ICAIIC 2024

Q-learning-based Traffic-aware Parent Selection for Wireless Powered Sensor Networks

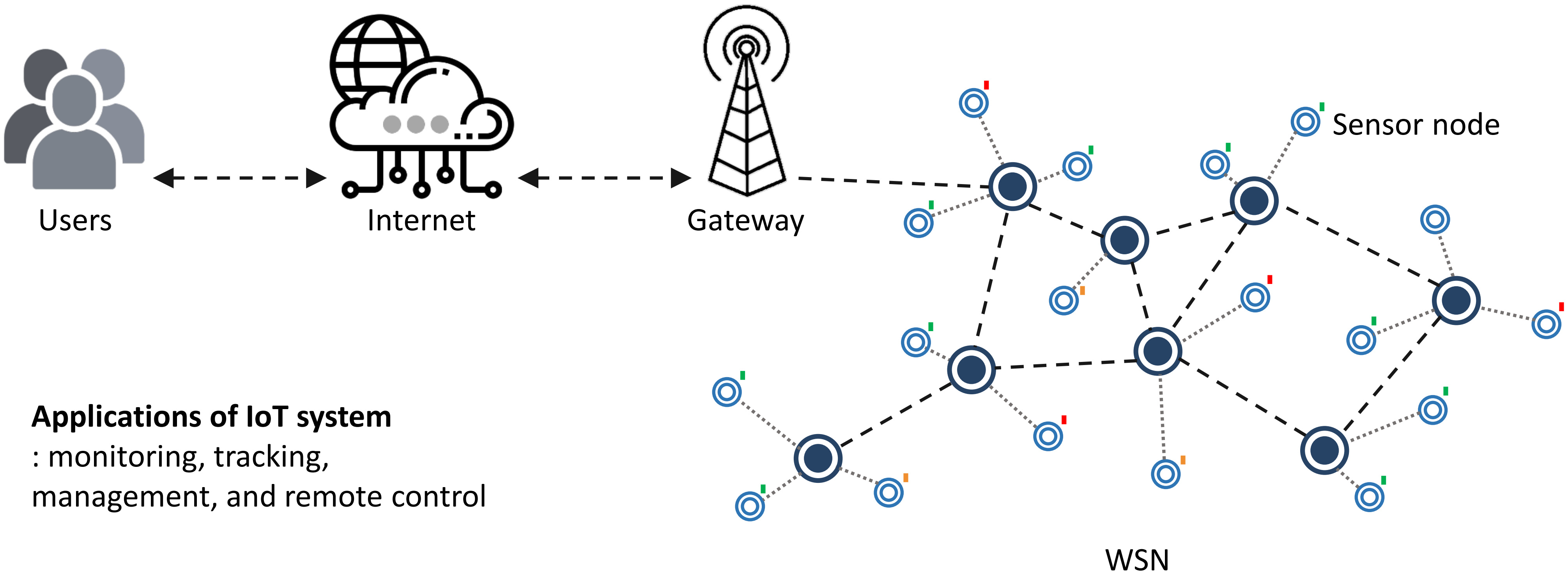
Sol-Bee Lee, **Jung-Hyok Kwon**, and Eui-Jik Kim
Hallym University

2024.02.21

Contents

Introduction

- Over the decades, the wireless sensor network (WSN) has played a key role in the Internet of Things (IoT) system.
- In WSNs, a number of low-power and battery-powered **sensor nodes require frequent recharging and replacement of battery** due to **the limited battery capacity**.



Applications of IoT system
: monitoring, tracking,
management, and remote control

Fig. 1. IoT System.

