmBC face challenges related to interference, control, and communication stability. This work presents the GlobalAmBC-DRL Control Module, a centralized control architecture that leverages deep reinforcement learning over a global aggregated network state. The proposed approach enables the reduction of inter-device interference while improving communication stability and energy efficiency. The module is implemented as a reusable extension in OMNeT++ and evaluated under different network densities. Results indicate reduced variability in communication success rate and improved energy balance compared to static strategies, providing a reproducible architectural basis for sustainable IoT systems.*

Resumo. *Redes IoT densas compostas por dispositivos sem bateria, baseadas em AmBC, enfrentam desafios de interferência, controle e estabilidade. Este trabalho apresenta o GlobalAmBC-DRL Control Module, uma arquitetura de controle centralizada que utiliza aprendizado por reforço profundo sobre um estado global agregado da rede. A abordagem permite reduzir a interferência entre dispositivos e aumentar a estabilidade da comunicação e a eficiência energética. O módulo é implementado como extensão reutilizável em OMNeT++ e avaliado sob diferentes densidades de rede. Os resultados indicam menor variabilidade na taxa de sucesso e melhor equilíbrio energético em comparação com estratégias estáticas, fornecendo uma base arquitetural reprodutível para IoT sustentável.*

1. Introdução

Redes de Internet das Coisas (IoT) densas e de larga escala têm sido amplamente empregadas em aplicações que demandam monitoramento contínuo, baixo custo de manutenção e operação sustentável. Nesse contexto, dispositivos IoT sem bateria emergem como uma alternativa promissora para reduzir custos operacionais e impactos ambientais. Entre

as tecnologias habilitadoras, a Comunicação por Retroespalhamento Ambiente (*Ambient Backscatter Communication* – AmBC) permite a transmissão de dados por meio da reflexão de sinais eletromagnéticos presentes no ambiente, eliminando a necessidade de fontes internas de energia [Huynh et al. 2018b, Memon et al. 2019].

Apesar de seus benefícios, redes IoT densas baseadas em AmBC impõem desafios significativos em nível de sistema. As severas restrições energéticas e computacionais dos dispositivos inviabilizam a execução local de mecanismos complexos de controle, enquanto a alta densidade de nós intensifica problemas de interferência, instabilidade de comunicação e degradação do desempenho coletivo da rede. Esses fatores tornam necessária a adoção de estratégias de gerenciamento capazes de capturar o comportamento global do sistema de forma eficiente.

Nesse cenário, técnicas de Aprendizado por Reforço (RL) e Aprendizado por Reforço Profundo (DRL) têm sido exploradas para otimização adaptativa de redes sem fio e sistemas IoT. Trabalhos recentes demonstram que o uso de DRL em comunicações por retroespalhamento pode melhorar a eficiência energética, o acesso ao espectro e a estabilidade da comunicação, incluindo extensões para cenários mais complexos, como redes assistidas por superfícies inteligentes e arquiteturas emergentes de 6G [Khan et al. 2023, Shoeibi et al. 2024, Imran et al. 2025].

Entretanto, grande parte dessas abordagens concentra-se na proposição de modelos ou algoritmos específicos, frequentemente baseados em decisões locais, estados detalhados por nó ou capacidades computacionais incompatíveis com dispositivos AmBC sem bateria. Além disso, muitas soluções carecem de integração explícita ao nível arquitetural de sistemas de simulação de redes, dificultando a reutilização, a comparação entre abordagens e a reprodutibilidade experimental [Jamshed et al. 2025, Sooriarachchi et al. 2025].

Diante dessas limitações, este trabalho propõe o *GlobalAmBC-DRL Control Module*, um módulo arquitetural de controle centralizado para redes IoT densas e sem bateria baseadas em AmBC. A proposta integra aprendizado por reforço profundo ao nível de sistema, deslocando a inteligência dos dispositivos finais para um controlador lógico externo que opera sobre um estado global agregado da rede. Essa abordagem permite mitigar a interferência entre dispositivos e aumentar a estabilidade da comunicação e a eficiência energética, ao mesmo tempo em que reduz a complexidade do espaço de estados e favorece a escalabilidade do controle.

Embora a centralização proporcione maior capacidade de processamento e aplicação de políticas adaptativas, essa abordagem pode introduzir desafios relacionados à latência do controle e à dependência de um ponto lógico de coordenação, aspectos que são analisados experimentalmente e discutidos de forma crítica ao longo do trabalho.

A Figura 1 apresenta uma visão geral da arquitetura proposta, evidenciando a separação entre a rede AmBC, composta por dispositivos IoT sem bateria, e o controlador centralizado responsável pela agregação do estado global e pela aplicação de políticas de controle adaptativas. Diferentemente de abordagens existentes, o aprendizado por reforço profundo é empregado como elemento central da arquitetura do sistema, e não apenas como um mecanismo auxiliar, sendo concebido com foco em reutilização e reprodutibilidade.

