IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A REVIEW ON EFFICENT METHODS OF DATA COLLECTION IN WIRELESS SENSOR NETWORKS

Christian Munyakayanza Gilles[1*], Ge Bin[2]

[1]Department of computer science and engineering, Anhui University of Science and Technology, Huainan, Anhui Province, 232001, China.

[2]Department of computer science and engineering, Anhui University of Science and Technology, Huainan, Anhui Province, 232001, China.

[*]mchrisgilles@gmail.com

Abstract. A wireless sensor network is a group of sensors nodes placed strategically in a field of interest to extract particular information in a determined range and then transmit this information to a base station also known as the sink node then feed into the internet. The data collection process is the main goal in a WSN. To achieve the aforementioned goal, we will consider sensing methods such as one-hop transmission, multi-hop transmission, or clustering. These methods are effective in various application demands for a collection of reliable data from our sensor nodes. The WSNs sensor nodes either battery-powered or harvested energy-powered, energy consumption still play a crucial role in the inaccuracy of data sensed. Therefore, we have to design better network communication protocols and network organizations to conserve energy. This paper highlights several adaptive methods by reviewing their approaches solutions and performance in data handling that contribute to numerous issues and challenges such as power consumption constraints, node failure, packet collision, and time delay.

Keywords: WSN; Data collection; Network communication; Power consumption constraints;

1. Introduction

The rise of the internet of things has improved our lives with the use of WSNs. They are the interface of the IoT system to the physical world[1], widely used in many application areas like in health monitoring, precision agriculture, environment (e.g., pollution), intelligent transportation, smart home, smart city, as well as other areas of human activity. The expansion of IoT has the potential to reshape the future of the industry. According to [2] [3], connected devices will reach 50 billion by 2020. This target is to reach by continuing to simplify how things connect and communicate today to enable a better human-computer interaction. Nowadays, machines with human sensory capabilities can make decisions without human input, in particular, the 5G era we are entering with wireless network latency promising to be as low as 1 to 2 milliseconds[4].

In the past decade, WSN's became smaller, affordable, and practical. Which makes them usable in diverse environments and locations difficult to access. However, since the sensor nodes batteries are impossible to replace after exhaustion, an additional source of energy harvesting mechanism for charging the node batteries is considerable although it still has various limitations. Sensor nodes closer to the sink node consume more energy compared to other nodes, which cause energy holes problems and data loss through data transmission. To minimize energy consumption in the network, we can reduce the distance data travels from the source sensor node to the sink with data funneling a data-gathering method. For a proper data collection, we will consider several approaches like mobile data collectors, mobile sink, or rechargeable sensor nodes.

In this paper, we present a review of data collection efficient methods in wireless sensor networks. i.e., balancing energy consumption with an organized network configuration to avoid individual node failure that results in lossless data gathering. Indeed, they are many issues to be discussed among them, like the issue of maintaining network connectivity while maximizing coverage [5]. Thus, recent years researchers have been focusing more on energy optimization methods with concepts like data aggregation that we will also consider in our review as well as data routing and energy harvesting. We will review researches that proposed data collection schemes with various algorithms and protocols especially the ones that emphasize energy-efficient models.

The rest of the paper is organize as follows. Section II presents a brief review of energy-efficient data compression methods in WSN data collection. Section III reviews energy-efficient routing protocols. Section IV describes the capabilities of the energy harvesting systems to increases the lifetime of the sensor nodes and power overall in the network. Finally, Section V conclude and discusses the reviewed data collection methods.

