Localised Energy Based Clustering with Incentives for Efficient M2M Communications

Raymond W. Juma, Anish M. Kurien, and Thomas O. Olwal
Department of Electrical and Electronic Engineering, Tshwane University of Technology, Pretoria, South Africa
Email: {rjwekesa, thomas.olwal}@gmail.com, kurienam@tut.ac.za

Abstract — In the existing literature, multi-hop communication-based clustering techniques for Machine-type Devices (MTDs) have been extensively studied to ensure energy-efficient Machine to Machine (M2M) communications. The techniques presented demonstrated advantages such as improved scalability and reliability performance in large scale M2M communication networks. However, significant waste in energy has been noted with some of the techniques during cluster formation and due to the inherent selfish behaviours of some of the MTDs when routing traffic from the edge to the sink regions. To mitigate selfish behaviours, encourage cooperation, and improve efficient energy performance amongst MTDs, this paper proposes a new method of clustering MTDs using local energy parameters augmented incentive, referred to as Local Energy based Clustering with Incentive Algorithm (LECIA). In this work, probing signals from the MTDs are considered to partition the network into regions. Local energy parameters are identified and then applied to cluster the MTDs in the partitioned regions. Centralised relay selection and incentive management system (CRSIMS) are invoked for relay device selection and stimulation of multihop transmissions respectively. Simulation results have indicated that the proposed approach has on average 5% and 37% more number of surviving devices, and 6% and 55 % more amount of remaining energy than the closely related conventional approaches, namely, the Hybrid Energy Efficient Distributed (HEED) and the Low Energy Adaptive Clustering Hierarchy-Centralised (LEACH), respectively.

Index Terms—Clustering, energy efficient, local energy, M2M, Partition, incentive

I. Introduction

The introduction of M2M communications in the recent past has made it possible for devices to communicate amongst themselves. This is the backbone of the Internet of Things (IoT) [1]. The noteworthy applications of M2M communications are observed in areas such as e-health, smart homes, habitat monitoring, and traffic management [2]. To support the autonomous M2M devices in prolonging their operating time, an energy efficient design is necessary. The M2M communication devices have distinctive features that include; massive transmissions from many MTDs, low data rates, low mobility and low energy [3]. The consumption of energy in M2M devices occurs during sensing, processing and data transmission. A large amount of energy is consumed during the data transmission phase [4], especially if the distance of transmission is quite significant. To reduce energy consumption that takes place during data transmission process, the devices are organized into clusters that are supervised by cluster heads (CHs) [5]-[7]. Other advantages of the clustering technique include high scalability, energy efficiency, privacy preservation, social networking, and prolonged life time of devices [8], [9]. In addition, clustering brings about efficient dynamic routing amongst devices and from devices to the sink [5].

To ensure that M2M communications are energy efficient, and are able to operate for longer periods of time, different clustering techniques have been proposed [5], [9]. Using clustering, the CHs aggregate the data packets of the cluster members, through multi hop transmission data are delivered to the destination [10], [11]. Nevertheless, the energy consuming techniques during cluster formation, the selfish conduct of some of the devices along the route to the destination, the depletion of energy for devices that are used as relay devices, are some of the problems faced by clustering. To mitigate these problems, this paper proposes a novel approach, Localised Energy based Clustering with Incentive Algorithm (LECIA) approach. In this approach, the device (s) probing signal is applied to partition the devices in the network into regions. The devices in the farthest region from the BS are organised into clusters. The formation of clusters is based on local energy parameters, with the device of low energy consumption ratio within a given radius being selected as a CH. The devices in the region closer to the BS are not organized into clusters. This helps them to conserve their energies that are normally lost during cluster formation. A centralized relay selection and incentive management system (CRSIMS) is invoked for the selection of a suitable relay device in the near region and the provision of incentives to the selected relay devices respectively. Data packets for the devices located in the far region are delivered to the BS/sink through the selected relay devices from the near region. The summary of the main contributions of this paper are given below:

1. A novel approach, referred to as LECIA, is proposed and designed, which utilises local energy parameters and incentives to minimise energy consumption in the M2M communications.

Manuscript received May 5, 2019; revised November 8, 2019.
This work was supported by Tshwane University of Technology.
Corresponding author email: rjwekesa@gmail.com.
doi:10.12720/jcm.14.12.1168-1179

2. Through Simulation, it is shown that the proposed, LECIA approach, can improve energy efficiency in M2M communications. LECIA displays a higher number of surviving devices, residual energy in the network and delivers higher number of data packets compared to the traditional clustering approaches HEED [9], [12] and LEACH [9], [13].

The remainder of this paper is organised as follows: Section II. discusses the related work regarding clustering approaches in M2M communications. In Section III, the system model is presented. In section IV, the details of the proposed technique is presented. In Section V, the performance and simulation results are evaluated. Finally, Section VI, highlights the conclusions and presents future work.

II. Related Work

The previous research works have proposed several clustering algorithms for energy efficiency in M2M communication networks [9]-[11]. The taxonomy of these clustering algorithms is discussed in this section in the following order: distributed clustering followed by centralised clustering.

A. Distributed Clustering Algorithms

Higher energy consumption in M2M communications is quite costly to operators. As a solution to high energy consumption, the distributed clustering algorithm (DCA) is examined in the existing literature as the key approach to reduce high energy consumption in M2M communications [14], [15]. The distributed clustering algorithms are associated with networks where devices are not centrally controlled in making decisions [16]. The routing decisions are determined based on the internal information. In this class of distributed clustering algorithms, is the low energy adaptive clustering hierarchy (LEACH) [13]. LEACH is an algorithm where devices are assumed to be homogeneous. The devices are organised into clusters without any control from the central point. Any member of the cluster can be elected as a Cluster Head (CH) to receive, aggregate and forward data packets of its members to the sink. It is an adaptive clustering algorithm that utilises a random approach in its even distribution of energy load in the communication network. The formation of clusters that are supervised by the CHs in LEACH brings about energy saving among the devices in the communication network. However, the probabilistic approach in selecting the CHs among the devices within a cluster can result into a device of low energy being selected; the low energy device selected will be eliminated too early and will end the life of that cluster. When a CH is located far from the BS, larger amount of energy is consumed during data packet transmissions. In [Liu, 08], the author proposed the LEACH DCHS approach which is founded on the deterministic techniques in the CH selection. However, unlike the LEACH approach [13], in the LEACH DCHS approach, the continuance of the steady state phase of the algorithm is prolonged and the concept of selecting new CHs in each round is maintained as it is in the LEACH approach. The generation of clusters in a new round only takes place when the maintained clusters reaches half the figure of the original clusters.

Stable Election Protocol (SEP) is proposed in [17]. In SEP devices are autonomous and can elect themselves as CHs depending on their energy levels. It is assumed that a certain percentage of devices have more battery power (advanced nodes) than others (normal nodes). Those with higher battery power stand a higher chance of being elected as CHs in the communication network. However, the method is probabilistic. As an improvement to SEP, a modified SEP (M-SEP) was proposed in [18]. M-SEP is characterised with the election of the CHs that are distributed in two, three up to the ninth hierarchy in a communication network. Both SEP and M-SEP are proposed to reduce high energy consumption in WSNs. However, for SEP, the election of CHs from the two categories of devices, advanced and normal nodes are probabilistic, and the approach is not dynamic. In the event of a normal node being selected as a CH, the performance of the cluster might be interrupted in case the device's energy is exhausted before another device is selected as a CH. In M-SEP, the random number approach utilised in the process of selecting the CHs does not guarantee that a device of high battery energy is being selected. In case of a device of lower battery energy is selected, the expected performance of the cluster and the whole network is lowered.