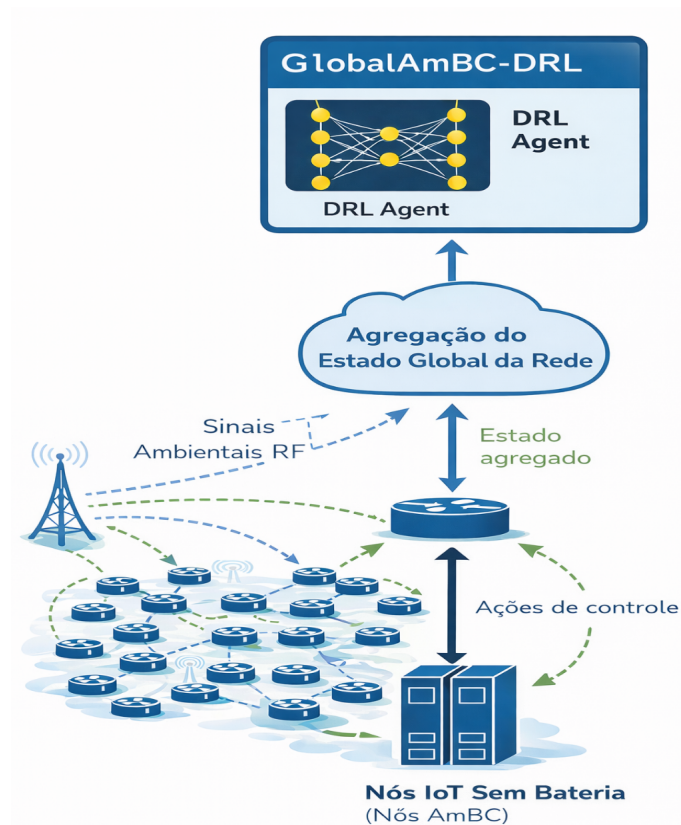


Figura 1. Visão geral da arquitetura proposta. Dispositivos IoT sem bateria operando com AmBC compõem a camada inferior, enquanto um controlador centralizado baseado em DRL agrega o estado global da rede e aplica ações de controle adaptativas em nível de sistema.

As principais contribuições deste trabalho são:

- a proposição de um módulo arquitetural de controle centralizado para redes IoT densas e sem bateria baseadas em AmBC;
- a integração de aprendizado por reforço profundo ao nível de sistema, permitindo controle adaptativo baseado em estado global agregado;
- a redução da complexidade computacional nos dispositivos finais por meio da externalização da inteligência de controle;
- a disponibilização de uma infraestrutura reutilizável e reprodutível para avaliação de políticas de controle em redes AmBC.

2. Trabalhos Relacionados

A Comunicação por Retroespalhamento Ambiente (AmBC) tem sido amplamente investigada como uma alternativa viável para a implementação de redes IoT sustentáveis e sem bateria. Trabalhos fundacionais e surveys consolidados apresentam os princípios do AmBC, bem como seus principais desafios físicos, de enlace e de escalabilidade em cenários densos [Huyh et al. 2018b, Memon et al. 2019]. Embora esses estudos demonstrem a viabilidade da operação energeticamente neutra, eles evidenciam que a gestão eficiente de redes compostas por um grande número de dispositivos permanece um desafio em nível de sistema.

Avanços mais recentes estendem o uso do AmBC para ambientes inteligentes e arquiteturas alinhadas a paradigmas 5G e 6G [Liu et al. 2021, Zheng et al. 2024]. Entretanto, essas contribuições concentram-se majoritariamente em aspectos de camada física ou modelagem analítica, assumindo configurações estáticas de rede e oferecendo suporte limitado a mecanismos de controle adaptativos em larga escala.

Em paralelo, técnicas de RL e DRL vêm sendo aplicadas à otimização de redes sem fio e sistemas IoT, incluindo alocação de recursos e gerenciamento energético. No contexto do AmBC, alguns trabalhos exploram o uso de RL para decisões dinâmicas em sistemas alimentados por energia ambiente [Huynh et al. 2018a, Perez-Ramirez et al. 2021]. Abordagens mais recentes demonstram o potencial do DRL em cenários mais complexos, como redes assistidas por superfícies inteligentes, comunicação bistática e integração com colheita de energia [Khan et al. 2023, Shoeibi et al. 2024, Zargari and Larsson 2024].

Apesar desses avanços, a maioria das propostas trata o aprendizado como um mecanismo pontual de otimização, frequentemente baseado em decisões locais ou modelos acoplados a cenários específicos. Essa limitação reduz a capacidade de capturar o comportamento coletivo da rede e compromete a escalabilidade em cenários densos. Além disso, poucos trabalhos integram explicitamente o aprendizado ao nível arquitetural de sistemas de simulação, dificultando a reutilização, a comparação entre abordagens e a reprodutibilidade experimental [Jamshed et al. 2025, Sooriarachchi et al. 2025].