2. Data compression

Data compression is effective in consuming less energy during data transmission and storing data. Normally, compressing data packets during data transmission is a practical way of eliminating redundancy in data sensed. First, we analyze the original data to eliminate the redundant information then we code the data. Hence, we will consider lossy compression to achieve higher compression ratio and lossless compression to maximize the lifetime of sensor nodes in our methods, since both are necessary for energy-efficient WSNs. Therefore, some of the efficient data compression methods are introduced in [6][7][8][9][10][11] using the most popular technique known as Huffman coding to encode the data from the source nodes and decoded on arrival at the sink node. The recent breakthroughs in achieving a better compression method suitable for WSNs were the Marcelloni Static Huffman algorithm. It is a lossless compression method, suitable for reducing storage space and low computational resources because it leaves more resources for the nodes to do more tasks compared to other compression algorithms. It uses the high correlation of data collected by the sensor nodes to achieve high compression ratios. It achieves the compression performance of 66.9% and 67.3% for temperature and relative humidity values collected by the sensor node. Another data compression method using Huffman coding that is found to be more efficient is the modified adaptive Huffman algorithm. It combines effectively the advantages of both static and adaptive Huffman algorithms by grouping the data into sets as in static Huffman algorithm and the Tree construction concept from the adaptive Huffman algorithm to increase the compression ratio.

The compression metric to compute the efficiency of this algorithm in Equation (1)

Compression ratio = [1 - {(compressed no. of bits)/original bits}] * 100 **(1)**

Table 1. Pseudocode of the modified adaptive Huffman algorithm

<table>
<tr><th colspan="2">Algorithm 1: Pseudocode for Huffman encoding</th></tr>
<tr><td>module modifiedhuffman_encode(diff , btree)
// Creation of root in transmitter.
createroot ( )
// Traversal of tree in search of node.
search (diff , btree)
IF (node present)
prefix = traverse (node)
suffix = arrayindex (diff)
increment node->wt
ELSE</td><td>prefix = traverse (nyt)
IF(diff >= 0)
suffix = $diff_4$|
IF(diff < 0)
suffix = $diff_5$|
nyt->lchild = createnyt( )
// Insertion of node in the tree.
nyt->rchild = createnode (diff)
nyt = nyt ->lchild
code = <<prefix,suffix>>
update (btree)
balance(btree)
endmodule</td></tr>
</table>

Algorithm 2: *Pseudocode for Huffman Decoding*

```
Module
modifiedhuffman_decode( code
[ ], btree )
    // Creation of root in the
receiver.
    createroot( )
    // Traversal of tree in
search of node.
    traverse ( btree )
    {
             IF( code[ i ] =
0 )
             curr = curr ->
lchild
             IF( code[ i ] =
1 )
            curr = curr ->
rchild
    }
   IF( curr = node )
             read suffix
           data = data
[ suffix
           increment node->wt
 // Encoding of data and
insertion of node in the
tree.
 IF ( curr = nyt )
            read suffix
      IF( suffix = 5bits)
         data = suffix4|
      IF(suffix = 4 bits)
         data = suffix
         nyt->lchild =
createnyt()
   nyt->rchild =
createnode(diff)
        nyt = nyt->lchild
     update(btree)
    balance(btree)
endmodule
```

The data compression method in [12] proposed a low-cost lossy compression approach with error bound guarantee. It reduces data congestion and minimizes energy dissipation by the utilization of a spatiotemporal correlation between data samples. Based on a compressing neural network a machine-learning algorithm automatically predicts human activities and environmental conditions. The performance evaluation of the proposed algorithms uses meteorological datasets and produced better compression results than other methods and better signal reconstruction.

Recently, research in IWSNs been focusing on mining the correlation of sensory data to reduce the amount of data transmission. In [13], the authors proposed a layered adaptive compression design for efficient data collection (LACD-EDC) in industrial wireless sensor networks(IWSNs). It focuses on multilayer network architecture to support the exploration of spatiotemporal correlations in sensory nodes and cluster heads nodes then compresses its data. Later at the sink node, is the decoding of the compressed data to recover the original data and achieve an approximate data collection. LACD-EDC also proposed an adaptive dictionary that enhanced data recovery quality. The results of the proposed method achieved an efficient data collection with approximate data recovery with high accuracy.

