International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

A survey on energy-efficient clustering techniques and mobile-agent based data aggregation techniques in WSN

Nishant Kumar Pathak[1],Dr.Harmaninderjit Singh Sidhu[2]

Abstract:*Data aggregation is one of Wireless Sensor Networks (WSN) popular techniques in reducing data duplication and increasing the energy efficiency. Sensor node energy constraints include methods of energy intensive data compression in order to extend the lifespan of the network. Among the different methods proposed to improve the efficiency of data aggregation were clustering and mobile agents in WSN. This paper presents a systematic review of the energy-efficient clustering schemes and mobile agent-based schemes used by the data aggregation protocols in WSN. The survey will then present a comparative analysis of clustering schemes and mobile agent-based schemes with a focus on their goals along with their strengths and limitations. This survey facilitates the researchers to select the appropriate clustering schemes and mobile agent based schemes employed by protocols for energy-efficient data aggregation in WSN.*

1. INTRODUCTION

Recent developments in low-power electronics, enabling technology, and omnipresent smart sensors have made wireless sensor network (WSN) one of the most important innovations in the last decade. A WSN combines automated communication systems for sensing, processing, and wireless transmission into small devices known as sensor nodes. These self-governing sensor nodes scatter randomly and densely over the geographic areas, monitor the surrounding environment or detect events, digitize the data collected, and route information to the resource-rich electronics device, referred to as the base station (BS) for further processing and analysis [1]. Sensor nodes however have limited bandwidth, power, memory, resource processing and limited lifetime.

Sensor node's main task is to sense the target phenomenon such as heat, light, temperature and then submit the data to the BS or sink as a query response. For WSNs, compared with data transmission, energy consumption is less for computation. Data generated from neighboring sensors are therefore often redundant and highly correlated. Therefore, instead of sending the sensed data individually to the sink node each time, if data is first collected and aggregated using aggregate functions such as sum), (avg) (etc. and then transmitted to the sink, a significant amount of redundant data is eliminated and thus a lot of energy is saved. Therefore, instead of sending the sensed data individually to the sink node each time, if data is first collected and aggregated using aggregate functions such as sum), (avg) (etc. and then transmitted to the sink, a significant amount of redundant data is eliminated and thus a lot of energy is saved. The aggregation of data can be seen as a fundamental processing procedure for reducing energy consumption and saving limited resources.

An efficient data aggregation strategy will minimize the amount of network traffic in WSN environments resulting in significant energy savings [2]. Throughout the literature, we examined that clustering [3-4] and mobile agents [5-6] are the most commonly used data aggregation techniques in WSN. Clustering the nodes in WSNs is an essential mechanism that provides consolidation of data in a hierarchical network by dividing the

[1,2]*DESH BHAGAT UNIVERSITY, MANDI GOBINDGARH*

network into small clusters. In order to effectively spread the management tasks among the nodes, some of them are chosen as the head of each cluster, commonly referred to as the cluster heads (CHs) and non-CHs nodes called member nodes join their closest CH as shown in Figure 1.The member nodes send their data to the corresponding CHs,

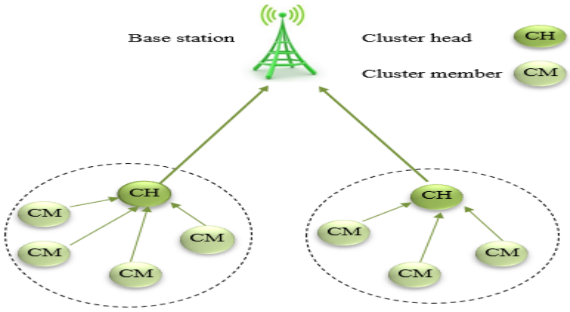


Figure 1: Cluster architecture

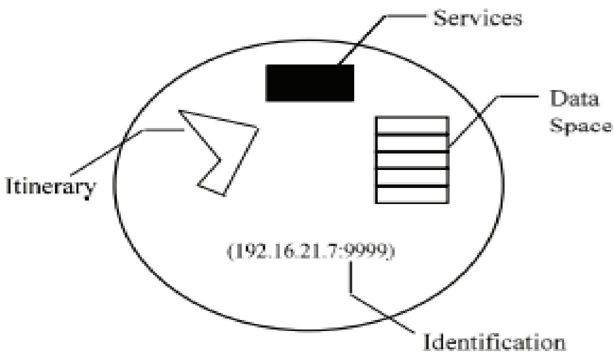


Figure 2: Structure of Mobile Agent

then CHs send aggregated data to the BS. Since a large amount of data generated by the sensors in WSNs is identical, the clustering uses the similarity between the data and then, by aggregating it, reduces the load on the network, resulting in a more effective energy consumption [7]. Another most widely used model for data aggregation in WSN is mobile agent. Mobile agent performs the function of aggregating data at node level rather than sink level, reducing the overhead of redundant network. The mobile agent has described the entity in the sense of WSNs as having four components: mobile agent Identification (ID), itinerary, data space, and services, where mobile agent ID is used to classify the mobile agent uniquely, data space is the data buffer of the agent carrying partially aggregated results, itinerary is the order of sensor nodes to be visited during mobile agent migration and service is the processing code carried with the agent as shown in Figure 2. The computing paradigm based on mobile agents will dramatically reduce bandwidth consumption by transferring the computational process to the position of the sensed data; otherwise its raw type transmission will consume more node energy [8].Thus, this survey explores the recently proposed energy-efficient clustering schemes and mobile-agent based schemes for data aggregation in WSN. The major contributions of the paper are as follows:

1. A detailed description on various energy-efficient clustering schemesemployed in WSN for data aggregation has been introduced.
2. A detailed description on various mobile agent-based schemes for energy-efficient data aggregation in WSN has been presented.

3. A comparative analysis of energy-efficient clustering-based schemes and mobile agent-based schemes for data aggregation with emphasis on their objectives along with their strengths and limitations has been provided.

The rest of the paper is organized as follows. Section 2 describes the energy-efficient clustering schemes employed in WSN for data aggregation. In Section 3, various mobile agent based schemes for data aggregation are discussed. Section 4 provides a comparative analysis of the energy-efficient clustering schemes and mobile agent-based schemes along with their advantages and limitations, and the conclusion is made in Section 5.

