Data Aggregation in Wireless Sensor Networks (WSs)- A review

[1]Sukhdeep Kaur and [2]Abhinav Hans

[1, 2]*DAV Institute of Engineering and Technology*
Jalandhar, Punjab.
sukhdeepkaur343@gmail.com, abhinavhans@gmail.com

Abstract

This paper present a review on the data aggregation techniques in WSNs. Data aggregation techniques have become an major part of WSNs to overcome the problem of flooding at the base station. Also data aggregation techniques reduce the communication cost of WSNs; because it reduces the redundant data by aggregation. Various data aggregation techniques have been considered in this paper. The data aggregation results in reduction of redundant data therefore reduce the amount of data transmission packets size which will reduce the energy consumption in each round. This paper ends with the shortcoming of data aggregation techniques.

Keywords: Data aggregation, Energy harvesting, Wireless sensor networks, Network lifetime

1. Introduction

A wireless sensor network (WSN) consists of hundreds to thousands of low-power multi-functional sensor nodes, work in an unattended environment, and have sensing, computation and communication capability. The fundamental components of a node are a sensor unit, an ADC (Analog to Digital Converter), a CPU (Central processing unit), a power unit and a communication unit. Sensor nodes are micro-electro-mechanical systems (MEMS) that construct a computable response to a change in various physical conditions like temperature and pressure. Sensor sense or measure the physical data of the area to be monitor. The repetitive analog signal sensed by the sensors is digitized by ADC and sent to controller for further processing. Sensor nodes are of extremely small size, use very low energy, are operate in high volumetric densities, and can be independent and adaptive to the environment.

While a sensor node has limited sensing and computation capacities, communication performance with power, a large number of sensor devices are distributed over an area of interest for collecting information (temperature, humidity, motion detection, *etc*.). These nodes can communicate with each other for sending or getting information either directly or through other intermediate nodes and therefore form a network, so each node in a sensor network act as a router within the network. In direct communication routing protocols (single hop), each sensor node communicates directly with a control centre called Base Station (BS) and sends gathered information. The BS is fixed and located far away from the sensors. BS is able to communicate with the end user either directly or through some accessible wired network. The topology of the sensor network change very repetitively. Nodes may not have global identification. As the distance involving the sensor nodes and base station in casing of direct communication is large, they use energy rapidly. In another approach (multi hop), data is routed through intermediate nodes to the base station and hence saves sending node energy.
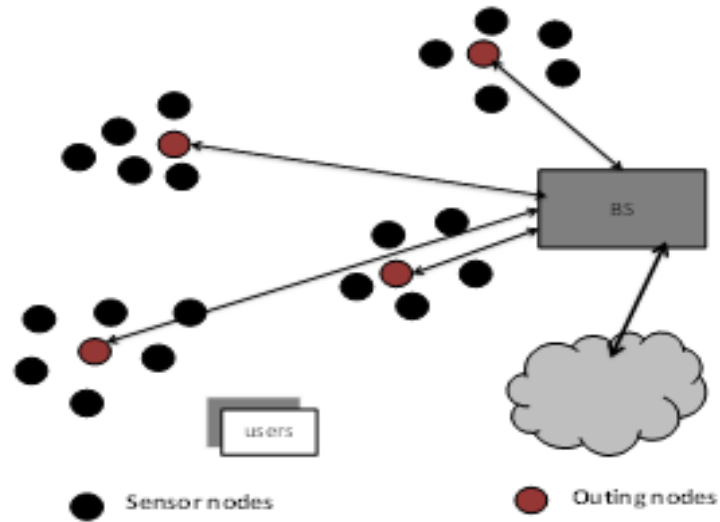
Figure 1. A Wireless Sensor Network Structure

A routing protocol is a protocol that specifies how routers (sensor nodes) communicate with each other; disseminate information that enables them to select routes between any two nodes on the network, the selection of the route being done by routing algorithms. Each router has a priori knowledge only of the networks attached to it directly.A routing protocol shares this information first among immediate neighbours, and then throughout the network. This way, routers gain knowledge of the topology of the network. A Wireless Sensor Network structure is shown in Figure 1.

