Table 2. Summary of clustering approaches

Protocol	Clustering type	Metrics addressed
Energy efficient fault-tolerant clustering algorithm[16]	Distributed energy-efficient and fault-tolerant clustering	Distance, energy, and cluster cardinality
WSN clustering for UAV-based data gathering[19]	Set-covering problem-based clustering	Minimum hops between CH and CMs
Energy-efficient and low packet loss clustering[20]	Kmeans++ and fuzzy logic -based clustering	Energy and storage of sensor nodes
Cluster-based communication topology selection[21]	PSO-based clustering	BER and UAV travel time
Bio-inspired clustering[22]	Dragonfly algorithm for cluster formation and management	Cluster building time and successful delivery probability
Clustering for sensor networks with mobile agents[23]	Optimal cluster size-based clustering	Energy consumption and latency
Distributed clustering in UWSNs[24]	Dynamic and distributed clustering	Energy consumption and stable and well-connected CHs
Self-adaptive energy-efficient operation[25,26]	Minority games and reinforcement learning-based clustering	Energy and UAV position
Socio-spatial resource management in wireless powered public safety networks[27]	Chinese restaurant process-based cluster formation	Communication interest, physical distance, and energy
Heuristic algorithm and cooperative relay for UWSNs[28]	Cooperative data relay and optimization scheme-based clustering	Average energy and flight distance
LEACH[17,32]	Distance-based distributed clustering	Network lifetime and latency

CH: cluster head; CM: cluster member; UAV: unmanned aerial vehicle; WSN: wireless sensor network; PSO: particle swarm optimization; BER: bit error rate; UWSN: UAV-aided WSN; LEACH: low-energy adaptive clustering hierarchy

are involved. In our proposed work, a WSN consists of homogenous sensor nodes distributed randomly in a ground area with the BS located at a distance away from the sensor nodes. The sensor nodes are static and identified with a distinct identifier. Sensor nodes use the same transmission power throughout the communication process. We considered a fixed-wing UAV for our work, which flies at constant height (h) and velocity (v) to collect data from CHs. The UAV is equipped with a directional antenna with flare angle $\theta = 60°$. The UAV is responsible for collecting the sensed data from the sensor nodes using the predefined path and then transferring the data to the BS. We have narrowed down our work towards the clustering approach in UWSNs, which is conducted after the deployment of sensor nodes and before the arrival of the UAV. Clustering is divided into CH selection and cluster formation. Sensor nodes are capable of sensing, storing the sensed data, and transmitting the data to their CH. However, the network topology may change because of the death of nodes due to energy depletion. According to a simplified energy consumption model in[32], the energy spent on transmitting and receiving k bits of data over distance d can be represented as

$$E_{Tx}(k,d) = \begin{cases} k \times E_{elec} + k \times E_{fs} \times d^2, d < d_o \\ k \times E_{elec} + k \times E_{amp\times} d^4, d \geq d_o \end{cases} \quad (1)$$

and

$$E_{Rx}(k) = k \times E_{elec} \quad (2)$$

respectively, where E_{elec} is the electrical energy for transmitter/receiver, E_{fs} is the free-space energy loss, E_{amp} is the energy for amplifier, and d_o is the threshold value of distance. The framework of a cluster-based UWSN is shown in Figure 1.

Cluster formation

There is a set of n sensor nodes, $S = \{s_1, s_2, s_3, \ldots, s_n\}$, and each sensor node s_i has z attributes (i.e., residual energy, energy consumption rate, *etc*.) represented by a vector $A = \{a_1, a_2, a_3, \ldots, a_n\}$. A sensor node s_i can be in one of the two states of CH and CM. The traffic load for CH and each CM is different, leading to unbalanced energy consumption in sensor nodes. As a result, few nodes may die very soon, which