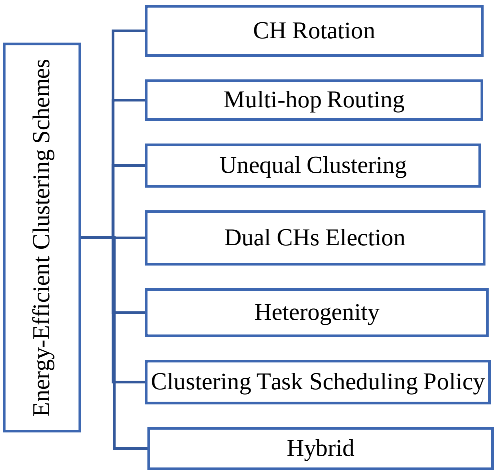


Figure 3: Energy-efficient clustering schemes in WSN for data aggregation

2. ENERGY-EFFICIENT CLUSTERING SCHEMES FOR DATA AGGREGARION

This section describes the various clustering schemes employed in WSN to provide data aggregation support in resource-constrained WSN.Each of the mechanisms has the ability to improve the energy efficiency of the network. Energy-efficient clustering schemes included in Figure 3are discussed in detail in the following subsections.

2.1 CH Rotation

CH rotation is a commonly used technique that aims to reduce the number of cluster head elections and the creation of clusters by rotating the position of CHs among network nodes. CH rotation is a collective event in which all cluster participants participate in the selection of a new CH for the next round. Usually when CH's residual energy falls below a threshold, the current CH selects new CH for the next round of communication based on some predefined CH-selection criteria such as residual energy, inter-cluster size, node-degree etc. The existing CH is a permanent member of the cluster, after choosing the new CH. Therefore, CH rotation balances the distribution of energy between the member nodes and also reduces the amount of control messages [9].

Pachlor et al. [10] proposed a load-aware cluster-head rotation (LAR-CH) solution, which sets a dynamic CH-rotation threshold to reduce the premature death of CH nodes. The LAR-CH uses the current CH energy charge to determine the CH-rotation dynamic threshold. The CH-rotation dynamic threshold decreases the premature death of CH nodes. The LAR-CH ensures a node is only selected as a CH when its residual energy is greater than the energy load for one communication round.

Micheletti et al. [11] proposed a novel approach called CER-CH where the concept of the CH routing tree is combined with the CH rotation to balance node energy consumption and produce more energy efficient CH routing trees. The CER-CH rotation heuristic combines the node residual energy with a node consumption model which estimates the node energy consumption by using local node information (i.e., communication rate, hardware and initial node energy), Data at cluster level (i.e. energy for intra-cluster communication) and knowledge on routing paths (i.e. energy for inter-cluster communication). The CH rotation cycle in CER-CH thus reduces the expense of intra- and inter-traffic energy and thus maximizes the lifespan of the network.

Wang et al. [12]proposed an Energy-Efficient Compressive Clustering Routing (EECSR) protocol that reduces energy consumption by rotating the CH position between the current CH and the Backup CH (BCH) protocol. To prevent an early exhaust of the CH's energy due to its heavier energy consumption load, the method to rotate the CH position was introduced to make energy use more equal.

Therefore, the present CH picked a node with the second highest residual energy value in a cluster as the Backup CH. The present CH informed the BCH of its position through unicast. This unicast rotational mechanism decreases the energy depletion during CH rotation induced by the broadcast. In addition, the optimal cluster size, the optimal CH allocation were proposed to further improve the energy efficiency.

Essa et al [13]proposed a novel energy-efficient protocol, known as the Weight and Energy-Efficient Rotating Clustering Protocol for WSNs (WRCS), which uses a weighting-based approach to elect CHs by considering the background of the node; i.e., its transmission range, degree, remaining energy and neighborhood centrality. To qualify for final CH, the nodes with the highest weight are chosen as CH candidates. Thus optimal CHs are chosen among the network sensor nodes. In addition, this method removes the re-clustering cycle for the entire network in each round, by rotating CH 's position within each cluster in the first step of setup. The CH rotation occurs when the current CH energy has drained below the threshold and then choose the highest weight node from among the members of the cluster. For conventional clustering algorithms, this weight-based CH rotation removes the regular re-clustering process, which has a major impact on energy consumption.

2.2 Unequal Clustering

The CHs close the BS consume more energy in multi-hop communication network, and drain energy faster than the CHs farther from BS. It is because CHs closer to BS are filled with heavy traffic for relaying data to BS due to intra-cluster traffic from their own cluster members, data aggregation, and inter-cluster traffic from other CH's. A CH node failure interrupts network connectivity not only with its member nodes but with neighboring CHs as well. To prevent the rapid depletion of CH 's resources, unequal clustering techniques can be used to balance the load between CH's. Unequal clustering decreases the cluster size closest to BS, and increases the cluster size as the difference between BS and CH increases. Smaller clusters close to the BS have fewer cluster members and less intra-cluster traffic. The smaller clusters thus use less resources for intra-cluster traffic, and focus more on inter-cluster traffic. Similarly, larger clusters further away from BS suggest more members of the cluster and expend more energy on traffic inside the clusters. As a result, it is investing less money on inter-cluster traffic and therefore no need to invest more resources on inter-cluster routing. Thus, unequal clustering allows all CHs to spend the same amount of energy such that the CHs near BS spend the same amount of energy as the CHs farther away from BS and thus effectively balances the load between the sensor nodes [14].

Logambigaiet al. [15]proposed a Fuzzy Logic Based Unequal Clustering Protocol called FBUC which is a distributed clustering algorithm focusing on the method of adding cluster members to the CH community. In FBUC the selection of tentative CHs is based on a probabilistic approach. Fuzzy logic is used after collection of preliminary CHs to determine radius of competition. For the calculation of competition radius the fuzzy input parameters are: residual energy, distance to BS and node degree. Based on the CH degree and distance to CH the nodes join the CH to efficiently use the energy and extend the lifespan of the network. Therefore, the CHs closer to the BS have smaller cluster size to conserve their battery energy for high traffic relay intercluster load.

Baranidharanet al. [16]proposed a Distributed Load Balancing Unequal Clustering termed as DUCG in WSN utilizing Fuzzy approach. DUCF balances the load between the clusters in order to prevent imbalanced clustering, which in effect increases the lifespan of the network. Smaller clusters are created for CHs closer to BS, and larger cluster sizes for CHs further from BS, for unequal clustering. Thus DUCF achieves load balancing by reducing the number of cluster members in a cluster.