a) Energy Harvesting

Energy harvesting refers to connect energy from the environment or other energy sources (body heat, foot strike) and convert it to electrical energy. The connected electrical energy powersthe sensor nodes. If the harvested energy source is large and occasionally/continuously available, a sensor node can be powered continuously.In this, all energy sources be supposed to be exploited to extract the accessible energy; among the others, the solar one is usually the most effective in outdoor applications for the high power density provide and exploitable throughout solar cells. Figure 1.2 shows the energy harvesting.
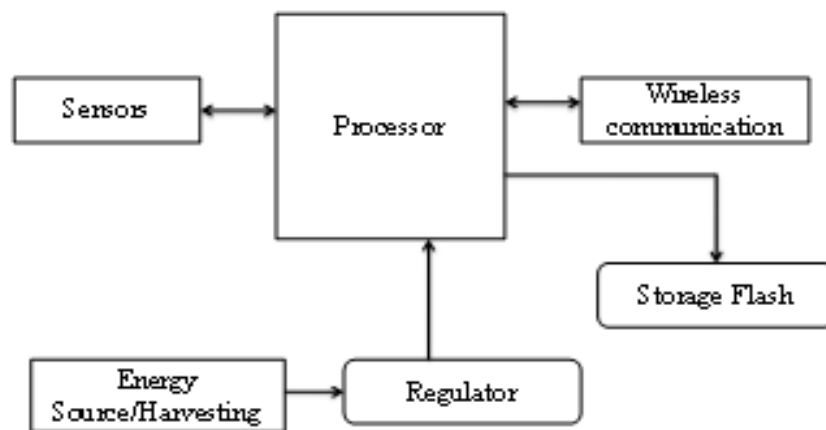
Figure 2. Energy Harvesting

b) Solar Sensor Nodes

Solar sensor nodes are chargeable nodes i.e. they are charged by solar energy. Solar power is the most common and matured amongst the different forms of energy harvesting. Although, it has the drawback of being able to make energy only when there is sufficient sunlight or artificial light; also, existing systems were not designed for use with low power WSNs, timely new research efforts.

c) Data Aggregation techniques

In typical WSNs, sensor nodes are usually resource-constrained and battery-limited. In order to save resources and energy, data must be aggregated to avoid overpowering amounts of traffic in the network. There has been extensive work on data aggregation schemes in sensor networks.

The aim of data aggregation is that eliminates redundant data transmission and enhances the lifetime of energy in WSNs. Data aggregation is the process of one or several sensors then collects the detection result from other sensor. The collected data must be processed by sensor to decrease transmission burden before they are transmit to the BS or sink.

The WSN consist of three types of nodes: Simple regular sensor nodes, aggregator node and querier. Regular sensor nodes sense data packet from the environment and send to the aggregator nodes. These aggregator nodes collects data from multiple sensor nodes of the network, aggregates the data packet using some aggregation functions like sum, average, count, max, min and then sends aggregates result to upper aggregator node or the querier node who generates the query.

It can be the BS or sometimes an external user who has permission to interact with the network. Data transmission between sensor nodes, aggregators and the querier consumes a lot of energy in WSNs.
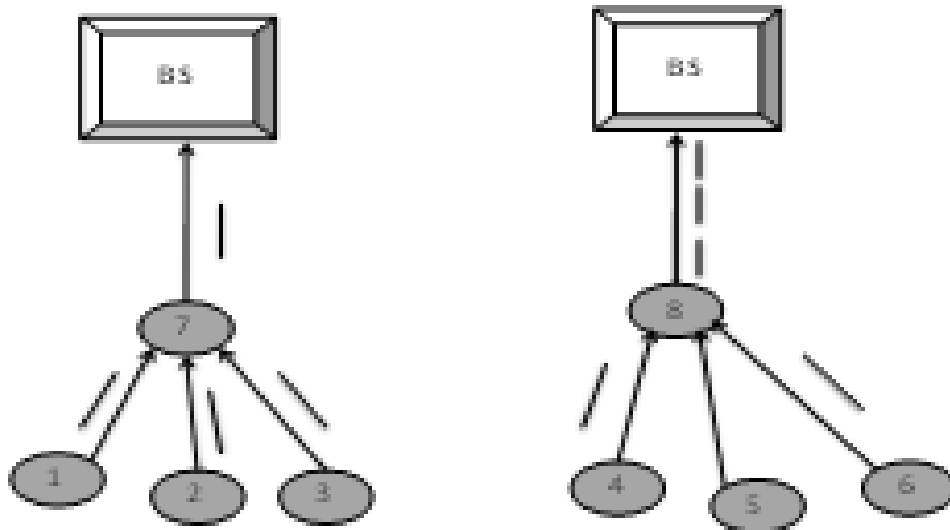
Figure 3. Data Aggregation Model and Non-data Aggregation Model

