

Energy-efficient and Distributed Data-aware Clustering Protocol for the Internet-of-Things

Chinmaya Mahapatra*, Zhengguo Sheng[†] and Victor C.M. Leung*

*Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada

[†]Department of Engineering and Design, University of Sussex, UK

Email:{chinmaya, vleung}@ece.ubc.ca, Z.Sheng@sussex.ac.uk

Abstract—Internet-of-Things (IoT) is a new paradigm of wireless technology, where smart sensors and machines communicate through amalgamating multivariied protocols and devices such as Zigbee, Bluetooth, Radio Frequency Identification (RFID) and Wireless Sensor Network (WSN). IoT systems, requires sensing, gathering, storing, processing and transmitting of data from real time sensors as well as virtual online sensors. Thus, energy efficiency and robust data delivery to users are the key requirements. In this paper, we introduce active RFID tags based cluster head (CH) selection, data-awareness and energy harvesting in the IoT systems. The results show that the IoT based WSN heterogeneous systems are better equipped to deal with energy efficiency and data delivery problems. Simulation results support our proposed method and show significant improvement over state-of-the art techniques.

I. INTRODUCTION

The Internet-of-Things (IoT) started in 2009 with a vision of connecting devices to devices and persons to devices. Technologies like Wireless sensor networks (WSNs) form the backbone of such interactions. The industrial sector estimates that by 2020 more than half billion devices will be connected with each other [1] [2]. When virtually every device is connected with each other and all connectivity is replaced by intelligent machines and automation, the system will be enormous and complex spanning across a varied range of protocols and standards. Since portable and battery operated systems like smartphones, tablets, and cameras will always be connected to the IoT, enormous amounts of user data will be generated and their energy consumption will dramatically increase.

IoT plays an important role by bringing together people, process, data, and things to make networked connections more relevant and valuable. Its technologies, including heterogeneous WSNs, are used to monitor many aspects of an ecosystem ranging from a small office space to a city, in real time. Routing is one of the critical technologies in IoT as opposed to traditional ad-hoc WSNs. It is more challenging due to constrained resources in terms of energy supply, processing capability, frequent topology changes and reliable data delivery within a limited time period. Based on network structure, routing protocols can be sub-divided into two categories, flat routing and hierarchical routing. In a flat topology, all nodes perform the same tasks and have the same functionalities in the network. Whereas, in a hierarchical topology, nodes perform different tasks and are typically organized into lots of clusters according to specific metrics. In clustering, members of the

clusters elect a cluster head (CH) [3]. All nodes belonging to the same cluster send their data to CH, where, CH aggregates data and sends aggregated data to base station (BS).

Clustering algorithms in the literature are divided based on their energy efficiency in two types of networks i.e., homogeneous and heterogeneous WSNs. Homogeneous WSNs considers that the all sensor nodes in the system have the same energy level and all the nodes takes turn according to a given probability to become CH. Low-Energy Adaptive Clustering Hierarchy (LEACH) [4], Power Efficient Gathering in Sensor Information Systems (PEGASIS) [5] and Hybrid Energy-Efficient Distributed Clustering (HEED) [6] are examples of cluster based protocols which are designed for homogenous WSNs. However, these techniques perform poorly in heterogeneous WSNs scenario as nodes having less energy expire faster than higher energy nodes.

Heterogeneous WSN topology takes into account that the nodes have different initial energy. Thus they perform better than homogeneous WSNs in a real application scenario with variety of sensors such as warehouses, home monitoring and surveillance. Distributed Energy Efficient Clustering (DEEC) [7], Developed DEEC (DDEEC) and Enhanced DEEC (EDEEC) [8] are some of the heterogenous WSN protocols. These distributed clustering algorithms for heterogeneous WSNs have similar topological structure to an IoT system. Although multi-hop routing and residual energy for selecting CHs are considered, they neither incorporate the intricacies nor the benefit of a diversified and event driven IoT system. The main pitfalls of the algorithms delineated w.r.t energy consumption and network lifetime are, energy consumed in cluster head (CH) selection at each round, assuming nodes in always ON state [9], and limited battery capacity of energy constrained sensors [10]. Thus it is required to come up with a protocol specifically for IoT systems.