Xia et al.[17]proposed an Energy-Efficient Routing Algorithm based on Unequal Clustering and Connected Graph called UCCGRA in WSN.UCCGRA is a distributed solution that increases energy efficiency in two ways: cluster head selection and cluster routing.

UCCGRA uses the voting system to create unequal size clusters that take into account the outcome of the vote and the average transmission power to the cluster members of a potential sensor node. Therefore the range of competition for a node decreases as its distance to the sink decreases. As a result it is predicted that clusters closer to the sink would have smaller cluster sizes. Throughout intra-cluster data processing, they can consume less energy, and conserve more resources for inter-cluster relay traffic.

Kaur et al. [18]proposed a Particle swarm optimization (PSO)-based clustering protocol with unequal and fault tolerance referred to as PSO-UFC in WSN. The protocol suggested addresses problems of imbalanced clustering and fault tolerance for the network's long-run service. The PSO-UFC protocol uses unequal clustering mechanism to balance intra-cluster and inter-cluster energy consumption between the Master CHs (MCHs) to solve the imbalanced clustering problem. The PSO-UFC therefore elects more MCHs in area closer to the BS, so that the MCHs closer to the BS have smaller cluster sizes to save their energy for high inter-cluster relay traffic load. Through using unequal clustering method, the PSO-UFC effectively balances the intra-cluster and inter-cluster energy consumption between the sensor nodes and maximizes the existence of the network.

2.3 Dual CHs Election

The CH bears a lot heavier traffic burden in clustering process than the member of the cluster. Therefore, a CH dissipates energy much more rapidly than its corresponding CMs. Dual CHs are chosen in each cluster in which one performs aggregation and the others perform transmission in order to achieve the energy consumption balance between the cluster nodes. The load is therefore distributed among the dual CHs and hence, the network lifetime is increased.

Jesudurai et al. [19] proposed an Improved Energy Efficient Cluster Head Selection protocol known as IEECHS in WSN. Two CHs are chosen in a separate cluster in the CH election process in IEECHS, and their work in different functions in which one is used for data fusion and another for data forwarding. This two CH selection approach is used to lower the communication costs between two clusters and also to reduce the energy consumption of IoT applications.

Panag et al. [20] proposed a novel dual head static clustering algorithm labeled as DHSCA to balance energy consumption by the sensor nodes and maximize the network lifetime. In DHSCA nodes are split into static clusters depending on their location to reduce the cluster reformation overhead in dynamic clustering. For each communication round two CHs are chosen for each cluster using a weighted function based on their residual energy and their distance from the sink and other cluster nodes, one for data aggregation and the other for data transmission. A node with a smaller value of its average distance from the other cluster members has a greater chance of being selected as an aggregating CH. This reduces the cost of contact in the intracluster. A node that is positioned closer to the sink is more likely to be chosen and function as a transmitting CH. That reduces the cost of contact between clusters. The aggregating cluster head collects, aggregates, and sends the data from the cluster members to the head of the transmitting cluster. The transmitting head of the cluster sends the aggregated data into the sink. If the sink distance from the transmitting cluster head is high, then the aggregated data is directly transmitted to the sink. Otherwise it is transmitted using a multi-hop technique that requires the transmission of other cluster heads. While DHSCA decreases energy consumption during contact between clusters and intra-clusters.

Vhatkar et al. [21] proposed a PDCH protocol based on the Power-Efficient Collection of Sensor Information Systems (PEGASIS) protocol [22] where it forms a chain between communication nodes. Double CH PEGASIS (i.e. PDCH uses the idea that reduces the overhead of cluster forming and also reduces the distance that the sensor node must travel in order to transmit data to the central node and compress the data to be transmitted. Unlike PEGASIS, for contact known as main CH i.e., it uses two CH instead of One CH in a single chain.

MCH, and CH secondary i.e. SCH—- Main chain node is chosen as MCH while branch chain node is chosen as SCH. MCH collects data from other nodes within the same chain in PDCH and needs to send it to SCH while secondary CH collects information from main CH and sends it to the BS. Thus, the job of CH is disseminated between two CHs, this increases the efficiency of energy-use and the balance the load among nodes, and prolong the lifetime.

2.4 Heterogeneity

Increasing the number of sensor nodes increases network capacity, but the cost is very high because installing an additional sensor requires the sensor's cost, which is ten times more than the batteries' cost. Therefore, that the lifespan of the network by installing some sensors with high battery is more acceptable and economical. The sensor networks with these characteristics are called heterogeneous wireless sensor networks, i.e., sensor nodes with different energy levels. Research [116- 120] shows that heterogeneous nodes can extend the life of the network and boost the efficiency of the network without significantly increasing the cost. The heterogeneous nodes are more able to provide data filtering, merging and transport. A heterogeneous node may have one or more heterogeneous resource types, e.g., increased energy capacity or communication capability. They may be equipped with more powerful microprocessor, or more memory, or both, compared to standard nodes. The introduction of heterogeneous nodes increases the capacity in the network and hence the lifespan of the network.

Javaid et al. [24] proposed an enhanced developed distributed energy-efficient clustering system known as EDDEEC for heterogeneous WSN. The EDDEEC protocol heterogeneous network model is based on three energy node levels: standard, advanced, and super. Super and advanced nodes have more power than regular ones. Therefore, for the initial transmission rounds, the super and advanced nodes are generally chosen to be selected as CHs, and when their energy decreases to the same level as normal nodes, these nodes will have the same CH option likelihood as normal nodes. Thus, in the EDDEC protocol, energy is spread efficiently over the network.

Singh et al. [25]proposed a multilevel network model known as HHEDML for heterogeneous wireless sensor networks. The authors introduced five levels of heterogeneity in this work, namely: 0-level, 1-level, 2-level, 3-level, and 4-level energy heterogeneity for both fuzzy and non-fuzzy implementations. The HEEDML-0 assumes that all sensor nodes in a WSN have equal capacity, for which the original HEED is being introduced. The heterogeneity of 1-level, 2-level, 3-level, and 4-level means that the sensor nodes in a WSN are equipped with two , three, four, and five energy levels , respectively. In the case of fuzzy implementation, in addition to the residual energy and node density for selecting the cluster heads, they have found two more parameters of distance and mean energy. HHEDML thus ensures that that the degree of heterogeneity increases the lifetime of the network in a significant proportion as opposed to the increase in the energy of the network.

