Journal of King Saud University – Computer and Information Sciences 34 (2022) 36–46

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Enhanced Zone-Based Energy Aware Data Collection Protocol for WSNs (E-ZEAL)

Aya H. Allam *, Mohamed Taha, Hala H. Zayed

Computer Science Department, Faculty of Computers & Informatics, Benha University, Egypt

ARTICLE INFO

Article history:
Received 23 May 2019
Revised 23 September 2019
Accepted 28 October 2019
Available online 2 November 2019

Keywords:
Internet of Things (IoT)
Wireless Sensor Networks (WSNs)
Routing protocols
Zone-based Energy-Aware data coLlection (ZEAL)

ABSTRACT

In the era of IoT, the energy consumption of sensor nodes in WSN is one of the main challenges. It is crucial to reduce energy consumption due to the limited battery life of the sensor nodes. Recently, Zone-based Energy-Aware data coLlection (ZEAL) routing protocol is proposed to improve energy consumption and data delivery. In this paper, an enhancement to ZEAL is proposed to improve WSN performance in terms of energy consumption and data delivery. Enhanced ZEAL (E-ZEAL) applies the K-means clustering algorithm to find the optimal path for the mobile-sink node. As well, it provides better selections for sub-sink nodes. The experiments are performed using the ns-3 simulator. The performance of E-ZEAL is compared to ZEAL. E-ZEAL reduces the number of hops and distance by more than 50%, Leading to speed up the data-collection phase by more than 30% with complete delivery of data. Moreover, E-ZEAL improves the lifetime of the network by 30%.

1. Introduction

By 2050, approximately 70% of the people in the world will live in urban areas, according to (Bureau, 2016). This quick urban growth is already placing pressure on the current infrastructure. To overcome this new demand, urban communities around the world focus on smart cities and IoT applications to reduce costs, improve interaction and communication, and enhance their services. There are many IoT applications for smart cities such as smart infrastructure, smart transportation, smart health care, and smart energy grid (Mohanty et al., 2016), etc. Fig. 1 illustrates some applications of IoT in smart cities (Arasteh et al., 2016).

IoT innovation makes it easier for urban buildings to improve their sustainability by saving energy consumption. For example, smart energy management systems use IoT devices to connect distinct, cooling, heating, lighting, and fire safety systems to central management software. The energy management application highlights areas of high use and energy floats so staff can address them. Research discovers that business buildings waste about 30% of the energy they use (Stauffer, 2013). Hence, using a smart building energy management system can be significant to save energy consumption. As more smart city buildings use energy management systems, the city will become more sustainable as a whole.

In (Abate et al., 2018), IoT can reform how urban communities expend water. Smart meters can improve leak detection and information integrity, avoid lost revenue because of wastefulness. In (Talukder et al., 2017), as more and more individuals move to urban communities, traffic crowding, which is already an enormous issue, is just going to get worse. Fortunately, the IoT is well-situated to make enhancements to this area. Data to be collected from traffic cameras, vehicles, smartphones, and street sensors to screen traffic incidents in real-time. Drivers can be alarmed of accidents and directed to roads that are less blocked. The potential outcomes and benefits of IoT applications in smart cities are enormous, and the effect will be considerable.

IoT is a network of physical objects such as devices, vehicles, buildings, and other items. These objects embedded with electronics, software, sensors, and network connectivity that enables data collection and exchange between them (Ngu et al., 2017). Data collection and transfer are performed through the internet without any help of humans (Yassen et al., 2016). Along these lines, IoT changes the world from real inanimate objects to intelligent virtual

* Corresponding author at: Department of Computer Science, Faculty of Computers and Informatics, Benha University, Benha Mansoura Road, next to Holding Company for Water Supply and Sanitation Benha, Qalyubia Governorate, Egypt.

E-mail addresses: aya.allam@fci.bu.edu.eg (A.H. Allam), mohamed.taha@fci.bu.edu.eg (M. Taha), hala.zayed@fci.bu.edu.eg (H.H. Zayed).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2019.10.012
1319-1578/

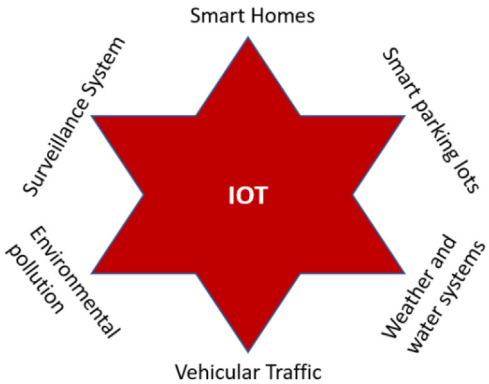

Fig. 1. IoT Applications.

objects. Also, it gives human the privilege of controlling every object around (Madakam et al., 2015). IoT contains four main components: 1) sensors and devices, 2) user interface, 3) network connectivity and 4) data processing as shown in Fig. 2 (S et al., 2009). Based on the IoT application, the architecture is defined to connect these components (Devi Kotha and Mnssvkr Gupta, 2018).

Network connectivity is one of the essential components of IoT technology. WSN represents the primary technique to perform network connectivity in IoT applications. WSN consists of sensors and gateway nodes. Sensor nodes communicate with each other to deliver the data to the gateway node. Then, the gateway node collects the data from sensor nodes and delivers to the operator over the internet. Fig. 3 illustrates the architecture of WSN.

One of the main challenges in WSN is saving the energy consumption of sensor nodes due to the short lifetime of sensor nodes (Alduais et al., 2016). Recently, Zone-based Energy-Aware data coLlection routing Protocol (ZEAL) is proposed to solve the energy consumption problem in WSN (Gallegos et al., 2018). On the first cycle, ZEAL forms independent routing zones (sets of nodes). Zones are independent of each other based on the trajectory of one or more mobile sinks. Each zone has one or more mobile-sink, sub-sink nodes, and member nodes. Mobile sinks move at constant speed back and forth between a start point and an endpoint (cyclical movements). In subsequent cycles, the mobile-sink moves among various zones to collect data from member nodes through sub-sink nodes. In this paper, an Enhanced ZEAL, named (E-ZEAL), is presented to improve WSN performance in terms of energy consumption and data delivery. E-ZEAL applies the K-means clustering algorithm to find the optimal path for the mobile-sink node. Also, it provides better selections for sub-sink nodes. The experiments are performed using the ns-3 simulator.

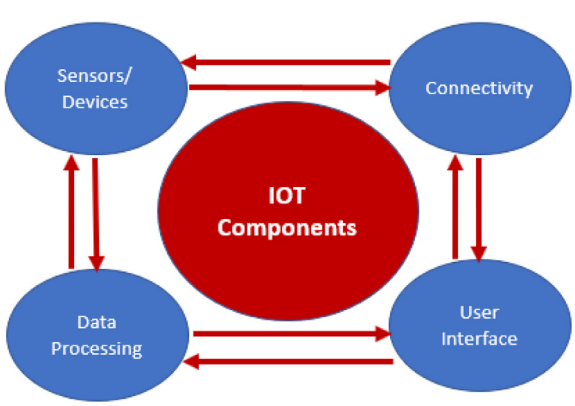

Fig. 2. IoT Components.

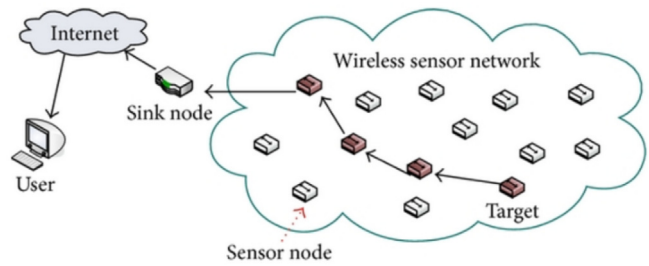

Fig. 3. WSN Architecture (Ryu et al., 2015).

The performance of E-ZEAL is compared to ZEAL. E-ZEAL reduces energy consumption and end-to-end delay compared with ZEAL. While E-ZEAL increases the lifetime, average throughput, and data delivery of the network compared to ZEAL. The rest of the paper is organized as follows; Section 2 provides a simple introduction of WSN routing protocols and ZEAL protocol. Next, Section 3 presents the proposed work. Section 4 demonstrates the results and discusses them. Finally, Section 5 concludes the paper.