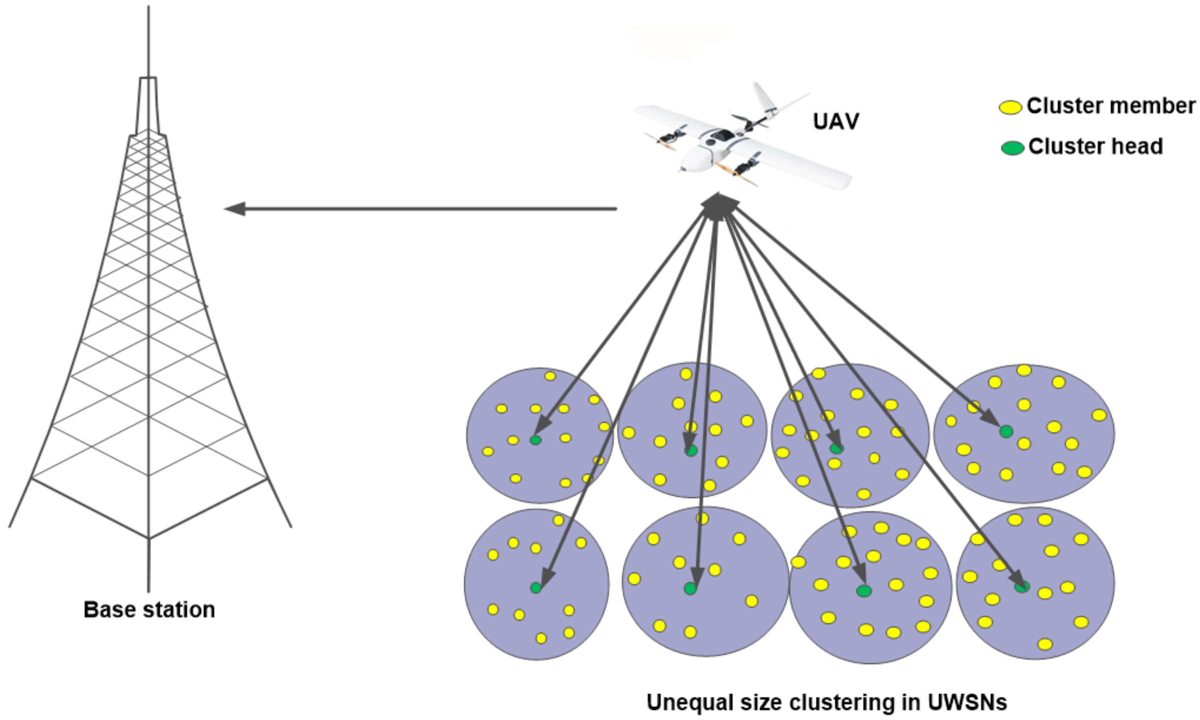

Figure 1. The architecture of a cluster-based UWSN. UAV: unmanned aerial vehicle; UWSN: UAV-aided wireless sensor network

ultimately affects the overall network operation. Owing to the energy constrained, the number of clusters and the number of CMs in a cluster should be proportional to the residual energy in CHs.

To resolve the issue mentioned above in UWSNs, a residual energy-based clustering scheme is proposed in this paper, in which each cluster has a different cluster size. Note that the cluster size means the number of sensor nodes in a cluster in this paper. After the sensor nodes are deployed in their application area, they start the cluster formation process. During the process, every node having energy higher than the minimum threshold energy (E_{thres}) transmits its status to all other nodes within its transmission range. Other nodes whose energy is below E_{thres} will not compete for the CH selection process. The status message includes node identifier, residual energy, and location coordinates. For any sensor node, the probability to be elected as CH for a particular round t is given by

$$P(t) = \left\{ \frac{E_{rem}(t)}{\sum_{i=1}^{n} E_i(t)} \times N_s \right\} \tag{3}$$

where $E_{rem}(t)$ is the remaining energy of CH at time t, N_s is the number of nodes competing for CH, and $\sum_{i=1}^{n} E_i(t)$ is the aggregated network energy. After CH selection in a particular region, no cluster is selected with the competition range ($Comp_{Ri}$) of that CH, which is calculated as

$$Comp_{Ri} = \alpha \frac{Max_{(d)} - d(C_i, C_k)}{Max_{(d)}} + (1 - \alpha) Range_{Ci} \tag{4}$$

where α is a constant whose value ranges between 0 and 1. $Range_{Ci}$ is the communication range of CH, $Max_{(d)}$ denotes maximum distance to select another CH, and $d(C_i, C_k)$ is the distance between two clusters.

Based on the energy level of CH, the size of the cluster (i.e., the number of sensor nodes in a cluster) is decided as

$$Cluster\_size = \frac{E_{rem}(t)}{E_{thres}} \times n \tag{5}$$

Algorithm 1. The algorithm is run on every sensor node before the arrival of UAV

```
Input: A set of n sensor nodes, S = {s1, s2, s3, ..., sn}, and the attributes of n sensor nodes, A = {a1, a2, a3, ..., an}
Output: CHs selection and cluster formation
// Cluster head election
    Every SN having energy higher than the threshold value broadcasts its energy status within its neighborhood. // broadcast
message includes node ID, residual energy, and location coordinates.
    SNs with highest energy are elected as CHs in a distributed manner.
      if SN is elected as CH // node with the highest energy
         CH calculates Cluster_size and broadcasts CH status within its neighborhood.
      else
         SN receives CH status from CHs and includes them in CH candidate list.
      end if
// Operation of cluster head
      for every CH do
         if Cluster_size limit is reached then
              Broadcasts Cluster_size limit status
         else if a join request is received from SN then
              Accept the join request message
              Send join reply message to the SN
         end if
      end for
// Operation of cluster member
      for every sensor node except CH do
         if receives Cluster_size limit status then
              Remove the CH from CH candidate list
              Find next CH with the lowest Comm_cost from CH candidate list
              Send join request to the CH
         else if receives join reply from CH then
              SN becomes a member of the CH.
         else
         end if
  end for
```

UAV: unmanned aerial vehicle; CH: cluster head; SN: sensor node

After calculating the *Cluster_size* value, CH broadcasts the CH status message. Depending on the signal strength, other nodes estimate the channel quality of the node-to-CH link and the distance away from the CH. The estimated values are used to find the communication cost between a sensor node (SN) and its CH as

$$Comm_{cost} = d_{SN-CH} \times \frac{1}{SNR_{SN-CH}} \quad (6)$$

where $d_{SN\text{-}CH}$ and $SNR_{SN\text{-}CH}$ are the distance from a node to CH and the signal-to-noise ratio between node and CH, respectively, which indicate the channel quality. The nodes request to join the CH that has the best value of communication cost. When the cluster size reaches the maximum, further requests are denied. Afterward, nodes request to join the next best CH on the basis of the next lowest value of communication cost. The cluster formation process is illustrated in Algorithm 1 and detailed with the help of Figure 2.