Figure 3 contain two models, one is data aggregation model and the second is non-data aggregation model in which sensor nodes 1, 2, 3,4,5,6 are regular nodes that collects data packet and report them back to the upper nodes where sensor nodes 7, 8 are aggregators that perform sensing and aggregating at the same time. In this aggregation model, 4 data

packet travelled within the network and only one data packet is transmitted to the BS and other non-data aggregation model also 4 data packet travelled within the network and all data packets are sent to the BS.

With the help of data aggregation process, decrease the number of data packet transmission and also save energy of the sensor node in the WSN. With the help of data aggregation, we improve the lifetime of WSN. Sink have a data packet with energy-efficient way with minimum data latency. So, data latency is very important in numerous applications of WSNS such as environment monitoring, health, monitoring, whereverthe freshness of data is also a significant feature.

It is vital to develop energy-efficient data-aggregation algorithms so that network lifetime is enhanced. There are various data aggregation techniques in WSN:

1. Cluster-Based Approach

In energy-constrained sensor networks of large size, it is inefficient for sensors to transmit the data directly to the BS. Cluster based approach is a hierarchical approach. In cluster-based approach, whole network is divided into several clusters. Each cluster has a cluster-head which is selected among cluster members. Cluster-heads do the role of aggregator which aggregates data received from cluster members locally and then transmit the result to BS.Several cluster-based network organization and data-aggregation protocols have been proposed for the WSN. Figure 4 shows a cluster-based sensor network organization. The cluster heads can communicate with the sink directly via long range transmissions or multi hopping through other cluster heads.
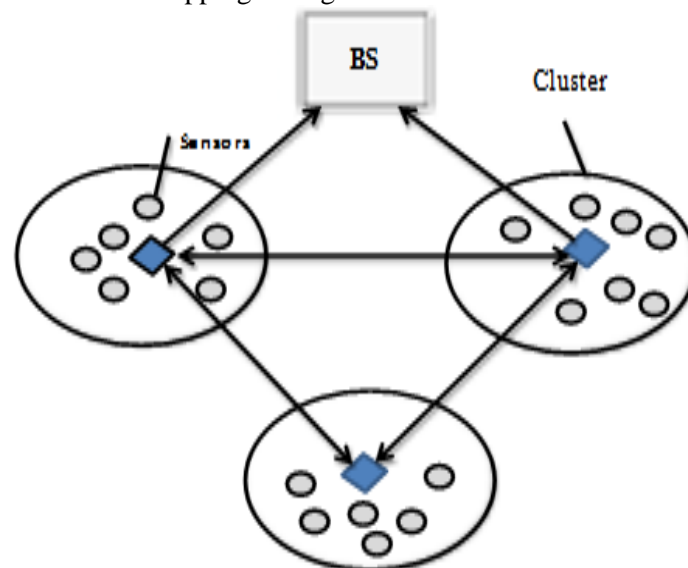
Figure 4. Cluster based Sensor Network. The Arrows Indicate Wireless Communication Links