IoT and heterogeneous WSNs systems are similar in being equipped with sensors, base station (data gathering and decision making node) and wireless transceivers. But IoT system is more diversified in involving some notable variations like interaction between multiple protocols, sensing systems having varied energy values, asynchronous event driven processing and gateway node in between sensors and BS to route data more efficiently. Moreover, due to the evolution of active RFID tags [11] with reading capability in the range of meters and various energy harvesting mechanisms [12] [13], prudent techniques in IoT systems using them are better equipped to handle the energy efficiency and network lifetime problem. Hence we propose a distributed data-aware energy-efficient clustering protocol for IoT (DAEECI) which includes data awareness, RFID based CH selection and RF energy harvesting. Our main

This work is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), Institute of Computing, Information, and Cognitive Systems (ICICS), Telus Inc. and other industry partners.

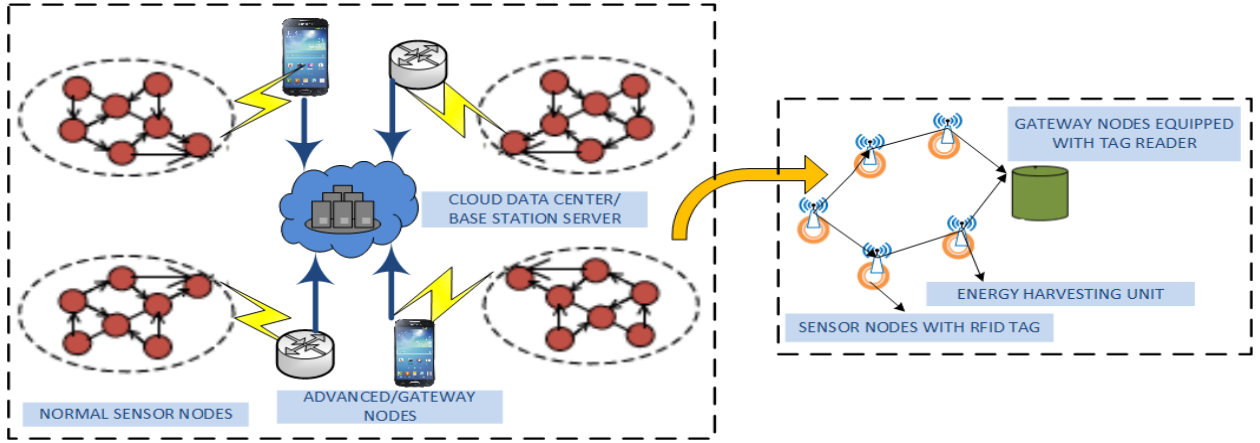


Figure 1. Data-aware RFID tag based IoT Architecture.

contributions are as follows:

- (1) The system is distributed in two-levels based on their initial energy as normal nodes with standard battery energy and advanced nodes with a times more energy than normal nodes [8]. We use the RFID tagging and reading mechanism to reduce the energy consumption during the cluster head (CH) selection phase till all the advanced nodes (also called gateway nodes) have their energy exhausted. Thereby prolonging lifetime of the network.
- (2) We propose data awareness by dividing the sensor based on urgent and regular data demand and switch nodes between high/low power state based on data requirement at the user side. The solution expects to save energy in the nodes and improving battery life.
- (3) We additionally incorporate RF energy harvesting for normal nodes to further improve network lifetime.

The rest of the paper is organized as follows. In Section II, we describe IoT network model. Section III describes our DAEECI protocol. Section IV analyzes our simulation results. Section V provides concluding remarks.