2. Related work

Routing protocols play an essential role in the efficiency of WSN performance. Routing protocol works to find the optimal path of the data movement from the source node to the destination node (Singh et al., 2010). The optimal routing path is defined by terms of energy consumption and quality of services. Based on the network structure of WSN, routing protocols are classified to flat protocols, hierarchical (clustering) protocols, and location-based protocols.

Hierarchical routing protocols are the most famous schemes (Singh and Sharma, 2015). Hierarchical routing protocols have many advantages such as network scalability, load balance, and energy distribution between various nodes (Al-Shalabi et al., 2018). Hierarchical routing protocols divide WSN into clusters (groups of nodes). Every cluster has one or more cluster heads. The cluster head is responsible for data collection from sensor nodes within the cluster and to deliver it to the gateway node. LEACH (Heinzelman et al., 2000), PEGASIS (Lindsey and Raghavend, n.d), TEEN (Manjeshwar and Agrawal, n.d.), and HEED (Younis and Fahmy, 2004) are the most popular hierarchical routing protocols.

On another classification, based on the scheme utilized to establish the route from the source to the destination, routing protocols can be categorized to:

- Proactive routing protocol: where routes are established in advance for all nodes in the network.
- Reactive routing protocol: where routes are established upon request from source to destination.
- Hybrid routing protocol: combines proactive and reactive routing protocols.

Heinzelman et al. (2000) proposed the mobility of base stations (mobile-sinks) to save energy consumption and increase the lifetime of sensor networks. Path-constrained protocols are proposed to determine the path of the mobile sink to move through and collect data from network nodes. The path can be a straight line, a circular path, or a random path to visit only a selected subset of nodes in the sensor field. Tang et al. (2012) proposed Multiple Enhanced Specified-deployed Sub-sinks (MESS) protocol. It utilizes path-limited trajectory and applies reliable and efficient data retrieving mechanism. The sub-sinks are deployed at equal distances along

the path. Chen et al. (2013) proposed Virtual Circle Combined Straight Routing (VCCSR) protocol to collect data from members in each cluster using the shortest paths. VCCSR finds an efficient virtual structure to route the data effectively. The virtual structure consists of a combination of virtual circles and straight lines. In (Dhamdhere and Guru, 2014), In (Dhamdhere and Guru, 2014), the Optimal Terminal Assignment based Path (OTABP) protocol is proposed. It is a proactive routing protocol which has one or more mobile sinks move in a constrained path to collect data from sub-sinks. Sub-sinks collect data from sensor nodes and send all data packets to the mobile sink. Sub-sinks are selected depending on the distance from the other network nodes. Mobile sink controls its speed depending on the amount of data queued in the sub-sinks.

Mitra and Sharma (2018) proposed a hierarchical routing protocol based on a virtual grid structure. Where data is delivered to mobile sink in three stages: 1) sensors send data to candidate sub-sinks, then 2) candidate sub-sinks send to sub-sink nodes, finally 3) sub-sinks send to the mobile sink. The path selection of the mobile sink depends on hop counts and data generation rates. Vahabi et al. (2019) proposed the Integration of Geographic and Hierarchical Routing (IoGHR) to increase the network lifetime. IoGHR divides the network to virtual grids, each grid consists of four nodes including the Cluster Head (CH). In each grid, CH collects the data from the nodes and deliver to the mobile sink. Mobile sink passes through a predefined path to collect the data from CHs. Table 1 illustrates the comparison of different characteristics of the Path-constrained WSN protocols.

Recently, Zone-based Energy-Aware data coLlection (ZEAL) routing protocol is proposed to improve energy consumption and data delivery (Gallegos et al., 2018). ZEAL is a hybrid and hierarchical routing protocol. It is used for scenarios of mobile nodes that collect data from sensor nodes. ZEAL path is constrained to a straight line. Since our enhancements are proposed to (ZEAL), we discuss it in details. Then in the next section, we present our modifications to enhance the performance of ZEAL in terms of data delivery and energy consumption.

ZEAL is based on the Maximum Amount Shortest Path (MASP) protocol (Gao et al., 2009). MASP was developed and evaluated using the OMNET++ simulator. MASP classifies network nodes into three types: ***Mobile-sink node***: moves in constant path and speed to collect data from sub sink nodes, ***Sub-sink node***: receives data packets from sensor nodes and send them to the mobile sink node, ***Member nodes***: sense data, create data packets and deliver to the sub-sink node. MASP is a proactive routing protocol optimizes the mapping between sensor nodes and sub sinks to maximize the amount of data collected by mobile sinks and also balance the energy consumption. In each network, at least there is one sink node. Only, the sub-sink nodes are allowed to establish direct communication with the mobile sink nodes. MASP Protocol is performed on two phases: ***Discovery phase***: that divides the network into zones and determines all candidates sub-sink nodes. Also, it fills the routing table to all nodes in the network. ***Data collection phase***: where sub-sink nodes aggregate data from sensor nodes and deliver it to the mobile-sink node, as illustrated in Fig. 4.

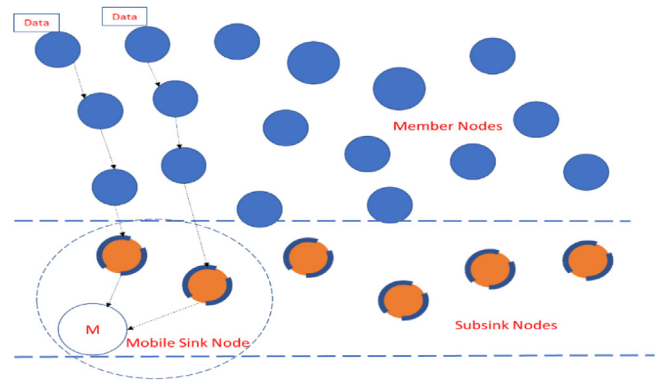

Fig. 4. Data collection phase based on MASP.

ZEAL has proposed enhancements over MASP to improve performance in terms of energy consumption and data delivery. One of these enhancements is to shorten the time of the setup phase. Another, that ZEAL doesn't depend on clock synchronization. Thus the setup phase of ZEAL is performed in one cycle instead of two cycles for MASP discovery phase. Similar to MASP, the network nodes are divided into *mobile-sink*, *sub-sink*, and *member nodes*. ZEAL processing is divided into two phases:

2.1. Setup phase

The setup phase is divided into two halves: *first half-cycle* and *second half-cycle*. ***First half-cycle***: mobile-sink moves in a steady path with a constant speed and send BCST1 messages with assigned zone ID. The time slot of each zone is calculated by dividing the time of half-cycle by a predefined number of zones. Every node receives BCST1 become a sub-sink candidate and sents UCST1 message to mobile-sink. After, mobile-sink receives UCST1 from each sub-sink and determines range time for each sub-sink. Finally, at the end of the first half cycle, the mobile-sink defines the communication time for each sub-sink. ***Second half cycle***: in the beginning, every candidate sub-sinks is assigned to a time slot. Mobile-sink calculates the member requirements parameter (***Mreq***) (Gao et al., 2009; Gallegos et al., 2016), as shown in Eq. (1). ***Mreq*** is the ideal number of member nodes for each sub-sink. Next, mobile-sink sends UCST2 message to each sub-sink with its ***Mreq*** value within its time slot. When sub-sink receives UCST2, each sub-sink sends BSCT2 message to all member nodes within the same zone. BSCT2 includes zone-ID, ***Mreq*** value and the number of hops in-between. Then, each member node updates its parameters based on BSCT2. At the end of the second half cycle, each member node establishes routes to all sub-sinks in the same zone. Fig. 5 shows the details of the *setup phase*.

$$Mreq = \frac{dt \times at}{ds \times mt} \tag{1}$$

Table 1
Characteristics comparison of path constrained WSN protocols.

Protocol	Year	Number of Sinks	Route Establishment	Nodes Characteristics	Limitations
MESS	2012	Single	Proactive	Heterogeneous	Manually deployment of sub-sinks required
VCCSR	2013	Single	Query driven	Heterogeneous	Manually deployment of sub-sinks required
OTABP	2014	Multiple	Proactive	Homogeneous	Nodes location awareness required
(Mitra and Sharma, 2018)	2018	Single	Proactive	Homogeneous	Nodes location awareness required
IoGHR (Vahabi et al., 2019)	2019	Multiple	Proactive	Homogeneous	Nodes location awareness required

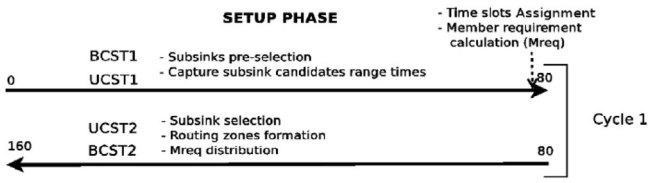

Fig. 5. Setup phase of ZEAL (Gallegos et al., 2018).

where ***dt*** is the data rate of each sub-sink to transmit data to mobile-sink, ***at*** is the assigned time slot to each sub-sink, ***ds*** is the application data rate used to transmit data to sub-sink nodes, ***mt*** is the total path time of mobile-sink.