Introduction

- To solve these problems, self-sustainable **<u>wireless powered sensor networks (WPSNs)</u>** have been extensively studied.
- WPSNs **<u>enable permanent operation of sensor nodes without battery replacement or interruption</u>** by using **<u>wireless power transfer (WPT) technology</u>**.
- Unlike traditional WSNs, WPSNs **<u>consist of multiple sensor nodes and hybrid access points (HAPs) that supply power to sensor nodes using radio frequency (RF) signals.</u>**

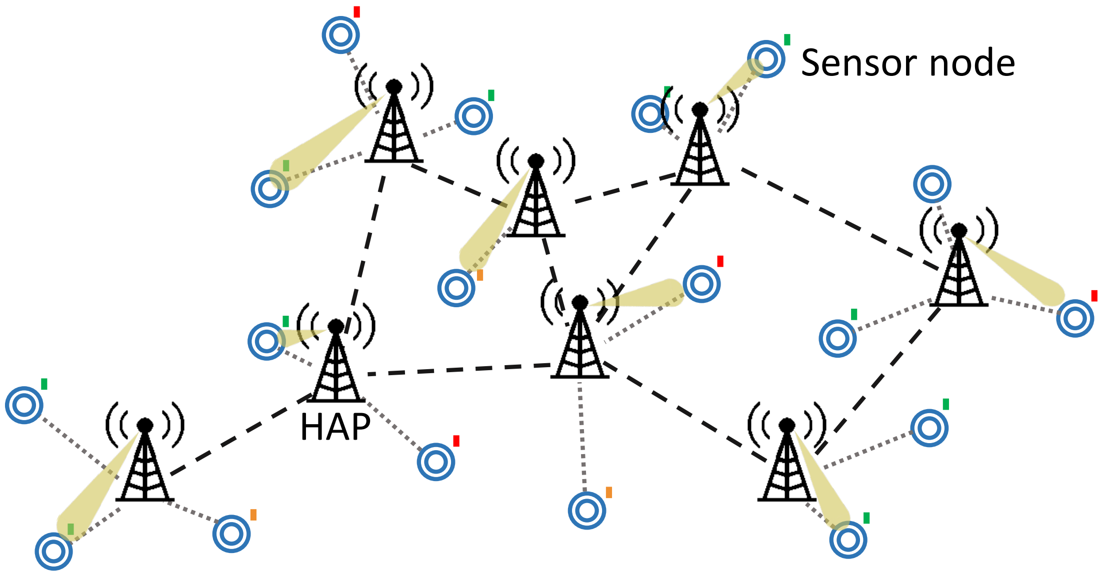


Fig. 2. WPSN.

Two types of traffic with different characteristics coexist in WPSN.
➔ Data traffic and Power traffic

