Dynamic Mobile Sink Path Planning for Unsynchronized Data Collection in Heterogeneous Wireless Sensor Networks

BY RAKESH BANOTH
CSE BTECH(22075018)

Recent years, wireless sensors and Internet-of-Things (IoT) devices have been deployed in remote locations for environmental monitoring, habitat monitoring, weather monitoring, and other monitoring applications for agricultural, military, research, and other purposes.

Since there are no base stations or internet services in wild areas and remote depopulated zones, mobile sinks (MSs), including robots, mobile vehicles (MVs), unmanned aerial vehicles (UAVs), and so on, are widely used to collect data from sensors, IoT devices, or disconnected sensor networks in these areas.

One MS can periodically visit some sensors, collect data from them, and carry and forward data to nearby base stations. However, MS has a limited energy supply and thus intelligently navigating to the most appropriate places for data collection becomes a critical but challenging problem.

HETEROGENEOUS WIRELESS SENSOR NETWORKS(HWSNs):

**The main problem that this article deals with is that of HWSN's (Heterogeneous Wireless Sensor Networks) .

**Heterogeneous sensors have different capabilities ,communications, power supplies, various power supplies.

**In homogeneous networks, we only have to deal with data delay generation, and not data generation time. But with HWSNs, we use data collection window to deal with this.

***DATA FUSION SCHEMES* to avoid problems of long delay and data non fresh in HWSNs.

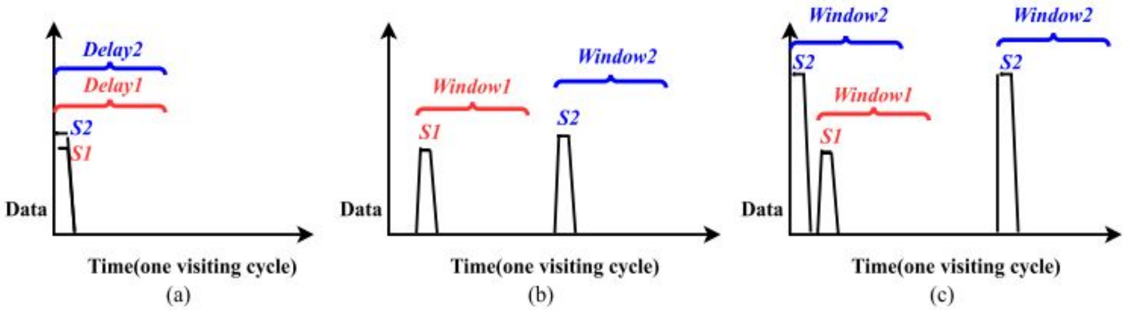


Fig. 1. Various data generations in homogeneous and heterogeneous sensors. (a) Sensors generate synchronized data. (b) Sensors generate unsynchronized data. (c) Sensors generate data with different rates.

**Always synchronized dat doesn’t exist.Data generation may be unsynchronized also, which often caused due to:

1.After long time running accumulated timing errors in sensors lead to generation of unsynchronized data.

2.Sensors may various types with data generation,sensors having different monitoring modes with different data generation rates.

MAIN GOAL:

Our main goal is to design the path of the MS with minimizing the total energy cost while satisfying life constraints, and the overall study design is provided in the previous image.

The article tells how not use the delay to definite data transmission requirements but adopt to a rigid collection window to represent the lifetime of data.

PROBLEM FORMULATION

A.SYSTEM MODEL:

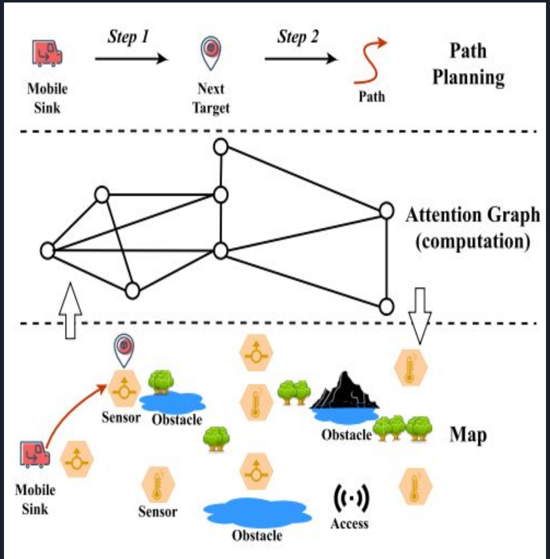
**This system consists of I wireless sensors, M MSs,

one MS depot, and one data station.This sensors are

Deployed in a remote areas with obstacles.

The model is built with following conventions and

assumptions:



1.*ASSUMPTION OF SENSORS:*

**Each sensor is connected to only one MS at a time.When sensor generate sensory data,they will release some collection tasks with time windows.