Singh et al. [26] proposed a heterogeneous distributed energy efficient clustering protocol known as hetDEEC which considers 3-level heterogeneous network model for WSNs to improve the network lifetime. It can define heterogeneity at 1-level, 2-level, and 3-level depending on the value of the model parameter. The hetDEEC heterogeneous network model selects CHs and their respective cluster members by the use of weighted choice likelihood and threshold function. The hetDEEC energy variability helps improve network capacity and longevity of the network.

2.5 Clustering Task Scheduling Policy

To optimize WSN 's durability, it is important to reduce the number of CHs choices in each series. This is because periodic re-clustering at the beginning of each round involves significant communication of messages between nodes; this imposes an extra overhead on the nodes of the network which leads to wasted energy. It is especially important because clustered WSNs are primarily used in applications for data collection (e.g. ecosystem monitoring), which involve periodic sensed data collection. This reduces the scheduling strategy of clustering tasks balances the load on network nodes and thus theoretically prolongs the life cycle of a network.

Neamatollahi et al. [27] proposed a fuzzy-based hyper roundpolicy (FHRP) to efficiently and flexibly schedule the clusteringtask. In FHRP, instead of every round, clustering is performed atthe beginning of every Hyper Round (HR), which is composed ofmany rounds as shown in Figure 4. During the network lifetime, the length of an HRis not fixed and is computed using a fuzzy inference system. Thenode’s residual energy and its distance from the sink are used asthe inputs of this fuzzy system and the HR length is its output.Therefore, with FIS, reclustering is triggered only when urgent, instead of at a fixed period of time by the use of round-based policy. Thus, FHRP decreases the clustering energy overhead by wisely schedulingthe clustering-task through FIS.Neamatollahi et al. [28] proposed an energy efficient distributed scheme for clustering a WSN, i.e. Dynamic Hyper Round Policy (DHRP), which schedules clustering activities to prolong the life of the network and the energy usage. DHRP schedules activating the clustering-task only at the necessary times to achieve load balance. As DHRP calculates the correct reclustering period taking the residual energy of the CHs into account, for most data delivery models, such as continuous, event driven, and query based, this policy is acceptable. Thus DHRP efficiently manages the clustering overhead over the network lifetime by dynamically controlling the re-clustering time.

2.6HybridWang et al.

[29] proposed an energy-efficient clustering routing algorithm that considers uneven cluster forming scheme and distributed rotational cluster head (CH) mechanism. The uneven cluster forming scheme reduces energy consumption and integrates each layer with the specific traffic loads. The CH rotation is based on residual energy and relative location in order to balance the energy consumption in each cluster. Finally, when the CH node rotates or the energy level changes, the routing path among CH nodes will be dynamically modified based on a proposed distance-and-energy-aware cost function to avoid the problem of the energy hole. The combination of irregular clusters, efficient CH node rotation and dynamic inter-cluster packet transmission has helped increase energy efficiency and extend the life of the network.Gajjar et al. [30] proposed FAMACROW which implements routing protocols for CH collection, clustering, and inter-cluster. For cluster head selection,

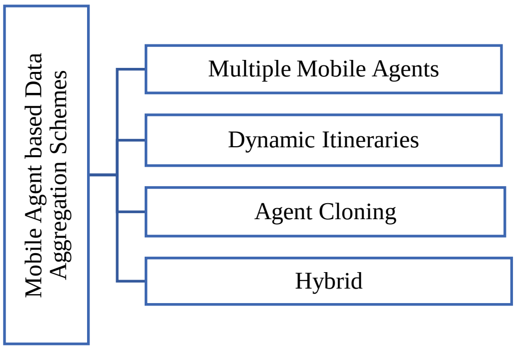
FAMACROW uses fuzzy logic with residual energy, number of neighboring nodes and communication connection quality as input variables. To prevent imbalanced clustering, FAMACROW uses an uneven method of clustering, with clusters closest to MS having smaller sizes than those further away. FAMACROW uses a technique based on Ant Colony Optimization for reliable and energy-efficient multi-hop inter-cluster routing from CHs to BS which takes into account distance from current CH and that from BS, residual energy, length of queue and probability of delivery. The clusters are then set for a number of upcoming rounds and cluster heads are rotated between cluster members according to their sorted PROFICIENCY (identified by order of data slots in TDMA schedule). This reduces the need of CH election in each round and. allocates energy load among all sensor nodes in the network.Sabor et al. [31] proposed a new clustering algorithm called the Unequal Multi-hop Balanced Immune Clustering Protocol (UMBIC) to increase the lifetime of WSN of different densities, small and high. The UMBIC protocol uses the Unequal Clustering Mechanism (UCM) and the Multi-Objective Immune Algorithm (MOIA) to change energy consumption in the clusters and inter-clusters. With relation to BS and residual energy, the UCM is used to partition the network into clusters of unequal size based on distance. While the MOIA constructs among them an optimum cluster and a routing tree centered on covering the entire sensor area, ensuring compatibility between nodes and minimizing the communication costs of all nodes. The UMBIC protocol only rotates the position of CHs among the nodes if the residual energy of one of the current CHs is less than the energy threshold, thus saving time and overhead.

3. MOBILE-AGENT BASED SCHEMES FOR DATA AGGREGARION

This section describes the various mobile agent based schemes for data aggregation in WSN. Each of the schemes has the ability to improve the energy efficiency of the network. Mobile agent-based schemes included in Figure 5 are discussed in detail in the following subsections.

3.1 Multiple Mobile-Agents

Through data aggregation the mobile agent fuses the information of each sensor node to achieve a certain degree of accuracy. Single mobile agent that carries out the task of data aggregation over the entire collection of sensorodes results in poor scalability. It is because when considering partial aggregation scenarios as the mobile agent,a increasing load of data from the sensor nod would need to be carried. Thus, to improve the efficiency of the network, multiple mobile agents may be utilized instead of a single mobile agent.



Gupta et al. [32]] proposed a data dissemination-based energy balanced mobile agents (EBMADD) protocol that uses multiple mobile agents for data dissemination tasks and each mobile agent is responsible for collecting sensed data from a specific group of sensor nodes. The approach proposed consists of three processes: the grouping of sensor nodes, the preparation of the itinerary of the mobile agent and the relocation of mobile agents. The entire monitoring area around the sink is divided into equal size wedges for the grouping of sensor nodes, and the sensor nodes of each wedge are arranged into a balanced tree. Source nodes are spread evenly along several itineraries to balance energy consumption in EBMADD. Static level-order traversal was employed at the sink on the balanced tree of sensor nodes for the planning of the itinerary of the mobile agent. The experimental results indicate that the EBMADD protocol can balance the power consumption and extend the life of the network.