In the setup phase, overlapping communication time and collision problem occur due to the closeness of sub-sink nodes to each other. In order to avoid this problem, Selective Time Assignment (SelectiveTA) algorithm was proposed (Gallegos et al., 2018). SelectiveTA applies two filters: ***Filter-1:*** removes the overlapped time slots, when many sub-sinks overlapped in time slots, sub-sink with largest time slot takes priority than other sub-sink candidates. ***Filter-2:*** eliminates a percentage of sub-sink candidates from each zone. This percentage is set by the network administrator. Removed sub-sinks time slots are distributed to the other sub-sinks leading to an increase in time slots for them. Also, eliminating sub-sink candidates preserves energy and reduces network traffic as no routes are created to the removed sub-sinks. But, it may affect the data delivery, that some member nodes lose communication to the mobile-sink.

2.2. Data collection phase

After the setup phase, the data collection phase is executed, as shown in Fig. 6. Where sub-sink nodes collect data and deliver to mobile-sink in specific time slots. Each member node calculates Priority (Pr) for all sub-sink nodes within the same zone using Eq. (2). Each member node selects the sub-sink node with the highest priority. The data aggregation process is performed every cycle as follows: mobile-sink sends POLL1 message to all sub-sink nodes. Next, the sub-sink node sends data packets to mobile-sink until the assigned time slot is ended or queue of data packets is depleted.

$$Pr = \alpha \times Mreq + \frac{(1-\alpha)}{Nhops} \tag{2}$$

where $\boldsymbol{\alpha}$ is a weight value between 0 and 1, ***Nhops*** is the number of hops between sub-sink and member nodes. ***Mreq*** is the ideal number of member nodes for sub-sink node.

ZEAL proposed a duty-cycle mechanism to reduce energy consumption. The duty-cycle mechanism applies wake-up/sleep time slots on member nodes, where the member node wakes up to send data and go to sleep mode until the sleeping time runs out. The sleeping time equals to [(half-cycle time − current time) × 2]. The mechanism is implemented as follows, while mobile-sink moves out from the current zone to move in the next one, it sends POLL2 message to the last sub-sink node in the current zone. Once the last sub-sink node receives POLL2, it reduces sleeping time with one second. Then it broadcasts SLP message containing zone-ID and the sleeping time to all member nodes in the same zone. Each member node receives SLP message rebroadcasts to other member nodes until all members within the zone go to sleep.

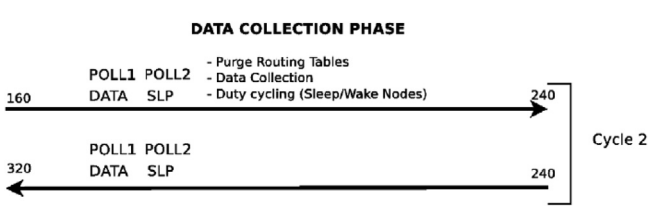

Fig. 6. Data Collection Phase of ZEAL (Gallegos et al., 2018).

3. Enhanced Zone-based Energy-Aware data coLlection (E-ZEAL)

ZEAL works to save energy by reducing the wake-up time of the member nodes utilizing two techniques: reduction of the setup-phase time and the duty-cycle mechanism. However, ZEAL doesn't handle the layout of the mobile-sink path, even though mobile-sink path affects WSN performance in terms of energy consumption and data delivery (Olariu and Stojmenovic, 2006). For instance, short-path increases the number of hops for communication, consumes more energy, and degrades the data delivery but reduces the data-collection phase time. On the other side, the long-path, moving near member nodes, reduces the number of hops for communication, saves energy, and improves data delivery but increases the data-collection phase time. The ZEAL does not handle the layout of the mobile-sink path. So the challenge is to find the optimal path which can achieve the balance between all parameters. In this context, the Enhanced-ZEAL (E-ZEAL) routing protocol is proposed to improve WSN performance in terms of data delivery and energy consumption. The proposed algorithm has three phases (*Pre-processing phase*, *Setup phase*, *Data collection phase*). The *Data collection phase* is identical to ZEAL without any modification, similar to the section above. *Pre-processing phase* details and the modifications to *setup phase* are discussed in the following sub-sections.

3.1. Pre-processing Phase

In the *Pre-processing Phase*, E-ZEAL finds the optimal path for the mobile-sink node. The optimal path is defined as the path which achieves the minimum number of hops and the minimum distance for the group of the sub-sink and member nodes. E-ZEAL applies the K-means clustering algorithm to find the optimal path similar to the scheme proposed in (Xing et al., 2008). The pseudo-code of the *Pre-Processing* phase is presented in Algorithm 1.

Algorithm 1: Pre-Processing Phase

N	Number of Nodes
L	Set of Locations of nodes {L1, L2. . .Ln}
K	Number of optimum clusters
K1, K2	Value around Optimum k
C	Set of Kmean Centroids {C1, C2,. . .., Ck}
D	Set of Distance between Centroids {D1, D2,., Dk}
Θ	Set of Angles between Centroids {Θ 1, Θ 2,. . .., Θ k}
T	Average Time of Kmean Path
INPUT L	
OUTPUT: D, T, Θ	
1 K, D, T, Θ, C=Ø	
2 K = Silhouette (L)	/* Function to get the best optimum k (Rousseeuw, 1987), used in the K-means method (de Amorim and Hennig, 2016) */
3 K1 < K < K2	
4 C = K-mean (L, K)	/* return centroids after calculate K-means*/
5 Θ = Calculate-angles (C)	/* return angles from K-means centroids */
6 D = Calculate-distance (C)	/* return distance between centroids vectors */
7 T = Calculate-Path-time (D)	/* return time for all K-means path */

In Algorithm 1, Silhouette method, proposed by Rousseeuw (1987), is utilized to find an optimum number of clusters (***k***). Consequently, the cluster number input is set for the K-means clustering algorithm (de Amorim and Hennig, 2016). The K-means clustering algorithm was proposed by Hartigan and Wong (1979). It is an unsupervised learning algorithm that works to group similar data points and finds out the mutual patterns. It searches for a fixed number of clusters (***k***) in the dataset. A cluster is defined as a collection of data points grouped due to certain similarities. ***k*** refers to the number of required centroids in the dataset. A centroid represents the center of the cluster. In WSN scenario, the locations of member nodes and sub-sink nodes represent the dataset. The centroids are the set of positions that have minimum hops and minimum distance from member nodes and sub-sink nodes; this approach introduced in Xing et al. (2008). The researchers designed a route passes by a set of positions, the rendezvous points, where mobile-sink collects data from sub-sink nodes located close to the designed route. Rendezvous points represent the centers of clusters of member nodes, as shown in Fig. 7. The proposed E-ZEAL algorithm applies the same methodology by implementing the K-means clustering algorithm to find the optimal path for the mobile-sink node presented in Algorithm 1. First, the distribution of the member nodes and sub-sink nodes are obtained. Then, the silhouette method is applied to get optimum ***k*** to be used as input to the K-means clustering algorithm. The centers of clusters and the locations of these centers are determined. After the centers are connected, the angles are calculated, and the path is set.