A Tabela 1 apresenta uma síntese comparativa entre trabalhos representativos da literatura e a proposta deste artigo, destacando diferenças quanto ao nível de controle, escopo arquitetural, suporte à reprodutibilidade e adequação a redes IoT densas sem bateria.

Tabela 1. Comparação entre trabalhos relacionados e a proposta deste artigo

Trabalho	Controle	DRL	Integração	Reprodutibilidade
Huynh et al. (2018)	Local	Sim	Não	Limitada
Memon et al. (2019)	Analítico	Não	Não	Não
Khan et al. (2023)	Local/híbrido	Sim	Parcial	Limitada
Shoeibi et al. (2024)	Local	Sim	Não	Limitada
Este trabalho	Global	Sim	Nível arquitetural	Alta

Diferentemente desses trabalhos, este artigo propõe um módulo arquitetural de controle centralizado que integra DRL ao nível de sistema em redes IoT densas e sem bateria baseadas em AmBC. Ao empregar um estado global agregado como base para a tomada de decisão, a proposta permite capturar o comportamento coletivo da rede, reduz a complexidade nos dispositivos finais e favorece a escalabilidade do controle.

Embora a centralização introduza desafios relacionados à latência do controle e à dependência de um ponto lógico de coordenação, essa abordagem possibilita a aplicação de políticas adaptativas mais robustas em comparação com estratégias puramente locais, constituindo uma alternativa promissora para o gerenciamento de redes AmBC densas.

3. Módulo GlobalAmBC-DRL de Controle Centralizado

Esta seção apresenta o *GlobalAmBC-DRL Control Module*, um módulo de controle centralizado para redes IoT densas e sem bateria baseadas em comunicação por AmBC. O

módulo integra simulação em nível de sistema com decisões baseadas em DRL, sendo projetado como componente reutilizável e independente de topologia.

Diferentemente de abordagens centradas apenas em algoritmos, a proposta introduz um elemento arquitetônico capaz de coordenar decisões globais sem impor sobrecarga computacional aos dispositivos finais, preservando a operação energeticamente neutra.

3.1. Fundamentação do Projeto

As limitações energéticas e computacionais de dispositivos AmBC inviabilizam a execução local de algoritmos de otimização. Além disso, a energia capturada do ambiente é inerentemente variável, o que reforça a necessidade de mecanismos de controle mais robustos.

Nesse contexto, adota-se um modelo de controle centralizado, no qual a inteligência é deslocada para um controlador externo. Essa abordagem reduz a complexidade dos dispositivos e favorece a estabilidade do sistema. Por outro lado, introduz desafios como dependência de um ponto único de controle e possíveis atrasos na aplicação das decisões, aspectos considerados na análise experimental.

3.2. Visão Geral da Arquitetura

O módulo é composto por três blocos principais:

- **Agregação de Estado:** coleta e compactação de métricas globais;
- **Interface de Decisão:** integração entre simulador e agente DRL;
- **Aplicação de Controle:** execução de ações globais na rede.

Essa estrutura modular reduz o acoplamento entre componentes e facilita a reutilização e a reprodutibilidade experimental.

3.3. Modelagem do Estado Global

O estado da rede é representado por um vetor agregado, evitando a observação individual dos nós e garantindo escalabilidade. No instante t :

$$\mathbf{s}_t = [\bar{E}_t, \bar{R}_t, \bar{I}_t, D_t, A_t], \quad (1)$$

onde \bar{E}_t é a energia média disponível, \bar{R}_t a taxa de sucesso, \bar{I}_t o nível de interferência, D_t a densidade de nós ativos e A_t a taxa de ativação.

Essa abstração captura o comportamento coletivo da rede com baixa complexidade.

3.4. Espaço de Ações

O agente atua sobre parâmetros globais, evitando controle individual. O vetor de ações é dado por:

$$\mathbf{a}_t = [P_t, \delta_t], \quad (2)$$

onde P_t define o nível de excitação do ambiente e δ_t regula a ativação dos dispositivos. A política busca maximizar o desempenho da rede respeitando restrições energéticas e a dinâmica do ambiente.

3.5. Reutilização e Consistência

O módulo é independente de topologia e parametrizado via simulador, permitindo sua aplicação em diferentes cenários. A interface de decisão é desacoplada do algoritmo, viabilizando a substituição do DDPG por outras técnicas de DRL sem alterações estruturais, o que favorece extensões e reprodutibilidade.