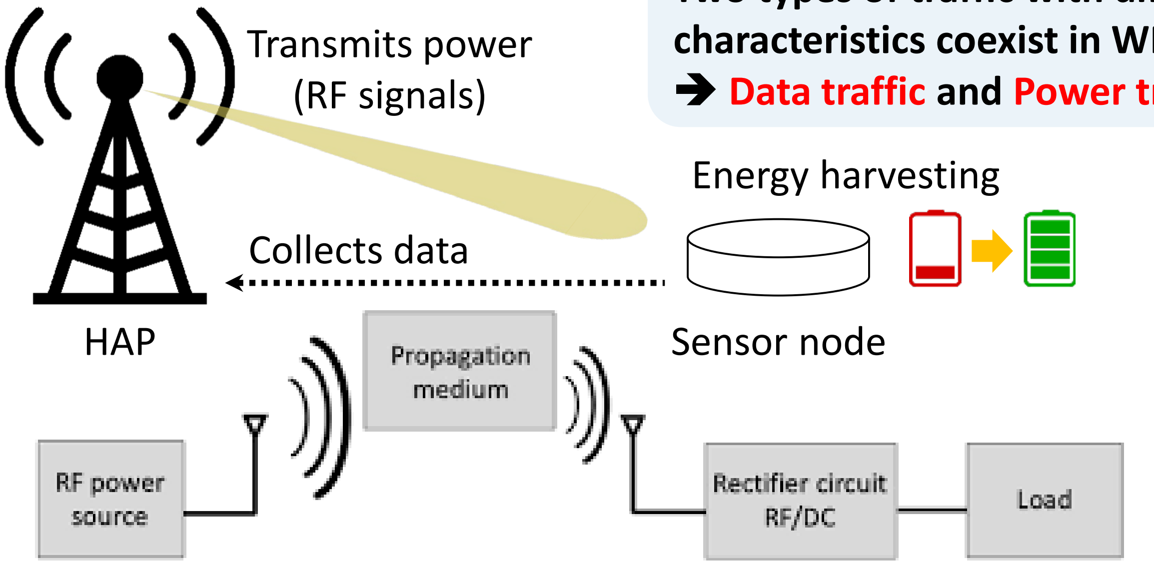


Fig. 3. WPT using RF signals

Introduction

- **Data traffic**
  - Data packets transmitted in the network and is transmitted through multi-hop from the sensor node to the root.
- **Power traffic**
  - RF signals transmitted in a single hop from the HAP to the sensor node to charge the sensor node.
  - Suffers from exponential power attenuation according to its propagation distance ➔ **a doubly near-far problem**

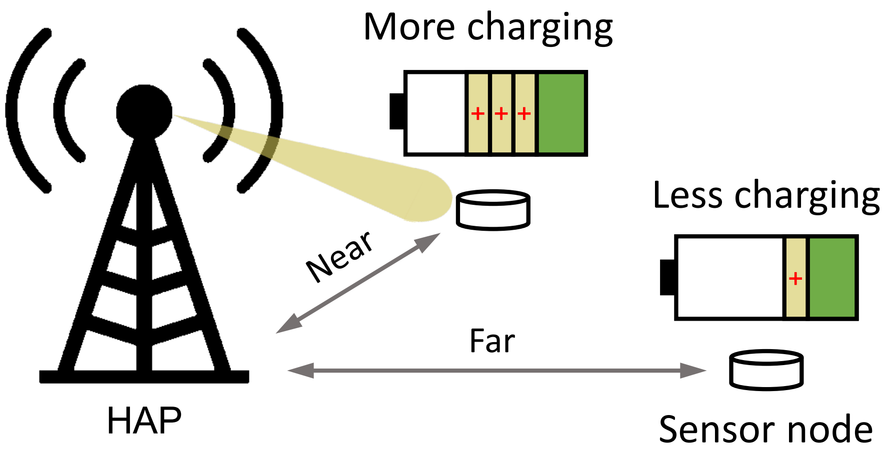


Fig. 4. Doubly near-far problem.

A doubly near-far problem results in **an imbalance in the amount of residual energy** between sensor nodes and **the inefficient use of channel resources** in WPSN.

Introduction

- **Limitations of existing resource allocation schemes**
  - **Do not consider changes in the channel state** between the sensor node and the HAP.
    ➔ Significantly impacts the amount of energy harvested by the sensor node.
  - **Do not consider the different characteristics of power traffic and data traffic** for resource allocation.
    ➔ Leads to unnecessary long energy harvesting and transmission delays of data packets.

Q-learning-based traffic-aware parent selection for WPSN (QTaPS)

- Aims to **mitigate the inefficient use of channel resources** and **reduce the transmission delay of data traffic**.
- Enables the **sensor node to adaptively select its data parent and power parent for energy harvesting and data transmission based on their respective traffic characteristics**.

System model

- **System architecture**
  - **Multi-hop WPSN** consisting of multiple HAPs and sensor nodes
    - ✓ Adopts a carrier sense multiple access with collision avoidance (CSMA/CA) channel access mechanism.
  - **HAP**
    - ✓ Does not require energy harvesting. ➔ only selects the data parent as its parent.
    - ✓ Supplies power to sensor nodes
    - ✓ Transmits data traffic received from the sensor nodes to its data parent.
    - ✓ One HAP can serve as both the power parent and data parent of the sensor node.
  - **Sensor node**
    - ✓ Selects two independent parents based on power traffic and data traffic.
      ➔ Power parent and data parent
    - ✓ Harvests energy from its power parent and then transmits its data packet to its data parent.

HAPs and sensor nodes can only select a HAP that supports both WPT and data communication as their parent.