3.2. Setup Phase

E-ZEAL is similar to the ZEAL in setup phase regarding the main functions. But we modify the priority equation (Eq.2) to provide a better selection for sub-sink nodes. As remarked in Eq.2, The priority depends on the number of members per sub-sink node (***Mreq***), the number of hops between the sub-sink and the member nodes (***Nhops***) and α, a weight parameter to achieve the balance between ***Mreq*** and ***Nhops***. However, ZEAL doesn't provide a method to determine the value of α. At first, we attempt to find an optimal value for α. But we found that, as long as the network architecture changes, α changes. Therefore, we modify the priority by applying the normalization on ***Mreq*** and ***Nhops*** parameters and removing the α parameter from the equation. Consequently, ***Mreq*** and ***Nhops*** have equal weights on the selection of sub-sink node. The priority is directly proportional to ***Mreq*** and reversely proportional to ***Nhops***. So normalization is performed to the maximum ***Mreq*** and the reverse of the minimum ***Nhops***, as shown in Eq. (3).

$$Pr = \frac{Mreq}{max(Mreq)} + \frac{\frac{1}{Nhops}}{\frac{1}{min(Nhops)}} \tag{3}$$

Also, we propose to modify the priority of the sub-sink selection by adding the distance between the sub-sink and the member nodes. It is not enough to depend only on ***Nhops*** because the member node may select a sub-sink node with a less number of hops but with a longer distance, which may lead to more energy consumption. Fig. 8 illustrates an example of the trade-off case between the number of hops and distance, node-1 selects sub-sink-3 instead of sub-sink-6 based on the less number of hops. Although the sub-sink-6 is closer to node-1. Hence, distance (***dist***) is an important parameter in the selection of sub-sinks beside ***Mreq*** and ***Nhops***. The distance between the sub-sink and the member nodes is calculated based on the Euclidean norm, as shown in Eq. (4).

$$\boldsymbol{dist}\ (\mathrm{m,s}) = \sqrt{(x2-x1)^2 + (y2-y1)^2} \tag{4}$$

where, $(x1, y1)$ represent the position of the member node **m** and $(x2, y2)$ represent the position of the sub-sink node **s**.

Similar to the number of hops, priority is reversely proportional to distance, as energy consumption increases to cover a longer distance. In order to achieve the balance between all terms (***Mreq***, ***Nhops, dist***), the distance parameter is normalized. Distance normalization is performed with respect to the reverse of the minimum ***dist***. Eq.5 illustrates the final form of the priority equation after normalization for ***Mreq*** and ***Nhops***, and the addition of the normalized distance parameter.

$$Pr = \frac{Mreq}{max(Mreq)} + \frac{\frac{1}{Nhops}}{\frac{1}{min(Nhops)}} + \frac{\frac{1}{dist}}{\frac{1}{min(dist)}} \tag{5}$$

Fig. 9 illustrates the flow chart of E-ZEAL setup phase in details. In summary, E-ZEAL is proposed to improve the performance of ZEAL in terms of data delivery and energy consumption. E-ZEAL modifies the priority equation (Eq. (2)), in the setup phase, to improve the selection of sub-sink nodes leading to enhancement in data delivery. Also, The E-ZEAL implements the K-means clustering algorithm, in

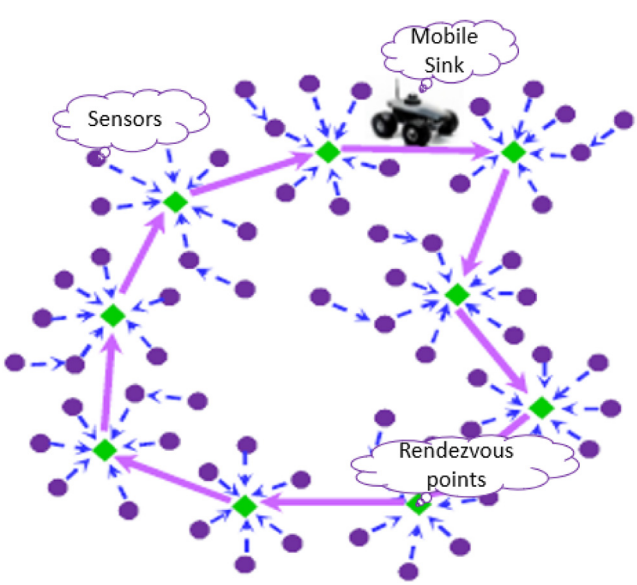

Fig. 7. Mobile-sink Path based on K-means Clustering Algorithm (Xing et al., 2008).

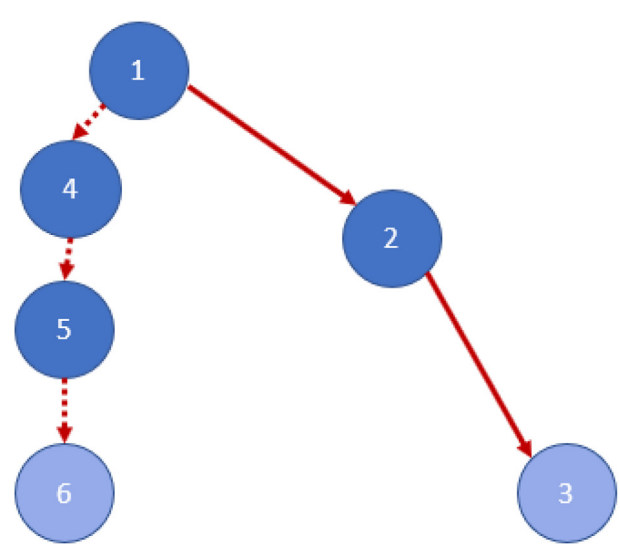

Fig. 8. Example of the selection process of sub-sinks based on the number of hops.

the pre-processing phase, to find the optimal path for the mobile sink to reduce energy consumption. In the next section, we evaluate our proposed E-ZEAL compared to ZEAL algorithm.

4. Results and Discussion

In Gallegos et al. (2018), the authors compared between ZEAL and MASP. MASP utilizes transmission time slots to send data from sub-sink nodes to the mobile sink. MASP provides more accurate data collection than ZEAL poll mechanism, but It requires a perfect synchronized clock between all nodes. This synchronization approach is the main drawback due to the complexity and the feasibility of real implementation. The experimental results, shown in Gallegos et al. (2018), demonstrated the effectiveness of ZEAL over MASP to increase network lifetime and improve data collection. Our proposed E-ZEAL is implemented and evaluated in comparison with ZEAL on ns-3 simulator utilizing energy model (Wu et al., 2011). E-ZEAL increases the network lifetime and improves data delivery than ZEAL by adding new criteria to select sub-sink nodes and to define the optimal path utilizing the K-mean algorithm. Table 2 illustrates the comparison between MASP, ZEAL, and E-ZEAL in details.

In all scenarios of the test cases, the network area is set to 400 m × 200 m, 120 nodes are distributed randomly. Moreover,