II. IOT NETWORK MODEL: CLUSTER HEAD SELECTION AND ENERGY COST FORMULATION

The IoT system taken here is depicted in Fig. 1. The network is a random distribution of N_{tot} sensor nodes in a square area of side X meters and gateway nodes K used for data aggregation and routing to BS. The nodes are differentiated based on their initial energy as advanced and normal sensor nodes. Advanced nodes have a times more energy than normal nodes and are also known as gateway nodes (K), as they route data to the base station. Thus the total nodes in the system are $(N_{tot} + K)$. Each cluster has one gateway node based on minimum distance, as their CH. Let E_B be the initial battery energy of the normal node. Let K be the number of distributed clusters that service all nodes and have one gateway node per cluster for data routing to base station. All the cluster heads send the aggregated data from sensor nodes they service to BS server. The BS server is user driven based on data request from different user generated applications. Let E_{tot-cl_s} is the total energy of the clusters given as

$$E_{tot-cl_s} = N_{tot} * E_B + K * E_B * (1 + a) \quad (1)$$

A. Active RFID tag based Cluster Head allocation

The cluster head selection is one of the major drawbacks of current clustering algorithms. In LEACH [4] and DEEC [7] algorithm, cluster head selection is divided into rounds, where each node randomly decides whether to become a cluster head based on a threshold $T_i(s)$ computed by apriori decided probability p_i .

$$T(s_i) = \begin{cases} \frac{p_i}{1 - p_i \cdot (r \bmod (\frac{1}{p_i}))} & , \text{if } S_i \in G \\ 0 & \text{Else} \end{cases} \quad (2)$$

where, r is the current round number, and G is the set of nodes that have not been cluster-heads in the last n_i rounds ($p_i = \frac{1}{n_i}$). Let the energy dissipated in a round (E_{round}) is adopted from the radio model in [7] as

$$E_{round} = L \left(\begin{array}{l} (N_{tot} + K) (E_{rx} + E_{tx}) + N_{tot} E_{DA} \\ + K \epsilon_{amp} d_{toBS}^4 + N_{tot} \epsilon_{fs} d_{toCH}^2 \end{array} \right) \quad (3)$$

where, E_{DA} is the data aggregation cost expended in CH, d_{toBS} is the average distance between CH to BS, d_{toCH} is the average distance between cluster members to CH, L is the number of bits to be transmitted, ϵ_{amp} is the energy consumption of transmitter amplifier circuit, E_{tx} is the transmitted energy consumed per bit and E_{rx} is the received energy per bit, ϵ_{fs} is the free space parameter. From (3), it can be inferred that $L * (N_{tot} + K) * (E_{tx} + E_{rx})$ and $L * N_{tot} * \epsilon_{fs} * d_{toBS}^2$ are the energy consumed for CH selection and routing data from nodes to CH, respectively.

Therefore, to save the energy consumed in CH selection, we propose to incorporate active RFID tags coupled to member nodes and a tag reader at the gateway node. The conceptual topology is depicted in the expanded view of the WSN in Fig. 1. RFID is an emerging automatic identification technology in which information is carried by radio waves. RFIDs are classified as passive, semi-passive, or active [11]. Passive RFID tags function without a battery, has almost infinite lifetime but can operate in the range of only couple of *centimeters*. Whereas, an active RFID [14] can be read at distances of 100 m or more, greatly improving the utility of the device, but it is battery powered and has shorter life. The use of active tags with sensor nodes and a tag reader at the gateway will eliminate the need of choosing the CH till

the gateway nodes are exhausted of their energy. Nodes collect data from the environment and send them to the RFID reader which in turn sends it to the BS. From the BS data are sent to the cloud in order to provide it to the user through the services initiated by the cloud. With the evolution of tags like CC2650 SensorTag¹ which operate with 2.4GHz transmission and supporting technologies such as Bluetooth, ZigBee and IPv6, it is feasible to incorporate the model for IoT WSN systems. Using our proposed method, as the tag reader reads the sensed data from the tags, computation for routing data to the CH is not required. The energy consumed for CH selection becomes $L * (N_{tot} * E_{rx} + K * (E_{tx} + E_{rx}))$. This happens till all the gateway nodes die in which case the routing follows energy consumption in (3) again.

