Data Reduction and Cleaning Approach for Energy-saving in Wireless Sensors Networks of IoT

Ali Kadhum Idrees[a,1,*], Chady Abou Jaoude[b,2], and Ali Kadhum M. Al-Qurabat[a,3],

[a] *Department of Computer Science, University of Babylon, Babylon, Iraq*

[b] *TICKET Lab, Faculty of Engineering, Antonine University, Baabda, Lebanon*

Emails: [1]ali.idrees@uobabylon.edu.iq, [2]chady.aboujaoude@ua.edu.lb, [3]alik.m.alqurabat@uobabylon.edu.iq

[*] *Corresponding Author*

***Abstract*—The wireless sensor devices of the Internet of Things (IoT) networks will represent one of the most providers of the big data on the network because it is implemented in the widespread of real-world applications. The large volume of gathered data from the sensor devices leads to increase the communication overhead and thus decrease the limited lifetime of the sensor devices of IoT. Therefore, it is necessary to clean and reduce the redundant sensed data to minimize the cost of communication and save the energy of sensor devices. In this paper, a Data Reduction and Cleaning Approach (DaReCA) for Energy-saving in Wireless Sensor Networks (WSNs) of IoT is proposed. This approach is based on two-level of data cleaning and reduction: the sensor level and the aggregator level. In the latter, we implement a divide and conquer method to merge the near similar data sets which are received from the sensor devices and reduce the transmitted data sets to the sink. In the former, the sensor node will employ a cleaning algorithm based on the leader cluster algorithm to remove redundant data from the sensed data before sending them to the aggregator. The proposed approach is evaluated and implemented using real sensed data of wireless sensor devices with the OMNeT++ network simulator. The proposed DaReCA approach can clean and reduce the sensed data and save energy whilst keeping suitable data accuracy.**

***Index Terms*—Data Reduction and Cleaning, leader Clustering, Divide and Conquer technique, Internet of Things (IoT); Wireless Sensor Networks (WSNs); Energy-conservation.**

I. INTRODUCTION

Wireless IoT (Internet of Things) devices like smart objects, sensors, cameras, and wearables are implemented in widespread real-world applications such as smart home, security, agriculture, smart cities, smart, health care, military, smart transportation, and so on [1]. The IoT is the standard by which the smart objects devices connected to the Internet, permitting these objects without people intervention to gather, compute, and communicate sensed data [2].

The huge number of smart objects in the IoT network leads to a huge amount of sensed data generated by these objects and transmitted across the network. The data correlation either spatial or temporal represents a big challenge in the IoT network [3]. Since there are limited resources are available on these objects like bandwidth, memory, battery life, and computation power [4], [5]; since the transmission and receiving of packets represents the most consuming part inside wireless sensor devices [6], hence, it is necessary to employ power-saving data reduction and cleaning schemes to remove the data redundancy and reduce the volume of data transmitted on the network to save the power of smart devices thus improve the life of network whilst retaining an acceptable rate of data accuracy. The principal benefits of data cleaning and reduction are to remove the sensed data redundancy, minimize communication costs, decrease the congestion, save energy, and lifetime improvement on the WSNs of IoT. The paper includes the following contributions.

1) A Data Reduction & Cleaning Approach (DaReCA) for Energy-saving in Wireless Sensor Networks of IoT is suggested. The proposed DaReCA approach is periodic and it works into two levels: sensor and aggregator levels. The final goal is to eliminate redundant data, decrease the communication cost, conserve the energy whilst preserving a suitable level of accuracy.
2) We propose a lightweight temporal data reduction based on a leader cluster algorithm in the sensor level to clean the redundant data from the sensed data before sending them to the aggregator.
3) We implement a spatial data reduction based on a divide and conquer method in the aggregator to eliminate the redundant data sets that are received from the sensor devices to further decrease the transmitted data sets to the base station.
4) Many simulation experiments are achieved using OMNeT++ simulator and using real sensed data from the wireless sensor devices to illustrate the enhanced performance of theDaReCA technique. The comparison is performed between the proposed DaReCA technique and three existing schemes: ATP [7] and Harb method [8], and PFF [9]. The conducted comparison results demonstrate that the proposed DaReCA technique produces a better performance compared with other approaches.