Energy Efficient Clustering Scheme (EECS) is discussed in [19]. The algorithm through local radio communication selects nodes with high residual energy as CHs and balances their distribution in the network. A novel technique is introduced that balances the load among the CHs. EECS utilises a long range of transmission from CH to the sink/BS which is detrimental to energy saving principle in a large network. To solve the problem of high energy consumption caused by long range data transmissions, the HEED algorithm is proposed in [12]. HEED utilises multihop inter cluster communication. In the selection of CHs, the residual energy of the node is considered. Despite its fair performance compared to LEACH and EECS, nodes close to the sink are over utilised in passing data to the sink, and therefore get eliminated from the network early. HEED also assumes that the devices closer to the BS automatically accept to be used as relays in forwarding data packets, this might not be the case. Some of the devices closer to the sink/BS are selfish and need incentives for them to accept to cooperate with the devices located far from the sink/BS in relaying data packets.

B. Centralised Clustering Algorithms

As a way of reducing energy consumption during cluster formation, devices can be controlled from a central location. The approach is referred to as the centralised clustering algorithm (CCA). The decisions of all the devices in the network is made from the central location as is presented in [14], [16]. The application of centralised clustering algorithms is suitable for location aware sensors [20]. In this approach, the devices in the network forward their specifications that includes location, identification (ID), and residual energy to the BS. The BS plays the role of determining the number of clusters, size of clusters and cluster head location in the network. Among the members of this group is the Low Energy Adaptive Clustering Hierarchy-Centralised (LEACH-C) [21] and the Power-Efficient Gathering in Sensor Information Scheme PEGASIS [22]. The advantage of LEACH-C is observed during cluster formation which is performed by the BS. The devices do not suffer energy losses that occur during a process such as selection of CHs. However, greater energy is consumed during packet transmissions to the sink. This happens when the sink is located far from the nodes in the network. LEACH-C performance is based on the sink/BS location [23]. The chain based PEGASIS algorithm does not suffer from overhead problems that arises from overheads associated with clustering processes. The only undoing of the approach is that it suffers from a redundancy problem and latency that is observed in data packet transmission for the distant nodes.

As presented in the above paragraph, the CCA approach brings about a reduction in energy consumption in the communication network. However, the fact that all the processes are controlled by the BS makes it a time-consuming algorithm. This is most true when the devices are massive as is the case with the MTDs. CCA is not suitable for large sized network since it has weak scalability.

C. Remarks on Related Work and Proposed Method

The preceding section has presented samples of clustering techniques that are proposed to reduce energy consumption in MTDs. The literature introduced multi-hop communication within WSNs to foster long distance transmission of data to the BS/sink [6], [9] [12], [13], [16]-[19], [22]. However, multi-hop transmissions usually face the problem of selfish behavior among participating nodes in a network in which case some nodes, especially those close to the BS/sink do not assist others with forwarding data packets because of fear of losing energy. It has also been noted that most of the approaches assumed that all the nodes/devices in the network are homogeneous and that a random approach was applied in selecting the CHs. The former does not present a practical situation, while the later presents chances of selecting an inferior device in terms of residual energy to be a CH.

This paper proposes a new method of clustering devices using a local energy parameter augmented incentive (LECIA) approach for efficient M2M communications. The proposed approach considers partitioning of the network into regions and the selection of CHs during the formation of clusters based on probing signals and the local energy parameters respectively. To stimulate multi-hop transmissions, centralised relay selection and incentive management system (CRSIMS) is invoked. CRSIMS involves selection of relay devices through which data packets from the far region (FR) are delivered to the BS/sink and the provision of incentives in the form of energy credits to the selected relay devices.

III. System Model

This section defines the proposed system model and the energy consumption model for the LECIA technique.

A. Proposed System Model

Consider that we have a single base station (BS) serving several heterogeneous MTCDs that are randomly distributed in a geographical region. In this scenario, a set of devices is represented as $MTCD_n, n = \{1,2, \dots d \dots, D\}$. MTCDs are assumed to have short-range transceiver communications and use low power to be cost effective. Suppose it is further considered that these MTCDs are battery powered and that their batteries are assumed to be impractical to recharge and replace. Hence, there is a need to minimise the energy consumption with these sets of MTCDs that are assumed to transmit packets regularly in the uplink direction. It is assumed that once MTCDs are distributed in the network, they remain stationary. As presented in section II, this design explores the idea of cooperative multi-hop transmissions of data packets generated by the MTCDs located in the outer regions of the network through another interior node that travels to the destination. Following such a concept, the network is then partitioned into near and far regions (NR and FR) with the BS as the reference point. The devices in the FR are organised into clusters that are supervised by cluster heads (CHs). The CHs in the FR aggregate data packets from Cluster Members (CM) and forwards to the optimally selected relay devices in the NR which will subsequently be referred to as MTC gateway (MTCG) in this paper. The gateways finally connect to the BS. The devices in the NR not selected as relay devices reach the BS on an individual basis. The proposed system model is shown in Fig. 1.

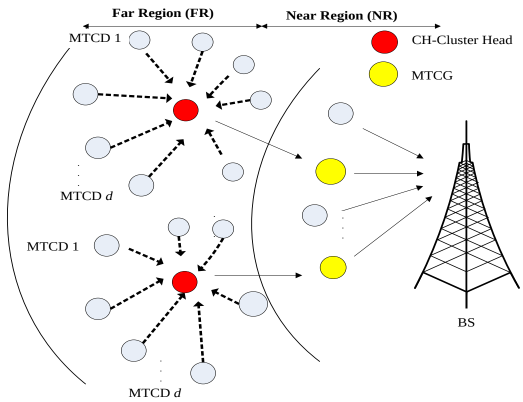

Fig. 1. Proposed system model

The energy consumption process as data packets are transmitted through a distance d to the BS/sink is described by the energy consumption model presented in section B.

B. Energy Consumption Model

In the computation of energy used during the transmission of data packets, an energy consumption model for radio propagation medium is considered [24]. Consider that L data packets need to be transmitted through a distance d, the energy consumed during the transmission can expressed as

$$E_{TX}(l,d) = \begin{cases} L \times E_{elec} + (L \times \epsilon_{fs} \times d^2), & d < d_0 \\ L \times E_{elec} + (L \times \epsilon_{mp} * d^4), & d \geq d_0 \end{cases} \quad (1)$$

The transmitter/receiver circuit are represented by E_{elec}, ϵ_{fs} and ϵ_{mp}. Depending on the distance between the transmitter and receiver, free space (ϵ_{fs}) or multipath (ϵ_{mp}) fading channel model is applied. Considering that L data packets are received, then the energy expended can be expressed as

$$E_{RX} = (L * E_{elec}) \quad (2)$$

In the next section, a complete description of the proposed approach is presented.

IV. A Local Energy Based Clustering with Incentive Algorithm

The proposed, Local Energy based Clustering with Incentive Algorithm (LECIA) is presented in three sub-sections: 1. Network partitioning and Cluster formation, 2. Centralised relay selection and incentive management system, and 3. Ilustrates the flow diagram of the proposed approach and pseudo codes.

