International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Impact Factor (2020): 6.733

Exploring Clustering Algorithms for Internet of Things: A Review

Deepak Jyoti

Assistant Professor Computer Science & IT, Shanti Devi Arya Mahila College Dinanagar Punjab, India
Email: erdjyoti346 [at]gmail.com

Abstract: ***In networking, the IOT is a major shift that helps to attach nearly all items on the world. Energy efficient routing will play a crucial part in the effective implementation of these systems, given the limited existence of smart devices. Clustering techniques organize the network nodes into groups or clusters, and the cluster head is accountable for the particular assigned node. In the sense of Wireless Sensor Networks (WSN), clustering techniques have been especially proposed, but their implementation can also answer related problems Internet of Things (IoT). By assigning big portions of coordination overhead to the cluster head, clustering could promote energy-efficient routing and configuration management. This article describes a comparison of existing clustering protocols depend on various parameters such as energy efficiency, scalability, etc of different IoT domains and proposes suitable cluster analysis for each domain.***

Keywords: Internet of things (IOT), Clustering, WSN, Networking.

1. Introduction

IoT is an evolving analysis field that integrates a variety of fields of study. The key IoT concept is to link all devices to the Internet, such as home appliances, mobile phones, vehicles, houses, robots, machines, and so on. The concept is to provide a virtual equivalent for all applications in the real world that detects essential data from the environment in order supply advanced end-user services [1]. Energy efficiency is one of the major concerns when using these devices, as communication and computation on the restricted device could rapidly discharge its battery capacity resources. In the scenario Nodes depend on self-organizing multi-hoping networking methodologies of WSN, which could operate in the absence of a base station and comparable strategies could be implemented for IoT. A main research problem is the creation of DR algorithms that could effectively observe paths among mobile nodes. Due to the obvious rise in computational burden on mobile nodes, dynamic networks do not use traditional algorithms. Connect performance and network topology could differ while a message packet has been routed [2].The maintenance of a high quality connection therefore involves regular measurement and upgrading of routing paths. Clustering is known to be the most efficient way to solve the performance issues of ad hoc networks and inevitably opposes their usage in the sense of IoT due to similar difficulties.

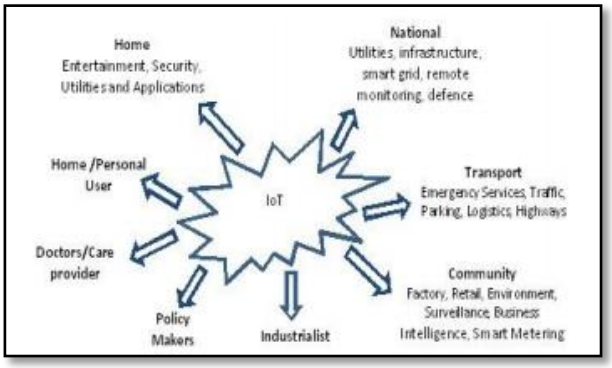

Figure 1: Internet of Things (IOT) [1]

Table 1: Evolution of IOT Technologies

Technology	Time Span	Description
RFID	1999	Passive identification, wireless networks
WSN	2005	WSN, Cloud computing, Web2.0, low energy communication
Smart Things	2012	Mobile computing, cooperating operation of objects, connecting devices
IOT	2017	Advanced sensor fusion, faster wireless connectivity, predictive analysis

Clustering [3] is the method of seeking a natural connection between certain particular nodes or the gathering of related objects. Three key forms of nodes are included in the clustering network, including 1) cluster heads, 2) regular nodes, and 3) gateway nodes. Every cluster is made up of one head of the cluster serving as a control unit. The node transmitted signal specifies the length of the cluster in the single hop cluster and the amount of hops is calculated in the multi-hop cluster. The information is retrieved by the cluster head from a regular node, which is transferred to the next hop. A connection among two clusters with distinct cluster heads is the action of the base stations. Together, they form the basis of the system, while there is no obligatory existence of a gateway node [4]. In addition, the base station (BS) acts as a medium of contact among the user and the sensor network. The volume of transmitted data is minimized by selecting a cluster head and combining neighboring nodes. Data aggregation is done to prevent duplication and coordination burden provided by various transmissions, Devices with improved computing abilities such as smart phones may be equivalent to numerous ubiquitous sensors such as temperature sensors, security cameras, home appliances, and so on in the IoT paradigm. In the Iot system t he transition in any node in the network causes a change in the topology that occurs in overhead topology improvement messages. Perfect alternatives for solving certain situations are the clustering algorithms. Various benefits are provided by clustering. Network