B. Data aware processing

Sensors in IoT systems are not always active. There are two types of data request from users, one is periodic monitoring type of application such as warehouses and industrial control and another is on demand processing such as home surveillance, temperature control, smoke and water detectors. Thus data awareness of sensors is critical to its longevity. Sensors that service users periodically have to be in *active* state all the time whereas the sensors sending data sporadically can be kept in *sleeping* state for most of the time. They can be woken up from sleep by asynchronous triggering on their pins when a certain threshold is crossed. An efficient approach to address this is *duty – cycling*, in which the receiver on-demand switches between active and sleeping states. Among the different categories of *duty – cycling*, namely *synchronous*, *pseudo – asynchronous* and pure *asynchronous*, latter provides the most efficient solution in terms of energy consumption [13]. In the *asynchronous* approach, the sensor device is in deep sleep mode and only wakes up when signalled by the BS or its neighbouring devices through an interrupt command generated by a low-power wake-up radio (WUR). Let the transmitted energy consumption of sleeping nodes is only α percent of E_{tx} , where $\zeta \leq \alpha \leq 1$, ζ is a small number close to 0. Let there be n_s number of sleeping nodes in the system. Therefore, the E_{round} is as follows

$$E_{round} = L \left(\begin{array}{l} (N_{tot} + K) (E_{rx} + \alpha E_{tx}) + N_{tot} E_{DA} \\ + K \varepsilon_{amp} d_{toBS}^4 + N_{tot} \varepsilon_{fs} d_{toCH}^2 \end{array} \right) \quad (4)$$

When $\alpha=1$, all the nodes are awake and transmitting data read by the tag reader. But when the data demand is low, the α value is small providing tremendous energy saving in the system.

C. RF Energy Harvesting

Energy harvesting is a promising remedy to cope with the energy challenge. A wireless node can harvest energy from different forms of environmental sources such as thermal, wind, solar, vibration [13]. Among these resources, wireless/RF energy harvesting is an attractive candidate and provides key advantages in virtue of being controllable, lower cost and smaller form factor implementation [13]. The recent technology trend in energy harvesting provides a fundamental method to prolong battery longevity of sensor devices [12]. In RF energy harvesting (EH) circuit, the antenna receives the transmitted

radio waves and converts the received RF energy into a stable direct current (DC) energy source to supply the sensor device. Energy harvesting depends on the distance from the harvesting source. If the EH circuit is deployed on the sensor devices with a power management unit, it can harvest RF energy from the transmitted electromagnetic waves of the transmitter circuit of its own as well as nearby nodes, gateway nodes and BS [12] [13]. In practice, the conversion from the received RF power to the usable DC supply comes with a certain amount of power loss in the matching circuit and in the internal circuitry of the power converter. The power conversion efficiency (η) of the converter is the ratio of the generated usable DC output power to the input RF power. State-of-the-art RF-to-DC converters (also known as rectifiers) can achieve high η values, up to 70% or more [12]. η is an indication of the amount of harvested energy that is available for the sensor device. Here, we assume that the energy harvested by the nodes vary randomly between $0 < \beta \leq 1$ of total harvested energy $E_H(t)$. $E_H(t)$ is the maximum harvested energy and is taken as η times the battery energy per unit time t .

III. DATA AWARE ENERGY EFFICIENT DISTRIBUTED CLUSTERING PROTOCOL FOR IOT (DAEECI)

In this section, we present the detail of our DAEECI protocol. DAEECI uses similar function of initial (E_{tot-cl}) and residual energy ($E_i(r)$) level as in [7] of the nodes to select the cluster-heads at each round. To avoid that each node needs to know the global knowledge of the networks, it estimates the ideal value of the network life-time to compute the reference energy ($\bar{E}(r)$) consumed by a node in a round. Our DAEECI divides the problem into different user cases based on data awareness (i.e. either α is 1 for periodic data sensing or $0 \leq \alpha < 1$ for sparse data sensing) and percentage of gateway nodes present in the IoT system (K is high or low). The normal nodes are assumed to have their dedicated RF energy harvesting circuit. The algorithm is summarized as in **Algorithm 1**.