2. Tree-Based Approach

The tree-based approach is defining aggregation from constructing an aggregation tree. The form of tree is minimum spanning tree, sink node consider as a root and source node consider as leaves. Information flowing of data starts from leaves node up to root means BS. Disadvantage of this approach, like WSNs are not free from failure, in case of data packet loss at any level of tree, the data will be lost not only for single level but for whole related sub tree as well. This approach is suitable for designing optimal aggregation techniques.
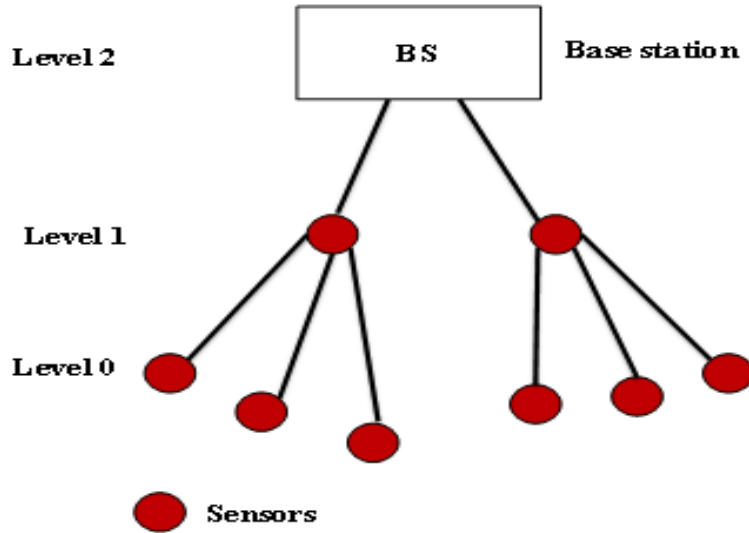
Figure 5. Tree-based Data Aggregation

3. Hybrid-Based Approach

Hybrid-based approach follow between tree and cluster based scheme. In this, the data aggregation structure can be adjusted according to particular network situation and to various performance statistics. There are three types of sensor nodes: normal, advance and super. To become cluster head in a round, normal nodes utilize residual energy base scheme. Advance and super nodes act as relay node to reduce the transmission load of a normal node cluster head when they are not cluster heads in a round.

The super nodes have two transceivers: one connects to the WSN, and the other connects to the super node network. The super node network provide enhanced QoS and is used to rapidly forward sensor data packets to the user. Super nodes could process sensor data before forward.
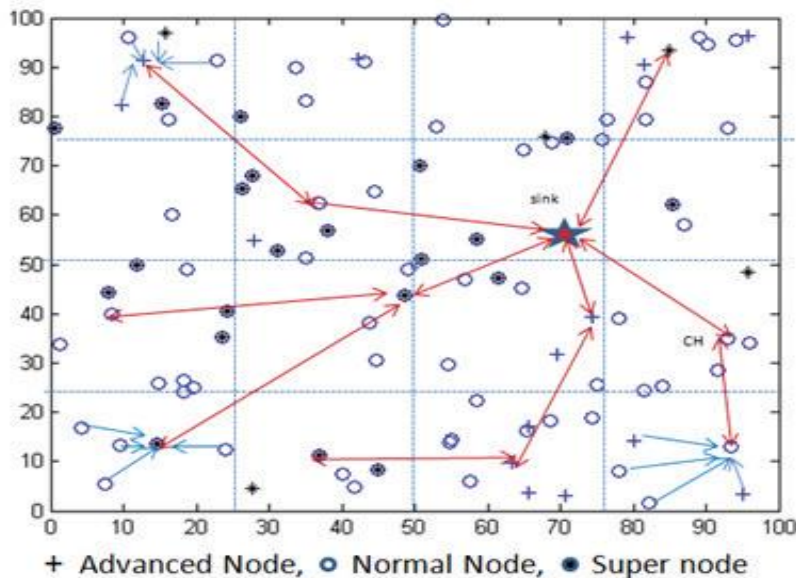

Figure 6. Hybrid based Approach (1)

The hybrid based approach shown in Figure 6 distributes nodes at random with different energy levels as normal node, advanced node and super node through mobile sink.

The threshold for the node to become CH is:

$$T_n = \frac{1}{1 - p(r \bmod \frac{1}{p})} \frac{E_r}{E_i} \frac{D_i}{D_{avg}} \partial$$

Where 'p' is the percentage probability of selecting cluster head, 'r' is the current round; (r mod1/p) is the representation of elected cluster heads. E_r is the residual energy after each round and E_i is the initial energy. D_i is the degree of connectivity and D_{avg} is connectivity of the network. ∂ is the density of node within the selected range of cluster before the election of cluster head.

Let C_i denote the cluster formed by CH node 'i' with neighbor nodes N_i, and then cost of aggregation with cluster C_i is define as:

$$C_i = \sum_{j \in N_i} \cos t(j,i) + \cos t(i, \sin k)$$

This indicate the cost of data aggregation in terms of energy consumption by the network, where cluster members send random packets to CH [cost (j, i)] and CH to sink [cost (i, sink)].

