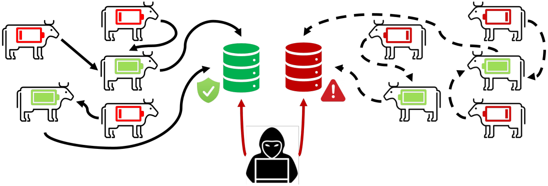
Energy-Adaptive Monitoring for Resilient Smart Farms



Motivation

- Lack of security-aware smart farm technologies under resource constraints
- Introducing serious food contamination when animal conditions are not properly monitored
- Potential high revenue loss of farmers due to the failure of protecting farms from attacks

Research Goal: Develop an uncertainty-aware MADRL-based monitoring system to achieve high monitoring quality under fluctuating energy, cyberattacks, and adversarial examples.

Problem Statement

Objective function: Minimize monitoring error ($\mathcal{ME}$) and system overload ($\mathcal{OL}$)

$$\underset{P=\{p_1,p_2,\dots,p_T\}}{\arg\max} \sum_{t=1}^{T} f(g_t(p_1, p_2, \dots, p_t)), \quad s.t. \quad \forall t \in [1, T], p_t \in \mathcal{P},$$

$$\begin{aligned} T &: \text{total monitoring step} \\ P &: \text{update policy} \\ p_t &: \text{monitoring action at time step } t \\ \mathcal{P} &: \text{action space} \\ g_t &: \text{sensor network at time step } t \\ f &: g_t \mapsto -\mathcal{ME}(g_t) - \mathcal{OL}(g_t) \end{aligned}$$

Network and Node Models

- **Network Model**: A wireless solar sensor network-based smart farm environment
  - Bluetooth Low Energy (BLE) protocol
    - $100m$ coverage
    - $2Mbps$ data rate
  - Long Range (LoRa) protocol
    - $> 1km$ coverage
    - $27kbps$ data rate
- **Node Model**
  - Temperature: temp_t^i
  - Heart beat: hb_t^i
  - Velocity: ma_t^i
  - Battery level: bl_t^i

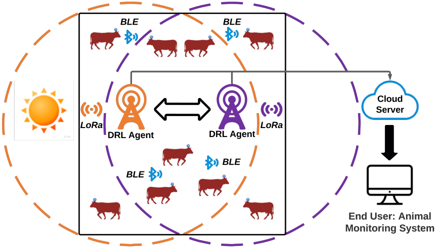


Figure 1: Wireless Solar Sensor Node-based Smart Farm Environment.

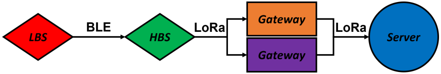
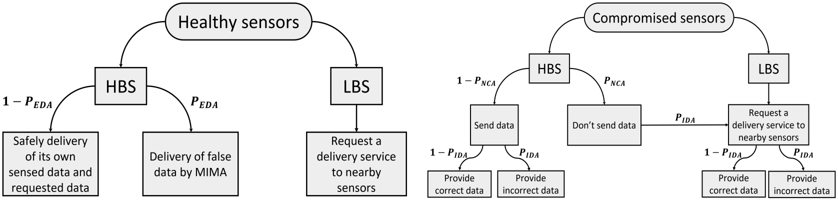


Figure 2: The overall data flow of the smart farm network.

Attack Model



HBS: high battery sensors; LBS: low battery sensors; MIMA: man-in-the-middle attackers

- **Non-compliance to the protocol:** P_{NCA}
  - Reject the data request
- **False data injection**
  - Inject from internal/external attacker: P_{IDA}/P_{EDA}
- **Denial-of-Service (DoS):** P_{IDA}
  - Send redundant data requests

Uncertainty-Aware Animal Monitoring

SL-based Formulation of a Multinomial Opinion

- A multinomial opinion X: $\omega_X = (\boldsymbol{b}_X, u_X, \boldsymbol{a}_X)$
- $\sum_{x \in \mathbb{X}} \boldsymbol{b}_X(x) + u_X = 1$

$\boldsymbol{b}_X$: *belief mass distribution* over $\mathbb{X}$

u_X: *uncertainty mass* representing *vacuity of evidence*

$\boldsymbol{a}_X$: *base rate distribution* over $\mathbb{X}$

- The dissonance $\boldsymbol{b}_X^{\mathrm{Diss}}$ of an opinion X:

$$\boldsymbol{b}_X^{\mathrm{Diss}} = \sum_{x_i \in \mathbb{X}} \left(\frac{\boldsymbol{b}_X(x_i) \sum\limits_{x_j \in \mathbb{X} \setminus x_i} \boldsymbol{b}_X(x_j) \mathrm{Bal}(x_j, x_i)}{\sum\limits_{x_j \in \mathbb{X} \setminus x_i} \boldsymbol{b}_X(x_j)} \right)$$

relative mass balance:

$$\mathrm{Bal}(x_j, x_i) = 1 - \frac{|\boldsymbol{b}_X(x_j) - \boldsymbol{b}_X(x_i)|}{\boldsymbol{b}_X(x_j) + \boldsymbol{b}_X(x_i)}$$