**Based on historical sensor data ,the data generation cycle of sensor is predictable so we can construct the basis of MS trajectory planning.

** The main disadvantage of this is we can predict the abnormol events and triggering modes.

2.ASSUMPTION OF MS:

**During the experiment, we install the MS's moving speed, energy consumption, and total available energy as v m/s, qm J/m, and Em J, respectively.

**Each MS can access only one sensor node to collect data at a time.

**This study does not discuss the details of MS types, related mechanisms, or environments but assumes that MS can collect data from sensors.

3.ASSUMPTION OF MS DEPOT:

In an area, there is only one depot with the location Ls as the start point of the MS trajectory. The depot can replenish energy for MS at any time.

TWO STAGE PATH PLANNING:

This planning scheme is proposed to collect data with MSs, which is designed to reduce the complexity of the planning problem in smaller state space.

In Step 1, we choose the next target from available sensors as the current visiting destination for the MS, while in Step 2, we train the MS to move to the target along an accessible path.

B.ENERGY COST OF MSs:

In this system, data transmitting and trajectory of each MS $m (m \in M)$ are the main energy consumption.

Energy consumption model of data transmission:

$$E_{\text{trans}} = \sum_{i=1}^{n} P_m^{\text{trans}} \frac{\Phi_i}{c_{mi}^{\text{trans}}}. \tag{1}$$

C.PROBLEM FORMULATION:

Since an MS’s energy budget is limited, and both trajectory and data transmission consume a large amount of energy, we try to minimize the total energy consumption while satisfying all sensors’ data transmission demands.

$$\max \sum_{m\in M}\sum_{(i,j)\in E} \lambda_1 x_{ij} z_{mij} \Phi_j - \lambda_2 E_{mij} \tag{4}$$

$$\text{s.t.} \sum_{m\in M}\sum_{(i,j)\in E} z_{mij} = 1 \tag{5}$$

$$\sum_{j\in I} z_{m0j} = 1 \tag{6}$$

$$\sum_{j\in I} z_{mij} - \sum_{i\in I} z_{mij} = 0$$

$$\sum_{i\in I} z_{mij} = 1$$

$$\sum_{(i,j)\in E} E_{mij} \leq E_m^{\text{init}}$$

$$t_{im}^{\text{arrive}} \leq t_i^{\text{end}}$$

$$t_{im}^{\text{arrive}} + w_{im} + t_{im}^{\text{trans}} + \frac{d_{ij}}{v} \leq t_j^{\text{end}}$$

$$w_{im} = \max\{0, t_i^{\text{start}} - t_i^{\text{arrive}}\}$$

where $\forall (i, j) \in E$, $\forall m \in M$, $\forall i \in I$, $x_{ij} \in \{0, 1\}$ and $z_{mij} \in \{0, 1\}$. $x_{ij} = 1$ indicates that an MS moves to sensor j from sensor i and accesses sensor j successfully; z_{mij} means an MS m is being used. $t^{trans}_{im} = (8_i / c^{trans}_{im})$ is the transmission time at the sensor i.

Equation 5 denotes that each sensor can only be connected with one MS at the same time.

Equations (6)–(8) are flow constraints. Equation (9) refers to the fact that the energy cost cannot exceed the energy capacity of the MS. Equations (10)–(12) are time constraints.

PROPOSED SOLUTION:

The greatest challenge for the combinatorial optimization problem is that the search space of solutions will be grown exponentially with the problem size, and it will result in hard to find the optimal solution, especially a large problem. DRL shows its potential to solve the NP-hard combinatorial optimization problem and has achieved several successes. Therefore, in this section, we first briefly introduce the proposed MSPP based on DRL and then demonstrate each component in detail.

**In the proposed MSPP, there are 2 key components : Task Selector and MS Controller.

**They both have similar pair of networks, actor and critic networks.

**A deep Q-learning network (DQN) with three GAT layers is used for Task Selector, employing PReLU activation. Actor and critic networks share a GAT-based architecture.

**Proximal policy optimization (PPO) is utilized for MS Controller training, featuring three linear layers with ReLU and one with Softplus activation in the actor network, and a five-layer critic network with shared parameters and a single output.

TASK SELECTOR:

$$a_i \in V_{\mathrm{mv}} \quad \forall i \in I \tag{13}$$

$$t_i^{\mathrm{end}} \geq t_{i-1}^{\mathrm{finish}} + \frac{d_{a_{i-1}a_i}}{v} + t_i^{\mathrm{trans}} \quad \forall i \in I \tag{14}$$

$$t_i^{\mathrm{finish}} \geq t_j^{\mathrm{end}}(j \in V_{\mathrm{mv}}) \quad \forall i \in I \tag{15}$$

Since the action aims at obtaining a sensor whose data can be successfully collected by the MS with minimum energy consumption, we use the feedback of the MS Controller as the reward function

$$R_s = \lambda_1 x_{im} \Phi_i - \lambda_2 \sum_{\tau=0}^{N-1} p_2. \tag{16}$$

EQUATION (13) represents a set of non visited sensors and EQUATION (14) is a time constraint: the MS must visit the sensor within the task time window, and remove all sensors, whose time windows exceed the current time from Vmv in EQUATION (15). A valid action must satisfy (13)–(15), and the invalid action will be masked.