To aggregate the data packets the aggregation functions at CH is defined as:

$$f(C_A) = \sum_{i=1}^{K} (X_i) + \frac{1}{M} \sum_{j=1}^{M} (Y_j)$$

Where X_i and Y_i are variables that represent the correlation of the number of data packets generated by participating nodes in the cluster.

2. Related Works

BHCDA use the absolutely compressible aggregationfunction [1] to enhance the bandwidth utilization and energysaving as compare with EECDA.The aggregation ratio [2] increases as the number of nodes and their density in the network increase compare to other solutions that are not affected by the parameters.Further cluster adjustment and routing update [3] with a small footprint running on the nodes which indicate EHGUC-OAPR can efficiently balance the energy consumption of the whole network and proficiently improve the data delivery ratio.

Reduced energy consumption in transmission [4] of aggregated data indicates profit of the increase in a lifetime of the network. EERDAT use that energy has been reduced effectively and reliability has been maintained as the size of the cluster is altered based upon the loss ratio [5]. MAC protocol has been use for an emergencyevent which takes the benefit of both contention [6] and TDMA alongside with the data aggregation approach. AOR use a regioning approach to group nodes [7] together. While compare to other non-OR routing protocols, AOR achieve high throughput, fairness and scalability. A multi-metrics protocol takes the trade-off [8] between data quality and energy consumption use to increase the lifetime of WSN.

E2IPAP considerably reduces the overhead [9] in disseminate off-path values and overall communication. E2IPAP has secure data aggregation protocols. SHM prototype demonstrates the advantages of performing flexible data aggregationfor energy-efficient and intelligent monitoring. Cluster-based data aggregation mechanism can save energy

and optimize [10] the sharing of computing tasks.An energy-aware routing mechanism is applied to select routes [11] with enough energy to transfer the data, therefore, extending battery lifetime, and enhancing the possibility of sustainable survival of nodes.

DAWN use a number of motion strategy [12] for mobile sinks to reduce the energy consumption beneathdifferent data collection necessities.DCDA-LEACH has the advantages of energy efficiency [13] and even distribution of energy consumption. The actual effect of DCDA-LEACH to expand the network life cycle depends on the importance of the measured data in the specific application.

PDA can protect the trend of private data [14] of a sensor node from being known by its neighboring nodes and data aggregators in a WSN. DAMS (Distributed data Aggregation active Monitoring System) above all designed for the wireless sensor network used in bridge diagnosis [15]. The DAMS can accurately monitor the nodes health status and link quality in the wireless sensor network.

Data redundancy is essential for secured transmission [16] and privacy in interference limited and noisy environment. A highly secured data aggregation method has been used which ensures that information of all live nodes inside the network is accessible inside the sink node with the minimum redundancy. In this, data of all live nodes stored in the sink are accessible inside the sink even if the wireless channel is polluted with noise and interference.

RMPRT algorithm use outperforms its competitors significantly in terms of sustainable packet injection rates [17]. R-MPRT-mod uses exclusively the stored energy in its cost function outperforms the other algorithms.

3. Research Gaps and Research Motivation

To evaluate the gaps in existing research; latest published papers of some well-known journals have been evaluated.

1. The survey has shown that the most of the existing data aggregation techniques are for limited battery WSNs. But in real time applications WSNs comes up with heterogeneous sensor nodes.[2, 8]
2. Many WSNs data aggregation at the base station by individual nodes causes flooding of the data which consequences in maximum energy consumption.[4]
3. Also most of data aggregation methods are either based upon the clustering or tree-based approach but the use of hybrid data aggregation has been ignored by the most of the researchers.[1,10,16]

So, using the hybrid data aggregation in energy harvesting based WSNs is the main motivation of this research work.

5. Conclusion and Future Scope

The energy and bandwidth of the sensors are valued resources along with important to consume proficiently. Data aggregation at the base station by individual nodes causes flooding of the data which consequences in most power consumption. To moderate this problem, numerous data aggregation techniques have been proposed so far. It has been found that the majority of the clustering base protocol has focused on intra cluster data aggregation. Furthermore the use of additive and divisible data aggregation has been neglect by many researchers.

In near future, it would be proposed a new hybrid data aggregation based clustering protocol to enhance the outcome of clustering protocol further. Moreover the use of data compression technique at CHs will also be done to diminish the data packets size to utilize the network bandwidth accurately.