At the beginning of CH selection, every SN broadcasts its status (node ID, residual energy, and the location coordinates) within its neighborhood. The SN with the highest value of residual energy is elected as CH for that region. The competitive range of that CH is then calculated using equation (2). No more CH is then selected within that competitive range, which helps to limit the number of CHs. The selected CHs compute their *Cluster_size* as in equation (3), which takes the residual energy of CHs into account. On the other hand, CMs calculate $Comm_{cost}$ using equation (4) and maintain the CH candidate list in the increasing order of $Comm_{cost}$. SNs send join requests to the CH having the least $Comm_{cost}$ value. The CH accepts join requests until cluster size limit is met. If *Cluster_size* exceeds the limit, the CH then broadcasts cluster limit status. The SN that receives cluster limit status removes the CH from its candidate list and then sends a join request to the next CH. After clustering, every sensor node including CH senses the attributes and transmits them to its CH periodically. In every round of CH-to-UAV communication, the residual energy of CH is checked. If the energy of CH is found to be less than or equal to E_{thres}, a new CH is selected accordingly.

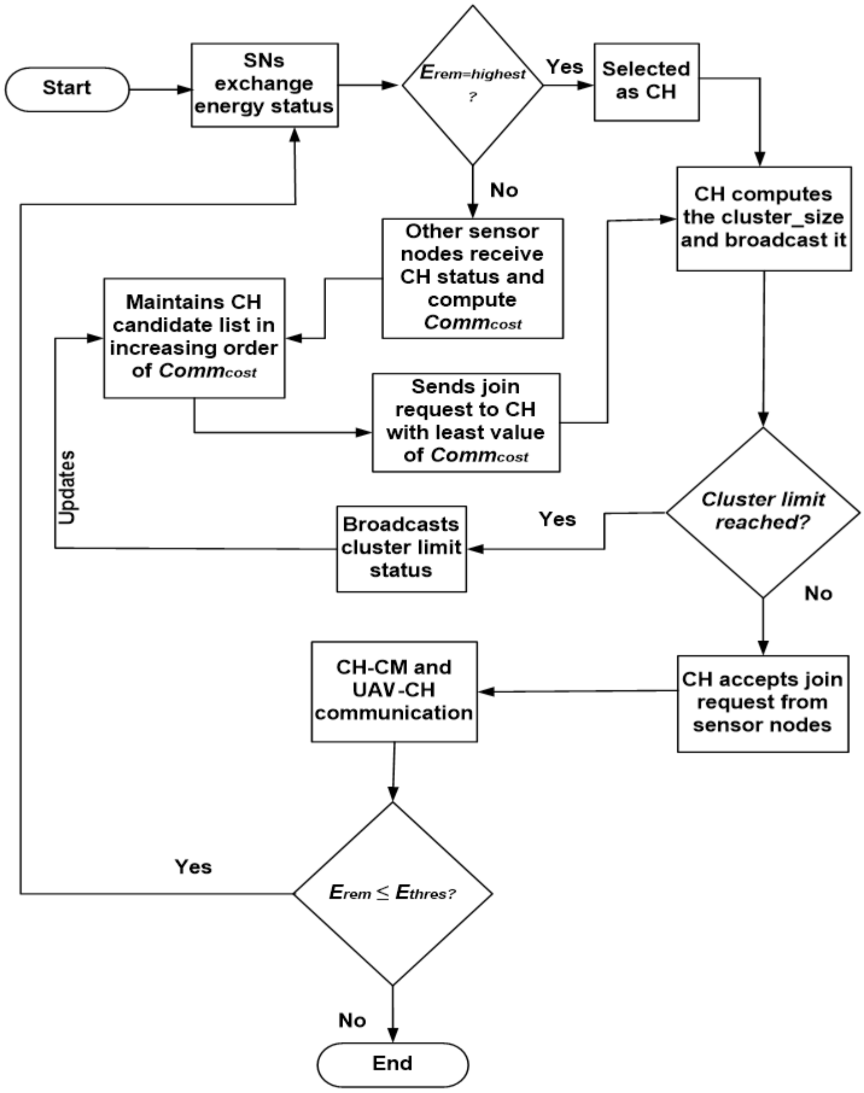

Figure 2. Flow diagram of the cluster formation process. UAV: unmanned aerial vehicle; CH: cluster head; CM: cluster member; SN: sensor node