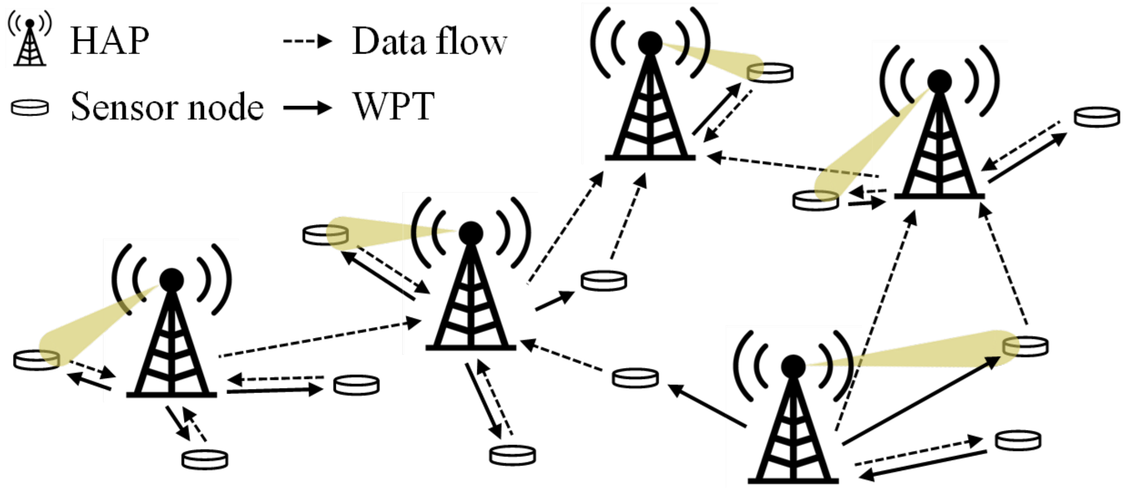


Fig. 5. System architecture of multi-hop WPSN.

System model

- **Energy model**
  - The received power of the sensor node from the HAP (i.e., power parent) in free space

➔ $P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d} \right)^2$

  - The amount of energy harvested by the sensor node for the duration of energy harvesting

➔ $P_{Rx} = \eta P_r T_{EH}$

  - The amount of energy consumed by the sensor node

➔ $E_{Tx} = E_{idle} T_{backoff} + E_{rx} T_{cca} + E_{tx} T_{data} + E_{rx} T_{ack}$

➔ $E_{Rx} = E_{rx} T_{data} + E_{rx} T_{ack}$

➔ $E_{Sleep} = E_{sleep} T_{sleep}$

➔ $E_{Idle} = E_{idle} T_{idle}$

Table 1. Description of variables.

Variable	Description
P_t	The transmission power of the power parent
P_r	The received power of the sensor node from the HAP
G_t	The antenna gain of the power parent
G_r	The antenna gain of the sensor node
λ	The wavelength of the RF signal
d	The distance between the sensor node and its power parent
α	The path loss exponent
P_{Rx}	The amount of energy harvested by the sensor node
T_{EH}	The duration of energy harvesting
η	The energy harvesting efficiency
E_{Tx}, E_{Rx}, E_{Idle}, E_{Sleep}	The amount of energy consumed by the sensor node during successful transmission, successful reception, idle, and sleep.
E_{tx}, E_{rx}, E_{idle}, E_{sleep}	The amount of energy consumed per second by the sensor node
$T_{backoff}$	The backoff period
T_{cca}	The clear channel assessment (CCA) period
T_{data}	The length of the data packet
T_{ack}	The length of the Ack packet

Design of QTaPS

- **QTaPS**
  - Is designed to enable the sensor node to adaptively select its power parent and data parent by considering parent selection metrics suitable for power and data traffic.
- **The multi-hop WPSN can be modeled as an environment.**
  - The environment includes **HAPs, sensor nodes, and power and data traffic** transmitted between sensor nodes and HAPs.
- The **sensor node selects its power parent and data parent separately**, utilizing **each Q-learning agent for the power parent and data parent**.

