RLBEEP: Reinforcement-Learning-Based Energy Efficient Control and Routing Protocol for Wireless Sensor Networks.

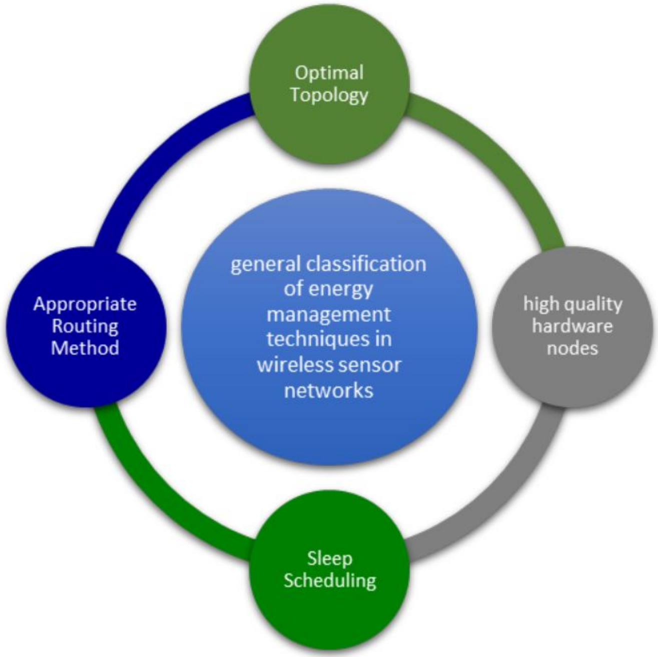
EL7044 - Conceptos Avanzados en Redes Inalámbricas
Sept 2024

Introduction

WSN are useful to monitoring environmental data.

Typical networks comprise several nodes, most of which embed a sensor, a battery, some memory, and a microcontroller.

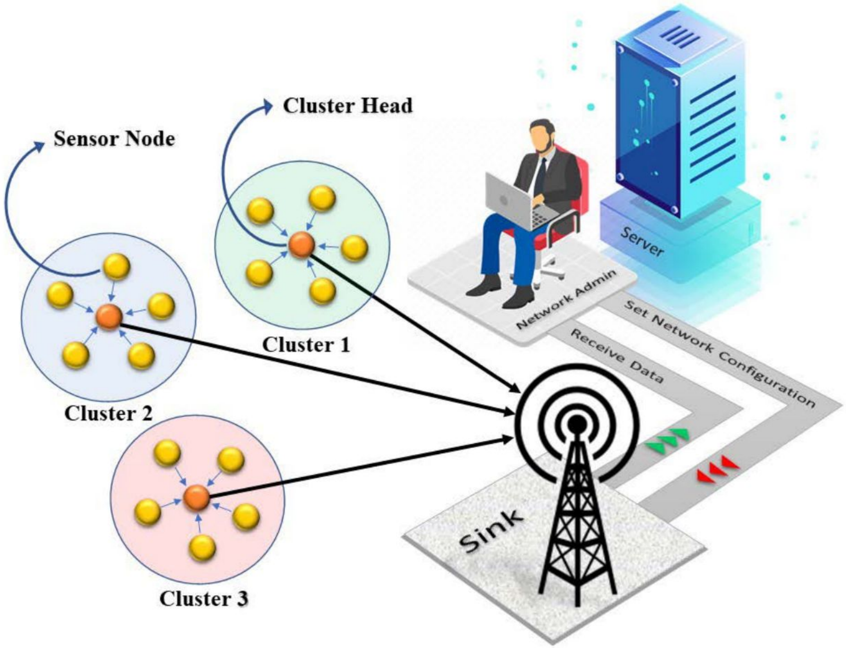
In general, four categories of techniques are used to improve the energy efficiency of sensor network systems.



Ways to achieve: control protocols

"control protocols are used to perform clustering, aggregation, compression, and sleep scheduling"

1. Clustering
2. Aggregation and compression reduce the amount of data sent over the network. This reduces data storage required in leaf nodes.
3. SSch disables unused nodes



RL - ML approach

Rt is the amount of the received reward from time t based on future rewards. rt+1 indicates the received reward at time t+1 and γ -> [0,1] is the discount factor.