The data compression method in [14] developed a novel energy-efficient framework for clustered WSN which combines data prediction, compression, and recovery techniques. The main goal was to minimize the amount of data transmission using data prediction and compression techniques while still guarantee data accuracy. In this method, they used public sensor dataset in the test. Results showed that the method was efficient for the continuous monitoring of WSN, idea for environmental application in clustered WSN.

3. Routing protocols

Routing protocols vary depending on their application. Various protocols serve the purpose but also drain the battery of the sensor nodes due to frequent data transfer when we use a less suitable protocol. WSN's employs some well-known routing tactics such as data aggregation and clustering.

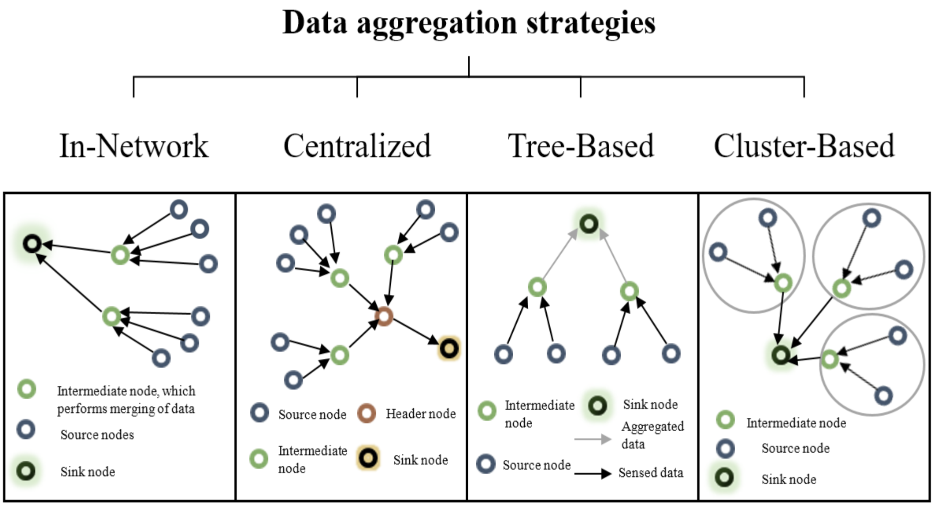

Fig. 1 Data aggregation

Data aggregation method in [15, 16] is used to reduce energy consumption and eliminate redundancy in data transmission while routing packets from the source node to the base station thereby reducing the bandwidth usage of the network. We have to handle sensor nodes in an energy-efficient manner throughout data aggregation to maximize the lifetime of the sensors network. For secure data aggregation, in this method we propose the Trust wEight Secure Data Aggregation algorithm (TESDA). During data transmission, a simulation proves that by using the data aggregation method TESDA, it was more secure and more energy-efficient than other methods like HEF or SDAF. From simulation results, the TESDA compared to SDAF reduces data aggregation deviation and achieves a reduction in latency by 14, 43%, and a reduction in energy consumption by 26.8%.

The clustering method in [17] is an important mechanism with a periodic rotation that can improve the lifetime of a network. It separates the communication process between sensor nodes and the base station into rounds. In [18] authors introduced Hierarchical distributed management clustering protocol(HDMC) in which each node tends to become a Cluster Head(CH), by the help of a judge who will compute tendency information sent by each node using formula Eq.(2) and select the suitable node as CH based on its history, its available resources, and the coverage area. The selected CH's in each round will be the judge in the next clustering round. The nodes are not aware of their neighbor's intention of becoming CH in the area when competing for the CH role. The HDMC compared to other popular protocols like HEED or LEACH gives a longer network lifetime.

Inclination (Node – x [Tn+1]) = (ß1 * Activity – history (Node – x [Tn+1])) + (ß2 * Energy (Node – x [Tn+1])) + (ß3 * Overlap (Node – x [Tn+1]). **(2)**

In this formula, *X* represents each node computing its tendency. ß1, ß2, and ß3 represent our three weight parameters multiplied by activity history, energy, and overlapping parameters. Each node shows its tendency in a range between zero and one. With all three parameters equal to one for each node, the node will not send the same tendency to all the CH requests because of the distance to the CH being different from the third parameter. Therefore, each node calculates the tendency for each request individually.