1. Network Partitioning and Cluster Formation

In order to ensure that devices embrace the localised energy clustering approach, the network is partitioned into regions. The partitions mark the boundaries of the clusters. In this work, the devices at the farthest region transmit to local cluster heads (CHs) through multi-hop transmissions, with data packets being delivered to the BS/sink. The devices in the near region are not organised into clusters. They communicate directly with the BS.

A. Network Partitioning

The partitioning process considers the work proposed in [25]. The source device (S_D) which is assumed to be at the furthest location from the BS sends a probing signal (Ps) to the destination device (D_D). The D_D is assumed to be at zero distance from the BS. When D_D can decode the message in the first phase and send an acknowledgement which is received by S_D, it means that S_D can establish a direct link with S_D with lower energy strain. The network will have no partitions. However, when the network is wide such that S_D and D_D are out of range, the probing signal from S_D is received by the devices within a close range. The devices within the close range will be able to decode the received signal and send acknowledgements to S_D and communicate the received SNR to the BS. The devices within this range are set to belong to the far region 1, denoted by $FR1 = \{FR_i\}$, where $i = \{1,2,\ldots,p\}$.

The same process is repeated in the second probing phase where, decode and forward $FR1$ devices concurrently broadcasts the probing signal. The receiver device applies maximal ratio combiner (MRC) explained in [26] to combine all the received copies of signals. It is assumed that the channel state information (CSI) is perfect. At this stage, if D_D can accurately decode the signal and send an acknowledgement that is received at$FR1$, the received SNR is reported to the BS at the same time. In this case, two regions will be marked and declared as near region denoted as NR. However, if the destination cannot be reached after the second probing phase, then a third probing phase is initiated. In the third probing phase, the devices that were able to successfully decode the signal from $FR1$ will be described as belonging to far region 2, denoted by $FR2 = \{FR_q\}$, where $q = \{1,2,\ldots,m\}$. The process continues such that $FR2$ devices concurrently broadcast the probing signal. In this study, the analysis is limited to two regions: the FR and IR. The second probing reaches the destination point. After partitioning the network into two regions, the devices in FR are organised into clusters. While some devices in the NR individually transmit their own data packets to the BS/sink, some are considered as relay devices based on a cost function criterion that is discussed in section IV (2A).

B. Cluster Formation

This sub - section explains the clustering procedure that takes place in the FR. It is the technique relies on the local energy consumption process to select the CH. The process considers local energy consumption ratio S_i. derivations presented in [27]. The process goes through three steps: (i) Device local energy parameter phase, (ii) Cluster head selection and (iii) Cluster formation with prediction times of T_1, T_2 and T_3 respectively.

1) Device local energy parameter phase

This phase is used to predict the energy consumption of a device at local level and ascertain its ability to be a CH. Each device broadcasts a device message (*Dev_Msg)* within a range of radius R_a. The broadcasted message comprises of device id and current energy (E_{cur}), and at the same time each device receives the same information from its neighbors. The computation of distance by each device to its neighbors is performed based on the received signal strength. From the result of the calculated distance to its neighbors, each device can predict its local energy consumption and thereafter, evaluate its local energy consumption ratio S_i. A device can predict whether it has the ability to be a CH or not depending on the computed local energy consumption ratio S_i . The mathematical derivation of S_i is illustrated in the proceeding paragraphs.

First, the energy consumed in the cluster topology involves the energy cost of the CH and its cluster members (CMs). For a CH, in one rotation, it broadcasts three messages: (1) Head Message (*H_Msg*) (2) Schedule Message (*S_Msg*) and (3) Route Message (*R_Msg*). In addition the CH receives *n* join messages (*nJ_Msgs*) and one route message (*R_Msg*). The derivation of energy expended in this process considers the work presented in [27], and is given by

$$\mathrm{E_{CH}} = 2\mathrm{L}(\mathrm{E_{elec}} + \epsilon_{fs}\, \mathrm{R_a^2}) + \mathrm{L}(\mathrm{E_{elec}} + \epsilon_{fs}\, (3\mathrm{R_a})^2) + \mathrm{nL}(\mathrm{E_{elec}} + 1)$$
$$= \mathrm{L}[\mathrm{E_{elec}}(\mathrm{n} + 4) + (\epsilon_{fs}\, \mathrm{R_a^2})] \quad (3)$$

Since in each round of the CH rotation, each device which is CM needs to send *n* join message (*nJ_Msg*) and receive control messages. The consumption of energy that occurs in these processes can be expressed as

$$\mathrm{E_{CM}} = \mathrm{n}\, (2\mathrm{LE_{elec}} + \mathrm{L}\, (\mathrm{E_{elec}} + \epsilon_{fs}\, \mathrm{d^2}_{\mathrm{to\ CH}}))$$
$$= \mathrm{nL}\, (3\mathrm{E_{elec}} + \epsilon_{fs}\, \mathrm{d^2}_{\mathrm{to\ CH}}) \quad (4)$$

Assuming $\mathrm{d_{to\ CH}}$ in equation (4) follows a uniform distribution over an interval $[0, \mathrm{R_a}]$, described as $U[0, R_a]$, then the expected value of $\mathrm{d^2}_{\mathrm{to\ CH}}$ is given as

$$\mathrm{E}[\mathrm{d^2}_{\mathrm{to\ CH}}] = \frac{\mathrm{R_a^2}}{2} \quad (5)$$

Substituting equation (5) into (4), gives

$$\mathrm{E_{CM}} = \mathrm{nL}\, (3\mathrm{E_{elec}} + \epsilon_{fs}\, \frac{\mathrm{R_a^2}}{2}) \quad (6)$$

Combining equation (3) and (6) gives the total energy expended during the rotation of the CH in one round ($\mathrm{E_{T(RCH)}}$), and is expressed as

$$\mathrm{E_{T(RCH)}} = \mathrm{E_{CH}} + \mathrm{E_{CM}}$$
$$= \mathrm{L}[\mathrm{E_{elec}}(\mathrm{n} + 4) + (\epsilon_{fs}\, \mathrm{R_a^2})] + \mathrm{nL}(3\mathrm{E_{elec}} + \epsilon_{fs}\, \frac{\mathrm{R_a^2}}{2})$$
$$= \mathrm{L}[\mathrm{E_{elec}}(4 + 4\mathrm{n}) + \epsilon_{fs}\, \mathrm{R_a^2}(11 + \frac{\mathrm{n}}{2})] \quad (7)$$

After the setup of a cluster, CMs collect data and transmit to the CH based on the Time Division Multiple Access (TDMA) schedule. The energy expended by all CMs (n) as they sense ($\mathrm{E_{sen}}$) and transmit data packets (L) to the CH in a cluster is expressed as

$$\mathrm{E_{CM-CH}} = \mathrm{nL}(\mathrm{E_{sen}} + \mathrm{E_{elec}} + \epsilon_{fs}\, \mathrm{d^2}_{\mathrm{to\ CH}}) \quad (8)$$