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \gamma^3 r_{t+4} + \cdots = \sum_{i=t}^{\infty} \gamma^{i-t} r_{i+1}$$

Paper approach

In this paper, an energy-efficient control and routing protocol in wireless sensor networks is presented. This algorithm is based on reinforcement learning for energy management approach in the network. This protocol seeks to optimize routing policies to maximize the long-term reward received by each node, using reinforcement learning, which is a machine learning approach. In addition, the proposed method uses the sleep scheduling technique and limits sending rate in nodes with low energy requirements. Notably, transmission can be restricted to changes in sensor values. The innovation of this paper is in fact the integration of three innovative techniques including sleep scheduling, data transmission restriction (data fusion) and packet routing using reinforcement learning.

Related works I - Traditional Approaches

Initial WSN protocols were largely static, limiting adaptability in dynamic environments.

- Data-Centric Protocols: Data aggregation methods to minimize energy use.
- Hierarchical Protocols: Use of clustering techniques to reduce transmission loads.
- Location-Based Protocols: Routing based on geographic information.
    - This is because IP protocols do not exist for sensor network.

Related works II - Adaptive and Machine Learning Approaches

- Neural Networks & Deep Learning: Back-propagation techniques for clustering (ELDC), degree-based node classification. -> finds and set ‘cluster rings’ nodes.
- Reinforcement Learning (RL) Approaches:
  - Q-Routing: Packets are transmitted to max Q value of each neighbour node
  - AdaR (Adaptive Routing): Combines Least Square Policy Iteration and Q-Learning.
    - searches for optimal policy in few steps
    - independently of the problem model
  - Learning-based Adaptive Routing Tree (ATP): Q-values are cost, selects min Q.
    - Robust to unpredictable link failures
    - Needs hyperparameter tuning

Related works III - Recent Developments

- RL with Feedback Mechanisms:
  - FROMS (Feedback Routing for Multiple Sink Optimization): Identifies optimal paths to sinks.
  - QELAR: Underwater WSN using machine learning for energy-efficient routing.
  - DACR: QoS-aware adaptive cooperative routing.
- Multi-Agent Systems:
  - MRL-SCSO: Self-configuration and optimization in WSN using multi-agent RL.

Related works IV - RLBR (Reinforcement-Learning-Based Routing)

- Q-Learning Approach: RLBR introduces Q-learning to optimize routing decisions by assigning Q-values to paths between nodes, based on factors like energy consumption and hop count.
- Advantages:
  - Simple and effective use of RL in routing decisions.
- Limitations:
  - The reward function is not fully optimized for dynamic network conditions, leading to suboptimal long-term performance.

Related works IV - DADF (Delay-Aware Data Fusion)

Data Fusion and Delay Awareness: DADF introduces hierarchical data fusion techniques and sleep scheduling to minimize redundant transmissions and optimize energy use.

- Key Features:
  - Combines hierarchical data fusion with sleep scheduling.
  - Uses Q-learning to manage routing while accounting for delay in data transmissions.
- Limitations:
  - DADF doesn’t optimize routing based on residual energy or the distance between nodes, which could further improve energy efficiency.

Related works - Summary

Methods	Sleep Scheduling	Data Fusion	Q-Learning
AdaR	✗	✗	✓
ATP	✗	✗	✓
FROMS	✗	✗	✓
QELAR	✗	✓	✓
DACR	✗	✗	✓
FTIEE	✗	✗	✓
MRL-SCSO	✗	✓	✓
RLBR	✗	✗	✓
DADF	✓	✓	✓
RLBEEP (proposed method)	✓	✓	✓

RLBEEP - Overview

Uses **RLBR** and **DADF** improved

Routing: Based on reinforcement learning.

Sleep Scheduling: Nodes switch between active and sleep states to conserve energy.

Data Transmission Restriction: Limits data transmissions to significant changes.

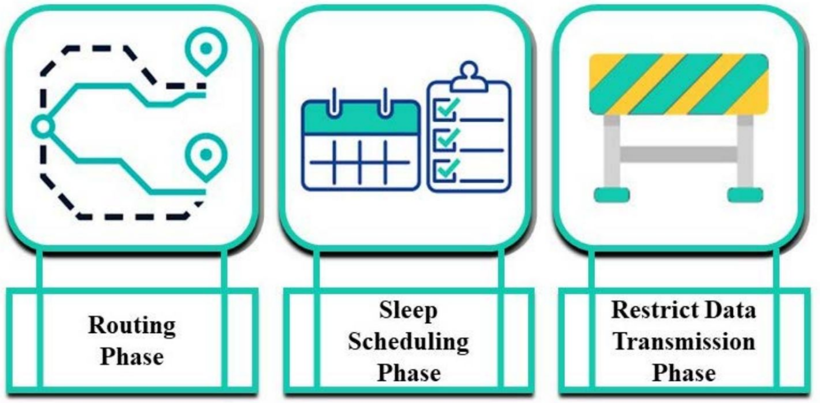


FIGURE 5. RLBEEP main phases.