The fixed clustering method in [19] proposed an Energy-Aware Multi-hop Routing protocol (EAMR) In which a sensor node is fixed to a cluster permanently to reduce excessive overhead and the number of cluster head changes. This protocol operation is composed of two major phases: the set-up phase and the steady-state phase. During the setup phase, first CHs are elected, then form cluster groups with cluster members, and to complete the phase each CH selects another CH as its relay nodes nearby the base station to help inter-cluster communication. The steady-state phase comes next after the setup phase; it consists of two parts: EAMR collects data from the sensor nodes to the base station, furthermore, execute the cluster head change if needed. In each round, cluster head change occurs only when its energy level is lower than the threshold value (ThV). Further, in the next round, a new cluster head will be selected among the cluster members of the previous one.

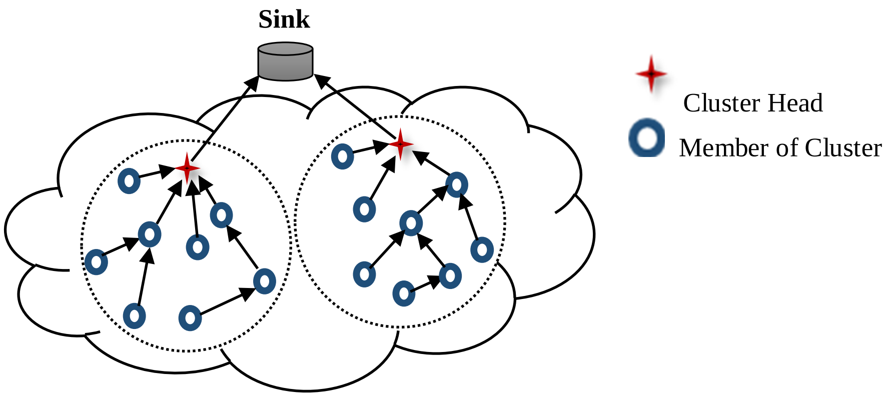

Fig 2. Clustered sensor network

4. Energy Harvesting

Energy harvesting is a technique that harvests, draws, or gathers a variety of unused surrounding energy sources(e.g., solar, thermal, vibration, and wind) and converts the harvested energy into electrical energy to recharge the batteries[20]. Through these energy sources, solar energy has the highest density and is the most used harvested energy source. Energy harvesting downside is low densities of power, periodical power supply in some situations (e.g., solar energy at night), and energy drain before the next harvesting cycle. During data collection, processing and sensing units consume more power. Further methods have been working on solving the aforementioned issues in joint data collection and energy harvesting.

Energy harvesting in WSN is only applicable in various areas. It applies mostly in areas like agricultural, bridge monitors, electricity usage meters, health monitors, gas sensors, HVAC controls, building automation, light switches, remote pipeline monitors, and water meters[21].

Energy harvesting is an important method that can help maximize the overall lifetime of the WSN and simultaneously strengthen the transmission coverage and sensing reliability of the sensor nodes. This harvesting method can use harvested energy to recharge our hardly accessible nodes battery or can use the same energy directly[1]. Among all aforementioned energy sources, RF harvesters provided features more suitable for wireless sensor networks. RF energy harvesting converts signals energy into DC power.