From (5), $\mathrm{E}[\mathrm{d^2}_{\mathrm{to\ CH}}] = \frac{\mathrm{R_a^2}}{2}$, substituting it in equation (6) results in

$$\mathrm{E_{CM-CH}} = \mathrm{nL}(\mathrm{E_{sen}} + \mathrm{E_{elec}} + \epsilon_{fs}\, \frac{\mathrm{R_a^2}}{2}) \quad (9)$$

Therefore, in each round, the energy expended by the CH ($\mathrm{E_{E-CH}}$) is given as

$$\mathrm{E_{E-CH}} = \mathrm{nE_{elec}L} + \mathrm{E_{sen}L} + (\mathrm{n} + 1)\mathrm{E_{agg}L} + 2\, (\mathrm{E_{elec}} + \epsilon_{fs}\, \mathrm{d^2}_{\mathrm{to\ NH}})\mathrm{L}$$
$$= [(\mathrm{n} + 2)\mathrm{E_{elec}} + \mathrm{E_{sen}} + (\mathrm{n} + 1)\mathrm{E_{agg}} + \epsilon_{fs}\, \mathrm{d^2}_{\mathrm{to\ NH}}]\mathrm{L} \quad (10)$$

Here, $\mathrm{d_{to\ NH}}$ represents the distance from the CH to its next hop selected device/CH, $\mathrm{E_{agg}}$ is the energy consumed during data aggregation from the CMs. In ensuring efficient connectivity of the CH to the next hope (NH) device in the cell, the maximum radius of communication is set at $\mathrm{R_{max}} = 3\mathrm{R_a}$.

Assuming $d_{to\ NH}$ follows a normal distribution over $[\mathrm{R_a}, 3\mathrm{R_a}]$ interval, then substituting this in equation (10) gives

$$\mathrm{E_{E-CH}} = [(\mathrm{n} + 2)\, \mathrm{E_{elec}} + \mathrm{E_{sen}} + (\mathrm{n} + 1)\mathrm{E_{agg}} + 9\, \epsilon_{fs}\, \mathrm{R_a^2}]\mathrm{L} \quad (11)$$

Therefore, in a cluster, the total energy consumed ($\mathrm{E_{CT}}$) is the sum of energies in equations (9) and (10), it is expressed as

$$\mathrm{E_{CT}} = \mathrm{E_{E-CH}} + \mathrm{E_{CM-CH}} \quad (12)$$

Expanding equation (12) and simplifying it, results in the expression

$$E_{CT} = \left[(\mathrm{n} + 2)E_{elec} + E_{sen} + (n + 1)E_{agg} + 9\, \epsilon_{mp}\, R_a^4\right]L + nL(E_{sen} + E_{elec} + \epsilon_{fs}\, \frac{R_a^2}{2})$$
$$= [(2 + 2\mathrm{n})\mathrm{E_{elec}} + (\mathrm{n} + 1)\left(\mathrm{E_{sen}} + \mathrm{E_{agg}}\right) + (\frac{\mathrm{n}}{2} + 9)\, \epsilon_{fs}\, \mathrm{R_a^2}]\mathrm{L} \quad (13)$$

The summation of equations (7) and (13) results in the total energy expended in a cluster ($\mathrm{E_{TEC}}$), which is expressed as

$$E_{TEC} = E_{CT} + E_{T(RCH)}$$
$$= \left[(2 + 2n)E_{elec} + (n + 1)\left(E_{sen} + E_{agg}\right) + \left(\frac{n}{2} + 9\right) \epsilon_{fs}\, R_a^2\right] L + L[E_{elec}(4 + 4n) + \epsilon_{fs}\, R_a^2(11 + \frac{n}{2})]$$
$$= [\{(6 + 6\mathrm{n})\mathrm{E_{elec}} + (\mathrm{n} + 1)\left(\mathrm{E_{sen}} + \mathrm{E_{agg}}\right) + (\mathrm{n} + 1)\, \epsilon_{fs}\, \mathrm{R_a^2}\,]\mathrm{L} \quad (14)$$

From equation (14), E_{TEC} in a cluster of devices $i = 1,2, \ldots, n$ is expressed as

$$E_{TEC} = \sum_{i=1}^{n} E_{con}(i) \quad (15)$$

The total current energy remaining E_{TCER} is expressed as

$$E_{TCER} = \sum_{i=1}^{n} E_{cur}(i) \quad (16)$$

The ratio (S_i) is the local energy consumption ratio, and is defined as the ratio of the total energy consumed in a cluster during sensing, broadcasting of messages, transmission of data to the CH, data aggregation and transmission to next hop to the current energy remaining. The defined ratio is expressed as

$$S_i = \frac{E_{TEC}}{E_{TCER}}$$

$$= \frac{\sum_{i=1}^{n} E_{con}(i)}{\sum_{i=1}^{n} E_{cur}(i)}$$

$$= \frac{\{(6+6n)E_{elec}+(n+1)(E_{sen}+E_{agg})+(n+21)\epsilon_{fs}R_a^2}{\sum_{i=1}^{n} E_{cur}(i)} \quad (17)$$

From equation (17), a device with lowest S_i within a defined maximum communication radius R_{max} is selected as a CH as explained in the next sub section.

2) Cluster head selection

After time T_1, CH selection process begins. During this phase, any device defined maximum communication radius R_{max} that fails to receive a *H_Msg* after the expiry of time t_i declares itself a CH by broadcasting a *H_Msg*. On receiving the *H_Msg*, the devices within defined maximum communication radius R_{max} send *J_Msg* to the nearest CH. The CH then prepares schedule messages and sends to other devices within R_a and enters the data transmission phase. The process is programmed to take T_2 prediction time. On expiry of T_2 prediction time, cluster phase formation phase starts.

3) Cluster formation phase

The devices choose the nearest CH by sending the *J_Msg* that contains the device ID and the device's energy status. On receiving the *J_Msg*, the CH prepares the list of TDMA schedules and broadcasts them to the devices in the cluster. The cluster phase formation takes prediction time T_3. Fig. 2 illustrates the flowchart for cluster formation in the FR.

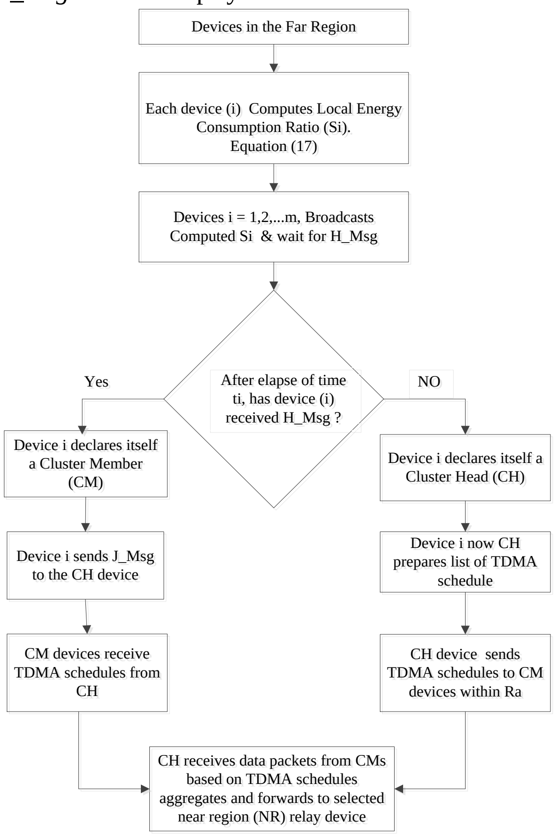

Figure 2. Flowchart for cluster formation

2. Centralised Relay Selection and Incentive Management System (CRSIMS)

This section discusses the factors upon which the relay device selection is based on, and the operation mechanism of the incentive management system. The two systems are termed as the Centralised Relay Selection and Incentive Management System (CRSIMS).