This paper is structured as follows. The next Section presents some most related works. Section III introduces the proposed DaReCA technique in more detail. The simulation results, comparison, and discussion are described in IV. Section V addresses the conclusions and future work of this paper.

II. Literature review

The data reduction and cleaning methods are widely considered by researchers in the last years. Several approaches are suggested to tackle this topic. This section introduces and addresses some existing related techniques which are performed to deal with this topic. The data reduction [10], [11] and compression [12], [13] are two main methods that are responsible for lowering the volume of the sensed data transferred across the network. The latter reduce and clean the sensed data before transmitting it to the sink and the former concentrates on compressing the sensed data before forwarding it to the upper device in the structure of the network [14]. In [15], the authors present a technique named DMLDA ("Message List Based Data Aggregation"). The dynamic list data structure is used by this technique. This data structure saves the previously received sensed data to use them in cleaning the redundant data. The work in [3] suggested a spatial-temporal correlation method to reduce the sensed data in cluster-based sensor networks. This approach minimizes the volume of sent data whilst keeping data quality. The authors in [16] suggested DADAC method for gathering the data measures and saving the energy of sensor devices. The sampling technique is employed by DADAC to eliminate the redundant measures thus prolong the lifetime of the network. The sampling strategy is based on PAA ("Piecewise Aggregate Approximation") and SAX ("Symbolic Aggregate approXimation") approaches to accomplish its aim. The work in [17] introduced a DaT technique for decreasing the volume of sensed data in wireless sensors. The authors proposed a K-nearest neighbor with a slight modification to group the data, after that, it transmits to the sink only one representative data measure from each group. The authors in [18] proposed a TLDA Protocol to improve the lifetime of the sensor networks. The time series methods are implemented by TLDA protocol on a network of two levels to remove redundant data based on the temporal and spatial similarity among collected data on the two levels. In [9], the PFF system is implemented within two levels in the network. Jaccard similarity is implemented at the sensor devices of the first level to decrease the data measures redundancy. In the second level of the network, PFF has further reduced the redundant data sets at the aggregator node. The PFF method is enhanced as introduced in node [19], [20] to reduce the redundancy of data before sending it to the sink. The work in [7] introduces a technique named ATP that employed at the sensor node to minimize the number of sensed readings transmitted to the base station. ATP operated on the two levels of the network. In the first level, it removes the redundant sensed data. The statistical strategies are utilized to eliminate the spatial data correlation at the aggregator of the second level. Harb et al. [8] introduced a periodic data aggregation approach to reduce the amount of collected data in sensor networks. This method is composed of two levels: aggregation and adaptation. The sensor level performs data cleaning to decrease the sent data volume from nodes to the cluster head. The functions of sets-similarity are executed in the second level of the network (cluster head) to eliminate redundancy in the data sets obtained from sensor nodes before delivering them to the sink. Some similar works are focused on the data gathering in WSNs are proposed in [6], [20], [21]. These suggested methods achieved a sampling approach based on discovering the temporal correlation between two data sequences periods. The proposed methods are lowered the data volume sent to the sink and increase the lifetime of the sensor device.

In this article, a Data Reduction and Cleaning Approach (DaReCA) for Energy-saving in Wireless Sensor Networks of IoT is suggested. Two-level of data cleaning and reduction are implemented by the DaReCA technique: the sensor and aggregator levels. In the former, the sensor node will operate a cleaning algorithm based on the leader cluster algorithm to discard redundant data from the sensed data before sending them to the aggregator. It is a simple and useful algorithm for IoT sensor nodes characterized by limited memory and computational resources. In the latter, the DaReCA technique executes a divide and conquer method to remove redundant data sets which are received from the sensor devices and decrease the transferred data sets to the sink.