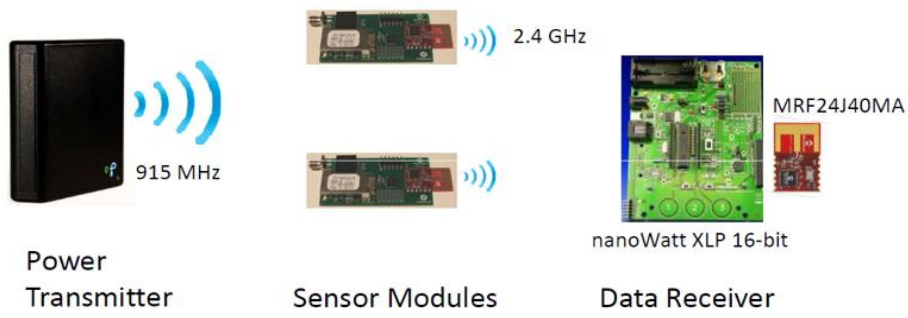

Fig. 3. RF energy transmission

RF energy is available in a wide array of frequency bands due to everyday technologies such as cellphone towers; TV broadcast stations, satellite, Wi-Fi routers, laptops, or any other communications networks. However, Radio Frequency (RF) has the lowest power density only a tiny amount of power can be harvested making it fitter for low power applications and short distance WSNs. In[22, 23] Devices with low-power consumption sensors can store the harvested power in supercapacitors to provide fast charging between cycles and can prolong WSNs lifetime by many years. Each communication cycle can last about a few days or weeks depending on the application. To enhance communication, simultaneous information transmission, and WPT (wireless power transfer) is proposed in this case because of the super-capacitor self-discharge between cycles. RF energy harvesting is demanding more research due to its potential to provide a long-lasting power supply.

Energy harvesting, for example, has been approached by mobile healthcare services. In [24] authors introduced body sensor networks(BSNs), low power wearable systems for continuous monitoring of patients in diverse environments outside the clinic for chronic respiratory disease. BSNs energy is harvested from thermal radiation or body heat and motion. It consists of three BSNs: wristband, chest patch, and spirometer consuming 0.83, 0.96, and 0.01 mill watts throughout data collection from sensor nodes to cloud base server storage.

Energy harvesting fair scheduling is another method proposed in [25] where authors study the problem in rescheduling optimization of energy harvesting mobile sensor nodes attached to cattle as a collar. Equipped with a solar panel and a wireless power transfer (WPT) receiver that harvest energy from the WPT transmitter nearby the animals, to improve the collection of the recorded biological data from each node to the base station. They demonstrate that the problem is NP-complete. They simulated the EHFS algorithm for the schedule optimization of data collection by prioritizing nodes depending on their radio link quality and energy harvested. It can maximize the amount of data collected and guaranteeing fair data reception.

Table 2. Comparison between the optimal solutions and the EHFS algorithm

Nodes	AMPL (Cplex)		EHFS	
	Packets	Runtime	Packets	Runtime
1	2499	1 s	2491	0.07 s
2	4999	12 s	4981	0.1 s
3	7499	28 s	7475	0.04 s
4	9998	63 s	9954	0.06 s
5	12498	1 m 27 s	12484	0.06 s
6	14998	5 m 15 s	14465	0.06 s
7	17498	l h 3 m	17353	0.08 s
8	19997	6 h 53 m	19808	0.08 s
9	22499	19 h 12 m	21793	0.22 s
10	24998	36 h 29 m	24583	0.22 s

The EHFS algorithm tests show no dead nodes. Compare to the optimal solution, it has a maximum difference of 706 when N= 9 with the number of packets less than the AMPL output by 1.16%. It outperforms Greedy-scheduling algorithms LE and HP by nearly 1.7 times. Also making it much more efficient than the optimization model on runtime.

Energy harvesting technology in rechargeable wireless sensor networks (RWSN) was proposed In [26] where the author designed a node-Gosper island-based scalable hierarchical cluster transmission method in conjunction with a wireless recharging plan for data collection in RWSNs. He provides an energy harvesting technology method using a robot car equipped with an electromagnetic resonance power transmitter, a high capacity battery, a wireless communication transceiver, and a GPS to recharge the relay sensor nodes that consume a lot of energy due to long-range data transmission to the sink node. A second robot car equipped with a low power radio frequency power transmitter was used to recharge normal sensors because of the low energy consumption with short-range data transmission to the relay nodes.