Khandnor et al. [33] proposed An energy-efficient multiple agent-based data aggregation (EEDA-MM) technique that takes account of aggregation ratio, network lifetime and energy consumption. The network is divided into four quadrants, and for each quadrant, a mobile agent is dispatched to obtain data from the quadrant assigned to it. The network is divided into four quadrants, and one mobile agent is assigned to one quadrant for data collection. After collecting data, the mobile agents begin their journey from the BS and return here. The energy used by the sensor nodes is equilibrated to increase the overall lifetime of the network.

Wang et al. [34] proposed a Multi-mobile agent energy efficient itinerary preparation called EMIP in WSN. The EMIP algorithm is an iterative partitioning of a WSN into a directional sector zone where sensor nodes within each sector are grouped together in a separate path. The nodes in a circular zone around the sink are chosen as the starting points for each route (i.e. the centers of the sector zones); the number of itineraries shall be determined by changing the circular zone size. The length of an itinerary is determined by the directional sector zone angle; these angles depend on the density of the nodes and may be different for each path A single itinerary planning algorithm is used when deciding the sectors to determine the visiting order for each sector, and the near-optimal routes for mobile agents can be obtained by efficiently selecting the angle in an adaptive manner.

Gavalas et al. [35] proposed a modern algorithmic method for energy-efficient preparation of mobile agents for data aggregation tasks. The algorithm adopts an iterated local search approach (ILS) in deriving multiple moving mobile agents' hop series over deployed source nodes that incrementally accumulate sensory data as they encounter SNs in a WSN.

ILS specifically integrates in the cost function used the effect of mobile agent size inflation along the path. This uses a detailed approach to measure the real energy costs of future mobile agent transfers taking into account the energy expended on intermediate nodes for data transmission. As a result, the ILS strategy produces low energy cost of the itineraries for mobile operators.

3.2 Dynamic Itineraries

The routes taken by mobile agents moving essentially decide the overall performance of applications for data aggregation. Itinerary preparation may be categorized as static or dynamic according to the position where decisions are made to route mobile agents. A dynamic itinerary scheme decides the route on the fly at each mobile agent hop, while a static scheme derives the route at the mobile agent dispatcher (i.e. sink) node before mobile agent is dispatched and is based on global network topology details. A mobile agent with a dynamic itinerary is more versatile and can adapt by adjusting its itinerary on the fly to faults during its traversal. Nevertheless, failures in the node or connection will invalidate the static itineraries that are defined centrally at the sink. In addition, in static schemes, sink nodes allow global network topology information to decide the itineraries of mobile agents. Although sink does not allow global network topology information to be held in dynamic scheme. Another advantage with the dynamic approach is that the agent packet size is much smaller than the static approach since an agent does not have a pre-computed route list [36].

Lohani et al. [37] proposed a dynamic mobile agent-based data aggregation solution that takes energy consumption, network lifespan, end-to - end latency and aggregation ration into consideration when making the decision to move agent to multihop sensor network. As our approach focuses on finding the most informative route by crossing comparatively fewer nodes, therefore, it takes less time for mobile agents to return to the processing dimension, resulting in lower delays. The proposed solution is therefore capable of balancing energy consumption among the nodes, which increases the overall lifetime of the sensor network. It also offers versatility in determining the impact of individual parameters when determining the MA migration route that can be modified by cost function weight factors. Hence, the cost function has been enhanced such that the maximum possible gain of information can be obtained thus reducing energy consumption, thereby increasing the accuracy of aggregated data.

Dong et al. [38] proposed a dynamic itinerary planning approach for an MA (DIPMA) to find an appropriate route that offers more versatile services to consumers using widespread WSNs. The authors create a new data search method for making the itinerary of a mobile agent according to defined requirements, and build data structures with minimal information stored in sensor nodes, where a mobile agent decides on the next destination based on the information. A mobile agent uses the mechanism to build its itinerary to collect sensory data to predict frost damage. By referring to the vectors of each node, which indicates the characteristics of the stored data in a node, a mobile agent determines its route to the next hop with searching neighbors at n-hop distance. Therefore, DIPMA achieves energy- and time-efficient selection to prolong network life and intelligent monitoring to fit different user requirements.

Xu et al. [39] proposed a Information-driven Mobile agent migration called IDMAP in WSN in which the next mobile agent hop is determined on the fly, based on the current situation of the network. When the mobile agent arrives at a node, it seeks the one with the lowest cost among its unvisited and alive neighbors so that it can decide a near-optimal itinerary. Because only local information is required, sensor nodes do not necessarily need a large transmission range, and thus can reduce the transmission energy consumption. In fact, this complex algorithm can be alerted of the broken connection by not getting the beacons from the neighbor and thereby bypassing the broken link. Hence, the dynamic algorithm has a better score in fault tolerance.

3.3 Agent Cloning

Growing mobile agent will bring a list of pre-computed itineraries which will increase their scale. In addition , mobile agents suffer from the bloated state issue, as they do not place any limits on the size of the agent's payload. Once an agent gathers aggregated data and migrates to the next node, its size increases with the accumulation of aggregated data at each node visited. Despite of this the agent will get bloated. A bloated mobile agent faces a pause in migration and absorbs more energy. Thus, the technique of agent cloning is used in which the mobile agent makes a clone of itself to match the payload in WSN as shown in Figure 6.