III. Proposed Data Reduction and Cleaning Approach

This section displays a Data Reduction and Cleaning Approach for Energy-saving in Wireless Sensors of IoT Networks. This approach can exploit the temporal and spatial similarity between the gathered sensed data to clean and reduce redundant data in the two levels of the network. This can lead to minimize the communication cost, save energy, and improve the lifetime of the network.

A. Wireless Sensor level data processing:

In this level, the periodic data sensing and collecting is achieved by every sensor device from the surrounding environment. In every time slot, one sensed data is gathered and the total time slots constitute one period. Every period inside the sensor device provides a vector of sensed data $V^p = \{v_1, \ldots, v_\rho\}$, where ρ is the total number of sensed data during one period. Most of the gathered data are alike or similar, particularly when there is no change in the condition of the environment for long periods. This increases in the collected redundant data that should be cleaned from the data vector V^p before sending it to the aggregator. Therefore, in this level of data processing, the sensor device will employ the leader cluster algorithm [22] with a slight modification to eliminate and clean the data redundancy in the accumulated sensed data of the vector V^p. The leader cluster algorithm will be applied to the gathered data V^p to group it into clusters. The main idea of this clustering method is to divide the dataset (V^p in our case) into groups, one leader data object for each group and all the other data objects which have a distance T between them and the leading data object will belong to the same group. It is fast, requiring only one pass through the dataset. Algorithm 1 shows the similar data cleaning based leader cluster algorithm. Since the gathered data are alike or close, hence, the difference function is identified to allow the leader cluster algorithm in each sensor device to verify if the two data items are within distance T.

Definition 3.1: Difference function. The Difference function between two data items d_i and $d_j \in D_s^p$ collected by the smart device is referred to as follow.

$$DataDif(v_i, v_j) \leftarrow \begin{cases} 1 & \text{if } |v_i - v_j| \leq T \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Where T is a distance threshold selected by the user application. Thus, if the data difference between v_i and v_j is less than or equal to T, then they are supposed to be in the same group.

Algorithm 1: *Similar Data Cleaning based Leader Cluster Algorithm*

Require: V^p, ρ, T
Ensure: *C: the final reduced vector*
$B_i \leftarrow false; //foralli, i \leftarrow 1, \ldots, \rho$
$k \leftarrow 1$;
$D_1^k \leftarrow V_1$;
$D_0^k \leftarrow 1$;
$B_1 \leftarrow true$;
Sort V in ascending manner;
for $j \leftarrow 2$ to ρ **do**
  if $(B_j = false)$ **then**
    if $(DataDif(D_1^k, V_j) = 1)$ **then**
      $D_0^k \leftarrow D_0^k + 1$;
      $D_{D_0^k}^k \leftarrow V_j$;
    else
      $k \leftarrow k + 1$;
      $D_0^k \leftarrow V_j$;
    end if
    $B_j \leftarrow true$;
  end if
end for
for $i \leftarrow 1$ to k **do**
  $C_k \leftarrow$ average of group D^i of lengthD_0^i;
end for
return C;

B. Aggregator level data processing:

This section shows the main function that will be achieved by the aggregator after receiving the cleaned data of the sensor devices in the network. This second level of data processing is responsible for further eliminating the unnecessary sensed data from the obtained data sets of sensor devices before transmitting it to the base station node. The data reduction at the aggregator level will exploit the similarity between the received data sets of sensor devices. The aggregator will employ the divide and conquer method to combine the most similar data sets in a one representative data set. The final results of this algorithm is a set of data vectors each of them represents one or more data set of sensor devices. The main idea of this algorithm is explained as follow. First, every received data set from every sensor device will be used to compute the variance of this data set. The received data set of sensor device k is referred as $D^k = \{d_1^k, d_2^k, ..., d_L^k\}$, where L is the data set length of of sensor device k. After receiving N data sets from N sensor devices, the variance of each data set would be calculated and the result is a set of N variance values for N sensor devices. This set is described as $V \leftarrow \{v_1, v_2, ..., v_N\}$.