A. Relay Selection

After the formation of clusters, the cluster members send their data packets to the CHs. A suitable relay device is selected to forward the data packets to the BS. The selection of the relay device consids the residual energy and link quality [28] and the betweenness centrality [29] factors of the devices in the network. The factors considered are formulated as a cost functions (C_F), and is expressed as

$$C_F = \propto_1 E_R + \propto_2 L_Q + \propto_3 B_C. \quad (18)$$

where, $\propto_1, \propto_2, and\ \propto_3$ are weights assigned to the factors of the cost function with - $\propto_1 + \propto_2 + \propto_3 = 1$. E_R, L_Q, and B_C represents the residual energy, link quality and the betweenness centrality respectively. The device's E_R, level indicates its ability to relay the data packets to the required destination. The higher the

E_R value is the larger the amount of data packets that can be forwarded and at a longer distance. The device that has the highest residual energy contributes a small value to the cost function. The link quality L_Q, is evaluated basing on the packet delivery ratio (PDR). The PDR ratio is the ratio of the packets correctly received at the receiver to the total number of packets send from the sender. A higher value of PDR ratio implies a higher value of L_Q, which returns a low value to the cost function.

The betweenness centrality is defined as a measure of centrality in a graph that is based on the shortest path. B_C is is defined as the shortest distance linking pairs of devices in the network [29]. In terms of a local routing table, it is normally expressed as

$$B_{C(i)} = \sum_{i\neq j\neq r\in v} \frac{g_{jr\,(i)}}{g_{jr}} \tag{19}$$

where, g_{jr} represents the total number of shortest paths from j to r and $g_{jr\,(i)}$ is the number of shortest paths from j to r going through i. A higher $B_{C(i)}$ means more number of packets passes through that particular relay device to the destination, it's strategically located in the network. A higher B_C should returns a low value to the cost function.

From the above discussion, a device is selected as a relay when the computed C_F of that device is the lowest amongst other devices in the region. A low C_F implies that the device has a high residual energy, the quality of the link is good, and the distance linking the devices is short. The C_F computation is performed in CRSIMS. In the next section, the operation of CRSIMS is presented.

B. CRSIMS Operation

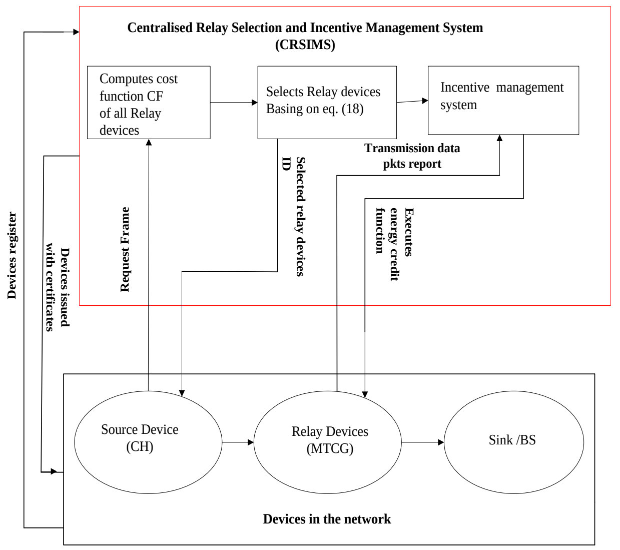

Fig. 3. Architectural layout of CRSIMS and how it links with devices in the network.

The work proposed in [30] is considered in the architectural arrangement of CRSIMS. In the arrangement, CRSIMS is linked to the devices through the internet or Wi-Fi hotspots. It is assumed that devices are equipped with network interfaces which enables them to send and receive messages in a wireless network [31]. Each device, once connected to the internet gets registered by CRSIMS which issues certificates which are unique for each device. Fig. 3 shows the architectural layout of CRSIMS.

In this arrangement, the source device which is the CH in the FR, adds on the energy credits after aggregating the data packets (L) of its cluster members, it also adds on the energy credits (E_C) and all are put under request frame. The CH, then sends a request frame to CRSIMS requesting for relay device R_D and waits for feedback. Through the application of equation (18), the cost functions of all the devices in the NR are computed. The relay device with the lowest cost function is identified, and the identity (ID) of the selected R_D is forwarded to the source device (CH). The CH forwards the data packets to the selected relay device, which transmits to the sink/BS. At the same time, R_D sends a report to CRISMS. The report is an indication that the forwarding of data packets has been implemented. The report is meant to stimulate CRSIMS to execute the E_C function depending on the required energy credit value that is given to R_D as a reward.

3. Flow Diagram of Proposed Approach - LECIA

The flow diagram of the proposed LECIA algorithm is shown in Fig. 4.

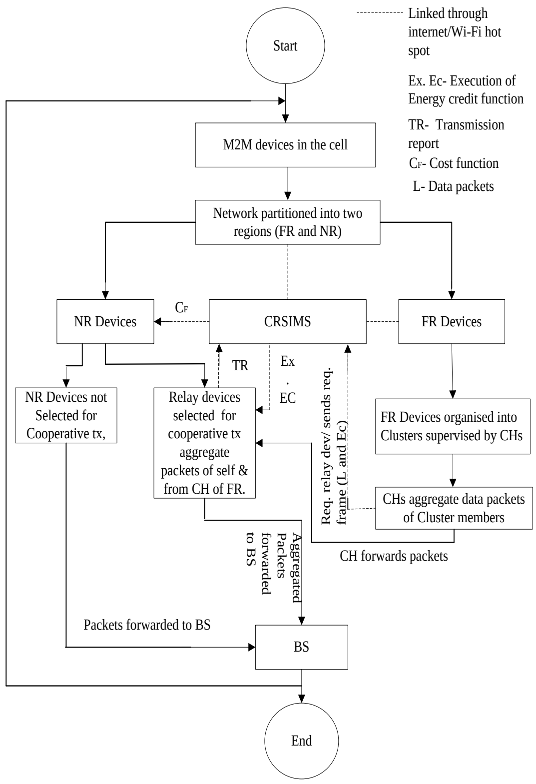

Fig. 4. Flow chart of the proposed LECIA algorithm

The flow diagram in Fig. 4 shows Network partitioning, cluster formation, relay selection, and incentive management as steps that define LECIA algorithm. The following section presents the pseudocodes that describe the steps of LECIA approach.

1) Pseudo code of network partitioning

The devices are randomly distributed in the network area, based on probing signal of the devices, partitioning of the network is performed as described by the following pseudo code.

Input: Devices in the network area.
Output: Network partitioned into regions (FR, NR)
Initialisation: Devices $\boldsymbol{D_{n,}}\ n = \{1,2,\dots,p,\dots D\}$, R-Region, S_D -Source device, $ACKD_D$ -Acknowledgement of Destination Device, P_S – Probing signal.