In each round of communication, CMs sense the environment and transfer the sensed data to its CH. The energy consumed by any CM during inter-cluster transmission is given as

$$E_{CM} = E_{sense} + k \times E_{elec} + k \times E_{fs} \times d_{CM-CH}{}^{2} + E_{Rx}(k_{Ack}) \tag{7}$$

where E_{sense} is the energy dissipated while sensing, d_{CH} is the distance between CM and CH, and K_{aCK} is the length of acknowledgment packet coming from CH in bits. After a CH receives information from its CMs, it aggregates the data and transmits the aggregated data to the UAV. If E_{agg} is the energy used for data aggregation, the energy consumed by the CH is given by

$$E_{CH} = E_{Rx}(k) \times cluster\_size + k_{Ack} \times E_{elec} + k_{Ack} \times E_{fs} \times d_{CH-CM(i)}{}^{2} + R^{t} \times k \times E_{elec} + k \times E_{fs} \times d_{(CH-UAV)}{}^{2} \tag{8}$$

where $d_{CH\text{-}CM(i)}$ is the distance between CH and CM(i), $d_{(CH\text{-}UAV)}$ is the distance between UAV and CH, and R^t is the number of re-attempts before transmitting the data to UAV. Because the number of CMs in the cluster and the number of CHs in the network are defined and limited by the residual energy and the competition

Table 3. Simulation parameters

Parameter	Value
Network simulator	MATLAB
WSN deployment area	100 m × 100 m
Number of sensor nodes	50-200
Probability of CH	0.1
Initial energy	2 J
Radio energy dissipation for Tx/Rx (E_{elec})	50 nJ
Amplification energy (E_{mp})	100 pJ
Idle state energy	50 nJ
Free space loss	10 pJ
UAV's speed	10 m/s
UAV's height	15 m
Data aggregation energy	5 nJ
Sensing energy	2.7 nJ
Number of UAVs	1
Packet size	150 bytes
UAV's range	500 m
Sensor's range	200 m
Data rate	1 Mb/s

UAV: unmanned aerial vehicle; CH: cluster head

range of a node as in equations (4) and (5), the energy consumed by the nodes will be proportionate, which results in prolonged network lifetime.

RESULTS

In this section, we evaluate the performance of the proposed algorithm in terms of network lifetime and the number of delivered packets via the MATLAB simulation. The proposed scheme is then compared with the conventional clustering scheme LEACH[32] in which CH is selected based on the Euclidean distance with other nodes. LEACH is a widely used clustering algorithm for wireless communication. In LEACH, CH is chosen in random rotation periods, but the node energy is not considered for CH selection, leading to the early death of some sensor nodes.

Simulation environment

We conducted the simulation in an area of 100 m × 100 m. We assume that 50-200 sensor nodes are randomly deployed in the network area. We also consider that all the sensor nodes have the same communication range and the same initial energy. We consider the fixed-wing UAV with a constant speed of 10 m/s and a height of 15 m. The flight path and schedule of UAV are predefined. Table 3 summarizes the simulation parameters used in our performance study.

Performance metrics

WSNs are particularly used in remote and inaccessible places where limited energy always remains a major challenge. In our proposed algorithm, a UWSN is applicable for surveillance and monitoring applications where the maximum data collection with the minimum possible energy is highly desired. The performance metrics considered for our simulation are as follows:

Network lifetime: Interest in energy-efficient communication is attributed to the limitations imposed by battery-powered sensor nodes. The performance of the network is significantly degraded with the death of the sensor nodes. To help the extension of the sensor nodes' lifetime, a clustering algorithm that balances the rate of energy consumption in CH and CM is proposed. Residual energy-based clustering prevents the early death of the nodes by adjusting the amount of traffic to the nodes. In our simulation study, the network lifetime is measured by the number of alive nodes, as shown in Figure 3.