Definition 3.2: Variance function. The Variance function for the data set of smart device k is defined as follows

$$V^k \leftarrow \frac{\sum_{i=1}^{L^k} (d_i^k - \mu^k)^2}{L^k - 1}. \quad (2)$$

where μ^k, is the mean of the data set of sensor device k.

Second, the Divide and Conquer (DandC) method is employed to the vector V that represents the values of variances of data sets of sensor devices which are received by the aggregator. The vector V is sorted in ascending way then it will be passed to the divide and conquer algorithm. This algorithm will divide the vector V into smaller vectors (subvectors), the division continues until each subvector consists of one variance value. After that, each similar two neighboring subvectors will be combined into one vector and passed to the upper level. The process is continued until getting the final reduced vector that represents the whole data sets which are received by the aggregator. After each combination for the close or similar variance values of the two sensor devices, the sensor devices identifications (ids), the average of the two variance values, and the average of the two data set values are saved. Finally, the aggregator will transmit one representative data set with the sensor devices identifications (ids) which are shared in this data set for each value in the resulted vector V after applying the divide and conquer method. This can reduce the redundant data sets at the aggregator level, save energy, and keeping an adequate accuracy level. The time and space requirements for Algorithm 2 are $O(\rho log_2 \rho)$ and $O(\rho log_2 \rho)$ respectively. The DataDif() function that employed in Algorithm 3 is the same function in Eq. 1 except replacing T by the parameter $VarMax$.

Algorithm 2: *Data Reduction based DandC Algorithm*

Require: V, B, s, ρ, $VarMax$
Ensure: V*: the final representative data sets vector*
if $(s > \rho)$ **then**
  $Midpoint \leftarrow (s + \rho)/2$;
  $DataReductionbasedDandC(V, B, s, Midpoint, VarMax)$;
  $DataReductionbasedDandC(V, B, Midpoint + 1, \rho, VarMax)$;
  $CleanAndMerge(V, B, s, Midpoint, \rho, VarMax)$;
end if
$returnV$;

Algorithm 3: *Clean and Merge Algorithm*

Require: V,B, s, $Midpoint$, ρ, $VarMax$
Ensure: V: *the final representative data sets vector*

for $i \leftarrow s$ to $Midpoint$ **do**
  for $j \leftarrow s$ to $Midpoint - 1$ **do**
    if $(B_i = true) and (B_{j+1} = true)$ **then**
      if $DataDif(V_i, V_{j+1}, VarMax) = 1$ **then**
        $V_i \leftarrow (V_i + V_{j+1})/2$;
        $B_{j+1} \leftarrow false$;
        Put the sensor id j+1 with the sensor id i;
        Save average of j+1 and i sensor data sets;
      end if
    end if
  end for
end for
for $i \leftarrow Midpoint + 1$ to ρ **do**
  for $j \leftarrow Midpoint + 1$ to $\rho - 1$ **do**
    if $(B_i = true) and (B_{j+1} = true)$ **then**
      if $DataDif(V_i, V_{j+1}, \delta) = 1$ **then**
        $V_i \leftarrow (V_i + V_{j+1})/2$;
        $B_{j+1} \leftarrow false$;
        Put the sensor id j+1 with the sensor id i;
        Save average of j+1 and i sensor data sets;
      end if
    end if
  end for
end for
for $i \leftarrow b$ to $Midpoint$ **do**
  for $j \leftarrow Midpoint + 1$ to **do**
    if $(B_i = true) and (B_j = true)$ **then**
      if $DataDif(V_i, V_j, \delta) = 1$ **then**
        $V_i \leftarrow (V_i + V_j)/2$;
        $B_j \leftarrow false$;
        Put the sensor id j with the sensor id i;
        Save average of j and i sensor data sets;
      end if
    end if
  end for
end for
return V;

IV. Experimental results

The performance of the proposed DaReCA is evaluated and illustrated with further discussion in this section. The OMNeT++ simulator [23] has been used to achieve extension simulation experiments and based on gathered real data sensor devices that are deployed in the Intel Berkeley Lab. [24]. The Lab consists of 47 sensor devices (see Figure 1) which are responsible for gathering climate measures every 31 seconds such as voltage, humidity, temperature, and light values.For the sake of simplicity, the measure of temperature is used in our experiments. The aggregator is placed at the center of the Lab in our network model. It will receive the measures from the sensor nodes periodically. DaReCA is employed the ”First Order Radio Model” as a consumed energy model in this paper that is proposed by Heinzelman [25]. The sensed reading represents 64-bits, hence the length of the packet is the total number of sensed readings multiplying by 64-bits.