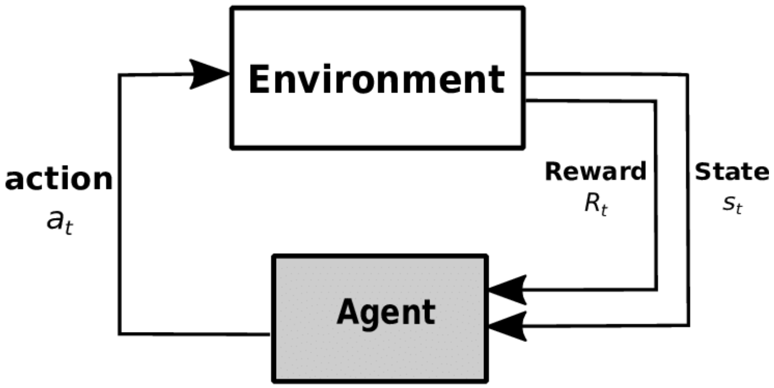
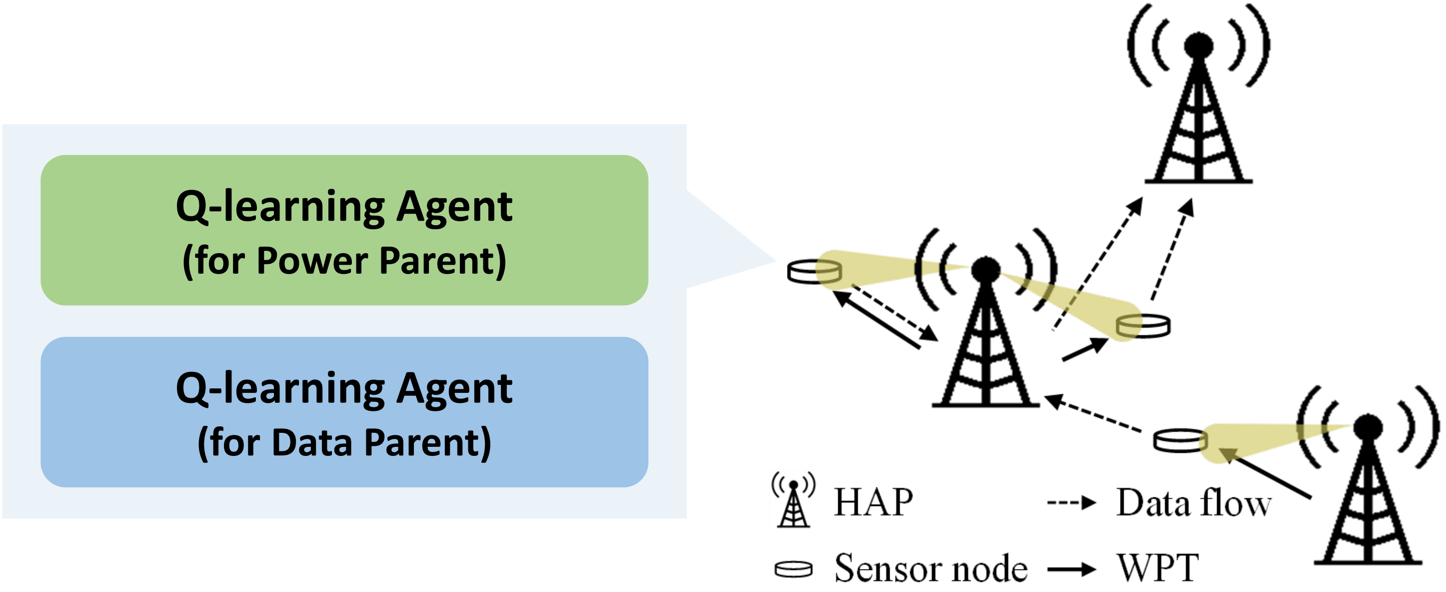


Fig. 6. Environment-Agent Interface.