4. Integração com o OMNeT++ e Fluxo de Execução

Esta seção descreve a integração do *GlobalAmBC-DRL Control Module* ao simulador OMNeT++ e o fluxo de execução adotado durante a simulação da rede AmBC. O objetivo é evidenciar como o módulo opera em nível de sistema, coordenando a coleta de métricas, a tomada de decisão baseada em DRL e a aplicação de ações globais de controle, sem violar as restrições energéticas dos dispositivos sem bateria.

Diferentemente de abordagens que utilizam aprendizado por reforço de forma pontual ou acoplada a decisões locais, a proposta integra explicitamente o processo de decisão inteligente ao ciclo de simulação da rede, permitindo a avaliação sistêmica do impacto das políticas de controle ao longo do tempo.

4.1. Posicionamento do Módulo na Arquitetura de Simulação

O *GlobalAmBC-DRL Control Module* é implementado como um componente lógico independente no ambiente OMNeT++, posicionado acima da camada de comunicação da rede AmBC. Esse posicionamento permite a observação de métricas globais da rede sem interferir diretamente nos protocolos dos nós individuais.

Enquanto trabalhos existentes integram aprendizado por reforço por meio de ajustes pontuais em parâmetros específicos, a proposta introduz um componente arquitetural explícito responsável por orquestrar decisões em nível de sistema, mantendo o desacoplamento entre a lógica de simulação e a lógica de decisão inteligente.

A Figura 2 ilustra o fluxo temporal de execução do módulo durante a simulação. Em cada instante de tempo t , o simulador fornece métricas locais da rede, que são agregadas em um estado global. Esse estado é processado pelo agente DRL, que retorna ações de controle aplicadas nos ciclos subsequentes ($t + 1, t + 2, \dots$), influenciando o comportamento da rede ao longo do tempo. Esse encadeamento temporal evidencia que existe um intervalo entre a observação do estado e a aplicação efetiva da ação, o que pode introduzir atrasos de controle em cenários de maior escala.

4.2. Fluxo de Execução da Simulação

Durante a simulação, o módulo opera de forma cíclica e sincronizada com o tempo do simulador. Em cada intervalo de decisão, o seguinte fluxo é executado:

1. coleta de métricas globais da rede AmBC;
2. agregação das métricas para construção do estado global;
3. envio do estado ao agente de aprendizado;
4. recebimento da ação global;
5. aplicação da ação nos parâmetros operacionais da rede.

Esse fluxo garante que as decisões de controle sejam tomadas com base em uma visão sistêmica da rede, em contraste com abordagens baseadas apenas em informações



Figura 2. Fluxo temporal de execução do módulo GlobalAmBC-DRL, destacando a interação entre o OMNeT++, o agregador de estado global e o agente DRL ao longo dos ciclos de simulação.

locais. Além disso, a natureza cíclica do processo permite avaliar o impacto acumulado das decisões ao longo do tempo, inclusive em cenários com variações na densidade de dispositivos ou dinâmica de ativação.

4.3. Integração com o Agente de Aprendizado por Reforço

A integração entre o OMNeT++ e o agente de aprendizado é realizada por meio de uma interface de decisão desacoplada, responsável pela troca de estados e ações. Essa interface permite a execução do agente externamente ao simulador, preservando a fidelidade da simulação em nível de rede.

Ao deslocar o processo de aprendizado para fora do simulador, a proposta viabiliza o uso de algoritmos mais sofisticados e facilita a substituição do agente, reforçando o caráter modular e reutilizável do módulo.

4.4. Aplicação das Ações de Controle

As ações produzidas pelo agente são aplicadas de forma global sobre a rede AmBC, ajustando parâmetros operacionais que influenciam o comportamento coletivo do sistema. Essa abordagem é mais estável e escalável em cenários de alta densidade, quando comparada a intervenções em nós individuais, além de reduzir a complexidade associada ao controle distribuído.

A Tabela 2 evidencia que o controle exercido pelo *GlobalAmBC-DRL Control Module* é realizado por meio de um conjunto reduzido de parâmetros globais. O parâmetro P_t representa o nível efetivo de potência associado à excitação ambiente, enquanto δ_t define políticas globais de ativação dos dispositivos. A atuação sobre variáveis agregadas mantém o alinhamento com as restrições energéticas dos dispositivos sem bateria e favorece a escalabilidade do sistema.

Tabela 2. Parâmetros globais de controle aplicados pelo GlobalAmBC-DRL Control Module

Parâmetro	Descrição
P_t	Nível global de potência associado à excitação ambiente
δ_t	Política global de ativação dos nós AmBC

4.5. Reprodutibilidade e Discussão em Relação à Literatura

A execução da simulação é inteiramente controlada por arquivos de configuração, permitindo a reprodução consistente dos experimentos em diferentes cenários. Parâmetros como densidade da rede, intervalos de decisão e métricas observadas são definidos de forma explícita.