RLBEEP

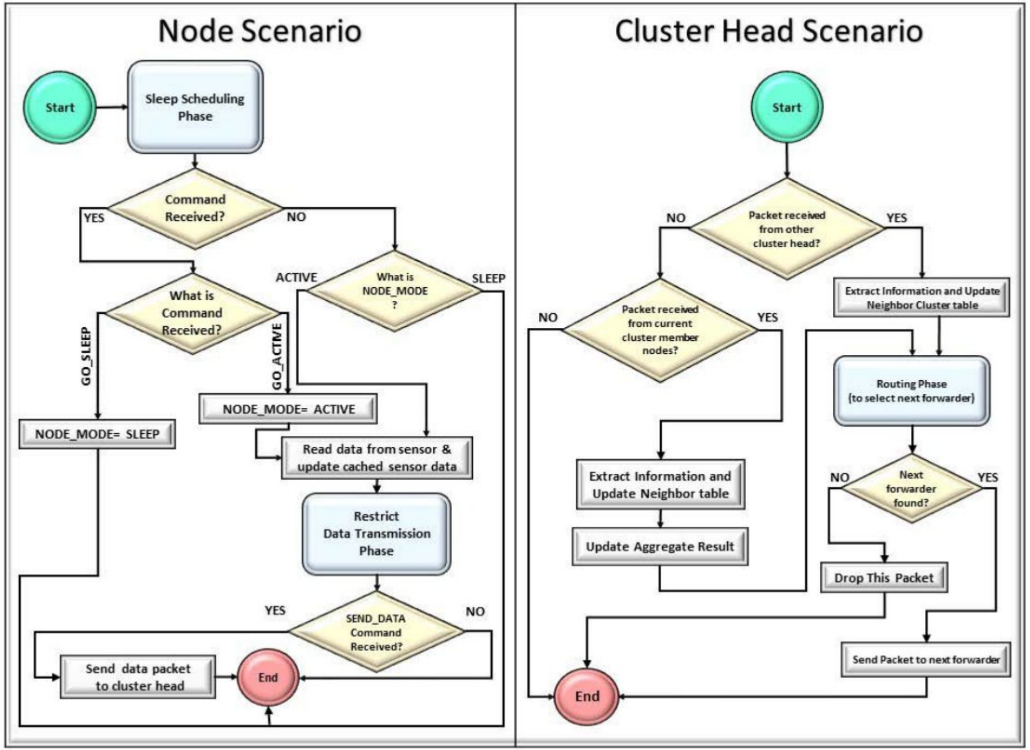


FIGURE 6. RLBEEP main scenarios in normal node and cluster head.

RLBEEP - Comparison with RLBR and DADF

- RLBR (Reinforcement-Learning-Based Routing)
  - Q-Learning Approach: RLBR uses Q-learning to assign Q-values to routes based on energy levels and hop count.
  - Limitations: RLBR’s reward function doesn't fully optimize long-term energy consumption or adapt to changing network conditions.
- DADF (Delay-Aware Data Fusion)
  - Hierarchical Data Fusion: DADF uses data aggregation and sleep scheduling to minimize redundant data transmission.
  - Drawbacks: It doesn’t sufficiently factor in residual energy and distance during routing decisions, leading to suboptimal energy use.

Improvements in RLBEEP

Enhanced Reward Function:

- In RLBEEP, the reward function is modified to optimize the routing path by considering:
  - Residual energy of neighboring nodes.
  - Hop count to the sink.
  - Distance between nodes. (new)

These modifications lead to more energy-efficient routing compared to RLBR.

Improvements in RLBEEP

Routing Phase (Q-Learning Update)

- RLBEEP updates the Q-value of the path between the current node and neighboring nodes using the formula:

$$Q_{new}(cur, nbr) = (1 - \alpha)Q_{old}(cur, nbr) + \alpha(R(cur, nbr) + Q(nbr)) \quad (3)$$

Here, R(cur,nbr) is a reward function based on the **residual energy** E(nbr), **hop count** h(nbr), and **distance** d(cur,nbr).