Design of QTaPS

- All HAPs within WPSN are defined as a set of states and actions.
  - ➔ $\mathbf{S} = \{s_1, s_2, \cdots, s_m\}$, $\mathbf{A} = \{a_1, a_2, \cdots, a_m\}$ (m: The number of HAPs in the WPSN)
- The **set of states** of each agent for the power parent and data parent
  - ➔ $\mathbf{S}_{PP} = \{s_1, s_2, \cdots, s_k\}$, $\mathbf{S}_{DP} = \{s_1, s_2, \cdots, s_k\}$ (k: The number of neighbor HAPs of the sensor node)
- The **set of actions** of each agent for the power parent and data parent
  - ➔ $\mathbf{A}_{PP} = \{a_1, a_2, \cdots, a_k\}$, $\mathbf{A}_{DP} = \{a_1, a_2, \cdots, a_k\}$ (k: The number of neighbor HAPs of the sensor node)

- **Each agent**
  - Transitions from one state to another by performing one action from its set of actions.
  - Interacts with the environment by performing actions (a_t^{PP} and a_t^{DP}) in the current state (s_t^{PP} and s_t^{DP}) at t.
  - Obtains the corresponding reward (r_t^{PP} and r_t^{DP}). (t: The number of executions of QTaPS for each agent)

Design of QTaPS

- **Reward functions**
  - Consider the parent selection metrics ➔ hop count, distance, link quality, and traffic load
  - The parent selection metrics reflect the characteristics of power and data traffic.
  - **Reward functions of agents for power and data parent**

➔ $r_t^{PP}\left(s_t^{PP}, a_t^{PP}\right) = hopCnt / \left(d \times tl\right)$

➔ $r_t^{DP}\left(s_t^{DP}, a_t^{DP}\right) = etx / \left(hopCnt \times tl\right)$

($hopCnt$: Hop count, the number of intermediate devices in the routing path from the selected HAP to root)

(tl: Traffic load, the total amount of data and power traffic of the selected HAP)

(etx : Link quality, the expected number of transmissions required by the sensor node to successfully transmit a data packet to the selected HAP)

Design of QTaPS

- Each agent maintains a Q-table to select the best parent, in which the action value function $Q(s_t, a_t)$ returns the expected sum of immediate and future rewards when an action a_t is selected in state s_t at .
- **Action value functions**
  - ➔ $Q\left(s_t^{PP}, a_t^{PP}\right) = (1-\alpha)Q\left(s_t^{PP}, a_t^{PP}\right) + \alpha\left\{r_t^{PP} + \gamma \max Q\left(s_{t+1}^{PP}, a_t^{PP}\right)\right\}$ (α: Learning rate, γ: Discount factor for future rewards)
  - ➔ $Q\left(s_t^{DP}, a_t^{DP}\right) = (1-\alpha)Q\left(s_t^{DP}, a_t^{DP}\right) + \alpha\left\{r_t^{DP} + \gamma \max Q\left(s_{t+1}^{DP}, a_t^{DP}\right)\right\}$
- Each agent selects the action with the highest Q-value in the Q-table for each state and receives the reward ➔ Updates the Q-values for each state and transitions to a new state.
- Therefore, the sensor node independently selects the power and data parent based on Q-tables updated by the two agents
  - By considering the different characteristics of power and data traffic in the dynamic network environment.

Simulation configuration

- **Simulation configuration**
  - PHY/MAC: Slotted CSMA/CA in IEEE 802.15.4
  - 25 HAPs and 2 to 20 sensor nodes
  - HAPs are arranged in a 5×5 grid
  - structure.
  - Sensor nodes are randomly deployed within a 120×120 m^2
  - A data packet generate rate: every 0.2s

  - **Performance metrics**
    - ✓ Aggregate throughput
    - ✓ Average end-to-end delay

  - **Simulation tool**
    - ✓ MATLAB simulator

Table 2. Simulation parameters.