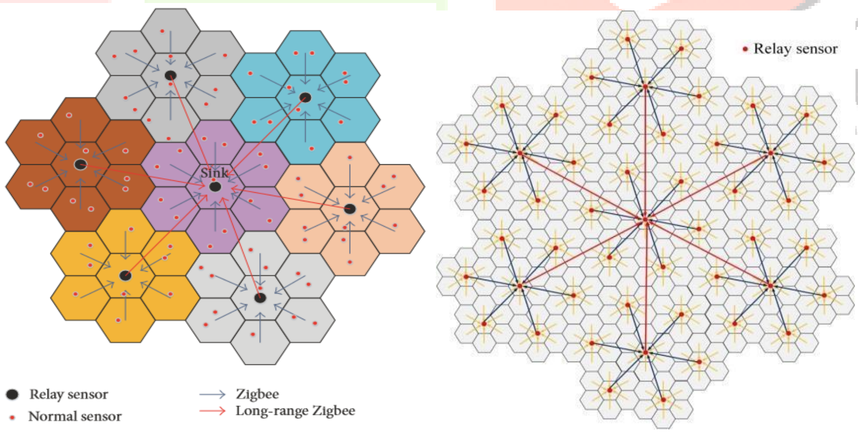

Fig.2. Data collection path in node-Gosper Island

The proposed method outperformed three well known traditional data collection methods (direct transmission, the LEACH, and multihop with minimum hop count) in terms of energy consumption. The wireless recharging method proposed using node-Gosper curve was also compared to other curve methods(SCAN, Hilbert curve, S-Curve (ad), and Z-Curve) and the results prove that the node-Gosper curve had the shortest charge path length, more alive nodes over the time and better traveling efficiency than other curves.

5. Conclusion

In this study, we reviewed three types of data collection methods of WSNs by introducing their purpose and the challenging tasks they are tackling. To deploy our sensors in diverse environments, we need proper protocols to handle energy levels in the sensor nodes. First, we provided data compression approaches suitable for various applications and that can conserve energy. Second, we discuss the routing protocols and secure data aggregation methods to solve packets collision, and latency that occurs during data transmission from the source node to the sink node. Further, we discussed the capabilities of energy harvesting methods and their applications. We also mentioned efficient conversion, storage of harvested energy, and other efficient methods that are still in development to meet the power requirement. In summary, although these recent research studies did achieve remarkable results, in future studies there are still great expectations through technological improvements like power harvesting technologies that have the potential to make our sensor node truly autonomous and self-sustainable in WSN.

REFERENCES

1. Geng, H.: Internet of things and data analytics handbook. (2017)

2. Chase, J.: The Evolution of the Internet of Things. Texas Instruments. (2013). https://doi.org/10.1109/ICCT.2015.7399787

3. Liu, X., Liu, W., Liu, Y., Song, H., Liu, A., Liu, X.: A trust and priority based code updated approach to guarantee security for vehicles network. IEEE Access. (2018). https://doi.org/10.1109/ACCESS.2018.2872787

4. Alsharif, M.H., Kim, S., Kuruoğlu, N.: Energy harvesting techniques for wireless sensor networks/radio-frequency identification: A review, (2019)

5. Boukerche, A., Sun, P.: Ad Hoc Networks Connectivity and coverage based protocols for wireless sensor networks. Ad Hoc Networks. 80, 54–69 (2018). https://doi.org/10.1016/j.adhoc.2018.07.003

6. Uthayakumar, J., Vengattaraman, T., Dhavachelvan, P.: A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications, (2018)

7. Medeiros, H.P., Maciel, M.C., Demo Souza, R., Pellenz, M.E.: Lightweight data compression in wireless sensor networks using Huffman coding. Int. J. Distrib. Sens. Networks. (2014). https://doi.org/10.1155/2014/672921

8. Marcelloni, F., Vecchio, M.: A simple algorithm for data compression in wireless sensor networks. IEEE Commun. Lett. (2008). https://doi.org/10.1109/LCOMM.2008.080300