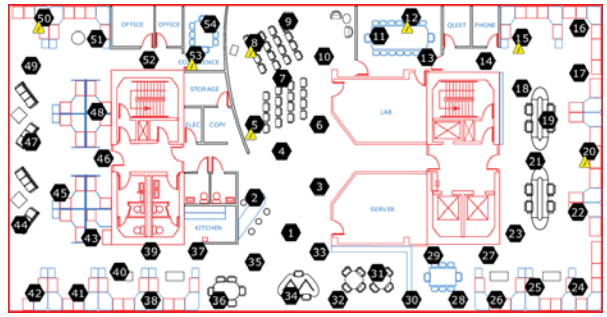

Fig. 1: The sensor network of Intel Berkeley Lab.

The DaReCA results are compared with ATP method [7], Harb technique [8], and the PFF scheme [9] results.

A. Data Percentage after Employing Reduction

One of the important jobs inside the sensor device is to clean and reduce the redundant data before forwarding it to the aggregator. In this section, we introduce the data percentage after employing a reduction algorithm of the proposed DaReCA technique inside smart device. Figure 2 shows Data Percentage after Employing Reduction. The proposed DaReCA technique can decrease the volume of transmitted data by each sensor device to aggregator from 75.44% up to 81.38% and from 94.1% up to 96.8% compared to ATP and PFF respectively. This comparison gives superiority for the DaReCA technique due to its efficiency in dealing with the unnecessary data by eliminating them to save energy and increase the performance of the network.

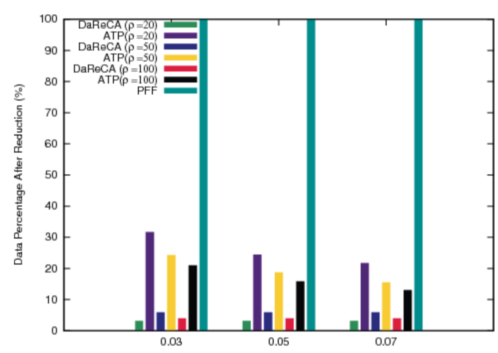

Fig. 2: Data Percentage after Employing Reduction.

B. Energy Consumption at Sensor Device

One critical factor that impacts on the lifetime of the network is the consumed power of the sensor device. This section presents an evaluation for the proposed DaReCA technique to see its capability in saving the power of the sensor device thus improve its lifetime. Hence, the energy consumption by the sensor device is illustrated in Figure

3 for various data sizes. The proposed DaReCA technique could decrease the consumed energy since it could reduce and remove redundant data as explained in Figure 2. It can be seen that the proposed DaReCA technique reduces the consumed power of the sensor device from 57.73% up to 73.91% and from 64.96% up to 76.92% compared to ATP protocol and PFF method respectively.

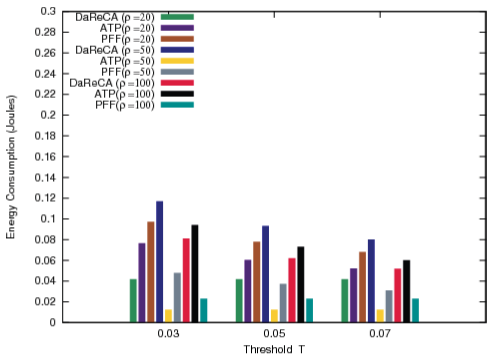

Fig. 3: Energy Consumption at Sensor Device.