Enquanto abordagens existentes utilizam aprendizado por reforço como um mecanismo de otimização pontual, o presente trabalho emprega o DRL como um elemento arquitetural central, integrado ao ciclo de execução da simulação. Essa concepção estabelece uma base sólida para extensões futuras, incluindo cenários dinâmicos com entrada e saída de dispositivos e possíveis variações na infraestrutura de controle.

5. Avaliação Experimental

Esta seção apresenta a avaliação experimental do *GlobalAmBC-DRL Control Module*, com o objetivo de demonstrar sua viabilidade prática, seu impacto no desempenho global da rede e sua adequação como componente arquitetural reutilizável. Diferentemente de trabalhos que buscam exclusivamente maximizar métricas isoladas, a avaliação conduzida neste estudo foca em validar o funcionamento do módulo em nível de sistema, evidenciando os benefícios do controle centralizado baseado em estado global agregado.

5.1. Ambiente de Simulação

Os experimentos são conduzidos utilizando o simulador OMNeT++, estendido com o framework AmBC desenvolvido neste trabalho. A rede simulada é composta por dispositivos IoT sem bateria que operam exclusivamente por retroespalhamento ambiente, compartilhando um meio sem fio comum e uma fonte de excitação ambiente.

O *GlobalAmBC-DRL Control Module* é integrado ao simulador como um componente lógico independente, responsável por coletar métricas globais da rede, construir o vetor de estado agregado e aplicar ações globais de controle ao longo do tempo de simulação.

A Tabela 3 resume os principais parâmetros utilizados nos cenários avaliados.

Tabela 3. Parâmetros principais do ambiente de simulação

Parâmetro	Valor
Nº de nós AmBC	Variável (cenários densos)
Topologia	Distribuição aleatória
Modelo energético	Operação energeticamente neutra
Fonte de excitação	Ambiente compartilhado
Intervalo de decisão	Definido por configuração
Duração da simulação	Tempo fixo por cenário

5.2. Dados Utilizados na Avaliação

A avaliação experimental utiliza dados gerados a partir da execução das simulações no OMNeT++, não dependendo de conjuntos de dados externos pré-existentes. Essa escolha segue práticas amplamente adotadas em estudos de redes IoT e retroespalhamento ambiente, nos quais o comportamento do sistema é avaliado por meio de métricas coletadas

dinamicamente durante a simulação, permitindo a análise controlada do impacto de diferentes políticas de controle e condições de rede [Liu et al. 2013, Xie et al. 2021, Zhao et al. 2022].

Os dados considerados incluem:

- energia média disponível na rede ao longo do tempo;
- taxa média de sucesso das transmissões;
- nível médio de interferência observado;
- densidade de nós ativos em cada intervalo de decisão.

Essas métricas são agregadas pelo módulo proposto e utilizadas tanto para alimentar o agente de aprendizado quanto para compor a análise quantitativa dos resultados. Adicionalmente, o processo de decisão considera implicitamente o intervalo entre a observação do estado global e a aplicação das ações, permitindo avaliar o impacto temporal do controle e os efeitos de atraso no ciclo de decisão sobre a dinâmica da rede.

Diferentemente de trabalhos que utilizam modelos analíticos ou dados sintéticos simplificados, esta abordagem permite avaliar diretamente o comportamento dinâmico do sistema sob diferentes políticas de controle.

5.3. Processo de Criação do Módulo

O *GlobalAmBC-DRL Control Module* é desenvolvido como uma extensão modular do framework AmBC no OMNeT++, seguindo princípios de baixo acoplamento e alta coesão. O processo de criação do módulo envolve três etapas principais:

1. Definição das interfaces de coleta de métricas globais da rede, garantindo independência em relação à topologia e ao número de nós;
2. Implementação do mecanismo de agregação de estado, responsável por construir o vetor de estado global a partir das métricas observadas;
3. Implementação da interface de decisão e da camada de aplicação de controle, permitindo a integração com um agente DRL externo e a aplicação consistente das ações globais.

Esse processo difere de abordagens existentes que frequentemente introduzem modificações pontuais em módulos específicos do simulador, sendo aqui concebido como um componente arquitetural autônomo e reutilizável.

O treinamento do agente considera hiperparâmetros padrão do DDPG, cuja escolha impacta diretamente a estabilidade e o tempo de convergência do processo de aprendizado. Esses parâmetros foram definidos empiricamente, buscando um equilíbrio entre estabilidade do treinamento e capacidade de adaptação do agente às dinâmicas da rede.

5.4. Cenários Avaliados

Para validar o comportamento do módulo, são considerados diferentes cenários de densidade de rede, variando o número de nós AmBC ativos ao longo das simulações. Em todos os cenários, o desempenho do sistema é comparado sob três estratégias de controle:

- **Sem controle:** operação da rede sem aplicação de políticas globais;
- **Controle estático:** utilização de parâmetros globais fixos;
- **GlobalAmBC-DRL:** controle adaptativo baseado em DRL.