**If the ith sensor is successfully served by the MS, then $x(im) = 1$, otherwise $x(im) = 0$. Coefficients $\lambda 1$ and $\lambda 2$ are introduced to standardize the reward. $\tau \in [0, N-1]$ is the time step of the MS Controller state update. Energy consumption p2 is calculated following.

MS CONTROLLER:

The objective of the MS Controller is to control the MS to move from the current position to the target position automatically.

The DRL agent needs to get all relevant information about the current state of the environment to fulfill the task successfully.

Therefore, the state Scontrol = {os, oc, oe} is made up

of three parts: starting coordinates os = (x, y,z), current

environmental information oc = {velocity:(vx , vy , vz);

coordinates:(x ′ , y ′ ,z ′)},

and target coordinates oe = (x ″ , y ″ ,z ″).

Algorithm 1 MS Controller

Input: Environment E, Action Space A, Initial Status $S = \{o_s, o_c, o_e\}$

Output: $< IsDone, \sum p_2 >$

repeat

  Generate corresponding action a_t by action network

  Execute action a_t and observe new observation o_c

  $\tau \leftarrow \tau + 1$

until Service success or failure

Use Equation (17) to compute reward R

Estimate the reward by $V(s) = F(s|\omega_V)$

Estimate advantages $\hat{A} = \sum R - V(s)$

Update the critic network: $\nabla_{\omega_V} E[(R^i - V^i)^2]$

Update the actor network: $\nabla_{\omega_G} E[min(r_\theta \hat{A}, clip(r(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A})]$

if the service successes **then**

  Compute feedback reward $\sum p_2$

  IsDone = True

else

  Compute feedback reward $\sum p_2$

  IsDone = False

end if

return $< IsDone, \sum p_2 >$

MSSP ALGORITHM:

Algorithm 2 MSPP Algorithm

Input: $< G, V_{mv}, V_{lv}, T_{cur} >$

Output: Best Path Found

for episode $k = 1, \ldots, K$ **do**
  Generate a random instance Q_p
  Get task set V
  $V_{mv} = V$ and $T_{cur} = \{0\}$
  while $V_{mv} \neq \emptyset$ **do**
    With probability $\varepsilon - greedy$ policy to select an action a_t based on the proposed constraints
    otherwise select $a_t = argmax_a Q(s, a)$
    Get destination $f(a_t)$
    Get information $<IsDone, \sum p_2>$ by Algorithm 1
    if IsDone == True **then**
      Reward $R_s = \lambda_1 \Phi_i - \lambda_2 \sum_\tau p_2$
      $V_{mv} = V_{mv} - [f(a_t)]$
      $V_{lv} = [f(a_t)]$
      $T_{cur} = T_{cur} + [f(a_t)]$
    else
      Reward $R_s = -\lambda_2 \sum_\tau p_2$
    end if
    Store transition tuple to replay memory $\mathcal{D}$
    Sample random mini-batch from $\mathcal{D}$
    Update actor network with gradient descent.
    Update critic network with TD-error.
  end while
end for

CONCLUSION:

In conclusion, this article introduces a novel data collection scheme tailored for real-time data generation scenarios. The proposed scheme establishes data collection windows for each data point, considering both the real-time data generation and maximum delay factors. Building upon this scheme, a two-stage Deep Reinforcement Learning (DRL)-based algorithm, named MSPP, is introduced for Mobile Sink (MS) trajectory planning in the context of unsynchronized data collection in Heterogeneous Wireless Sensor Networks (HWSNs).

The algorithm incorporates a unique exploration strategy to address various time windows associated with data sensors, aiming to enhance efficiency and optimality in data collection. Simulation results validate the effectiveness of the proposed method, showcasing its capability to ensure data transmission from sensors while minimizing energy costs. The generality of the framework, examined under realistic settings encompassing tasks, sensors, MSs, and networks, positions our method to offer insights into diverse monitoring applications such as temperature, sound, pollution levels, humidity, and more.

FUTURE WORK:

As a direction for future work, it would be intriguing to extend this method to more intricate network environments. Exploring scenarios involving distributed multi-MS cooperation could be a practical solution, providing a more comprehensive understanding of the challenges and opportunities in dynamic data collection systems. Overall, the proposed approach demonstrates promising results and opens avenues for further advancements in the field.