scalability is provided [5]. Routing choices within a cluster and manages the corresponding head of the cluster thereby minimizing the complexity of routing. Clustering saves communication bandwidth by testing Cluster Heads (CH) data, thereby preventing the repeated interchange of data packets among nodes [4]. The overlap for topology maintenance is significantly minimized by sustaining the network at the cluster level. The end devices are only concerned with the relation to their CHs without being influenced by inter-CH tier changes [6]. A CH increases the network lifespan by using advanced control techniques to conserve the power of the computer. In required to preserve that nodes are operating in low-power mode when not active to conserve battery resources, a CH could even applied on different scheduling algorithms.

The article is organized as obeys: Section II provides the various clustering algorithms. Section III show the related work of clustering in IOT. In the last, it is concluded in Section IV.

2. Clustering Algorithms

Here, explains different clustering methods introduced in the context of WSN:

- **LEACH and LEACH-C:** LEACH [7], which divides the network into clusters using a distributed approach. The probability of being a CH is determined by every node and it broadcasts its judgment. In order to achieve the CH, a node decides its cluster so that it needs the minimum communication energy. The method produces for the balance of energy between nodes by using random rotation between CHs. In order to fetch the information to the base station, CH nodes serve as routers and also hire data fusion and aggregation on the cluster sensed data. LEACH ensures energy conservation by allowing a node to select a cluster focused on minimum communication resources. In the choice of a cluster head, no overhead is also needed, while each node determines individually. Since CH has chosen is probabilistic in LEACH, there could be a circumstance where a very low energy node is identified as CH through compromising the entire cluster. In addition, even though nodes may interact directly with CH, LEACH could result in one-hop intra- and inter-cluster configuration. In LEACH-C [8], the setup and steady step is identical to LEACH, but the choice of CH is focused on energy knowledge and load balancing. In order to absorb more resources and minimize usability, the Base-Station is accountable for the overall network configuration and cluster design.

- **HEED**

It [9] is a multi-hop clustering method that selects CH for intra-cluster communication depend on residual battery and energy available. In a way close to other clustering techniques, an initial version of CHs is probabilistically selected depend on the residual energy. The value of communication in the intra-cluster serves as an implicit variable to provide an indication of the level or proximity of a node to neighboring nodes. This makes it possible for a node to determine why not to join a specific cluster. A cluster that uses low power rates for communication enables spatial reuse because of its long range because of reduced interference while high power is used for inter-cluster communication. HEED allows efficiently spread CHs around the network and offers sound load balancing.

- **Energy efficiency Load-Balancing clustering (EELBC):** EELBC [10] refers to a class of centralized networks in which sink executes an overall role by attaching sensor nodes to various distance-based gateways and constructing a min-heap on the set of sensors assigned to the gateway. The core of the minimal heap is created by the gateway and modification is done in a way that the lowest packed gateway is at the root.

- **Energy-efficient unequal clustering (EEUC):**

CHs in close proximity to BS tend to remove quicker as they transfer traffic from remote nodes in order to get data from their own cluster is an issue that exists often in WSNs. This is referred to as the hot-spot issue that EEUC [11] is trying to resolve between clusters by balancing resources. In order to save resources, clusters nearer to the sink are made lower than distant clusters.

- **EECS**

In this scheme [12], nodes broadcast their residual energy to neighboring nodes to appoint the node with highest residual energy as the CH for a defined round. EECS expands the creation of clusters in LEACH, which selects CH on the basis of average distance; by dynamic cluster fitting depend on the distance from the BS. This tends to address the issue that clusters far from the BS need high communication power than the neighboring clusters, leading to better use of limited storage resources and improved network life. Congestion in clusters close to BS can lead to early CH death.