1. Region (R) = 0
2. **begin** (network partitioning)
3. While (true)
4. S_D (n) Sends P_S
5. **If** S_D (i) receives $ACKD_D$ then
6. R = R+1
7. break, // successful
8. **else**
9. **If** S_D (i) receives $ACKD_p$ then
10. R = R+1
11. Dn = D_p
12. **end**
13. **end**
14. **end**

After network partitioning, the devices in the far region are organised into clusters. The process of cluster formation is based on local energy parameters. The pseudo code for cluster formation is described in (ii).

2) Pseudo code of cluster formation

1. current_state ← *1*
2. **begin** (cluster formation algorithm)
3. **switch** (current_state)
4. **case 1:** *state* ← *Candidate*
5. Broadcast the *Dev_Msg*
6. current_state ← *2*
7. ***break;***
8. **case 2:** **if** (*T*1 has not expired) **do**
9. Receive the *Dev_Msg*
10. Update neighbourhood table NT[ ]
11. *ti* ← broadcast waiting time
12. current_state ← *2*
13. **else**
14. current_state ← *3*
15. **end**
16. ***break*;**
17. **case 3:** **if** (*T*2 has not expired) **do**
18. **if** *CurrentTime* < *ti* **do**
19. **if** receive a *H_Msg* from the neighbour NT[*i*] **do**
20. *state* ← *P lain*
21. NT[*i*].*state* ← *Head*
22. **else**
23. Continue
24. **end**
25. **else if** *state* = *Candidate* **do**
26. *state* ← *Head*
27. *Ra* ← competing radius
28. Broadcast the *H_Msg*
29. **end**
30. **end**
31. current_state ← *3*
32. **else**
33. current_state ← *4*
34. **end**
35. **break;**
36. **case 4:** **if** (*T*3 has not expired) **do**
37. **if** *state* = *P lain* && has not sent the *J_Msg* **do**
38.
39. Send the *J_Msg* to the nearest CH
40. **else if** *state* = *Head* **do**
41. Receive *J_Msgs*
42. **end**
43. current_state ← *4*
44. **else**
45. current_state ← *null*
46. **end**
47. **break;**
48. **case default:**
49. **break;**
50. **end**
51. **end**

After the devices are organised into clusters which is supervised by the a device selected as the CH, members of the cluster forward their data packets to the CH. The received data packets are to be forwarded to the BS through a suitable relay device that is selected based on the cost function. The selected relay device, receive energy credits on completion of the data forwarding assignment. The pseudocode for relay selection and execution of energy credit function is presented in (iii) and (iv) respectively. The sequence of events starts by the source device which in this case is the CH making a request for relay device, this is followed by the computations at CRISMS for relay selection and concludes with data packet forwarding by the selected relay device and receiving of incentive.

3) Pseudo code for CH request for relay device

1. Initialise E_C - Energy Credits, L- Data payloads, Relay Devices (R_D) in the NR, D = 1,...p
2. Aggregate Ec and L into request frame
3. CH sends request frame (R) to CRISMS
4. Wait
5. **If** response is received from CRISMS **then**
6. Extract R_D ID from response
7. Forward data packets to R_D
8. **else**
9. Go to step (5)
10. **end**
11. Go to step (1)

4) Pseudo code for CRISMS and data packet forwarding

1. **If** request frame (R) is received from CH **then**
2. Extract E_C value
3. Extract L
4. Initialise C← *max value of cost function*
5. **For** D=1: p
6. **If** C_F (R_D) < C
7. C← C_F (R_D)
8. ID← ID of R_D
9. **else**
10. Continue
11. **end**
12. **end**
13. Send ID of R_D to CH
14. **If** REPORT received from R_D **then**
15. Execute E_C function based on the E_C value
16. **else**
17. Go To step (14)
18. **end**
19. Go To step (1)

Relay Device ($\boldsymbol{R_D}$) Data packet forwarding

20. **If** received data packets from CH **then**
21. Forward data to destination
22. Send REPORT to CRISMS
23. GET Reward from CRISMS
24. Go to step (20)
25. **else**
26. Go to step (20)
27. **end**

The preceding section has described the operation of the proposed algorithm- LECIA, the next section presents the performance evaluation of the algorithm.

V. Performance Evaluation of Proposed LECIA Algorithm

This section presents the simulation setup that was conducted to evaluate the performance of the proposed LECIA technique. The section also provides a discussion of the results obtained. As a bench mark, LECIA is compared with a closely related algorithms, the HEED algorithm [9] [12] and the LEACH algorithm [9] [13]

A. Simulation Setup

M2M devices in a single cell are considered in two regions as they transmit data packets to the BS. The partitioning is based on the range coverage of the probing signal of the devices. Based on the probing signal and starting with the devices located furthest from the BS, after two hops, it is assumed that the destination has been reached. The network is partitioned in a ratio of 50:50 (FR: NR). The rest of simulation parameters are presented in Table I.

TABLE I: SIMULATION PARAMETERS

Parameter	**Value**	**Description**
D	200	Number of devices
BS	(1000, 1000)	Location of BS Metres
BS_{RC}	14, 000 M	Radius of BS coverage
L	0 - 1000	Data payload- bits (varying)
$\boldsymbol{\varepsilon_{fs}}$	1*10^-12 J/bit/m^2	Amplification energy of free space.
$\boldsymbol{\varepsilon_{mp}}$	1.3*10^-15J/bit/m^4	Amplification energy of multipath
E/b.	50*10^-9J	Energy/bit
$E_{agg.}$	5*10^-9J	Energy of data aggregation
R_{Sim}	3000	Rounds of simulations
$E_{D(in)}$	1 – 10 J	Initial device energy
E_{inc}	0 – 10 J	Incentive energy

B. Simulation Results

This section presents a sample of the simulated results. The energy efficiency is described in terms of; (1) The number of devices that survive in the network versus the number of rounds of packet transmissions; (2) The remaining total energy in the network versus the number of rounds of packet transmissions; and (3) The number of data packets delivered to the BS in the network.

1) Number of Surviving devices versus rounds of packet transmissions

The display in Fig. 5 indicates that, as the number of packet transmissions increases, the number of devices remaining in the network decreases rapidly in the LEACH approach but is shown to be slower in the HEED approach and in the proposed LECIA. The mathematical analysis carried out by averaging the number of the remaining devices over the simulation period indicates that, the proposed scheme has on average 5% more number of surviving devices than its closely related traditional algorithm the HEED and 55% more number of surviving devices than the LEACH algorithm. This is attributed to the reduced energy consumption that is achieved through a more energy efficient cluster formation process and multi hop transmissions.