DRL-based Monitoring Update

- **States**:
  - Global critic state: $s_t^i = g_t(k_1, k_2, \ldots, k_t)$
  - Local actor state: $s_t^i = g_t^i(k_1, k_2, \ldots, k_t)$
  - g_t/g_t^i is represented by the history action sequence
- **Actions**:
  - n_t^i: the total number of LBS
  - Action space: $\mathbf{a}_t^i = \{0, \lfloor \frac{n_t^i}{2} \rfloor, n_t^i\}$
- **Rewards**:
  - $r_t^i = f(g_t(k_1, k_2, \ldots, k_t))$ based on $f(g_t) = -\mathcal{ME}(g_t) - \mathcal{OL}(g_t)$ where k_i is an action taken in step i

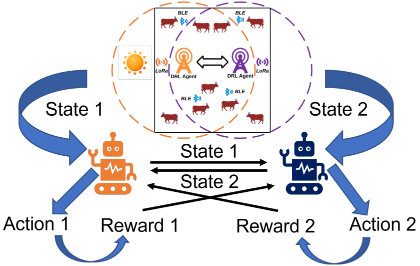


Figure 3: The proposed Multi-Agent Deep Reinforcement Learning (MADRL) framework.

Data Aggregation at LoRa Gateways

Uncertainty (Vacuity) Maximization

- Move belief mass $\boldsymbol{b}_X$ to uncertainty mass u_X
- Update uncertainty based on recent data

$$\omega_X = (\boldsymbol{b}_X, u_X, \boldsymbol{a}_X) \longrightarrow \ddot{\omega}_X = (\ddot{\boldsymbol{b}}_X, \ddot{u}_X, \boldsymbol{a}_X)$$

$$\ddot{u}_X = \min_i \left[\frac{\boldsymbol{P}_X(x_i)}{\boldsymbol{a}_X(x_i)} \right],$$

$$\ddot{\boldsymbol{b}}_X(x_i) = \boldsymbol{P}_X(x_i) - \boldsymbol{a}_X(x_i) \cdot \ddot{u}, \quad \text{for} \quad x_i \in \mathbb{X}$$

$$\text{Trigger condition:} \;\; u_X < \rho$$

Experimental Setup

Dataset: EmbediVet Devices (EVD)

Metric	Description
Serial	A unique animal identifier
Heart rate	Heart bits per min.
Average-temperature	Average body temperature in Celsius
Min-temperature	Minimum temperature in Celsius
Max-temperature	Maximum temperature in Celsius
Average-activity	Average activity recorded by the number of steps taken
Battery-level	Residual battery life
Timestamp	Date and time of transmission

Environmental Setup

- Modeling sun's movement in a day
- Consensus agreement: Ensure the maximum number of requests executed in the consolidated priority list based on Hopcroft–Karp algorithm [1]
- Gateway locations: Cover the whole farm with the same coverage of each gateway

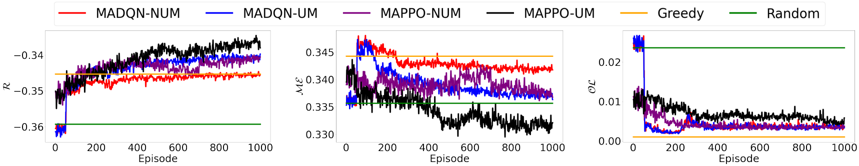
[1] Hopcroft et al., 1973

Asymptotic Complexity Analysis

Scheme	Complexity
MADQN/MAPPO	$O(n_e \times t_{train})$
Greedy	$O(n_{action})$
Random	$O(1)$
Data Mitigation (DM)	$O(1)$

- MADQN/MAPPO incurs smaller cost than Greedy with large action space
- Random and DM are the most efficient algorithms among all while showing poor performance

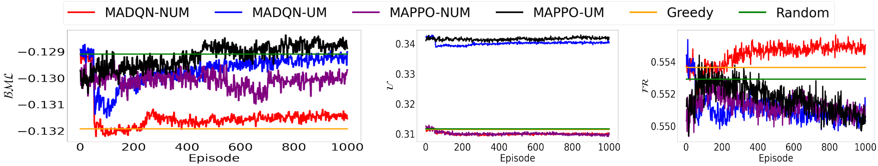
Performance Comparison $(\mathcal{R}, \mathcal{ME}, \mathcal{OL})$



(a) Accum. reward $(\mathcal{R})$ (b) Monitoring error $(\mathcal{ME})$ (c) Overload $(\mathcal{OL})$

- MAPPO-UM has a stationary decision process and achieves the best performance among all comparing schemes.
- DRL algorithms with uncertainty maximization (UM) outperform those without UM.
- The UM technique can update the uncertainty information from time to time, which reflects the sensor network status in a timely manner.