C. Accuracy of Data

On one hand, it is essential to decrease the volume of data before sending it to the aggregator to save the sensor device power, on the other hand, it is necessary to maintain a suitable level of accuracy in the received data at the aggregator from the sensor devices. This accuracy of data can be defined as the percentage of lost data in the received data at the aggregator. Figure 4 shows the main comparison results of the data accuracy between the proposed DaReCA technique and other methods. It can be remarked from the results that the proposed DaReCA technique could decrease the volume of lost data from 54.17% up to 99.77% and from 41.07% up to 99.78% compared to ATP and PFF methods respectively. Hence, the proposed DaReCA technique could be eliminating efficiently redundant data whilst keeping an acceptable level of integrity.

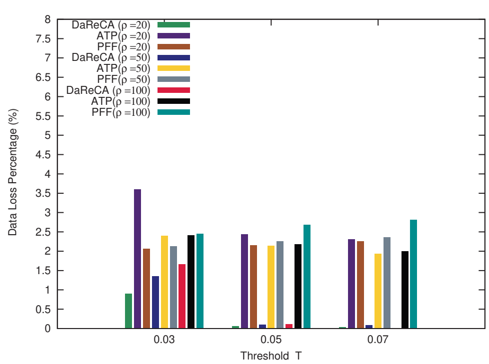

Fig. 4: Accuracy of Data.

D. Transmitted sets Percentage to the sink

The data correlation can appear among the data sets which are received by the aggregator. Hence, the proposed DaReCA technique has the responsibility to further clean and remove the redundant data before transmitting it to the base station. The transmitted data sets percentage to the base station is presented in Figure 5. It can be remarked from the results that the proposed DaReCA technique minimized the data sets from 20.66% up to 39.8%, from 29.5% up to 41.09%, from 23.9% up to 55.37%, and from 19.8% up to 49.91% compared to Harb (Tukey), Harb (Fisher), PFF 0.8, and PFF 0.75 respectively. In Harb (Bartlett), it can be seen that this method produces better results for one case. Nevertheless, the proposed DaReCA technique outperforms the other methods in most cases by further reducing the redundant data sets and save power.

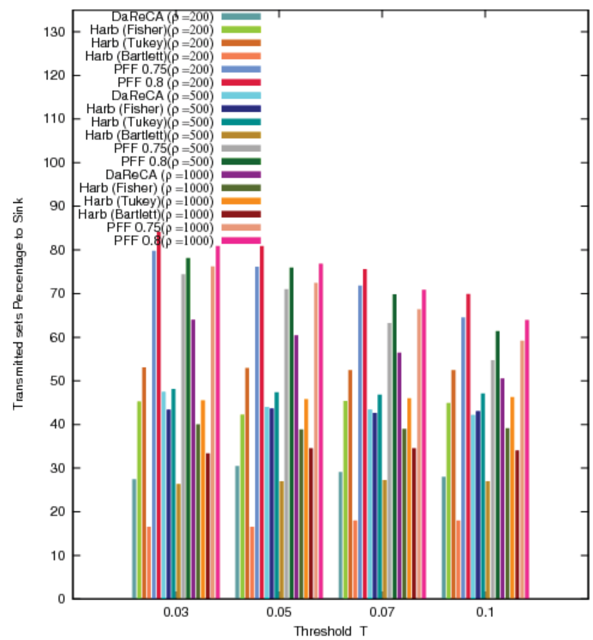

Fig. 5: Transmitted sets Percentage to the sink.

E. Energy Consumption at Aggregator

The consumed power at the aggregator is studied to show the performance of the proposed DaReCA technique compared to other approaches. These results of the energy consumption comparison are introduced in Figure 6. It can be noted from the results that the proposed DaReCA technique could decrease the consumed power from 71.05% up to 82.59%, from 56.8% up to 80.66%,from 68.75% up to 77.46%, from 80.42% up to 85.88%, and from 79.76% up to 85.01% compared to Harb (Tukey), Harb (Fisher), Harb (Bartlett), PFF 0.8, and PFF 0.75 respectively. The proposed DaReCA technique presents better achievement compared with other schemes. This improvement due to the ability of the DaReCA technique to reduce further redundant data sets before sending it to the base station.