References

[1] D. Mantri, N. R. Prasad and R. Prasad, "BHCDA: Bandwidth efficient heterogeneity aware cluster based data aggregation for Wireless Sensor Network", Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on IEEE, **(2013)**.

[2] A. M. S. Said, A. W. Ibrahim, A. Soua, and H. Afifi, "Dynamic aggregation protocol for wireless sensor networks", Advanced Information Networking and Applications (AINA), 2013 IEEE 27th International Conference, **(2013)**, pp. 356-361.

[3] Y. Wu and W. Liu, "Routing protocol based on genetic algorithm for energy harvesting-wireless sensor networks", IET Wireless Sensor Systems 3.2, **(2013)**, pp. 112-118.

[4] D. Mantri, N. R. Prasad and R. Prasad, "Grouping of clusters for efficient data aggregation (GCEDA) in wireless sensor network", Advance Computing Conference (IACC), 2013 IEEE 3rd International, IEEE, **(2013)**.

[5] B. S. Mathapati, S. R. Patil and V. D. Mytri, "Energy Efficient Reliable Data Aggregation Technique for Wireless Sensor Networks", Computing Sciences (ICCS), 2012 International Conference on. IEEE, **(2012)**.

[6] G. M. Tamilselvan and S. Kiruthika, "An energy efficient data aggregation based medium access control protocol using centre at nearest source approach for sensor networks in a lattice topology", Computing, Communication and Applications (ICCCA), 2012 International Conference on. IEEE, **(2012)**.

[7] Z. A. Eu and H. P. Tan, "Adaptive opportunistic routing protocol for energy harvesting wireless sensor networks", Communications (ICC), 2012 IEEE International Conference on IEEE, **(2012)**.

[8] G. T. Raju, D. K. Ghosh, T. S. Kumar, S. Kavyashree and V. Nagaveni, "Wireless sensor network lifetime optimization", In Advances in Recent Technologies in Communication and Computing (ARTCom 2011), 3rd International Conference, **(2011)**, pp. 244-248.

[9] L. Zhu and M. Li, "An energy efficient and integrity-preserving aggregation protocol in wireless sensor networks", Performance Computing and Communications Conference (IPCCC), 2011 IEEE 30th International, IEEE, **(2011)**.

[10] P. Nie and B. Li, "A cluster-based data aggregation architecture in WSN for structural health monitoring", Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International, IEEE, **(2011)**.

[11] H. C. Chu, W. T. Siao, W. T. Wu and S. C. Huang, "Design and implementation an energy-aware routing mechanism for solar wireless sensor networks", High Performance Computing and Communications (HPCC), 2011 IEEE 13th International Conference, IEEE **(2011)**, pp. 881-886.

[12] S. J. Tang, J. Yuan, X. Y. Li, Y. Liu, G. H. Chen, M. Gu, J. Zhao and G. Dai, "DAWN: energy efficient data aggregation in WSN with mobile sinks", In Quality of Service (IWQoS), 2010 18th International Workshop on, **(2010)**, pp. 1-9.

[13] J. Xibei, Z. Huazhong and Z. Jingchen, "Research of data aggregation routing protocol in wsn data-related applications", Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on IEEE, vol. 1, **(2010)**.

[14] R. Bista, K. H. Dae and J. W. Chang, "A new private data aggregation scheme for wireless sensor networks", Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on IEEE, **(2010)**.

[15] H. Xiao, T. Li, H. Ogai, X. Zou, T. Otawa, S. Umeda and T. Tsuji, "The health monitoring system based on distributed data aggregation for WSN used in bridge diagnosis", In SICE Annual Conference, pp. 2134-2138.

[16] M. B. Zanjani, R. Monsefi and A. Boustani, "Energy efficient/highly secure data aggregation method using tree-structured orthogonal codes for Wireless Sensor Networks", Software Technology and Engineering (ICSTE), 2010 2nd International Conference, IEEE, vol. 2, **(2010)**.

[17] D. Hasenfratz, A. Meier, C. Moser, J. Jia Chen and L. Thiele, "Analysis, comparison, and optimization of routing protocols for energy harvesting wireless sensor networks", Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), 2010 IEEE International Conference on, IEEE, pp. 19-26, **(2010)**.