Algorithm 1 Data aware energy efficient distributed clustering protocol

Initialize :

- Uniformly distributed region $X * X$.
- $N_{tot}, K, E_{DA}, E_{tx}, E_{rx}, \varepsilon_{fs}, \varepsilon_{amp}, E_h(t), L$.
- $d_{toCH} = \frac{X}{\sqrt{2 * K * \pi}}, d_{toBS} = 0.765 * \frac{X}{2}$.

Start :

- The average energy of r^{th} round is given as

$$\bar{E}(r) = \frac{1}{(N_{tot} + K)} E_{tot-cl} \left(1 - \frac{r}{R} \right) \quad (5)$$

where, R denotes the total rounds and is defined as

$$R = \frac{E_{tot-cl}}{E_{round}} \quad (6)$$

- If nodes have different amounts of energy, p_i of the nodes with more energy should be larger than p_{opt} (optimum probability of choosing a cluster head).

Continued.....

¹ Available[online]: <http://www.ti.com/lit/ug/tidu862/tidu862.pdf>

$$p_i = \begin{cases} \frac{(1+\alpha)p_{opt}E_i(r)}{\bar{E}(r)}, \sum_K E_K(r) > 0 \\ \frac{p_{opt}E_i(r)}{\bar{E}(r)}, \sum_K E_K(r) \leq 0 \end{cases} \quad (7)$$

• The energy dissipated in a round E_{round} , incorporating total cluster energy $\sum_K E_K(r)$, data awareness factor α and RF energy harvesting factor η is given as
 \rightarrow for $\sum_K E_K(r) > 0$

$$E_{round} = L \left(\frac{N_{tot}E_{rx} + K(E_{rx} + \alpha E_{tx})}{N_{tot}E_{DA} + K\varepsilon_{amp}d_{toBS}^4} \right) - N_{tot}\beta E_H(t) \quad (8)$$

\rightarrow for $\sum_K E_K(r) \leq 0$

$$E_{round} = L \left(\frac{N_{tot}(E_{rx} + \alpha E_{tx}) + K(E_{rx} + \alpha E_{tx})}{+N_{tot}E_{DA} + K\varepsilon_{amp}d_{toBS}^4} \right) - N_{tot}\beta E_H(t) \quad (9)$$

Thus we can find the lifetime of network R by putting (1), (8) and (9) in (6).

End

IV. RESULTS AND ANALYSIS

In this section we provide performance evaluation of our DAECCI algorithm. We define a network area of $100 \times 100 m^2$. The simulation parameters are provided in **Table I**. The

Table I. IOT SIMULATION PARAMETERS

Parameters	Value
Network Size	$100 \times 100 m^2$
Sensor nodes N_{tot} in each Cluster	100
Initial battery energy of nodes E_B	0.5 J
Packet Size L	4000 bits
E_{tx} and E_{rx}	50 nJ/bit
ε_{fs}	10 nJ/bit/ m^2
ε_{amp}	0.0013 pJ/bit/ m^4
E_{DA}	5 nJ/bit/signal
p_{opt}	0.1
α and β	$rand(0, 1)$
η	0.4

performance metrics taken in the simulations are *number of Alive nodes*, *Residual energy of nodes* and *Packets sent to BS*. We used Matlab for evaluating our algorithm with other known protocols. In our scenario, we have evaluated the system with four different cases based on α and K for 10000 rounds. For all the cases, we assume that the advanced nodes (gateway nodes) have $a = 3$ times the more energy than the sensor nodes. The cases are as follows:

Case 1 : $a = 3$, $N_{tot} = 100$, $K = 30$, $0.8 \leq \alpha \leq 1$, noE_h . Here, the data demand on sensor nodes is high with no energy

harvesting present.