V. Conclusion and Perspectives

In this article, a Data Reduction and Cleaning Approach (DaReCA) for Energy-saving in Wireless Sensors of IoT Networks is proposed. This approach is based on two-level

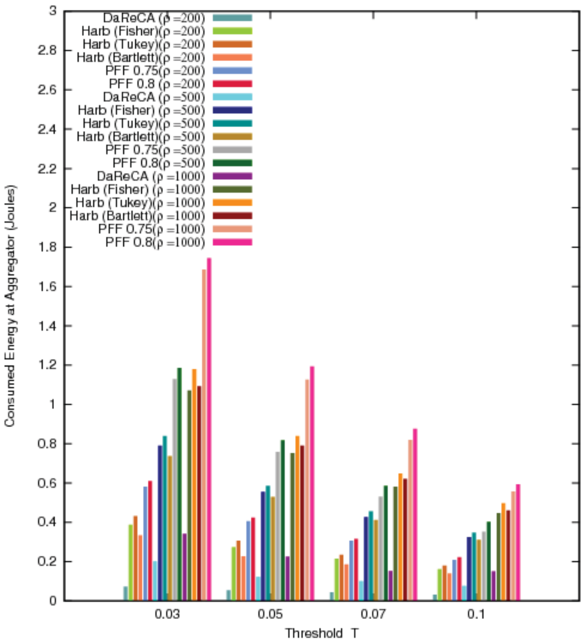

Fig. 6: Energy Consumption at Aggregator.

of data cleaning and reduction: the sensor level and the aggregator level. A cleaning and reduction method based on the leader cluster algorithm to eliminate repetitive data from the sensed data before sending it to the aggregator is employed at the smart device. The aggregator device will execute the divide and conquer technique to further reduce the redundant data before forwarding it to the sink device. The performance evaluation of the proposed DaReCA technique has shown better results compared with other approaches in terms of energy consumption at the sensor device and aggregator, data reduction, and accuracy. In the future, we plan to combine this proposed DaReCA technique with a prediction technique and implement the new system on a real sensor network.

References

[1] H. Harb, A. K. Idrees, A. Jaber, A. Makhoul, O. Zahwe, and M. A. Taam, "Wireless sensor networks: A big data source in internet of things," *International Journal of Sensors Wireless Communications and Control*, vol. 7, no. 2, pp. 93–109, 2017.

[2] S. Pattar, R. Buyya, K. R. Venugopal, S. Iyengar, and L. Patnaik, "Searching for the iot resources: Fundamentals, requirements, comprehensive review, and future directions," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 2101–2132, 2018.

[3] G. B. Tayeh, A. Makhoul, C. Perera, and J. Demerjian, "A spatial-temporal correlation approach for data reduction in cluster-based sensor networks," *IEEE Access*, vol. 7, pp. 50 669–50 680, 2019.

[4] A. K. Idrees, K. Deschinkel, M. Salomon, and R. Couturier, "Distributed lifetime coverage optimization protocol in wireless sensor networks," *The Journal of Supercomputing*, vol. 71, no. 12, pp. 4578–4593, 2015.

[5] A. K. Idrees, K. Deschinkel, M. Salomon, and R. Couturier, "Perimeter-based coverage optimization to improve lifetime in wireless sensor networks," *Engineering Optimization*, vol. 48, no. 11, pp. 1951–1972, 2016.

[6] A. K. Idrees and A. K. Al-Qurabat, "Energy-efficient adaptive distributed data collection method for periodic sensor networks," *International Journal of Internet Technology and Secured Transactions*, pp. 1951–1972, 2017.

[7] H. Harb, A. Makhoul, R. Couturier, and M. Medlej, "Atp: An aggregation and transmission protocol for conserving energy in periodic sensor networks," in *Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2015 IEEE 24th International Conference on*. IEEE, 2015, pp. 134–139.