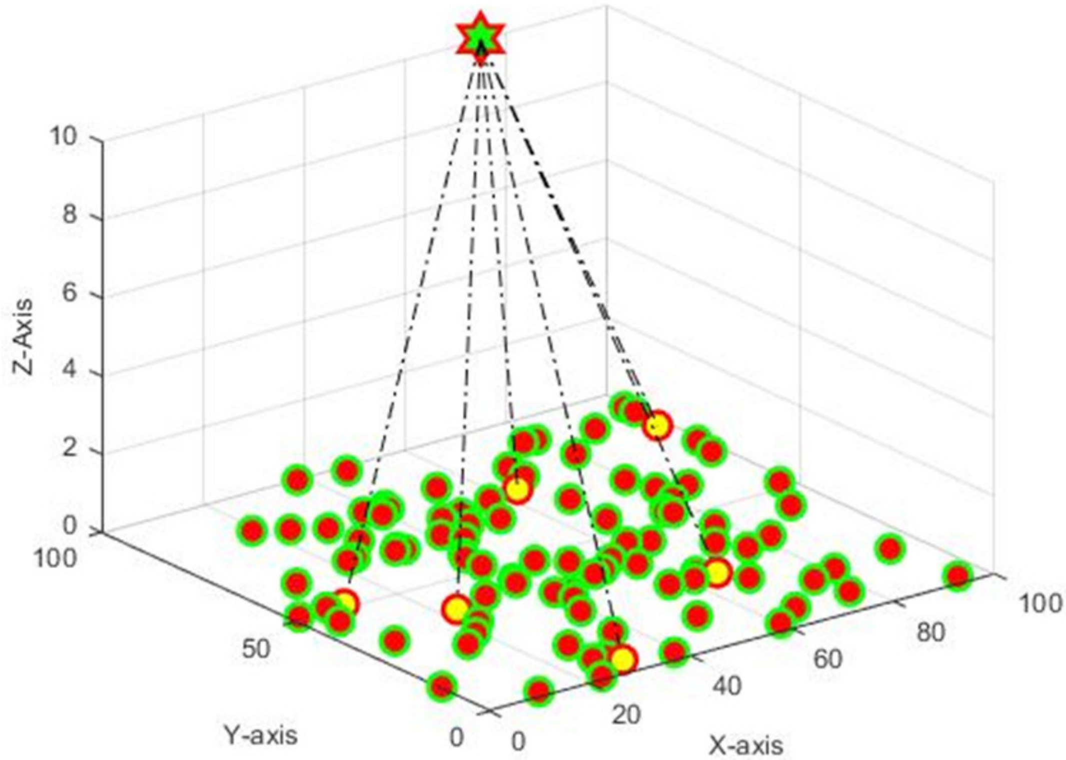

Figure 3. Unmanned aerial vehicle-aided data collection in wireless sensor networks

Number of delivered packets: Data collection efficiency is another essential metric that defines the performance of UWSNs. If the network experiences high packet loss due to the nodes' death, the BS may not make good decisions. This may result in the loss of enormous life and property in some critical applications. In our work, we have increased the number of collected packets by suppressing the early death of the sensor nodes due to energy depletion. The number of the delivered packet can also be used for the metric for data collection efficiency.

Simulation results and discussion

In this subsection, the simulation results are summarized and comparatively discussed. In Figure 3, UAV-based data collection from the ground sensor nodes is shown. For simplicity, the height and velocity of the UAV are kept constant. The UAV follows a simple predefined path and collects data from CHs during its flight. In the simulation, we varied the number of rounds and observed the results. Figure 4 represents the total number of alive nodes versus simulation time in rounds. The three different cases of network density are taken into simulation; i.e., 50, 100, and 200 nodes are deployed in the same network area. From Figure 4, we can infer that fewer sensor nodes are dead in the proposed clustering scheme compared to the conventional one[32] for the three different cases of network density. As a result, network lifetime is significantly prolonged with our residual energy-based clustering algorithm.

Figure 5 shows the number of packets that are delivered to the UAV versus simulation time in rounds. As in the network lifetime evaluation, the three different network densities (50, 100, and 200 nodes) are taken into simulation for the same network area. As the number of dead nodes is decreased in a network, the number of delivered packets (i.e., packet delivery ratio) is significantly improved. In other words, packet loss is markedly reduced. By comparing the performance of the proposed algorithm with that of the conventional one[32], it is easily inferred that the packet delivery ratio is significantly improved in the proposed clustering scheme.