Improvements in RLBEEP

Data Transmission Restriction

- Threshold-Based Approach: Only sends data when significant changes occur, determined by the change threshold.
- Algorithm: The algorithm tracks minimum and maximum sensor values, transmitting data only when the changes exceed a set threshold, avoiding unnecessary transmissions.

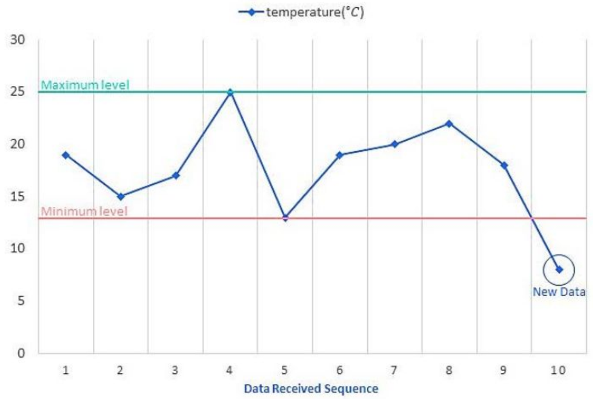
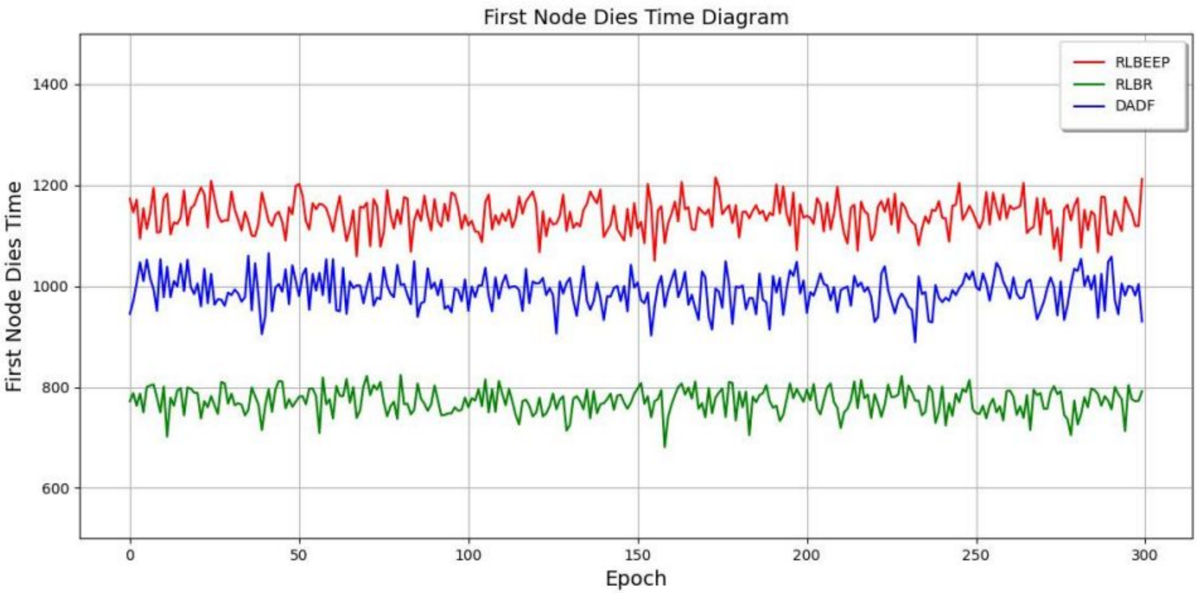


FIGURE 7. Temperature data received sequence.

TABLE 2. Simulation setup hyper-parameters value.