- **Algorithm for cluster establishment (ACE):**The method comprises of 2 phases: new clusters are spawned and existing clusters are migrated. Initially, clusters are in an un-clustered state and wait for the next phase to adjust the configuration and it decides to be a cluster member cluster head depending on the current state and surroundings [13].

- **GROUP**

A hybrid method that operates in four steps, i.e. creation of the CG, query forwarding, forwarding data as well as restoration of loss. One sink situated in the vicinity of the network is chosen as the primary sink and the sink elects one node as CH is focused on residual energy and distance value. Unicast and broadcast correspondence are used for the forwarding of requests. This method finds its use in large-scale WSN [14].

- **PEACH**

Due to the high workload of cluster creation, clustering algorithms mainly suffer. PEACH aims to resolve this issue in way to minimize the power cost as well as to enhance the lifespan of the network., PEACH decreases over-head clustering, by using wireless channel overhearing features to allow adaptive multilevel clustering.

- **Passive clustering (PC):** PC [15] is a clustering technique that only operates when the clusters are actively involved in communication in an on-demand manner,

forming clusters. The conditions of various nodes in the cluster between the communicating nodes are piggybacked and ex-changed. When the nodes receive a packet from them, they refresh the database of CH and Gateway (GW) nodes. In a distributed manner, the PC uses Gateway Selection Heuristics (GSH) to decide how node states adjust in just such a way that the minimum level of GW nodes is used to maintain connectivity. The method used is easy to applied and reduces overhead included in the creation of clusters.

To fix specific concerns facing WSN, the formulas mentioned above have been suggested. A comparison study of these protocols needs to be carried out in order to use these methods in IoT systems. A comparison study of the above mentioned methods is provided in Table 2 on the foundation of variables such as energy efficiency, data aggregation, mobility, form of network, etc. Such research will help determine which strategies for a specific IoT specification are most likely possible. Since IoT faces WSN-like problems such as energy efficiency, usability and data aggregation, suitable evolutionary techniques could be used in IoT to solve significant difficulties. Even so, concerns such as accessibility and the form of network, which may not be very important in the sense of WSN, reflect important issues for different IoT applications.

Table 2: Comparsion of clustering algorithms [7-15]

Algorithm	Control Manner	Energy Efficiency	Data Aggregation	Scalability	CH Selection	CH Rotation	Mobility	Network Type
LEACH	Distributed	Less	Never	Never	Probabilistic	Agreed	Agreed	Homogeneous
LEACH-C	Centralized	Less	Never	Never	Probabilistic	Agreed	Never	Homogeneous
EELBC	Centralized	Average	Never	Never	Probabilistic	Agreed	Never	Homogeneous
EEDC	Centralized	High	Never	Never	Probabilistic	Agreed	Never	Homogeneous
LCA	Distributed	Less	Never	Never	Id-Based	Agreed	Agreed	Homogeneous
PEACH	Distributed	Medium	Agreed	Agreed	Probabilistic	Agreed	Agreed	Homogeneous
GROUP	Mixed	High	Never	Never	Weight Based	Never	Never	Homogeneous
HEED	Distributed	Medium	Agreed	Agreed	Probabilistic	Agreed	Agreed	Homogeneous
EEHCA	Distributed	High	Agreed	Agreed	Probabilistic	Agreed	Never	Heterogeneous
EACLE	Distributed	Average	Agreed	Never	Connectivity	Agreed	Never	Heterogeneous
EECS	Distributed	Medium	Never	Agreed	Weight Based	Agreed	Never	Homogeneous
EEUC	Distributed	High	Agreed	Agreed	Probabilistic	Agreed	Never	Homogeneous
S-WEB	Distributed	High	Never	Agreed	Probabilistic	Never	Agreed	Homogeneous
MCEEC	Centralized	Agreed	Yes	Never	Probabilistic	Agreed	Never	Heterogeneous

3. Literature Survey

Liu et al., (2019) Introduces a DRL method for DNC with edge servers in IoT systems. The goal is to satisfy the specifications of either the IoT system and EC via improving the group of data clustering. The outcomes demonstrate that the suggested DQN design could produce strong performance as contrasted to the static benchmark solution [16].