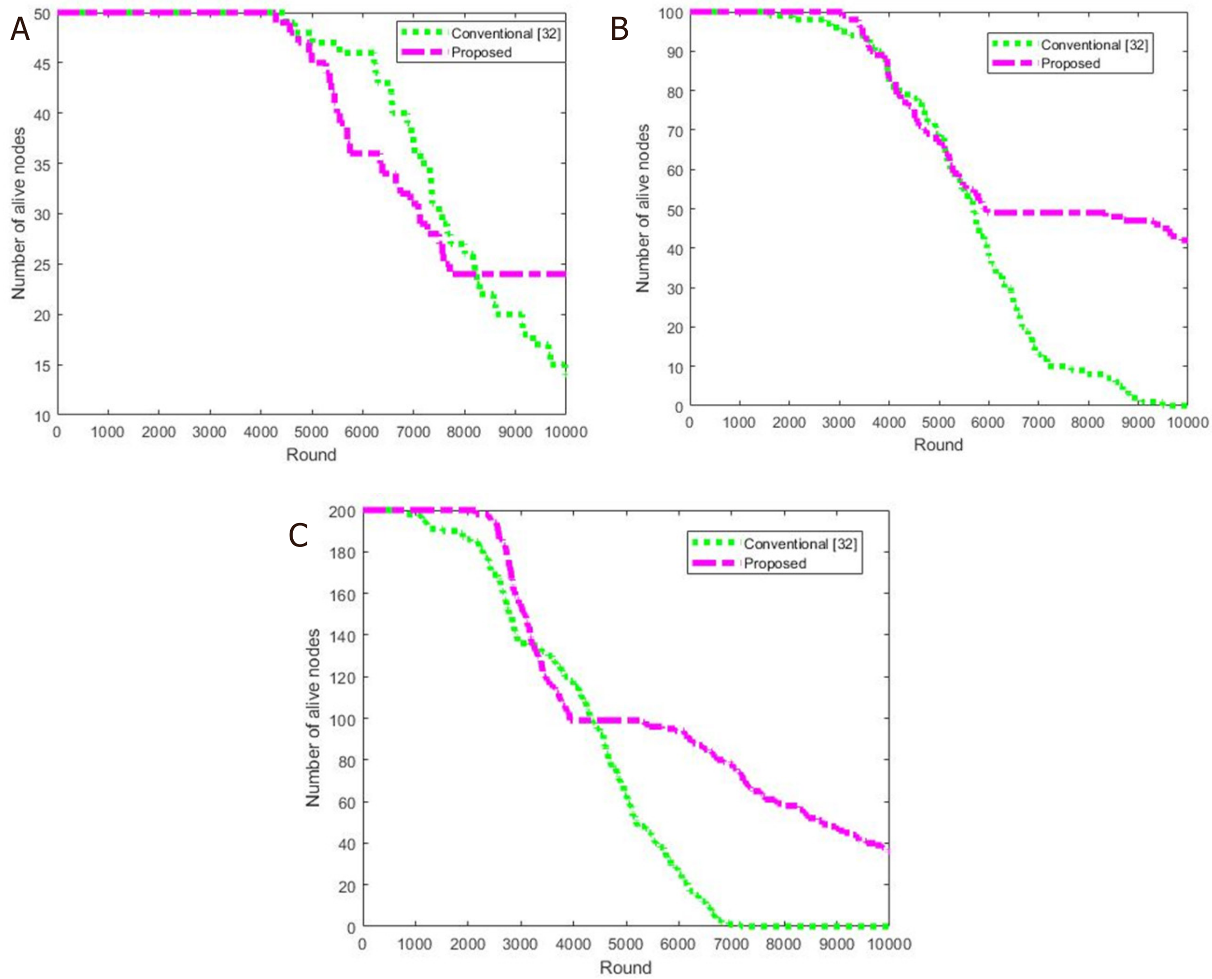

Figure 4. Network lifetime in rounds. A: the number of alive nodes when 200 nodes are used; B: the number of alive nodes when 100 nodes are used; C: the number of alive nodes when 50 nodes are used

In summary, the proposed clustering algorithm outperforms the conventional one in terms of network lifetime and data delivery ratio. This is mainly due to the fact that the residual energy is effectively exploited in clustering sensor nodes for energy-efficient sensor-to-UAV communication in the proposed clustering algorithm.

DISCUSSION

In UWSNs, the energy consumption is unbalanced due to the variable traffic load in the sensor nodes. As a result, some nodes die earlier, and thus, the network may not function properly. In this paper, this issue is addressed and resolved. That is, we proposed a residual energy-based clustering algorithm for sensor-to-UAV communication in UWSNs, in which cluster size is determined according to the available residual energy level. Our simulation results show that the proposed clustering scheme outperforms the conventional one in terms of network lifetime and data delivery ratio.

In this work, we narrowed our research towards the CH selection and cluster formation in UWSNs focusing on surveillance and monitoring applications, but an in-depth theoretical analysis and its validation are not addressed. We will come up with the in-depth analysis and its validation in a future work. UAV path planning for efficient data collection and other different issues such as secure communications and fault-tolerant mechanisms are left as another future work. Moreover, proper synchronization and cooperation