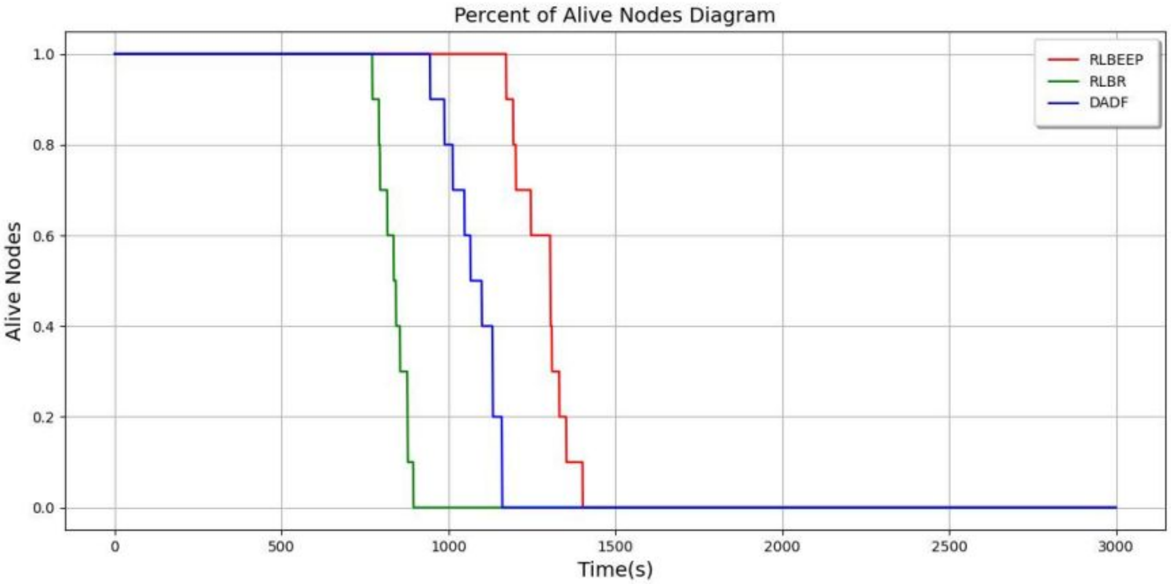
Hyper-Parameters	Value
Number of Nodes	10
Number of Clusters	4
Send Distance Range	10 m
Node Initial Energy	100 J
Alpha	0.5
Power consumption in send	0.3 J
Power Consumption in receive	0.2 J
Power Consumption in Active (Standby State)	0.1 J
Power Consumption in Sleep	0.05 J
Maximum Longitude	60.0 m
Maximum Latitude	60.0 m
DFR_{min}	5.0
DFR_{max}	55.0
Total Simulation Time	3000 s
Number of Epoch (Period)	300

Results



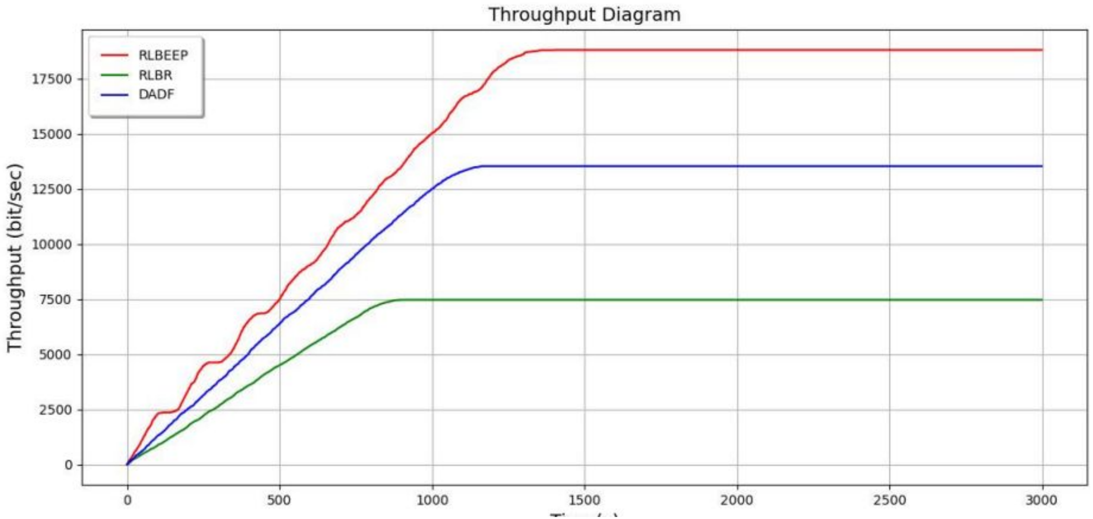
First Node Dies Time Diagram
RLBEEP
RLBR
DADF
First Node Dies Time
1400
1200
1000
800
600
0
50
100
150
200
250
300
Epoch

Results



Percent of Alive Nodes Diagram
RLBEEP
RLBR
DADF
Alive Nodes
Time(s)
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000
2500
3000

Results


Throughput Diagram
RLBEEP
RLBR
DADF
Throughput (bit/sec)
0
2500
5000
7500
10000
12500
15000
17500
0
500
1000
1500
2000
2500
3000