Qadori et al. [40] proposed a clone mobile agent route planning approach to reduce the length of the task while improving the event-to-sink efficiency in real-time applications, particularly when a large number of source nodes are visited by the mobile agent. The proposed CMIP adopts the mobile agent cloning principle to minimize the length of the task when the itinerary of the mobile agent has a large number of source nodes to be visited. The CMIP mitigates this issue by splitting the itinerary into sub-itineraries, each of which is assigned to a single MA. In addition, the itinerary of a reversed mobile agent is suggested to further the energy usage and length of the

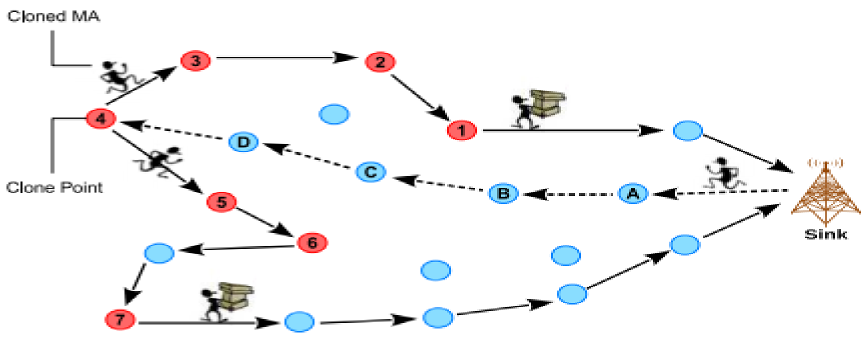


Figure 6: Agent cloning technique

Gupta et al. [41] proposed a clone-based dynamic and distributed agent migration (CDDAM) scheme where each host node uses local information to determine the next-hop for the agents. CDDAM operates in two phases: the rooted creation of tree spans and the clone-based migration of agents. In the clone-based agent migration, sink dispatches an agent to the root node of each spanning tree for data aggregation task. Dispatched MA continues its root node migration and moves to the nodes of the vine. Cloning of the agent is performed at each intermediate node with multiple children, and cloned agents are dispatched from unvisited nodes to tree branches for data aggregation tasks. This scheme also provides a method for local repair of the damaged nodes.

3.4 Hybrid.

Gupta et al. [42] proposed a complex, distributed scheme called energy and trust-conscious mobile agent migration (ETMAM) incorporating energy and confidence as selection criteria for flying routes to complete data aggregation tasks for the traveling agent. A trust assessment process for mobile agent-based WSNs in ETMAM tests the confidentiality of sensor nodes. ETMAM not only provides the moving agents with secure migration paths but also uses cloning mechanisms to optimize the migration path of agents and reduce the payload of agents.

Gupta et al. [43] proposed a Scalable and Load Balanced Mobile Agents Data Aggregation (SLMADA) protocol, in which the agent's itinerary is dynamically chosen at each hop. To avoid the issue of bloated state, SLMADA sets the agent's maximum data payload limit and uses the agent cloning method in which the agent makes a clone of itself according to the size of its payload. To have a load-balanced and scalable solution, an agent cloning technique is employed. SLMADA lets the mobile agent automatically adjust their itinerary according to the network's current state.Gupta et al. [44] proposed a multiple mobile agent with a data dissemination protocol based on dynamic itineraries (MMADIDD) that uses multiple mobile agents for data dissemination tasks and each mobile agent is responsible for collecting sensed data from a specific area and deciding its flying route at each hop using local data.

The proposed protocol is designed to track applications to collect the periodic sensed data from ambient conditions, such as temperature, humidity and pressure. The protocol dynamically specifies the order of sensor nodes that will be visited during the migration of mobile agents. The circular sensing region is divided into several equiangular wedges centered at the drain. The proposed protocol not only adjusts to unexpected node failures but also extends the network lifetime.

Rais et al. [45] proposed an Efficient Multi-agent route planning (OMIP) algorithm, where the source nodes are grouped into clusters. OMIP adopts Efficient Clustering Routing Protocol (ECRP) for partitioning the network into v-clusters, choosing a medoid node (MN) in each cluster, and then reducing the average distance between MN and other cluster nodes. The MN in OMIP serves as the CH, as well as the mobile agent's launch and arrival node. The sink in OMIP dispatches a mobile agent to any cluster in the network. The mobile agent begins the process of data collection from MN, roams the source nodes within the cluster and returns to MN. Once the MA finishes data collection, the aggregated data moves back to the sink. The MN in OMIP serves as the CH, as well as the mobile agent's launch and arrival node. The sink in OMIP dispatches a mobile agent to any cluster in the network. The mobile agent begins the process of data collection from MN, roams the source nodes within the cluster and returns to MN. Once the MA finishes data collection, the aggregated data moves back to the sink.

El Fissaoui et al. [46] proposed a Multi-mobile Power and Fault Aware Agent (MAEF) planning itinerary in WSN. MAEF groups the nodes of the source into clusters and each cluster has a CH. In MAEF the itineraries of the mobile agents are only arranged among the CHs and decided using MST. In addition, the number of routes equals the number of nodes within the range of transmission of the sink. In MAEF, once the mobile agent first enters the CH, it notifies the source nodes of the CH to forward their data to the CH. Upon hitting the last CH in its itinerary, the mobile agent starts collecting data from CHs on its way back to the sink. MAEF's itinerary duration is the shortest, this is because of the policy of only visiting CHs and not all sensor nodes. In the event of the failure of the first itinerary due to node failure as is often the case in WSNs, the mobile agents measure alternate paths and are ready to go.

Table 1 Mobile agent-based schemes for data aggregation

Mobile agent-based Schemes	**Protocols Employed**	**Objectives**	**Merits**	**Limitations**
Multiple Mobile agents	EBMADD [32] EEDA-MM [34] EMIP [35] Gavalas et al. [35] SLMADD [43] OMIP [45] MAEF [46]	• Uses multiple mobile agents to carry a growing load of data retrieved from sensor nodes to the sink	• Improves efficiency of network • Distributes load among the multiple agents	• High Computational complexity •
Dynamic Itineraries	Lohani et al. [37] DIPMA [38] IDMAPC [39] SLMADD [43] OMIP [45] MAEF [46] ETMAM [42]	• Determines the route on the fly at each hop of the mobile agent	• Flexible and adaptable to faults during its traversal by changing its itinerary on the fly • Sink does not require to maintain global information of network topology • The size of the agent packet is much smaller	• Consumes valuable node energy resource • Requires larger mobile agent
Agent Cloning	ETMAM [42] CMIP [40] CDDAM [41] SLMADD [43]	• Mobile agent makes clone of itself at the cloning point for balancing the payload in WSN.	• Avoid bloated problem in mobile agent	• Suffers from scalability issue

4. COMPARATIVE ANALYSIS

In this section, we summaries and compare all the energy-efficient clustering schemes and mobile agent-based schemes for data aggregation discussed in this survey. Table 1 and 2 provides a brief summary of energy-efficient clustering schemes and mobile agent-based schemes, respectively employed in WSN with emphasis on their objectives along with their advantages and limitations.