Bensaid et al., (2020)A new clustering method depend on FCM is being proposed for WSN-focused IoT apps. The algorithm considers an FCM approach to shape the clusters and minimize the total consumed energy in each cluster to find the optimal Cluster Head (CH) for every transmission round. A comparative with the LEACH algorithm is provided to evaluate the efficiency as well as to certify the suggested approach, Findings demonstrate the suggested FCM approach increases the lifespan of system that via improving the remaining energy including 50 percent [17].

Ganeshan et al., (2019)To efficiently make use of the available energy, a clustering scheme is introduced. The suggested approach would position an additional IoT system that functions as CH in the user's structure with strong loading power. The data would be distributed to CH by the cluster participants and the CH passes the aggregated information to the CP. The 2 efficiency metrics, such as the amount of data transmissions and the maximum set of information transmissions to the cloud network, are utilized for experimentation. The findings outcomes have confirmed that the implementation of a clustering method minimize power usage [18].

Aboalnaser et al., (2019) It examines energy usage in IoT networks as well as suggests a latest TA method depend on clustering solutions to manage communication energy among nodes. The engineered system tries to optimize the energy between all SD or continue the lifespan of the network. In means of energy efficiency and lifespan of nodes, the efficacy of the suggested method is assessed or the outcomes suggest better productivity of the suggested method [19].

Behera et al., (2019) The emphasis is on an effective CH election approach that vibrates the location of the cluster head between the greater energy nodes related to others. Initial energy, residual energy and the optimum conditions of cluster heads are considered in the method to pick the next category of system cluster heads appropriate for IoT applications, like weather forecasting, smart buildings and networks. The findings demonstrates that the improved design plays best as compared to the Leach algorithm through raising performance including 60%, lifespan by 66%, residual energy including 64% [20].

Fan et al., (2019) For the complex strategic situation of NB IoT, in which the system keyboards are grouped depending on the time from the BS and TTI, as well as assigned recycle preamble software to increase the ability of network access, an improvement access approach is applied to get high QoS. To test the modified entrance approach depend on CRPD, the capability is modeled and evaluated (ERACRPA). The

test results reveal that the ERA-CRPA could efficiently minimize the risk of collision with the random access preamble, thus enhance access productivity and enhance the NB-IoT network's equivalent capacity [21].

Xiu et al., (2018) Suggest a PADC for confidentiality and accessibility focused on the k-means method and original centre values as well as the process of distance measurement from another locations to the centre value. In addition, PADC efforts to minimize the deviation impact by identifying deviations through the clustering phase. Security review shows that device fulfils the aim of differential security as well as avoids leakage of confidential data. In the meantime, performance evaluation indicates that our method increases the accessibility of clustering outcomes at the similar phase of privacy contrasted to the current differential privacy approaches, indicating that the suggested PADC system outperforms for smart power facility in IoT [22].

Miead et al., (2019) It focused on directive transmission in urban outdoor situations and illustrates the value of correctly estimating the set of mm Wave radio channel clusters modeled with a ray-tracer software package. Studies show that clustering is a tough job, since one or a mixture of more CVIs does not always offer the optimum set of clusters. Use score fusion techniques to observe the optimum partitioning for the k-means method depend on the multipath rays' power and time of arrival or based on their angle of incidence. When the k-power-means technique is utilized, the intensity of each entering ray is the most critical clustering element, enabling the prevailing routes collected to drag the others into them to build a cluster. The collection of clusters is lower and it is quicker to carry a conclusion on CVIs or score fusion factors [23].

4. Conclusion

In this article, which examined numerous clustering techniques commonly utilized in the context of WSN, Techniques were categorized on the foundation of variables including energy efficiency, mobility, data aggregation as well as network form. Clustering algorithms provide a promising solution to topology management, energy needs, and data aggregation problems of IoT. Even so, problems such as fault tolerance, service efficiency and safe communication have to be addressed. With the aid of this document, different parameters will enable the best clustering method to be found.

References

[1] Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M.and Ayyash, M., “Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications, ” IEEE Communications Surveys & Tutorials, Vol. 17 No. 4, pp. 2347-2376, 2015.

[2] Yao, X., Wang, J., Shen, M., Kong, H., & Ning, H., “An Improved Clustering Algorithm and Its Application in IoT Data Analysis” Computer Networks, 2019.