Case 2 : $a = 3$, $N_{tot} = 100$, $K = 30$, $0.8 \leq \alpha \leq 1$, E_h . Here, the data demand on sensor nodes is high with energy harvesting present.

Case 3 : $a = 3$, $N_{tot} = 100$, $K = 30$, $0.2 \leq \alpha \leq 0.4$, E_h . Here, the data demand on sensor nodes is low with energy harvesting present.

Case 4 : $a = 3$, $N_{tot} = 100$, $K = 50$, $0.8 \leq \alpha \leq 1$, E_h . Here, the data demand on sensor nodes is high with energy harvesting present. Moreover, there are higher number of gateway nodes present compared to previous three cases.

Fig. 2 represents the number of nodes alive during the lifetime of the network. It clearly shows that by introducing RFID based cluster selection and data aware processing, the lifetime improves significantly of the IoT network. LEACH and DDEEC perform poorly as all its nodes are dead by the end of 4000 rounds. Our DAECCI algorithms performance without energy harvesting is comparable to the EDEEC algorithm. With the introduction of E_H , our method outperforms the EDEEC method as around 20% nodes are still alive at the end of 10000 rounds. It can also be inferred that the low data demand of the sensors in case of sparse sensor data requirement almost boosts up the lifetime of the system by 100%. Fig. 3 represents the sum of residual energy of all the

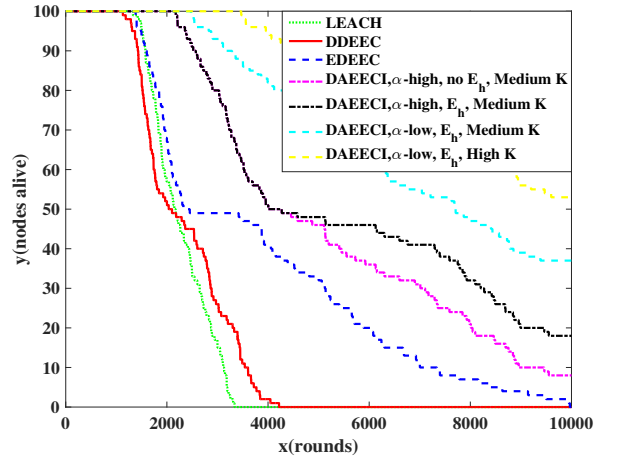


Figure 2. Number of alive nodes in an IoT system versus number of rounds.

nodes in the network. The DAECCI algorithm incorporating RFID tags and data awareness again allows nodes to have higher residual energy compared to the LEACH, DDEED and EDEEC methods. The EDEEC and higher data demand systems DAECCI almost perform similarly. Fig. 4 represents the packets sent to the BS from the cluster heads. The notable thing to infer is that low data demand reduces the amount of packet sent to the BS, whereas irrespective of the high data demand (high α), our algorithm still delivers more packets to the BS than other state-of-the-art methods.

V. CONCLUSION

In this paper, we have proposed a Data aware energy efficient distributed clustering protocol for IoT (DAECCI) by saving CH selection energy using active RFID tags, cutting processing energy by incorporating data awareness factor in

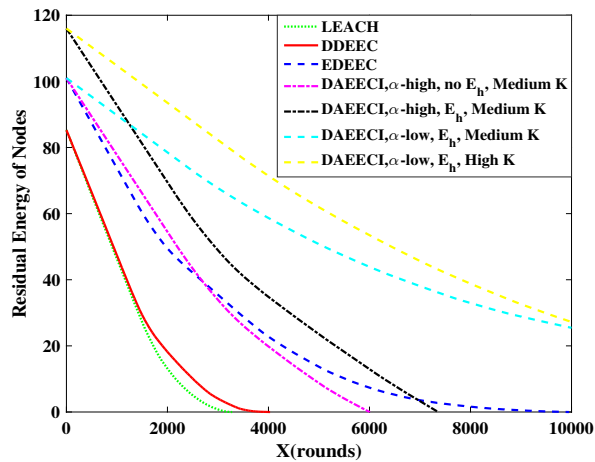


Figure 3. Total residual energy of nodes in the IoT architecture.