9. Journal, I., Engineering, O.F., Compression, D., Using, T., Code, H., Wireless, F.O.R., Network, S.: DATA COMPRESSION TECHNIQUE USING HUFFMAN CODE FOR WIRELESS. 4, 374–378 (2015)

10. Li, Y., Liang, Y.: Temporal lossless and lossy compression in wireless sensor networks. ACM Trans. Sens. Networks. 12, (2016). https://doi.org/10.1145/2990196

11. Tharini, C., Vanaja Ranjan, P.: Design of modified adaptive Huffman data compression algorithm for wireless sensor network. J. Comput. Sci. 5, 466–470 (2009). https://doi.org/10.3844/jcssp.2009.466.470

12. Abu Alsheikh, M., Lin, S., Niyato, D., Tan, H.P.: Rate-distortion balanced data compression for wireless sensor networks. In: IEEE Sensors Journal (2016)

13. Chen, S., Zhang, S., Zheng, X., Ruan, X.: Layered adaptive compression design for efficient data collection in industrial wireless sensor networks. J. Netw. Comput. Appl. (2019). https://doi.org/10.1016/j.jnca.2019.01.002

14. Wu, M., Tan, L., Xiong, N.: Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications. Inf. Sci. (Ny). (2016). https://doi.org/10.1016/j.ins.2015.10.004

15. Padmaja, P., Marutheswar, G. V.: Energy efficient data aggregation in wireless sensor networks. In: Materials Today: Proceedings (2018)

16. Pachlor, R., Shrimankar, D.: VCH-ECCR: A Centralized Routing Protocol for Wireless Sensor Networks. J. Sensors. (2017). https://doi.org/10.1155/2017/8946576

17. Fanian, F., Kuchaki Rafsanjani, M.: Cluster-based routing protocols in wireless sensor networks: A survey based on methodology, (2019)

18. Shahraki, A., Kuchaki Rafsanjani, M., Borumand Saeid, A.: Hierarchical distributed management clustering protocol for wireless sensor networks. Telecommun. Syst. (2017). https://doi.org/10.1007/s11235-016-0218-7

19. Cengiz, K., Dag, T.: Energy Aware Multi-Hop Routing Protocol for WSNs. IEEE Access. (2017). https://doi.org/10.1109/ACCESS.2017.2784542

20. Tan, Y.K.: Energy harvesting autonomous sensor systems: Design, analysis, and practical implementation. (2017)

21. Geng, H.: Internet of things and data analytics handbook. (2017)

22. Kang, X., Ho, C.K., Sun, S.: Full-Duplex Wireless-Powered Communication Network With Energy Causality. IEEE Trans. Wirel. Commun. (2015). https://doi.org/10.1109/TWC.2015.2439673

23. Tran, H., Åkerberg, J., Björkman, M., Tran, H.V.: RF energy harvesting: an analysis of wireless sensor networks for reliable communication. Wirel. Networks. (2019). https://doi.org/10.1007/s11276-017-1546-6

24. Dieffenderfer, J., Goodell, H., Mills, S., McKnight, M., Yao, S., Lin, F., Beppler, E., Bent, B., Lee, B., Misra, V., Zhu, Y., Oralkan, O.,

Strohmaier, J., Muth, J., Peden, D., Bozkurt, A.: Low-Power Wearable Systems for Continuous Monitoring of Environment and Health for Chronic Respiratory Disease. IEEE J. Biomed. Heal. Informatics. (2016). https://doi.org/10.1109/JBHI.2016.2573286

25. Li, K., Yuen, C., Kusy, B., Jurdak, R., Ignjatovic, A., Kanhere, S.S., Jha, S.: Fair Scheduling for Data Collection in Mobile Sensor Networks with Energy Harvesting. IEEE Trans. Mob. Comput. (2019). https://doi.org/10.1109/TMC.2018.2861384

26. Chen, C.C.: A novel data collection method with recharge plan for rechargeable wireless sensor networks. Wirel. Commun. Mob. Comput. (2018). https://doi.org/10.1155/2018/7419182