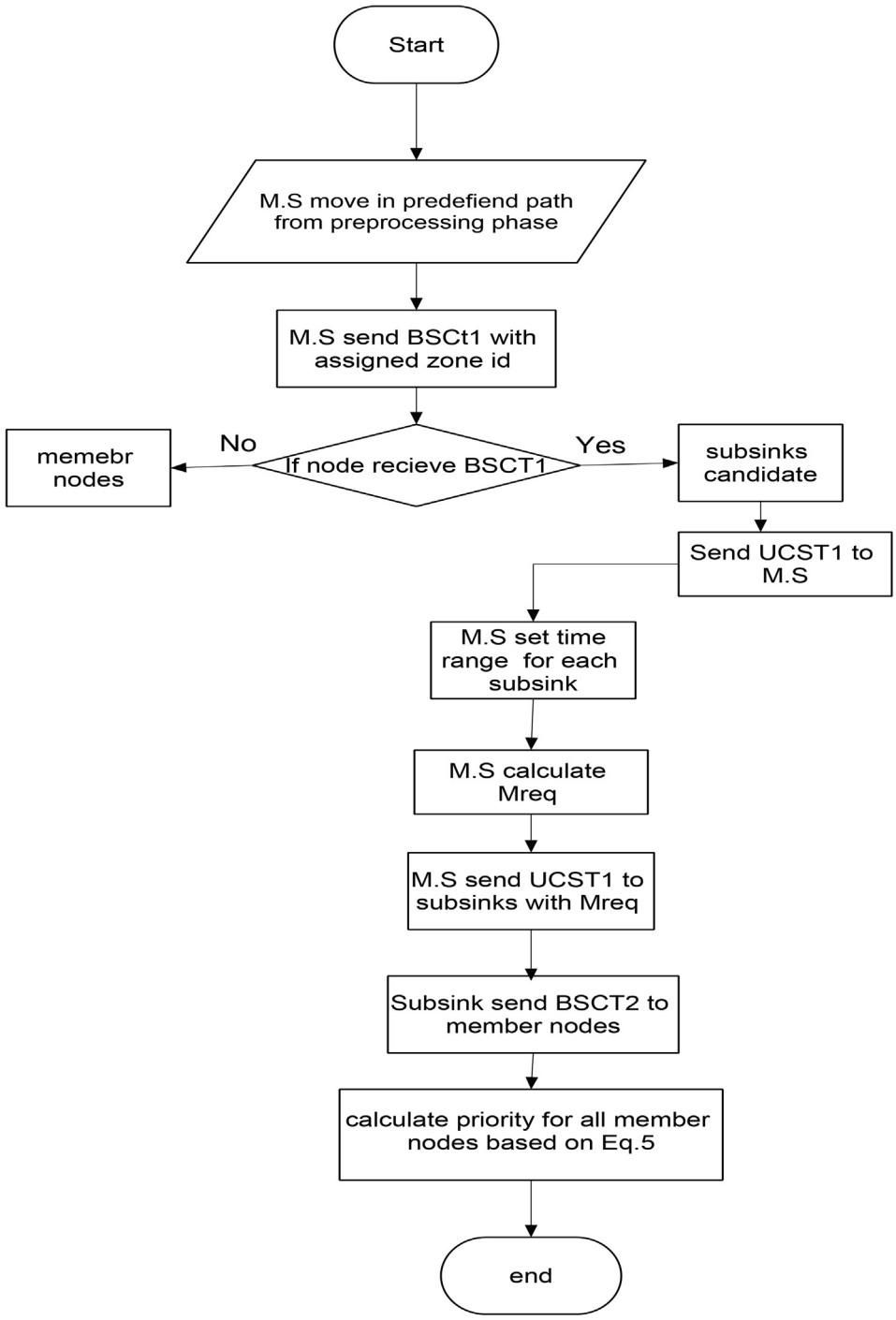

Fig. 9. Flow Chart of Setup phase of E-ZEAL.

Table 2
Comparison between MASP, ZEAL and E-ZEAL.

Characteristics	MASP	ZEAL	E-ZEAL
Number of phases	2	2	3
Automatic zone distribution	No	Yes	Yes
Energy saving function	No	Duty cycle mechanism	Duty cycle mechanism Distribution awareness path
Criteria to select sub-sink nodes	Member requirements Number of hops	Member requirements Number of hops	Member requirements Number of hops Distance between sub-sink nodes and sensor nodes
Ineffective sub sinks filter	Yes	No	No
Constrained path sink movement	Yes	Yes	Distribution awareness path
Routing Type	Proactive	Hybrid	Hybrid
Initial Energy of nodes	Homogenous	Homogenous	Homogenous
Method for selecting a path	No	No	K-mean
Implementation layer	Network layer	Network layer	Network layer
Data collection mechanism	Synchronized transmission	Poll mechanism	Poll mechanism

tests are repeated with different random seeds for verification. Table 3 provides all details of the simulation parameters for the various tests. In the *Pre-processing Phase,* the Silhouette method is applied to find optimum ***k to*** be used as input to the K-means clustering algorithm, as shown in Fig. 10. Then, the member and sub-sink nodes are divided into clusters. The centroid of each cluster is determined. After, the path is set by connecting the centroids of all clusters, as shown in Fig. 11. Quite the opposite, ZEAL sets a straight path at the bottom of the network area that mobile-sink moves forward and backward on a horizontal line. Therefore, E-ZEAL achieves a reduction in the distance (***dist***) and the number of hops (***Nhops***) between sub-sink and member nodes. To prove the reduction, an experiment is implemented to measure the maximum ***Nhops*** and the maximum ***dist*** in the whole network for all sub-sink and member nodes. Table 4 shows the result of E-ZEAL path in comparison with bottom-path in ZEAL. It is observed that E-ZEAL achieves a reduction in ***Nhops*** and ***dist*** by more than ***50%***.

The proposed approaches work to improve the performance of ZEAL in terms of data delivery and energy consumption. An experiment is implemented to assess the data delivery and the average remaining energy in the network. In general, all experiments compare between ZEAL and E-ZEAL (with optimal path and updated priority Eq.5). Fig. 12 shows the data delivery performance of ZEAL bottom-path and E-ZEAL path. It can be observed that E-ZEAL achieves complete data delivery by 240 data packets in two data-collection cycles for 120 nodes.

Similarly Fig. 13 shows the average remaining energy performance of ZEAL and E-ZEAL. E-ZEAL saves more energy for member nodes for two data-collection cycles. However, Fig. 14 illustrates the average remaining energy performance for 23 cycles. The E-ZEAL improves energy consumption along with time, leading to an increase in the lifetime of member nodes. After 23 cycles, the average remaining energies are 261 and 845 J for ZEAL and E-ZEAL respectively, representing 8% and 28% of the initial energy in turn. The results validate our proposed approach.

To evaluate the path applied in E-ZEAL, different paths (middle-path, diagonal-path, letter-V, and zigzag-path) are implemented besides the bottom-path implemented in ZEAL. The optimal path passes close to the most sub-sink and member nodes and reduces the overall number of hops and distance. Fig. 15 illustrates all different paths for the same network with 120 nodes. E-ZEAL path and other paths are evaluated in terms of average remaining energy, the length of the path, and the time and the maximum

Table 3
The simulation parameter.

Parameter	Value
Target Area	400 m × 200 m
Number of Nodes	120
Initial Energy of Nodes	3000 J (equivalent to a single AAA NiMH battery).
Maximum Communication Range of Nodes	52 m
Application Data Size	1029 bytes
Speed of Mobile-sink	5–12 m/s
Rx-Current	0.313 Amp
Tx-Current	0.380 Amp
Idle-Current	0.273 Amp
Busy-Current	0.273 Amp
Sleep-Current	0.033 Amp

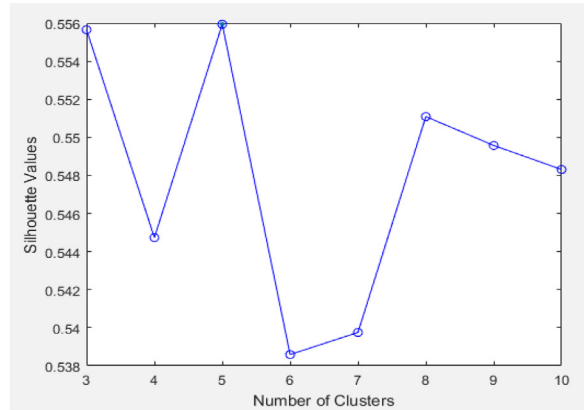

Fig. 10. Silhouette method to select the best K clusters.

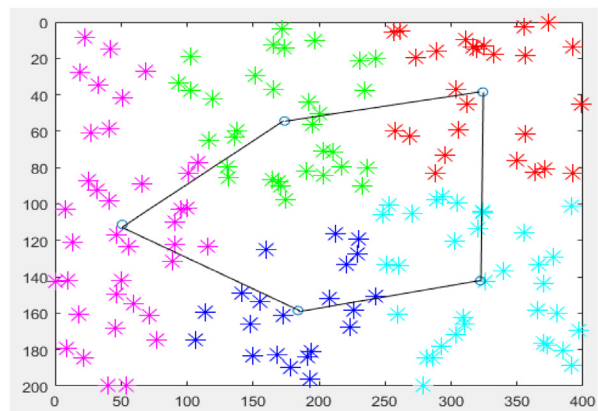

Fig. 11. E-ZEAL Path based on the K-means Clustering Algorithm.

Table 4
Maximum hops and the maximum distance of ZEAL and E-ZEAL paths.

Path	Max hops	Max distance
Bottom	9	320
E-ZEAL (K-means path)	4	146

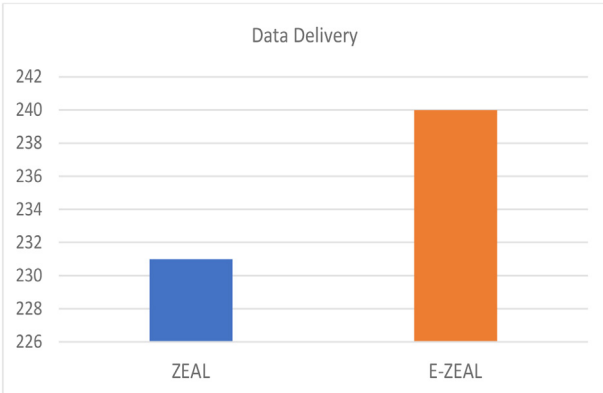

Fig. 12. Data delivery of E-ZEAL and ZEAL in 2 data-collection cycles.