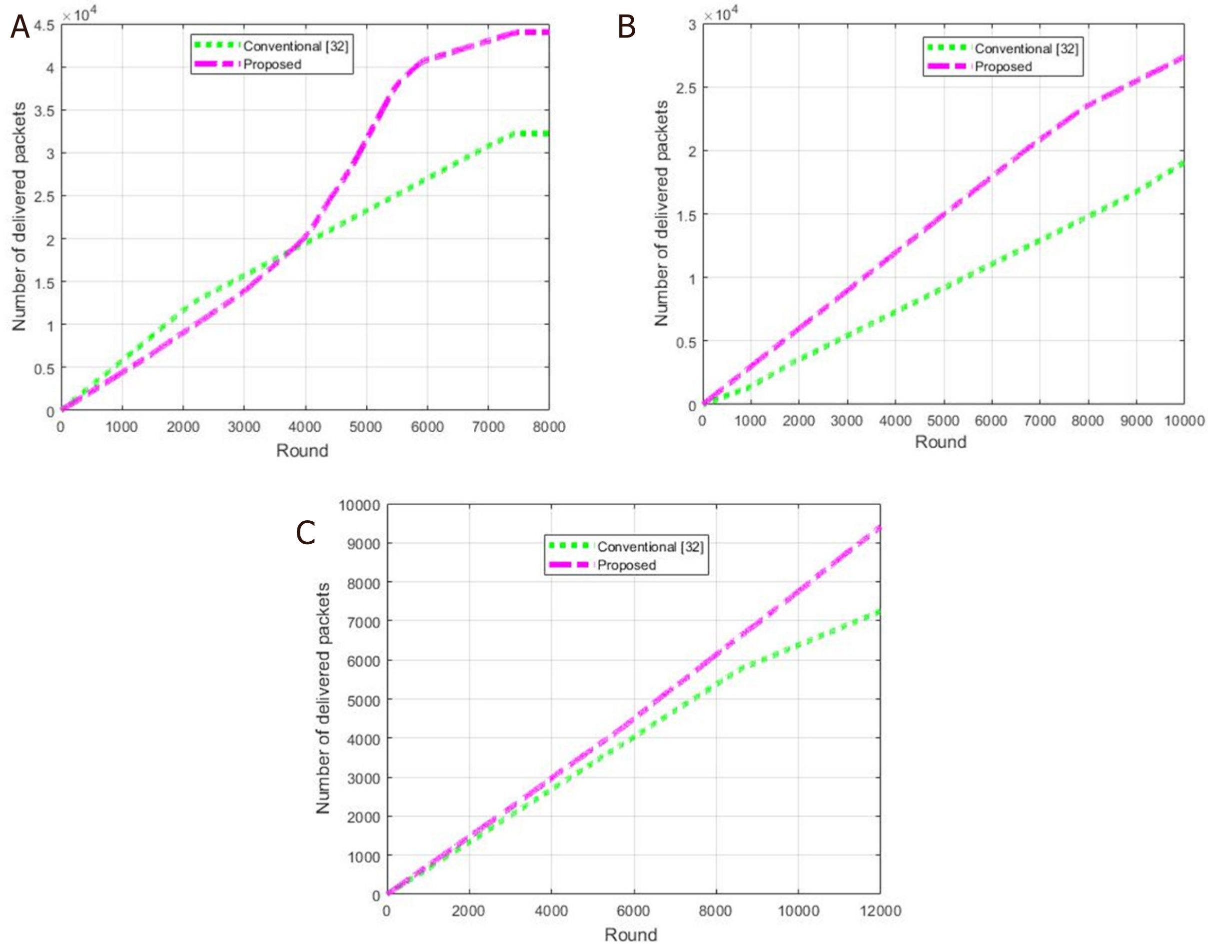

Figure 5. Packet delivery in rounds. A: the number of delivered packets when 200 nodes are used; B: the number of delivered packets when 100 nodes are used; C: the number of delivered packets when 50 nodes are used

between the flight time of UAV and the sleep/awake cycles of sensor nodes can be studied to reduce energy wastage during the periodic wake-up of sensors to check UAV's arrival.

DECLARATIONS

Acknowledgments

A preliminary version of this work was presented at the 8th International Conference on Smart Media and Applications, Guam, USA, December 2019[11]. The authors wish to thank the editor and anonymous referees for their helpful comments in improving the quality of this paper.

Authors' contributions

Developed and simulated the clustering algorithm: Poudel S
Directed the research and contributed to refining the algorithm and interpreting the simulation results: Moh S
Contributed to refining the algorithm in a formal form and reviewed the simulation results: Shen J
Drafted the paper: Poudel S
Revised and approved: Moh S, Shen J

Availability of data and materials

Not applicable.

Financial support and sponsorship
This work was supported in part by the National Research Foundation of Korea grant funded by the Korea government (MIST) (No. 2019R1F1A1060501).