Table 2 Energy-efficient clustering schemes for data aggregation

Clustering Schemes	Protocols Employed	Objectives	Merits	Limitations
CH Rotation	LAR-CH [10] CER-CH [11] EECSR [12] WRCS [13] Wang et al. [29] FAMACROW [30] UMBIC [31]	• Rotate the role of CHs amongst the member nodes in the cluster, usually when the residual energy of CH falls below a threshold, • The current CH selects new CH for next communication round based on some predefined CH-selection criteria such as residual energy, inter-cluster distance, node-degree etc.	• Reduces the communication overheads of CH election phase • Maximizes network lifetime	• CH rotation is a static process does not allow the addition of new nodes to the system and does not adjust its behavior for dead nodes. • The large amount of broadcast leads to a waste of energy and the decline of energy efficiency.
Unequal Clustering	FBUC [15[DUCF [16] UCCGRA [17] PSO-UFC [18] Wang et al. [29] FAMACROW [30] UMBIC [31]	• Generates cluster of smaller size in the area closer to BS and the cluster size increases as the distance between the BS and CH increases. • Smaller cluster near the BS to handle the large intra-cluster traffic load.	• Achieves load balancing • Maximizes network lifetime • Suitable for large-scale WSNs	• Effective computational intelligence technique is required for optimal CH election
Dual CHs Election	IEECHS [19] DHSCA [20] PDCH [21]	• Elects dual CHs in each cluster in which one performs aggregation and the others performs transmission.	• Distributes load among the dual CHs • Maximizes network lifetime	• Frequent re-clustering is required for CHs election
Heterogeneity	EDDEEC [24] HHEDML [25] HetDEEC [26]	• Deploy nodes possess one or more types of heterogeneous resources, e.g., enhanced energy capacity or communication capability	• Equally distributes energy among the sensor nodes in the network • Maximizes network lifetime	• Heterogenous nodes are more expensive than the homogeneous nodes • Requires effective deployment strategy
Clustering Task Scheduling Policy	FHRP [27] DHRP [28]	• Reduces the frequency of re-clustering by keeping the same CH for the next round, until its residual energy remains less than threshold	• Reduces communication overheads of clustering • Prolongs network lifetime	• Do not achieve the hyper round length due to its distributed nature

5. CONCLUSION

The collection of data in WSN has become a big concern for the researchers around the world. A novel taxonomic research on energy-efficient clustering schemes and mobile agent-based schemes employed by the

WSN protocols for data aggregation was presented in the current survey.The importance of data aggregation in resource-constrained WSN reflects a thorough study of various clustering schemes and mobile agent based schemes with focus on their goals. The comparative analysis enables the selection of suitable clustering schemes and mobile agent-based schemes used in WSN to enable energy-efficient aggregation of the data.

References

1. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., &Cayirci, E. (2002). Wireless sensor networks: a survey. Computer networks, 38(4), 393-422.
2. Qayyum, B., Saeed, M., & Roberts, J. A. (2015). Data aggregation in wireless sensor networks with minimum delay and minimum use of energy: A comparative study. In Accepted for publication in Electronic Workshops in Computing (eWiC). British Computer Society
3. Abbasi-Daresari, S., &Abouei, J. (2016). Toward cluster-based weighted compressive data aggregation in wireless sensor networks. Ad Hoc Networks, 36, 368-385.
4. Dhand, G., & Tyagi, S. S. (2016). Data aggregation techniques in WSN: Survey. Procedia Computer Science, 92, 378-384.
5. Lohani, D., & Varma, S. (2016). Energy efficient data aggregation in mobile agent based wireless sensor network. Wireless Personal Communications, 89(4), 1165-1176.
6. El Fissaoui, M., Beni-hssane, A., Ouhmad, S., & El Makkaoui, K. (2020). A Survey on Mobile Agent Itinerary Planning for Information Fusion in Wireless Sensor Networks. Archives of Computational Methods in Engineering, 1-12.
7. Afsar, M. M., &Tayarani-N, M. H. (2014). Clustering in sensor networks: A literature survey. Journal of Network and Computer Applications, 46, 198-226.
8. El Fissaoui, M., Beni-hssane, A., Ouhmad, S., & El Makkaoui, K. (2020). A Survey on Mobile Agent Itinerary Planning for Information Fusion in Wireless Sensor Networks. Archives of Computational Methods in Engineering, 1-12.
9. Xu, L., Collier, R., & O'Hare, G. M. (2017). A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios. IEEE Internet of Things Journal, 4(5), 1229-1249.
10. Pachlor, R., &Shrimankar, D. (2018). LAR-CH: A cluster-head rotation approach for sensor networks. IEEE Sensors Journal, 18(23), 9821-9828.
11. Micheletti, M., Mostarda, L., & Navarra, A. (2019). CER-CH: combining election and routing amongst cluster heads in heterogeneous WSNS. IEEE Access, 7, 125481-125493.
12. Wang, Q., Lin, D., Yang, P., & Zhang, Z. (2019). An energy-efficient compressive sensing-based clustering routing protocol for WSNs. IEEE Sensors Journal, 19(10), 3950-3960.
13. Essa, A., Al-Dubai, A. Y., Romdhani, I., &Esriaftri, M. A. (2017, May). A new weight based rotating clustering scheme for WSNS. In 2017 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6). IEEE.
14. Arjunan, S., &Pothula, S. (2019). A survey on unequal clustering protocols in Wireless Sensor Networks. Journal of King Saud University-Computer and Information Sciences, 31(3), 304-317.
15. Logambigai, R., & Kannan, A. (2016). Fuzzy logic based unequal clustering for wireless sensor networks. Wireless Networks, 22(3), 945-957.
16. Baranidharan, B., &Santhi, B. (2016). DUCF: Distributed load balancing Unequal Clustering in wireless sensor networks using Fuzzy approach. Applied Soft Computing, 40, 495-506.
17. Xia, H., Zhang, R. H., Yu, J., & Pan, Z. K. (2016). Energy-efficient routing algorithm based on unequal clustering and connected graph in wireless sensor networks. International Journal of Wireless Information Networks, 23(2), 141-150.
18. Kaur, T., & Kumar, D. (2018). Particle swarm optimization-based unequal and fault tolerant clustering protocol for wireless sensor networks. IEEE Sensors Journal, 18(11), 4614-4622.
19. Jesudurai, S. A., &Senthilkumar, A. (2019). An improved energy efficient cluster head selection protocol using the double cluster heads and data fusion methods for IoT applications. Cognitive Systems Research, 57, 101-106.
20. Panag, T. S., & Dhillon, J. S. (2018). Dual head static clustering algorithm for wireless sensor networks. AEU-International Journal of Electronics and Communications, 88, 148-156.