Parameter	Value	Parameter	Value
PHY	IEEE 802.15.4	d	0–2 m
MAC	Slotted CSMA/CA	P_t	100 mW
Number of HAPs	25	E_{tx}	20.98 mA
Number of sensor nodes	2–20	E_{rx}	17.96 mA
Data packet size	127 bytes	E_{idle}	0.001 mA
Ack packet size	5 bytes	E_{sleep}	0.001 mA
Superframe order (SO)	6	η	0.65
Beacon order (BO)	6	λ	0.125
SIFS	192 μs	α	0.8
LIFS	640 μs	γ	0.9

Simulation results

- **Aggregate throughput**
  - Increases until a certain number of sensor nodes is reached and then tends to decrease as the number of sensor nodes increases ➔ Higher power and data traffic, competition, collisions and delays of data packets
  - **QTaPS**
    - ✓ Enables the sensor node to select the HAP that is relatively close to the root and has a low traffic load among neighbor HAPs as its data parent.
  - **OF0 and MRHOF**
    - ✓ Use a single parent or both energy harvesting and data transmission.
    - ✓ As OF0 performs parent selection based on the hop count, it creates a shorter path from the sensor node to the root HAP compared to MRHOF.

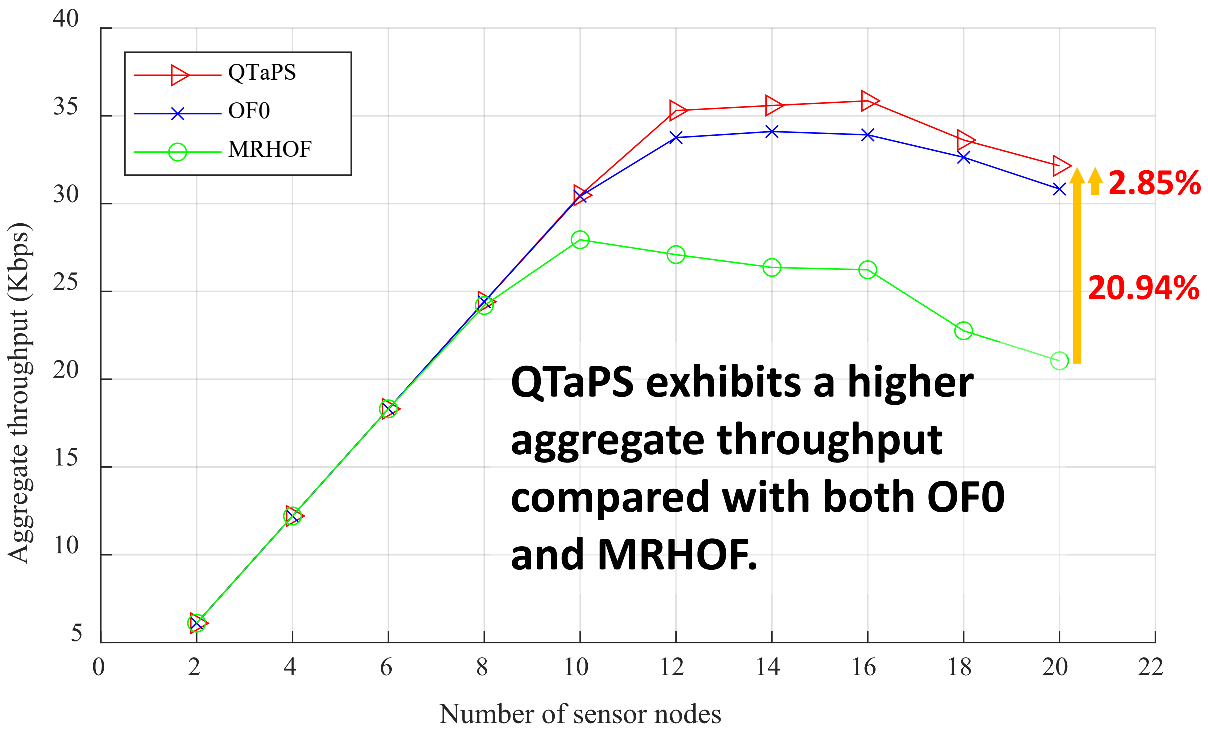


Fig. 7. Aggregate throughput.