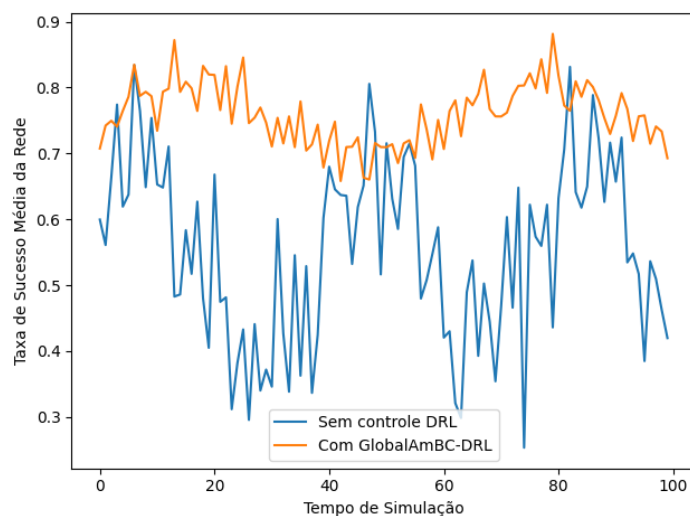


Figura 3. Comportamento temporal da taxa de sucesso média da rede, comparando cenários sem controle e com o módulo GlobalAmBC-DRL.

Essa comparação permite isolar o impacto do módulo proposto, evidenciando os ganhos proporcionados pela tomada de decisão dinâmica, inclusive em cenários com variações na densidade e ativação dos dispositivos.

A Figura 3 apresenta o comportamento temporal da taxa de sucesso média da rede ao longo da simulação. Observa-se que, na ausência de controle adaptativo, o desempenho apresenta elevada variabilidade, refletindo a instabilidade típica de redes densas. Em contraste, o uso do módulo proposto resulta em comportamento mais estável e previsível, com menor flutuação ao longo do tempo.

5.5. Resultados Quantitativos

Os resultados indicam que o *GlobalAmBC-DRL Control Module* melhora de forma consistente a estabilidade da comunicação em redes densas sem bateria, evidenciando ganhos de robustez e estabilidade em nível sistêmico. Em comparação com as abordagens sem controle e com controle estático, o módulo mantém taxas médias de sucesso mais elevadas e reduz flutuações associadas à interferência e à sobrecarga da rede.

Observa-se ainda que o comportamento do agente DRL tende à estabilização ao longo dos ciclos de simulação, refletindo a convergência da política aprendida para configurações mais eficientes de controle global.

Além disso, o controle centralizado contribui para a preservação da operação energeticamente neutra dos dispositivos, embora variações na disponibilidade de energia ambiente possam impactar o desempenho em cenários mais adversos.

Tabela 4. Comparação qualitativa do comportamento energético médio

Estratégia	Estabilidade Energética	Variabilidade	Uso de Potência
Sem controle adaptativo	Baixa	Alta	Não coordenado
Potência fixa (15 mW)	Alta	Baixa	Elevado
GlobalAmBC-DRL	Alta	Baixa	Adaptativo

A Tabela 4 resume o comportamento energético médio observado. Enquanto estratégias sem controle apresentam alta variabilidade, o uso de potência fixa implica maior consumo. O módulo proposto equilibra estabilidade e adaptação, regulando o uso de potência em nível sistêmico.

Esses resultados reforçam que decisões globais baseadas em estado agregado são adequadas para cenários de alta densidade, especialmente quando combinadas com mecanismos adaptativos.

5.6. Reprodutibilidade e Disponibilização do Código

Com o objetivo de garantir transparência e reprodutibilidade, todo o código-fonte do framework AmBC, bem como do módulo *GlobalAmBC-DRL*, foi disponibilizado em repositório público.

O repositório contém:

- implementação completa do módulo;
- cenários e arquivos de configuração do OMNeT++;
- scripts de coleta e análise de métricas;
- documentação para reprodução dos experimentos.

O código pode ser acessado em: <https://github.com/LABORA-INF-UFG/GlobalAmBC-DRL-SBRC2026>.

A disponibilização pública diferencia este trabalho de abordagens que apresentam apenas resultados numéricos, permitindo validação, reprodução e extensão por outros pesquisadores.

Esses resultados fornecem evidências empíricas que sustentam a análise apresentada na Seção 6, na qual são discutidas as implicações arquiteturais e limitações da abordagem proposta.