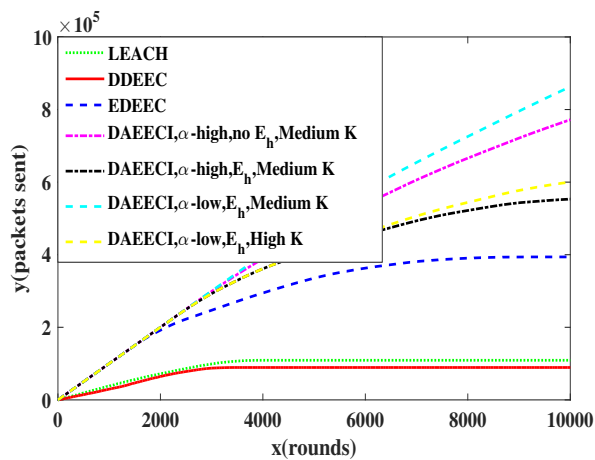


Figure 4. Total packets delivered to the base station server from nodes .

the system and improving lifetime by inculcating RF energy harvesting. We formulate energy consumption models in each round data is sent from sensor nodes to BS through gateway nodes. Our simulation depict substantial improvement in lifetime of network and data delivery to the BS.

REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *Internet of Things Journal, IEEE*, vol. 1, no. 1, pp. 22–32, Feb 2014.
- [2] Z. Sheng, S. Yang, Y. Yu, A. Vasilakos, J. McCann, and K. Leung, "A survey on the ietf protocol suite for the internet of things: standards, challenges, and opportunities," *Wireless Communications, IEEE*, vol. 20, no. 6, pp. 91–98, December 2013.
- [3] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: a survey," *Wireless communications, IEEE*, vol. 11, no. 6, pp. 6–28, 2004.
- [4] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *Wireless Communications, IEEE Transactions on*, vol. 1, no. 4, pp. 660–670, 2002.
- [5] S. Lindsey and C. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *Aerospace Conference Proceedings, 2002. IEEE*, vol. 3, 2002, pp. 3–1125–3–1130 vol.3.

- [6] O. Younis and S. Fahmy, "Heed: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 3, no. 4, pp. 366–379, 2004.
- [7] L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer communications*, vol. 29, no. 12, pp. 2230–2237, 2006.
- [8] T. Qureshi, N. Javaid, M. Malik, U. Qasim, and Z. Khan, "On performance evaluation of variants of deec in wsns," in *Broadband, Wireless Computing, Communication and Applications (BWCCA), 2012 Seventh International Conference on*, Nov 2012, pp. 162–169.
- [9] D. Kumar, "Performance analysis of energy efficient clustering protocols for maximising lifetime of wireless sensor networks," *Wireless Sensor Systems, IET*, vol. 4, no. 1, pp. 9–16, March 2014.
- [10] J. Yu, Y. Qi, G. Wang, and X. Gu, "A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution," *{AEU} - International Journal of Electronics and Communications*, vol. 66, no. 1, pp. 54 – 61, 2012.
- [11] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939–1954, 2012.
- [12] X. Lu, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "Wireless networks with rf energy harvesting: A contemporary survey," *Communications Surveys Tutorials, IEEE*, vol. 17, no. 2, pp. 757–789, 2015.
- [13] P. Kamalinejad, C. Mahapatra, Z. Sheng, S. Mirabbasi, V. Leung, and Y. L. Guan, "Wireless energy harvesting for the internet of things," *Communications Magazine, IEEE*, vol. 53, no. 6, pp. 102–108, June 2015.
- [14] M. Hossain and V. Prybutok, "Consumer acceptance of rfid technology: An exploratory study," *Engineering Management, IEEE Transactions on*, vol. 55, no. 2, pp. 316–328, May 2008.