Simulation results

- **Average end-to-end delay**
  - Increases rapidly after the number of sensor nodes reaches a certain number ➔ Both power and data traffic also increase ➔ More collisions and transmission delays of data packets
  - **QTaPS**
    - ✓ The sensor node using QTaPS utilizes two independent parents to harvest energy from a nearby HAP and transmit data packets to the HAP closer to the root.
    - ✓ The closer the distance between the sensor node and its power parent, the less time the sensor node spends harvesting the required energy.
  - **OF0 and MRHOF**
    - ✓ Sensor nodes only use a single parent regardless of the traffic type.
    - ✓ HAPs more frequently suffer from bottlenecks due to the concentration of power and data traffic.

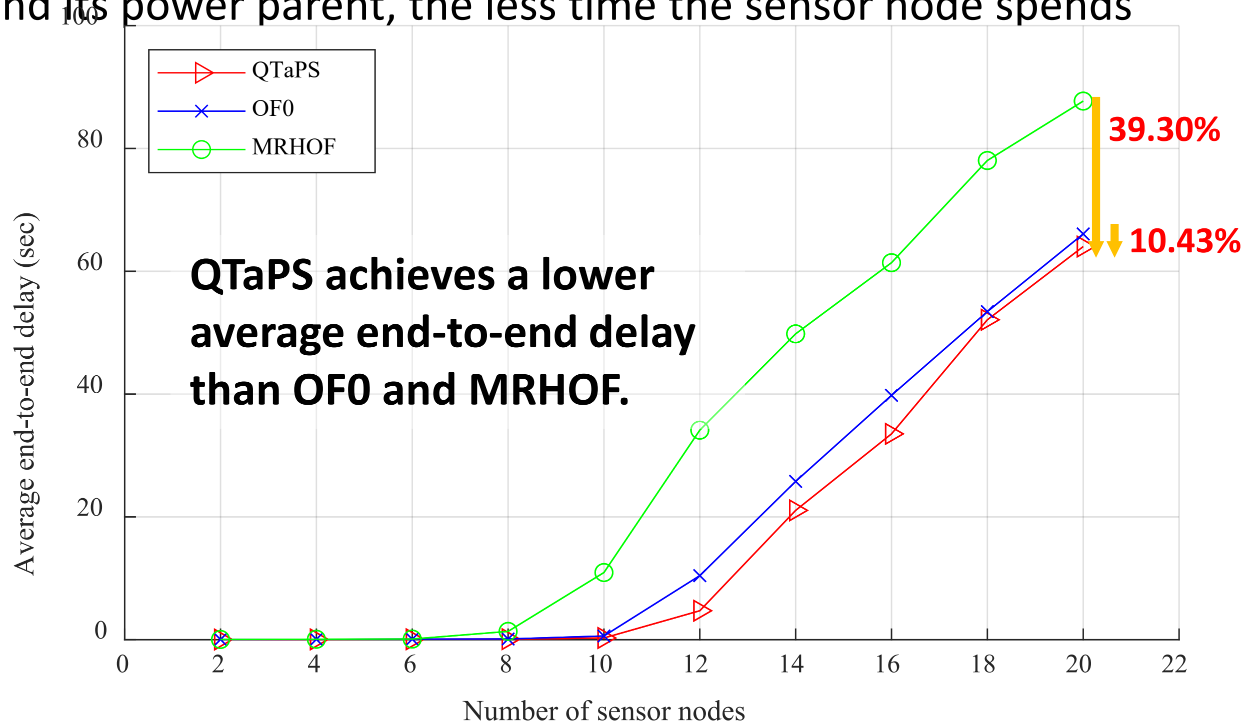


Fig. 8. Average end-to-end delay.

Conclusion

- We presents the **Q-learning-based traffic-aware parent selection for WPSN (QTaPS)** for WPSNs.

- **QTaPS**
  - Aims **to reduce the transmission delay of data packets** and **efficiently utilize channel resources** by considering the characteristics of power and data traffic.
  - Enables **the sensor node to select two separate parents for data and power traffic by employing two Q-learning agents.**

- The simulation results demonstrated that **QTaPS selects a more suitable parent for energy harvesting and data transmission compared with the existing parent selection schemes**.

- Quantitatively, QTaPS achieved 2.85% and 20.94% higher aggregate throughput and 10.43% lower average end-to-end delay compared with OF0 and MRHOF, respectively.

Thank you