6. Discussão

Os resultados obtidos evidenciam que o controle centralizado baseado em aprendizado por reforço profundo contribui para a redução da interferência entre dispositivos e, consequentemente, para o aumento da estabilidade da comunicação e da eficiência energética em redes IoT densas e sem bateria baseadas em retroespalhamento ambiente. Ao deslocar a inteligência para um controlador externo, a arquitetura reduz a complexidade dos nós AmBC, preservando sua operação energeticamente neutra mesmo em cenários de alta densidade.

O uso de um estado global agregado permite capturar o comportamento coletivo da rede, favorecendo a escalabilidade e reduzindo o custo computacional. Essa visão sistêmica possibilita mitigar efeitos como interferência acumulada e degradação progressiva do desempenho, comuns em abordagens baseadas em decisões locais, além de permitir uma adaptação mais consistente a variações no ambiente de comunicação.

A arquitetura modular proposta permite a utilização de diferentes algoritmos de aprendizado por reforço sem modificações estruturais, possibilitando extensões futuras, como abordagens hierárquicas, multiagente ou distribuídas, mantendo a consistência do modelo arquitetural.

Apesar dos benefícios, a centralização introduz desafios relevantes. A dependência de um controlador único caracteriza um potencial ponto único de falha e pode expor o sistema a vulnerabilidades, indisponibilidade ou degradação do serviço em caso de falhas no componente central. Além disso, em cenários reais, aspectos relacionados à segurança e à confiabilidade do controlador tornam-se críticos, especialmente em aplicações sensíveis.

Adicionalmente, o ciclo de controle — envolvendo coleta de estado, inferência do agente e aplicação da ação — introduz um intervalo temporal entre observação e atuação, que pode impactar o desempenho em cenários de larga escala ou com elevada dinâmica de rede. Esse efeito é particularmente relevante quando há variações rápidas no estado do sistema.

A modelagem adotada assume operação energeticamente neutra como aproximação, desconsiderando variações abruptas na disponibilidade de energia ambiente, como flutuações na excitação por RF. Em cenários reais, essa intermitência pode influenciar diretamente a capacidade de comunicação dos dispositivos e afetar a estabilidade do sistema.

Adicionalmente, os resultados foram obtidos sob densidade relativamente estável, não contemplando explicitamente cenários dinâmicos com entrada e saída de dispositivos (*churn*). A avaliação do comportamento da arquitetura sob variações abruptas na densidade e na ativação dos nós representa uma direção relevante para trabalhos futuros.

Por fim, destaca-se que a configuração dos hiperparâmetros do agente DRL influencia diretamente o desempenho e o tempo de convergência do sistema. Parâmetros como taxa de aprendizado e fator de desconto afetam a estabilidade do treinamento e a capacidade de adaptação do agente, sendo aspectos importantes para investigações futuras.

Ainda assim, os resultados demonstram que a integração de aprendizado por reforço profundo ao nível de sistema constitui uma direção promissora para redes IoT sustentáveis, oferecendo um compromisso entre desempenho, escalabilidade e viabilidade energética, mesmo diante das limitações identificadas.

7. Conclusão e Trabalhos Futuros

Este trabalho apresenta o *GlobalAmBC-DRL Control Module*, um módulo arquitetural de controle centralizado projetado para suportar redes IoT densas e sem bateria baseadas em retroespalhamento ambiente. Diferentemente de abordagens centradas exclusivamente em algoritmos de otimização, a proposta enfatiza uma visão de sistema, introduzindo um componente reutilizável e reproduzível que integra aprendizado por reforço profundo ao ciclo de execução da simulação de redes em nível arquitetural.

A adoção de um controle centralizado fundamentado em estado global agregado demonstra-se adequada ao contexto de dispositivos AmBC, cujas restrições energéticas e computacionais inviabilizam decisões locais complexas. Os resultados experimentais indicam que o módulo proposto contribui para maior estabilidade da comunicação, preservação da operação energeticamente neutra dos dispositivos e melhor adaptação a cenários de alta densidade, quando comparado a estratégias sem controle ou baseadas em políticas estáticas.

Do ponto de vista arquitetural, a separação entre o módulo de controle, o simu-

lador de rede e o agente de aprendizado por reforço reforça a robustez e a generalização da proposta. Essa modularidade permite a substituição de algoritmos de aprendizado, a variação de cenários experimentais e a extensão do framework sem modificações estruturais, favorecendo a reutilização por outros pesquisadores. A disponibilização do código-fonte e dos cenários de simulação em repositório público reforça o compromisso com a reprodutibilidade e a transparência científica.

Além de validar a viabilidade prática do módulo, este trabalho contribui para preencher uma lacuna identificada na literatura ao oferecer uma infraestrutura arquitetural explícita para o controle inteligente de redes IoT sem bateria baseadas em retroespalhamento ambiente. Ao empregar o aprendizado por reforço como elemento central da arquitetura do sistema, a proposta permite avaliar impactos globais e dinâmicos que não são capturados por abordagens pontuais ou puramente analíticas.