21. Vhatkar, S., Shaikh, S., &Atique, M. (2017, February). Performance analysis of equalized and double cluster head selection method in wireless sensor network. In 2017 Fourteenth International Conference on Wireless and Optical Communications Networks (WOCN) (pp. 1-5). IEEE.
22. Lindsey, S., & Raghavendra, C. S. (2002, March). PEGASIS: Power-efficient gathering in sensor information systems. In Proceedings, IEEE aerospace conference (Vol. 3, pp. 3-3). IEEE.
23. Rostami, A. S., Badkoobe, M., Mohanna, F., Hosseinabadi, A. A. R., &Sangaiah, A. K. (2018). Survey on clustering in heterogeneous and homogeneous wireless sensor networks. The Journal of Supercomputing, 74(1), 277-323.
24. Javaid, N., Rasheed, M. B., Imran, M., Guizani, M., Khan, Z. A., Alghamdi, T. A., &Ilahi, M. (2015). An energy-efficient distributed clustering algorithm for heterogeneous WSNs. EURASIP Journal on Wireless communications and Networking, 2015(1), 1-11.
25. Singh, S., Chand, S., & Kumar, B. (2016). Energy efficient clustering protocol using fuzzy logic for heterogeneous WSNs. Wireless Personal Communications, 86(2), 451-475.
26. Singh, S., Malik, A., & Kumar, R. (2017). Energy efficient heterogeneous DEEC protocol for enhancing lifetime in WSNs. Engineering Science and Technology, an International Journal, 20(1), 345-353.
27. Neamatollahi, P., Naghibzadeh, M., &Abrishami, S. (2017). Fuzzy-based clustering-task scheduling for lifetime enhancement in wireless sensor networks. IEEE Sensors Journal, 17(20), 6837-6844.
28. Neamatollahi, P., Abrishami, S., Naghibzadeh, M., Moghaddam, M. H. Y., & Younis, O. (2017). Hierarchical clustering-task scheduling policy in cluster-based wireless sensor networks. IEEE Transactions on Industrial Informatics, 14(5), 1876-1886.
29. Wang, Z., Qin, X., & Liu, B. (2018, April). An energy-efficient clustering routing algorithm for WSN-assisted IoT. In 2018 IEEE Wireless Communications and Networking Conference (WCNC) (pp. 1-6). IEEE.
30. Gajjar, S., Sarkar, M., & Dasgupta, K. (2016). FAMACROW: Fuzzy and ant colony optimization based combined mac, routing, and unequal clustering cross-layer protocol for wireless sensor networks. Applied Soft Computing, 43, 235-247.
31. Sabor, Nabil, Mohammed Abo-Zahhad, Shigenobu Sasaki, and Sabah M. Ahmed. "An unequal multi-hop balanced immune clustering protocol for wireless sensor networks." Applied Soft Computing 43 (2016): 372-389.
32. Gupta, G., Misra, M., & Garg, K. (2011, December). An energy balanced mobile agents based data dissemination protocol for wireless sensor networks. In Proceedings of the 1st International Conference on Wireless Technologies for Humanitarian Relief (pp. 89-95).
33. Khandnor, P. (2019). Energy Efficient Data Aggregation Using Multiple Mobile Agents in Wireless Sensor Network. In Smart Innovations in Communication and Computational Sciences (pp. 279-287). Springer, Singapore.
34. Wang, J., Zhang, Y., Cheng, Z., & Zhu, X. (2016). EMIP: energy-efficient itinerary planning for multiple mobile agents in wireless sensor network. Telecommunication Systems, 62(1), 93-100.
35. Gavalas, D., Venetis, I. E., Konstantopoulos, C., &Pantziou, G. (2016). Energy-efficient multiple itinerary planning for mobile agents-based data aggregation in WSNs. Telecommunication Systems, 63(4), 531-545.
36. El Fissaoui, M., Beni-hssane, A., Ouhmad, S., & El Makkaoui, K. (2020). A Survey on Mobile Agent Itinerary Planning for Information Fusion in Wireless Sensor Networks. Archives of Computational Methods in Engineering, 1-12.
37. Lohani, D., & Varma, S. (2016). Energy efficient data aggregation in mobile agent based wireless sensor network. Wireless Personal Communications, 89(4), 1165-1176.
38. Dong, M., Ota, K., Yang, L. T., Chang, S., Zhu, H., & Zhou, Z. (2014). Mobile agent-based energy-aware and user-centric data collection in wireless sensor networks. Computer networks, 74, 58-70.
39. Xu, Y., & Qi, H. (2007). Dynamic mobile agent migration in Wireless Sensor Networks. International Journal of Ad Hoc and Ubiquitous Computing, 2(1-2), 73-82.
40. Gupta, G. P., Misra, M., & Garg, K. (2014). Energy and trust aware mobile agent migration protocol for data aggregation in wireless sensor networks. Journal of Network and Computer Applications, 41, 300-311.

41. Qadori, H. Q., Zukarnain, Z. A., Alrshah, M. A., Hanapi, Z. M., & Subramaniam, S. (2018). CMIP: Clone Mobile-Agent Itinerary Planning Approach for Enhancing Event-to-Sink Throughput in Wireless Sensor Networks. IEEE Access, 6, 71464-71473.
42. Gupta, G. P., Misra, M., & Garg, K. (2015). An Energy Efficient Distributed Approach-Based Agent Migration Scheme for Data Aggregation in Wireless Sensor Networks. Journal of Information Processing Systems, 11(1).
43. Gupta, G. P., Misra, M., & Garg, K. (2017). Towards scalable and load-balanced mobile agents-based data aggregation for wireless sensor networks. Computers & Electrical Engineering, 64, 262-276.
44. Gupta, G. P., Misra, M., & Garg, K. (2012, January). Multiple mobile agents based data dissemination protocol for wireless sensor networks. In International Conference on Computer Science and Information Technology (pp. 334-345). Springer, Berlin, Heidelberg.
45. Amine, R., Khalid, B., & Mohamed, O. (2018). Determination of itinerary planning for multiple agents in wireless sensor networks. International Journal of Communication Networks and Information Security, 10(1), 99-109.
46. El Fissaoui, M., Beni-Hssane, A., &Saadi, M. (2018). Multi-mobile agent itinerary planning-based energy and fault aware data aggregation in wireless sensor networks. EURASIP Journal on Wireless Communications and Networking, 2018(1), 92