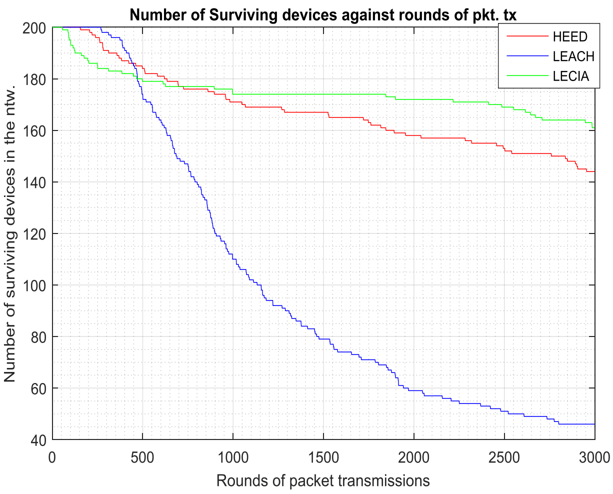

Fig. 5. Number of Surviving Nodes versus Number of Packet Transmissions

In the formation of clusters, CHs were selected based on local energy parameters. By invoking CRSIMS, a suitable relay device is selected and incentivized with energy credits. The devices in the proposed algorithm embrace multi-hop transmissions and therefore, consume less energy during data packet transmissions. Many devices survive when compared to traditional algorithms such as the HEED and LEACH. The LEACH algorithm adopts the clustering approach but due to the lack of multi hop transmission, devices located far from BS consume more energy and end up failing. The HEED algorithm utilises clustering and multi hop transmissions. However, due to the lack of energy incentives such as that offered in the proposed LECIA algorithm, devices that are near the BS exhausted their energy early.

2) Trend of total energy remaining versus number of rounds of packet transmissions

Fig. 6 shows the trend of remaining energy in the network versus the rounds of packet transmissions. It is shown that the proposed LECIA scheme has 6% more energy compared to HEED and 37% more energy than LEACH. The introduction of Clustering based on local energy availability parameters in the FR and the engagement of CRSIMS enhances multi-hop transmissions. Devices transmit data through short distances and consume less energy. Long distance transmission in LEACH works against its energy conservation principle. The devices closer to the BS in HEED use more energy as they relay data packets and since there is no way of energy replenishment, more become eliminated from the network.

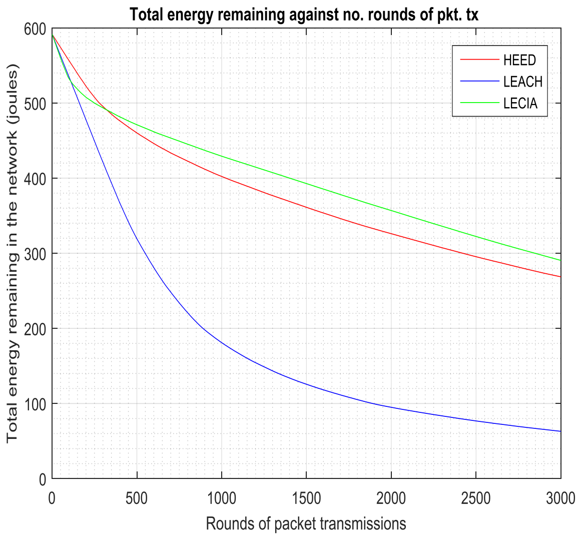

Fig. 6. System energy remaining versus number of packet transmissions

3) Data packets delivered to the BS

Fig. 7 displays the trend of the number of packets, delivered to the BS versus the number of rounds of packet transmissions. Until upto about 1500th round of packet transmissions, HEED and LECIA quantity of packets delivered to the BS is almost the same. However, above 1500th round of packet transmissions, The proposed LECIA algorithm delivers more data packets to the BS than HEED. The engagement of the energy incentive credits to relay devices present in LECIA supports their continuous operation in the M2M communications. In the LECIA approach, energy efficient routes are selected, and assist in delivering a higher number of data packets to the BS compared to the HEED algorithm. In the HEED approach, lack of rewards makes potential relay devices to get exhausted and die. The death of potential relay devices lowers the number of data packets delivered to the BS with longer routes being used for the far located devices. In the LEACH approach, there is no multi-hop transmissions with devices located furthest transmitting data packets over long distances and less number compared to LECIA and HEED reach the BS.

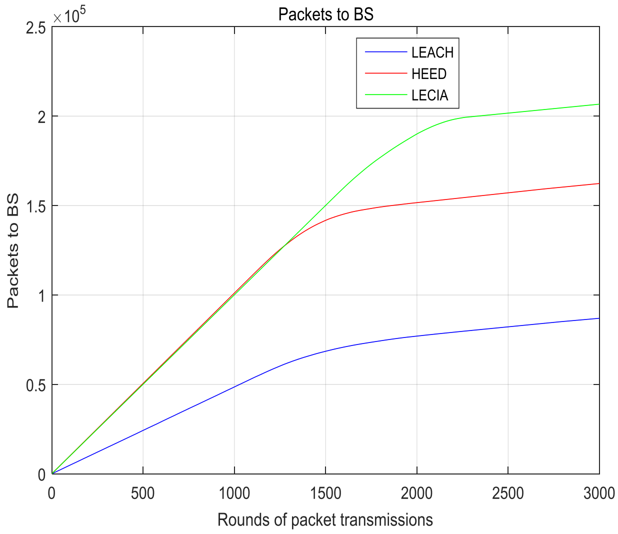

Fig. 7. Number of packets, delivered to the bs versus the number rounds of packet transmissions

VI. Conclusion and Future Work

In this paper a Local Energy-based Clustering with an Incentive Algorithm (LECIA) was proposed for M2M communications. The algorithm utilises local energy parameters with the partition of the network into regions being based on the range at which the probing signal of device(s) can be received. In the formation of clusters in the far region of the partitioned zone, the selection of the CHs is based on local energy consumption ratios. A device of low energy consumption ratio within a given range is selected as the CH. The invoking of a cost function technique culminated with the selection of the most suitable relay device that is rewarded with energy credit after assisting to forward the data packets to the BS. The proposed algorithm has shown to have better performance with 5% and 55% more number of remaining devices, and 6 % and 37% more residual energy in the network when compared to the traditional closely related algorithms such as HEED and LEACH algorithms respectively. This shows that the proposed algorithm enables more devices in the network to utilise less energy during the data transmission process and operate for a longer period. Based on this, it can be concluded that the proposed LECIA scheme is an energy efficient design for M2M communications.

Since this study assumed that the devices were stationary after they were randomly distributed in the network, future work could consider the dynamic nature of M2M devices during analysis of energy efficiency. A further challenge that needs further consideration is the effect of interference. Since one cell was only considered, interference was excluded, for simulation to reflect practical and realistic situations; it will be worthy to be considered for future work.

REFERENCES

[1] B. N. Gopalsamy, "Communication trends in internet of things," in *Developments and Trends in Intelligent Technologies and Smart Systems*, IGI Global, 2018, pp. 284–305.

[2] M. Muntjir, M. Rahul, and H. A. Alhumyani, "An analysis of Internet of Things (IoT): Novel architectures, modern applications, security aspects and future scope with latest case studies," *Int. J. Eng. Res. Technol*, vol. 6, pp. 422–447, 2017.

[3] K. Zheng, F. Hu, and W. Wang, "Radio resource allocation in LTE-Advanced cellular networks with M2M communications," vol. 50, no. 7, pp. 184–192, 2012.

[4] M. Elshrkawey, S. M. Elsherif, and M. E. Wahed, "An enhancement approach for reducing the energy consumption in wireless sensor networks," *J. King Saud Univ. Inf. Sci.*, vol. 30, no. 2, pp. 259–267, 2018.

[5] A. Zeb, *et al.*, "Clustering analysis in wireless sensor networks: The ambit of performance metrics and schemes taxonomy," *Int. J. Distrib. Sens. Networks*, vol. 12, no. 7, p. 4979142, 2016.