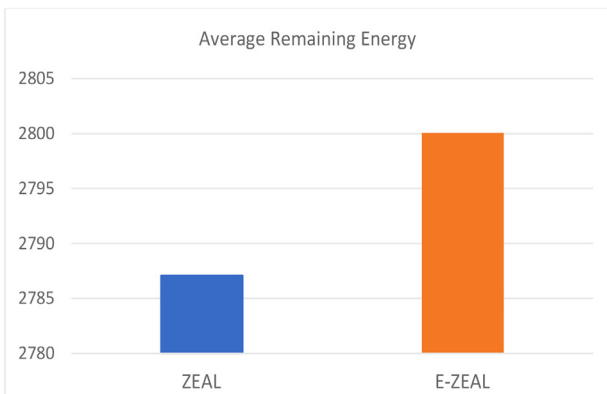

Fig. 13. Average Remaining Energy of E-ZEAL and ZEAL in 2 data-collection cycles.

speed taken by the mobile-sink node to complete one-cycle (forward and back).

Fig. 16 illustrates a path length comparison between bottom-path, middle-path, diagonal-path, letter-V, zigzag-path, and E-ZEAL path. The E-ZEAL path is one of the longest paths, while the bottom\middle paths are the shortest ones. The length of the E-ZEAL path is around time-and-one-half of the bottom-path

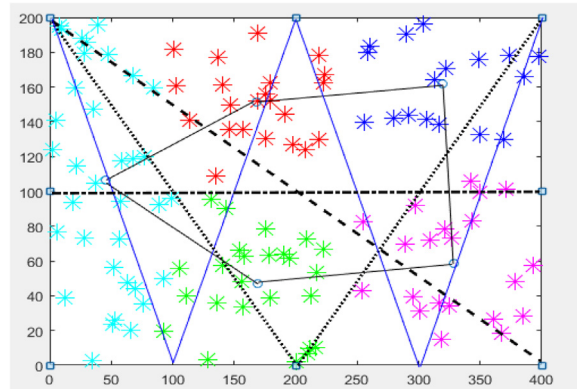

Fig. 15. Bottom-path, Middle-path, Diagonal-path, Letter-V, Zigzag-path, and E-ZEAL path.

length. Nevertheless, the length of the path does not necessarily assess the path. That, as long as the path length increases, there are more opportunities for mobile-sink to move close to member nodes and sub-sinks nodes accordingly, the number of hops and distance decrease leading to speed-up *data-collection phase*. Fig. 17 shows the maximum mobile-sink speed of the different paths. As noticed, mobile-sink speed of the E-ZEAL path is the fastest, around more than double speed on the bottom-path. As a result, the E-ZEAL, utilizing the E-ZEAL path, achieves the full-cycle of data collection in a shorter time. Fig. 18 compares the path time of all paths. E-ZEAL path is ***30%*** less than ZEAL bottom-path.

Fig. 19 illustrates a comparison between E-ZEAL path and all other paths regarding the average remaining energy. The results are for two-cycles of data collection using an application data rate of 8232 bps (1 packet send by nodes in 1 cycle). In short, the evaluation criteria depend on the time that mobile-sink can accomplish the data-collection cycle without any loss in data delivery.

The E-ZEAL overcomes the ZEAL regarding the average remaining energy, as shown in Figs. 13, 14, and 19. The reason is that E-ZEAL completes the whole data-collection cycle in a shorter time leading to improvement in the duty-cycle of all member nodes (decrease the overall wake-up time and increase the overall sleeping time). The results confirm that the proposed E-ZEAL achieves the shortest time with full data delivery, as shown in Fig. 12.

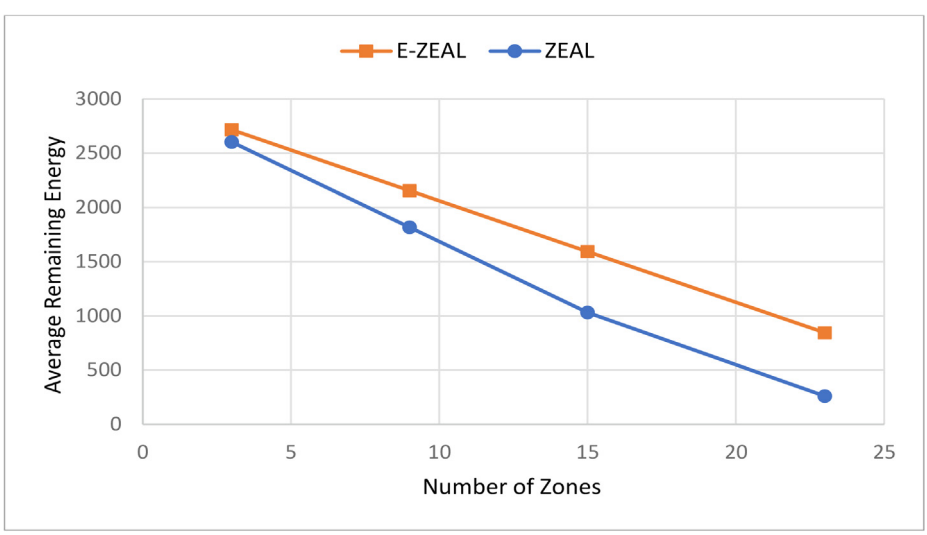

Fig. 14. Average Remaining Energy of E-ZEAL and ZEAL in 23 cycles.

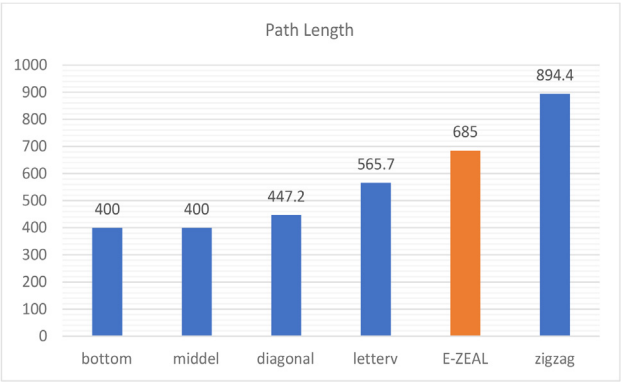

Fig. 16. Path Length of Bottom-path, Middle-path, Diagonal-path, Letter-V, Zigzag-path, and E-ZEAL path.

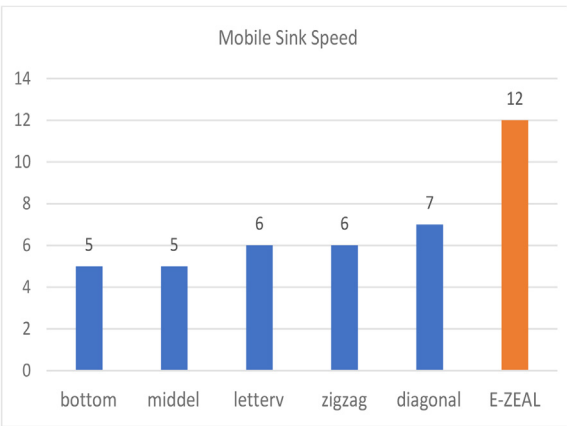

Fig. 17. Maximum mobile-sink speed of Bottom-path, Middle-path, Diagonal-path, Letter-V, Zigzag-path, and E-ZEALpath.

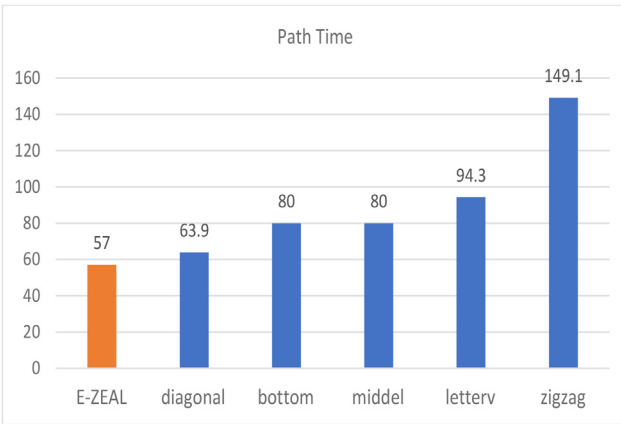

Fig. 18. The Time of the Path of Bottom-path, Middle-path, Diagonal-path, Letter-V, Zigzag-path, and E-ZEAL path.