Conflicts of interest
All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

REFERENCES
1. Zhang L, Wan C. Dynamic path planning design for mobile sink with burst traffic in a region of WSN. Wireless Comm Mobile Comput 2019;2019:1-8.
2. Khan AW, Abdullah AH, Anisi MH, Bangash JI. A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks. Sensors (Basel) 2014;14:2510-48.
3. Nazib RA, Moh S. Routing protocols for unmanned aerial vehicle-aided vehicular Ad Hoc networks: a survey. IEEE Access 2020;8:77535-60.
4. Ali ZA, Masroor S, Aamir M. UAV based data gathering in wireless sensor networks. Wireless Pers Commun 2019;106:1801-11.
5. Arafat MY, Moh S. Routing protocols for unmanned aerial vehicle networks: a survey. IEEE Access 2019;7:99694-720.
6. Poudel S, Moh S. Energy-efficient and fast MAC protocol in UAV-aided wireless sensor networks for time-critical applications. Sensors (Basel) 2020;20:2635.
7. Anjum SS, Noor RM, Anisi MH. Review on MANET based communication for search and rescue operations. Wireless Pers Commun 2017;94:31-52.
8. Arafat MY, Moh S. Location-aided delay tolerant routing protocol in UAV networks for post-disaster operation. IEEE Access 2018;6:59891-906.
9. Fragkos G, Tsiropoulou EE, Papavassiliou S. Disaster management and information transmission decision-making in public safety systems. 2019 IEEE Global Communications Conference (GLOBECOM). 2019 9-13 Dec. Waikoloa, HI, USA: IEEE; 2019. pp. 1-6.
10. Anisi MH, Rezazadeh J, Dehghan M. FEDA: fault-tolerant energy-efficient data aggregation in wireless sensor networks. Proc. of 16th Int. Conf. on Software, Telecommunications and Computer networks; 2008 Sep 25-27. Split, Dubrovnik, Croatia: IEEE; 2008, pp. 188-92.
11. Poudel S, Moh S. Unequal-size clustering approach for unmanned aerial vehicle-aided wireless sensor networks. Proc. of 8th Int. Conf. on Smart Media and Applications; 2019 Dec 5-7. Guam, USA; 2019. pp. 133-6.
12. Poudel S, Moh S. Medium access control protocols for unmanned aerial vehicle-aided wireless sensor networks: a survey. IEEE Access 2019;7:65728-44.
13. Anisi MH, Abdullah AH, Razak SA. Energy-efficient and reliable data delivery in wireless sensor networks. Wireless Netw 2013;19:495-505.
14. Hwang S, Su Y, Lu K, Dow C. A cluster-based approach for efficient multi-source multicasting in MANETs. Wireless Pers Commun 2011;57:255-75.
15. Sirsikar S, Wankhede K. Comparison of clustering algorithms to design new clustering approach. Procedia Computer Science 2015;49:147-54.
16. Azharuddin M, Kuila P, Jana PK. Energy efficient fault tolerant clustering and routing algorithms for wireless sensor networks. Comput Electr Eng 2015;41:177-90.
17. Abu Salem AO, Shudifat N. Enhanced LEACH protocol for increasing a lifetime of WSNs. Pers Ubiquit Comput 2019;23:901-7.
18. Wang G, Lee B, Ahn JY, Cho G. A UAV-aided cluster head election framework and applying such to security-driven cluster head election schemes: a survey. Secur Commun Netw 2018;2018:1-17.
19. Mi J, Wen X, Sun C, Lu Z, Jing W. Energy-efficient and low package loss clustering in UAV-assisted WSN using Kmeans++ and Fuzzy Logic. Proc. of International Conference on Communications Workshops in China (ICCC Workshops); 2019 Aug 11-13, Changchun, China. IEEE Access; 2019; pp. 210-5.
20. Tazibt CY, Bekhti M, Djamah T, Achir N, Boussetta K. Wireless sensor network clustering for UAV-based data gathering. Proc. of WD 2017: Wireless Days 2017; 2017 March 29-31, Porto, Portugal. IEEE; 2017. pp. 245-7.
21. Ho DT, Grøtli EI, Sujit PB, Johansen TA, Sousa JB. Cluster-based communication topology selection and UAV path planning in wireless

sensor networks. In proc. of International Conference on Unmanned Aircraft Systems (ICUAS); 2013 May 28-31, Atlanta, GA, USA. IEEE; 2013. pp. 59-68.

22. Aftab F, Khan A, Zhang Z. Bio-inspired clustering scheme for Internet of Drones application in industrial wireless sensor network. Int J Distrib Sens N 2019;15:155014771988990.
23. Lotfinezhad M, Liang B. Energy Efficient Clustering in Sensor Networks with Mobile Agents, proc. of the IEEE Wireless Communications and Networking Conference (WCNC'05); 2005, Mar. 13-17, New Orleans, LA USA: IEEE; 2005. pp. 1872-7.
24. Okcu H, Soyturk M. Distributed clustering approach for UAV integrated wireless sensor networks. IJAHUC 2014;15:106.
25. Sikeridis D, Tsiropoulou E, Devetsikiotis M, Papavassiliou S. Self-adaptive energy efficient operation in UAV-assisted public safety networks. Proc. of 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC); 2018, June 25-28. Kalamata, Greece: IEEE; 2018. pp. 1-5.
26. Sikeridis D, Tsiropoulou EE, Devetsikiotis M, Papavassiliou S. Wireless powered public safety IoT: a UAV-assisted adaptive-learning approach towards energy efficiency. J Netw Comput Appl 2018;123:69-79.
27. Sikeridis D, Tsiropoulou E, Devetsikiotis M, Papavassiliou S. Socio-spatial resource management in wireless powered public safety networks. Proc. of MILCOM 2018 - 2018 IEEE Military Communications Conference (MILCOM) 2018, Oct. 29-31 Los Angeles, CA, USA. IEEE; 2019. pp. 810-5.
28. Ho DT, Grøtli EI, Johansen TA. Heuristic algorithm and cooperative relay for energy efficient data collection with a UAV and WSN. Prof of International Conference on Computing, Management and Telecommunications (ComManTel); 2013, Jan 21-24, Ho Chi Minh City, Vietnam. IEEE; 2013. pp. 346-51.
29. Arafat MY, Moh S. Comparison of clustering algorithms based on weighted clustering metrics for unmanned aerial vehicle networks. In Proc. 18th Int. Conf. Electron. Inf. Comm. 2019 January 22-25, Auckland, New Zealand, (ICEIC). pp. 1-4.
30. Arafat MY, Moh S. A survey on cluster-based routing protocols for unmanned aerial vehicle networks. IEEE Access 2019;7:498-516.
31. Jan B, Farman H, Javed H, Montrucchio B, Khan M, et al. Energy efficient hierarchical clustering approaches in wireless sensor networks: a survey. Wirel Commun Mob Com 2017;2017:1-14.
32. Heinzelman W, Chandrakasan A, Balakrishnan H. An application-specific protocol architecture for wireless micro sensor networks. IEEE Trans Wireless Commun 2002;1:660-70.