Performance Comparison $(\mathcal{BML}, \mathcal{U}, \mathcal{FR})$



(d) Batt. maintenance level $(\mathcal{BML})$ (e) Uncertainty $(\mathcal{U})$ (f) Freshness $(\mathcal{FR})$

- MAPPO-UM achieves the best battery maintenance level $(\mathcal{BML})$ compared to other schemes.
- Different schemes could have very different policies due to two conflict goals in our multi-objective function.
- The overall performance order of the considered schemes is: MAPPO-UM $\geq$ MADQN-UM $\approx$ MAPPO-NUM $\geq$ MADQN-NUM$\geq$ Greedy $\geq$ Random.

Sensitivity Analysis for Adversarial Attacks

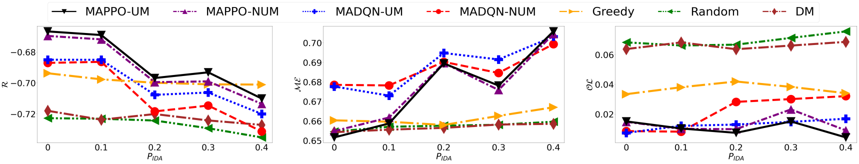
P_{ADV} \ Attack	*FGSM*	*MIM*	*PGD*	*BIM*
0.1	2	2	8	3
0.2	7	9	14	9
0.3	12	8	16	8
0.4	13	13	21	9

FGSM: Fast Gradient Sign Method
MIM: Momentum Iterative Method
PGD: Projected Gradient Descent
BIM: Basic Iterative Method
P_{ADV}: Adversarial attack probability

- MAPPO-UM is evaluated with respect to accumulated reward.
- The results are shown in the relative performance decline *permillage* [2].
- PGD is the strongest while BIM is the weakest.

[2] a rate or proportion per thousand

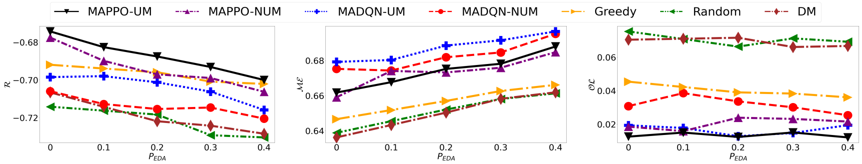
Effect of Varying the Internal Attack Probability (P_{IDA})



(a) Accumulated reward ($\mathcal{R}$) (b) Monitoring error rate ($\mathcal{ME}$) (c) Overload ($\mathcal{OL}$)

- The overall performance order with respect to the three metrics is: MAPPO-UM $\geq$ MAPPO-NUM $\geq$ Greedy $\geq$ MADQN-UM $\geq$ MADQN-NUM $\geq$ DM $\geq$ Random.

Effect of Varying the External Attack Probability (P_{EDA})



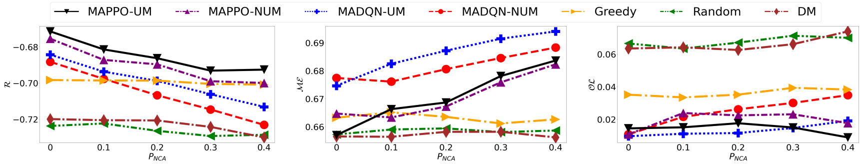
(a) Accumulated reward ($\mathcal{R}$)

(b) Monitoring error rate ($\mathcal{ME}$)

(c) Overload ($\mathcal{OL}$)

- The overall performance order with respect to the three metrics is: MAPPO-UM $\geq$ MAPPO-NUM $\geq$ Greedy $\geq$ MADQN-UM $\geq$ MADQN-NUM $\geq$ DM $\geq$ Random.

Effect of Varying the Attacker's Non-Compliance Probability (P_{NCA})



(a) Accumulated reward ($\mathcal{R}$)

(b) Monitoring error rate ($\mathcal{ME}$)

(c) Overload ($\mathcal{OL}$)

- The overall performance order with respect to the three metrics is: MAPPO-UM $\geq$ MAPPO-NUM $\geq$ Greedy $\geq$ MADQN-UM $\geq$ MADQN-NUM $\geq$ DM $\geq$ Random.

Key Contributions & Findings

Our proposed MAPPO-UM achieves:

- A strong resilience against attacks by achieving the best monitoring quality and minimum system overload.
- The best energy maintenance level by intelligently leveraging the uncertainty information.
- The enhanced monitoring quality and energy-adaptive operation with the uncertainty maximization (UM) technique using more recent evidence.
- Different optimal monitoring policies in different scenarios due to the different payoffs to monitoring updates.
- Strong robustness under harsh environments as demonstrated via extensive sensitivity analyses.