[8] H. Harb, A. Makhoul, D. Laiymani, O. Bazzi, and A. Jaber, "An analysis of variance-based methods for data aggregation in periodic sensor networks," in *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXII*. Springer, 2015, pp. 165–183.

[9] J. M. Bahi, A. Makhoul, and M. Medlej, "A two tiers data aggregation scheme for periodic sensor networks." *Adhoc & Sensor Wireless Networks*, vol. 21, no. 1, 2014.

[10] A. K. M. Al-Qurabat, C. A. Jaoude, and A. K. Idrees, "Two tier data reduction technique for reducing data transmission in iot sensors," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2019, pp. 168–173.

[11] A. K. Idrees, W. L. Al-Yaseen, M. A. Taam, and O. Zahwe, "Distributed data aggregation based modified k-means technique for energy conservation in periodic wireless sensor networks," in *2018 IEEE Middle East and North Africa Communications Conference (MENACOMM)*. IEEE, 2018, pp. 1–6.

[12] J. Azar, R. Darazi, C. Habib, A. Makhoul, and J. Demerjian, "Using dwt lifting scheme for lossless data compression in wireless body sensor networks," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2018, pp. 1465–1470.

[13] J. Azar, A. Makhoul, R. Darazi, J. Demerjian, and R. Couturier, "On the performance of resource-aware compression techniques for vital signs data in wireless body sensor networks," in *2018 IEEE Middle East and North Africa Communications Conference (MENACOMM)*. IEEE, 2018, pp. 1–6.

[14] D. Kaur, G. S. Aujla, N. Kumar, A. Y. Zomaya, C. Perera, and R. Ranjan, "Tensor-based big data management scheme for dimensionality reduction problem in smart grid systems: Sdn perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 10, pp. 1985–1998, 2018.

[15] T. Du, Z. Qu, Q. Guo, and S. Qu, "A high efficient and real time data aggregation scheme for wsns," *International journal of distributed sensor networks*, vol. 11, no. 6, p. 261381, 2015.

[16] A. K. Idrees and A. K. M. Al-Qurabat, "Distributed adaptive data collection protocol for improving lifetime in periodic sensor networks." *IAENG International Journal of Computer Science*, vol. 44, no. 3, 2017.

[17] R. Alhussaini, A. K. Idrees, and M. A. Salman, "Data transmission protocol for reducing the energy consumption in wireless sensor networks," in *International Conference on New Trends in Information and Communications Technology Applications*. Springer, 2018, pp. 35–49.

[18] A. K. M. Al-Qurabat and A. K. Idrees, "Two level data aggregation protocol for prolonging lifetime of periodic sensor networks," *Wireless Networks*, vol. 25, no. 6, pp. 3623–3641, 2019.

[19] H. Harb, A. Makhoul, R. Tawil, and A. Jaber, "A suffix-based enhanced technique for data aggregation in periodic sensor networks," in *Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International*. IEEE, 2014, pp. 494–499.

[20] A. K. Idrees, H. Harb, A. Jaber, O. Zahwe, and M. A. Taam, "Adaptive distributed energy-saving data gathering technique for wireless sensor networks," in *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2017, pp. 55–62.

[21] T. Wang, J. Zeng, Y. Lai, Y. Cai, H. Tian, Y. Chen, and B. Wang, "Data collection from wsns to the cloud based on mobile fog elements," *Future Generation Computer Systems*, 2017.

[22] A. R. Webb and K. D. Copsey, "Statistical pattern recognition," 2011.

[23] A. Varga, "nesc language manual," *https://github.com/tinyos/nesc/blob/master/doc/- ref.pdf?raw=true*, 2003.

[24] S. Madden, "Intel berkeley research lab," *http://db.csail.mit.edu/labdata/labdata.html*, 2004.

[25] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *System sciences, 2000. Proceedings of the 33rd annual Hawaii international conference on*. IEEE, 2000, pp. 10–pp.