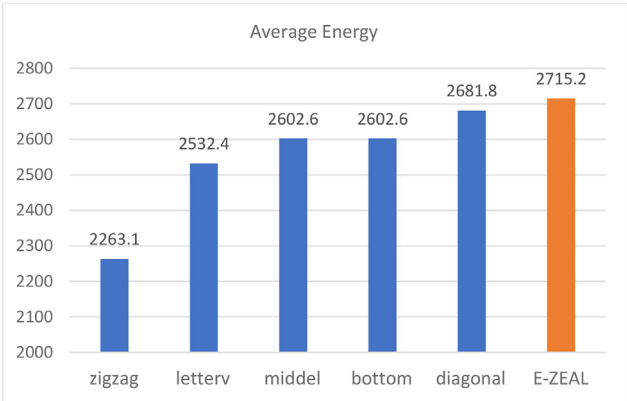

Fig. 19. Average Remaining Energy of Bottom-path, Middle-path, Diagonal-path, Letter-V, Zigzag-path and E-ZEALpath.

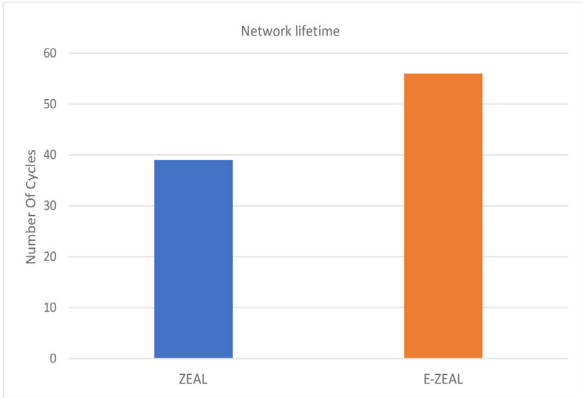

Fig. 20. The average network lifetime of E-ZEAL and ZEAL.

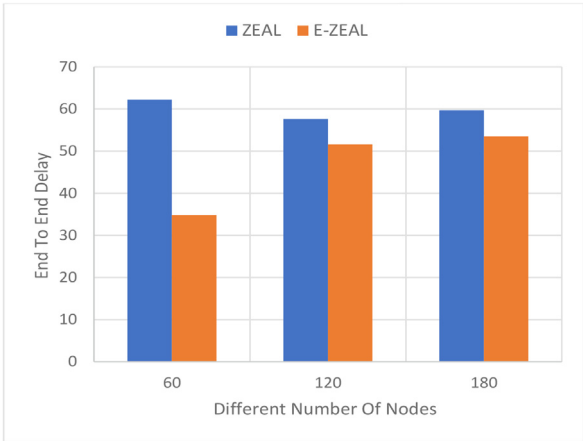

Fig. 21. End to End delay in E-ZEAL and ZEALwith different network densities.

Fig. 20 illustrates the average network lifetime of E-ZEAL and ZEAL, E-ZEAL increase the lifetime of nodes in the network with 30% than ZEAL in Figs. 21–23 show Quality of Service (QoS) parameters such as (end-to-end delay, throughput, and the average of remaining energy) with different network densities in terms of sensor nodes (60, 120 and 180 nodes). The experiments are deployed utilizing the flow monitor module in the ns-3 simulator (Carneiro et al., 2009). Fig. 21 shows that E-ZEAL decreases the end-to-end delay time in comparison with ZEAL by 30%. Fig. 22 illustrates the improvement to the average throughput of the network. On average, throughput is improved by 20% all over the network densities. The throughput is calculated by the amount of the transmitted data per second [bps] (Sawai et al., 2016). Fig. 23 confirms our results regarding the saving of average remaining energy but for different network densities. Hence, E-ZEAL supports network scalability.

The duty-cycle mechanism is one of the advantages implemented by the ZEAL protocol. As stated above, the sleeping time

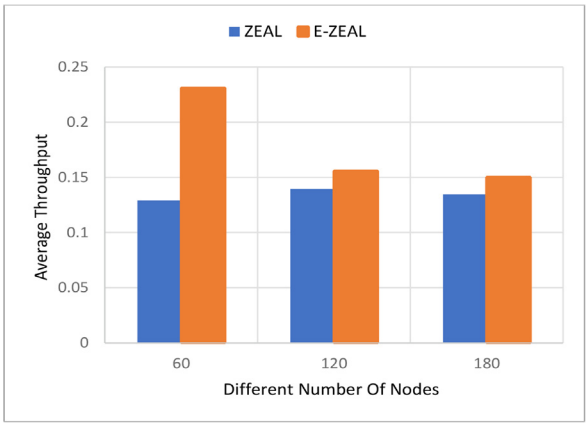

Fig. 22. Average Throughput with different network densities.

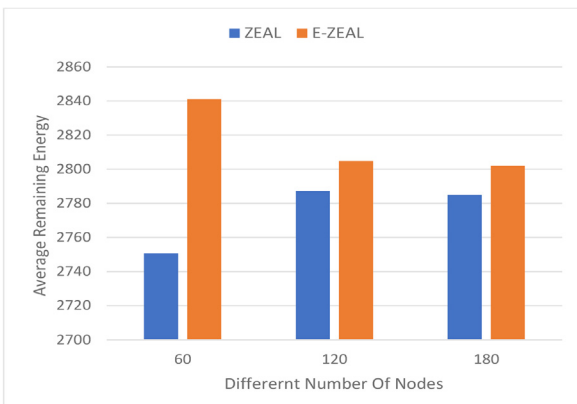

Fig. 23. Average Remaining energy with different network densities.

is the main parameter to reduce the energy consumption of member nodes. ZEAL divides member nodes into zones. The number of zones in ZEAL is a parameter defined by the network administrator (Gallegos et al., 2018). Our experiments demonstrate that the number of zones affects the average time of the duty-cycle (wake-up \sleeping time), leading to an effect on the average remaining energy. Figs. 24 and 25 illustrate the average sleeping time and average remaining energy for different zone numbers, respectively. As observed, the results show that there is an optimum number of zones. For instance, a five-zone setup achieves better performance. The number of zones affects the duty-cycle and broadcasting performance. When the number of zones increases, the average sleeping time increases, leading to an improvement in average remaining energy. On the other hand, the number of member nodes sleep reduces due to missing reception of SLP (zone-ID and the sleeping time) message. Our experiments show that the total number of member nodes (1 2 0) go to sleep with five-zone setup. However, only 80 member nodes go to sleep with the ten-zone setup. So about one-third of member nodes do not sleep. Accordingly, we believe that the number of zones is a vital parameter. In future work, we will try to add criteria to E-ZEAL to find the optimum number of zones which achieves better average remaining energy.

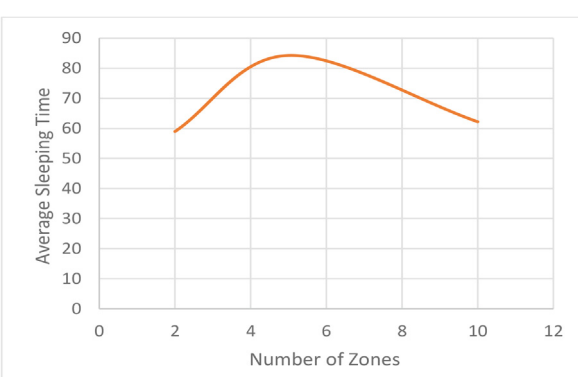

Fig. 24. Average Sleeping Time for ZEAL with the different zone number.

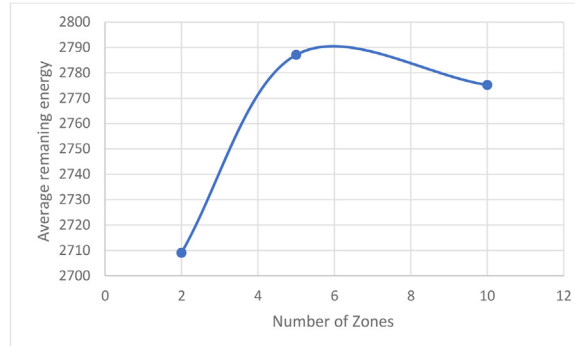

Fig. 25. Average Reaming Energy for ZEAL with the different zone number.