[6] S. Wang, J. Yu, M. Atiquzzaman, H. Chen, and L. Ni, "CRPD: A novel clustering routing protocol for dynamic wireless sensor networks," *Pers. Ubiquitous Comput*., vol. 22, no. 3, pp. 545–559, 2018.

[7] C. Kalita and P. K. G. Thakurta, "Energy efficient routing to improve lifetime in MANET: A clustering approach," *J. Commun*., vol. 13, pp. 679–684, 2018.

[8] P. Patil, U. Kulkarni, and N. H. Ayachit, "Some issues in clustering algorithms for wireless sensor networks," in *Proc. Special Issues on 2nd National Conference-Computing, Communication and Sensor Network*, 2011.

[9] J. W. Raymond, T. O. Olwal, and A. M. Kurien, "Cooperative communications in Machine to Machine (M2M): Solutions, challenges and future work," *IEEE Access*, 2018.

[10] E. E. Tsiropoulou, G. Mitsis, and S. Papavassiliou, "Interest-aware energy collection & resource management in machine to machine communications," *Ad Hoc Networks*, vol. 68, pp. 48–57, 2018.

[11] S. A. El-Feshawy, W. Saad, M. Shokair, and M. I. Dessouky, "An efficient clustering design for cellular based machine-to-machine communications," in *Proc. 35th National Radio Science Conference (NRSC)*, 2018, pp. 177–186.

[12] O. Younis and S. Fahmy, "HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Trans. Mob. Comput*., vol. 3, no. 4, pp. 366–379, 2004.

[13] I. Journal, W. C. Vol, A. Gupta, S. Malik, M. Goyal, and P. Gupta, "Clustering approach for enhancing network energy using LEACH protocol in WSN," vol. 2, no. 1, 2012.

[14] S. Kassan, P. Lorenz, and J. Gaber, "Low energy and location based clustering protocol for Wireless Sensor Network," in *Proc. IEEE International Conference on Communications*, 2018, pp. 1–6.

[15] Z. Liu, Q. Zheng, L. Xue, and X. Guan, "A distributed energy-efficient clustering algorithm with improved coverage in wireless sensor networks," *Futur. Gener. Comput. Syst.*, vol. 28, no. 5, pp. 780–790, 2012.

[16] M. M. Zanjireh and H. Larijani, "A survey on centralised and distributed clustering routing algorithms for WSNs," in *Proc. IEEE 81st Vehicular Technology Conference*, 2015, pp. 1–6.

[17] N. Nisan, *et al.*, "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks," *Wirel. Pers. Commun*., vol. 1, no. 4, pp. 1–5, 2014.

[18] D. Singh and C. K. Panda, "Performance analysis of modified stable election protocol in heterogeneous wsn," in *Proc. International Conference on Electrical, Electronics, Signals, Communication and Optimization*, 2015, pp. 1–5.

[19] M. Ye, C. Li, G. Chen, and J. Wu, "EECS: An energy efficient clustering scheme in wireless sensor networks," in *Proc. 24th IEEE International Performance, Computing, and Communications Conference*, 2005, pp. 535–540.

[20] O. Younis, M. Krunz, and S. Ramasubramanian, "Location-unaware coverage in wireless sensor networks," *Ad Hoc Networks*, vol. 6, no. 7, pp. 1078–1097, 2008.

[21] I. A. Lekhal and S. Chouraqui, "Routing optimization using an enhanced protocol for wireless sensor networks with chain structure," *J. Commun.*, vol. 13, pp. 691–701, 2018.

[22] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, "A centralized energy-efficient routing protocol for wireless sensor networks," *IEEE Commun. Mag*., vol. 43, no. 3, pp. S8-13, 2005.

[23] C. H. Lung and C. Zhou, "Using hierarchical agglomerative clustering in wireless sensor networks: An energy-efficient and flexible approach," *Ad Hoc Networks*, vol. 8, no. 3, pp. 328–344, 2010.

[24] D. Tudose, L. Gheorghe, and N. Tapus, "Radio transceiver consumption modeling for multi-hop wireless sensor networks," *UPB Sci. Bull. Ser. C*, vol. 75, no. 1, pp. 17–26, 2013.

[25] R. I. Ansari, S. A. Hassan, and C. Chrysostomou, "Energy efficient relay selection in multi-hop D2D networks," in *Proc. International Wireless Communications and Mobile Computing Conference*, 2016, pp. 620–625.

[26] U. Singh, S. Solanki, D. S. Gurjar, P. K. Upadhyay, and D. B. da Costa, "Wireless power transfer in two-way AF relaying with maximal-ratio combining under nakagami-m fading," in *Proc. 14th International Wireless Communications & Mobile Computing Conference*, 2018, pp. 169–173.

[27] J. Yu, L. Feng, L. Jia, X. Gu, and D. Yu, "A local energy consumption prediction-based clustering protocol for

wireless sensor networks," *Sensors*, vol. 14, no. 12, pp. 23017–23040, 2014.

[28] J. Katiravan, D. Sylvia, and D. S. Rao, "Energy efficient link aware routing with power control in wireless ad hoc networks," *Sci. World J.*, vol. 2015, 2015.

[29] Z. Liu, C. Jiang, J. Wang, and H. Yu, "The node importance in actual complex networks based on a multi-attribute ranking method," *Knowledge-Based Syst*., vol. 84, pp. 56–66, 2015.

[30] R. Kaushik and J. Singhai, "Enhanced node cooperation technique for outwitting selfish nodes in an ad hoc network," *IET Networks*, vol. 4, no. 2, pp. 148–157, 2014.

[31] M. Stemm and R. H. Katz, "Vertical handoffs in wireless overlay networks," *Mob. Networks Appl*., vol. 3, no. 4, pp. 335–350, 1998.

Raymond Juma received Master of. Technology Degree in Electrical Engineering from Tshwane University of Technology (TUT) Pretoria South Africa in 2012. He is currently a Doctor of Engineering student at TUT. His research interests include energy efficient resource management in wireless networks and mobile network optimisation.

Anish M. Kurien (IEEE**,** M'04) received the D.Tech. Degree in electrical engineering and the Ph.D. Degree in computing (through co-tutelle) from Tshwane University of Technology (TUT), Pretoria, South Africa, and the University Paris-Est, Champs-sur-Marne, France, in 2012. He is an Associate Professor and he is currently the Node Director for the French South African Institute of Technology, TUT and is responsible for post- graduate programs in the Department of Electrical Engineering, TUT. He is currently involved in research projects in wireless communications, radio resource management, and mobile network optimization. He is also is involved in several industrial projects related to wireless networks and technology development. His research interests include feature extraction and pattern recognition algorithms applied to mobile network subscriber classification.

Thomas O. Olwal (IEEE, S'16) received the Ph.D. Degree in computer science from the University of Paris-EST, Champs-sur-Marne, France, in 2010, and the D.Tech. Degree in electrical engineering from Tshwane University of Technology (TUT) (through the co-tutelle Programme), Pretoria, South Africa, in 2011. He is currently lecturing at the TUT as an Associate Professor and previously worked as a Senior Researcher with Wireless Computing and Networking R&D, Council for Scientific and Industrial Research (CSIR) of South Africa. His research interests include analysis and design of the spectrum, energy-efficient radio resource management, Internet of Things, autonomous robotics, and automation and their emerging applications. He serves as a reviewer in several ACM/IEEE conferences and journals