Entretanto, a abordagem centralizada introduz desafios que devem ser considerados, como a dependência de um controlador único e possíveis impactos associados à latência do ciclo de controle em cenários de maior escala. Além disso, a modelagem adotada assume condições médias de disponibilidade energética, podendo não capturar totalmente a intermitência da energia ambiente em cenários reais.

Como trabalhos futuros, destacam-se diversas extensões naturais da arquitetura proposta. Entre elas, a investigação de arquiteturas hierárquicas com múltiplos controladores cooperativos, a incorporação de mecanismos de controle distribuído ou híbrido, e a consideração de cenários com mobilidade, variações abruptas de densidade de nós e maior variabilidade temporal da energia ambiente. Adicionalmente, a integração do módulo com dados reais e a validação em testbeds experimentais representam passos importantes para aproximar a proposta de aplicações práticas.

Do ponto de vista computacional, o custo de treinamento do agente de aprendizado por reforço é amortizado pelo uso de um controlador centralizado, executado externamente aos dispositivos IoT. Além disso, a agregação do estado global reduz a dimensionalidade do espaço de observação, contribuindo para a escalabilidade da abordagem em cenários densos. Dessa forma, o *GlobalAmBC-DRL Control Module* demonstra viabilidade computacional, escalabilidade e adequação prática para o controle adaptativo de redes IoT densas e sem bateria, consolidando-se como uma base arquitetural promissora para aplicações futuras.

Referências

- Huynh, N. V. et al. (2018a). Optimal and low-complexity dynamic spectrum access for rf-powered ambient backscatter system with online reinforcement learning. *arXiv pre-print*. Early work applying RL to backscatter communications.
- Huynh, N. V., Thai, H. D., Lu, X., Niyato, D., and et al. (2018b). Ambient backscatter communications: A contemporary survey. *IEEE Communications Surveys & Tutorials*. Overview of ambient backscatter communications and challenges.
- Imran, M., Ahmad, I., and Chen, Y. (2025). Simulation-based optimization of energy harvesting systems for 6g networks using deep reinforcement learning. *Journal of Communications and Networks*. DRL-based optimization of energy harvesting systems targeting future 6G networks.

- Jamshed, M. A. et al. (2025). Artificial intelligence, ambient backscatter communication. *arXiv preprint*. Survey on AI techniques applied to AmBC.
- Khan, M. A., Rehman, A. U., Khan, Z., and Niyato, D. (2023). Deep reinforcement learning for backscatter communications: Augmenting intelligence in future internet of things. *IEEE Internet of Things Journal*. Survey and analysis of DRL techniques applied to backscatter communication systems and IoT.
- Liu, Q. et al. (2021). Ambient backscatter communication-based smart 5g iot network. *EURASIP Journal on Wireless Communications and Networking*. Example of AmBC in smart IoT networks.
- Liu, V., Parks, A., Talla, V., Gollakota, S., Wetherall, D., and Smith, J. R. (2013). Ambient backscatter: Wireless communication out of thin air. *ACM SIGCOMM Computer Communication Review*, 43(4):39–50.
- Memon, M. L. et al. (2019). Backscatter communications: Inception of the battery-free era. *Electronics*. Review on backscatter communications and battery-free systems.
- Perez-Ramirez, D. F. et al. (2021). Deepgantt: A scalable deep learning scheduler for backscatter networks. *arXiv preprint*. Deep learning approaches for scalable backscatter communications.
- Shoeibi, A., Gholami, A., and Mosavi, A. (2024). A novel drl-based scheme for ris-assisted energy harvesting in batteryless iot networks. *Future Internet*, 16(12):460. DRL-based optimization for energy harvesting and performance in batteryless IoT networks.
- Sooriarachchi, V. P. et al. (2025). Ambient backscatter- and simultaneous wireless information and power transfer. *Applied Sciences*. Hybrid mechanisms combining power transfer with backscatter communications.
- Xie, L., Xu, Y., and Wang, R. (2021). Resource allocation in ambient backscatter communication networks using deep reinforcement learning. *IEEE Internet of Things Journal*, 8(6):4494–4506.
- Zargari, S. H. and Larsson, E. G. (2024). Refined deep reinforcement learning for mimo bistatic backscatter resource allocation. *arXiv preprint arXiv:2405.14046*. DRL-based resource allocation in MIMO bistatic backscatter communication systems.
- Zhao, Y., Li, B., and Wang, Y. (2022). Deep reinforcement learning for resource optimization in iot networks. *IEEE Transactions on Network and Service Management*, 19(2):1234–1246.
- Zheng, K. et al. (2024). Ambient iot towards 6g: Standardization, potentials, and challenges. *arXiv preprint*. Recent directions in Ambient IoT and backscatter communications.