5. Conclusion

In this paper, E-ZEAL is proposed to enhance the performance of ZEAL routing protocol in terms of data delivery and energy consumption. E-ZEAL has three phases: pre-processing phase, setup phase, and data collection phase. In the pre-processing phase, E-ZEAL implements the optimal path for the mobile sink to reduce energy consumption utilizing the K-means clustering algorithm. In the setup phase, E-ZEAL improves the selection of sub-sink nodes in purpose to enhance the data delivery. Experiments are implemented in the ns-3 simulator. Quality of services parameters such as (lifetime, end to end delay, throughput, and remaining energy) are evaluated for performance comparison between E-ZEAL and ZEAL. The results illustrate that E-ZEAL provides better performance in all aspects. E-ZEAL succeeds to reduce the number of hops and distance between sub-sink nodes by more than ***50%,*** leading to speed up the data-collection phase by more than ***30%*** and to decrease the end-to-end delay by ***30%***. Moreover, E-ZEAL improves the network lifetime by ***30%***. After 23 data-collection cycles, the average remaining energy of member nodes utilizing E-ZEAL is ***28%*** of the initial energy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abate, F., Carratu, M., Liguori, C., Ferro, M., Paciello, V., 2018. Smart meter for the IoT. In: 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). IEEE, pp. 1–6. doi: 10.1109/I2MTC.2018.8409838.

Al-Shalabi, M., Anbar, M., Wan, T.-C., Khasawneh, A., 2018. Variants of the low-energy adaptive clustering hierarchy protocol: survey, issues and challenges. Electronics 7, 136. https://doi.org/10.3390/electronics7080136.

Alduais, N.A.M., Abdullah, J., Jamil, A., Audah, L., 2016. An efficient data collection and dissemination for IOT based WSN. In: 2016 IEEE 7th Annual Information

Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, pp. 1–6. doi: 10.1109/IEMCON.2016.7746084.
Arasteh, H., Hosseinnezhad, V., Loia, V., Tommasetti, A., Troisi, O., Shafie-Khah, M., Siano, P., 2016. Iot-based smart cities: a survey. In: EEEIC 2016 – Int. Conf. Environ. Electr. Eng. 2–7. doi: 10.1109/EEEIC.2016.7555867.
Bureau, P.R., 2016. Population Reference Bureau [WWW Document]. Population Ref. Bur.
Carneiro, G., Fortuna, P., Ricardo, M., 2009. FlowMonitor – a network monitoring framework for the Network Simulator 3 (NS-3). In: Proceedings of the 4th International ICST Conference on Performance Evaluation Methodologies and Tools. ICST. doi: 10.4108/ICST.VALUETOOLS2009.7493.
de Amorim, R.C., Hennig, C., 2016. Recovering the number of clusters in data sets with noise features using feature rescaling factors. doi: 10.1016/j.ins.2015.06.039.
Devi Kotha, H., Mnssvkr Gupta, V., 2018. IoT application, a survey. Int. J. Eng. Technol. 7, 891. https://doi.org/10.14419/ijet.v7i2.7.11089.
Dhamdhere, S.D., Guru, S.K., 2014. Robust Data Collection in Wireless Sensor Networks with Mobile Sinks.
Gallegos, A., Noguchi, T., Izumi, T., Nakatani, Y., 2016. Simulation study of Maximum Amount Shortest Path routing in Wireless Sensor Networks using Ns-3. In: 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN). IEEE, pp. 198–204. doi: 10.1109/ICUFN.2016.7537016.
Gallegos, A., Noguchi, T., Izumi, T., Nakatani, Y., 2018. Zone-based energy aware data collection protocol for WSNs. IEICE Trans. Commun. E101.B, 750–762. https://doi.org/10.1587/transcom.2017EBP3133.
Gao, S., Zhang, H., Das, S., 2009. Efficient data collection in wireless sensor networks with path-constrained mobile sinks. In: 2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops. IEEE, pp. 1–9. doi: 10.1109/WOWMOM.2009.5282492.
Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a K-means clustering algorithm. Appl. Stat. 28, 100. https://doi.org/10.2307/2346830.
Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H., 2000. Energy-Efficient Communication Protocol for Wireless Microsensor Networks.
Lindsey, S., Raghavendra, C.S., n.d. PEGASIS: Power-efficient gathering in sensor information systems. In: Proceedings, IEEE Aerospace Conference. IEEE, pp. 3-1125-3–1130. doi: 10.1109/AERO.2002.1035242.
Madakam, S., Ramaswamy, R., Tripathi, S., 2015. Internet of things (IoT): a literature review. J. Comput. Commun. 03, 164–173. https://doi.org/10.4236/jcc.2015.35021.
Manjeshwar, A., Agrawal, D.P., n.d. TEEN: a routing protocol for enhanced efficiency in wireless sensor networks. In: Proceedings 15th International Parallel and Distributed Processing Symposium. IPDPS 2001. IEEE Comput. Soc, pp. 2009–2015. doi: 10.1109/IPDPS.2001.925197.
Mitra, R., Sharma, S., 2018. Proactive data routing using controlled mobility of a mobile sink in Wireless Sensor Networks. Comput. Electr. Eng. 70, 21–36. https://doi.org/10.1016/J.COMPELECENG.2018.06.001.
Mohanty, S.P., Choppali, U., Kougianos, E., 2016. Everything you wanted to know about smart cities: the Internet of things is the backbone. IEEE Consum. Electron. Mag. 5, 60–70. https://doi.org/10.1109/MCE.2016.2556879.
Ngu, A.H., Gutierrez, M., Metsis, V., Nepal, S., Sheng, Q.Z., 2017. IoT middleware: a survey on issues and enabling technologies. IEEE Internet Things J. 4, 1–20. https://doi.org/10.1109/JIOT.2016.2615180.
Olariu, S., Stojmenovic, I., 2006. Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting. In: Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications. IEEE, pp. 1–12. doi: 10.1109/INFOCOM.2006.296.
Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.
Ryu, J.H., Irfan, M., Reyaz, A., 2015. A review on sensor network issues and robotics. J. Sensors 2015, 1–14. https://doi.org/10.1155/2015/140217.
S, H., Karnouskos, S., Schroth, C., 2009. The Internet of Things in an Enterprise Context, Future Internet. Futur. Internet Symp. 152–169. doi: 10.1007/978-3-642-00985-3.
Sawai, K., Peng, J., Suzuki, T., 2016. Throughput measurement method using command packets for mobile robot teleoperation via a wireless sensor network. Int. J. Adv. Comput. Sci. Appl. 7. https://doi.org/10.14569/IJACSA.2016.070446.
Singh, S.K., Singh, M.P., Singh, D.K., 2010. Routing protocols in wireless sensor networks-a survey. Int. J. Comput. Sci. Eng. Surv. 1. https://doi.org/10.5121/ijcses.2010.1206.
Singh, S.P., Sharma, S.C., 2015. A survey on cluster based routing protocols in wireless sensor networks. Procedia Comput. Sci. 45, 687–695. https://doi.org/10.1016/J.PROCS.2015.03.133.
Stauffer, N.W., 2013. Reducing wasted energy in commercial buildings [WWW Document].
Talukder, M.Z., Towqir, S.S., Remon, A.R., Zaman, H.U., 2017. An IoT based automated traffic control system with real-time update capability. In: 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, pp. 1–6. doi: 10.1109/ICCCNT.2017.8204095.
Vahabi, S., Eslaminejad, M., Dashti, S.E., 2019. Integration of geographic and hierarchical routing protocols for energy saving in wireless sensor networks with mobile sink. Wirel. Networks 25, 2953–2961. https://doi.org/10.1007/s11276-019-02015-5.
Wu, H., Nabar, S., Poovendran, R., 2011. An Energy Framework for the Network Simulator 3 (ns-3).
Xing, G., Wang, T., Jia, W., Li, M., 2008. Rendezvous design algorithms for wireless sensor networks with a mobile base station. In: Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing – MobiHoc '08. ACM Press, New York, New York, USA, p. 231. doi: 10.1145/1374618.1374650.
Yassen, M.B., Aljawaerneh, S., Abdulraziq, R., 2016. Secure low energy adaptive clustering hierarchal based on internet of things for wireless sensor network (WSN): Survey. In: 2016 International Conference on Engineering & MIS (ICEMIS). IEEE, pp. 1–9. doi: 10.1109/ICEMIS.2016.7745310.
Younis, O., Fahmy, S., 2004. HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. IEEE Trans. Mob. Comput. 3, 366–379. https://doi.org/10.1109/TMC.2004.41.
Tang, B., Wang, J., Geng, X., Zheng, Y., Kim, J.U., 2012. A novel data retrieving mechanism in wireless sensor networks with path-limited mobile sink. Int. J. Grid Distrib. Comput. 5, 133–140.