[3] Xu, Lina, Gregory MP OHare, and Rem Collier, “A smart and balanced energy efficient multihop clustering algorithm(smart-beem) for mimo iot systems in future networks”, Sensors, vol. 17, no. 7, pp. 1574, 2017.

[4] S. Bhandari, S. K. Sharma and X. Wang, “Cloud-assisted device clustering for lifetime prolongation in wireless IoT networks, ” in IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1-4, 2017.

[5] A. Abbasi, M. Younis, “A survey on clustering algorithms for wireless sensor networks, ” Computer Communications, Vol. 30, pp. 2826-2841, 2007.

[6] D. Bandyopadhyay and J. Sen, “Internet of things: Applications and challenges in technology and standardization”, Wirel. Pers. Commun., vol. 58, no. 1, pp. 49–69, 2011.

[7] J. Duan, D. Gao, D. Yang, C. H. Foh, and H.-H. Chen, “An energyaware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications”, IEEE Internet Things J., vol. 1, no. 1, pp. 58–69, 2014.

[8] H. Junping, J. Yuhui, and D. Liang, “A time-based cluster-head selection algorithm for LEACH, ” in Computers and Communications ISCC 2008, IEEE Symposium on, pp. 1172–1176, 2008.

[9] K. Maraiya, K. Kant, and N. Gupta, “Efficient cluster head selection scheme for data aggregation in wireless sensor network, ” Int. J. Comput. Appl., vol. 23, no. 9, pp. 10–18, 2011.

[10] B. Singh and D. K. Lobiyal, “A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks, ” Human-Centric Comput. Inf. Sci., vol. 2, no. 1, p. 13, 2012.

[11] P. C. S. Rao, P. K. Jana, and H. Banka, “A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks, ” Wirel. networks, vol. 23, no. 7, 2017.

[12] H. Farman et al., “Multi-criteria based zone head selection in Internet of Things based wireless sensor networks, ” Futur. Gener. Comput. Syst., 2018.

[13] S. Kallam, R. B. Madda, C.-Y. Chen, R. Patan, and D. Cheelu, “Low energy-aware communication process in IoT using the green computing approach, ” IET Networks, vol. 7, no. 4, pp. 258–264, 2017.

[14] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, “An information framework for creating a smart city through internet of things, ” IEEE Internet Things J., vol. 1, no. 2, pp. 112–121, 2014.

[15] T. M. Behera, U. C. Samal, and S. K. Mohapatra, “Energy Efficient Modified LEACH Protocol for IoT Application, ” IET Wirel. Sens. Syst., vol. 8, no. 5, pp. 223 – 228, 2018.

[16] Liu, Q., Cheng, L., Ozcelebi, T., Murphy, J., & Lukkien, J., “ Deep Reinforcement Learning for IoT Network Dynamic Clustering in Edge Computing”, 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), 2019.

[17] Bensaid, R., Ben Said, M., & Boujemaa, H., “Fuzzy C-Means based Clustering Algorithm in WSNs for IoT Applications”, International Wireless Communications and Mobile Computing (IWCMC), 2020.

[18] M.Ganesan, Dr.N.Sivakumar, “An energy efficient IoT based Healthcare System based on clustering technique”, Third International Conference on

Electronics Communication and Aerospace Technology ICECA, 2019.
[19] Aboalnaser, S. A., " Energy–Aware Task Allocation Algorithm Based on Transitive Cluster-Head Selection for IoT Networks", 12th International Conference on Developments in eSystems Engineering (DeSE), 2019.
[20] Behera, T. M., Mohapatra, S. K., Samal, U. C., Khan, M. S., Daneshmand, M., & Gandomi, A. H., "Residual Energy Based Cluster-head Selection in WSNs for IoT Application", IEEE Internet of Things Journal, 2019.
[21] Fan Wu, Sijia Huang, "An Enhanced Random Access Algorithm Based on the Clustering-Reuse Preamble Allocation in NB-IoT System", IEEE, Vol. 7, 2019.
[22] Xiong, J., Ren, J., Chen, L., Yao, Z., Lin, M., Wu, D., & Niu, B., " Enhancing privacy and availability for data clustering in intelligent electrical service of IoT", IEEE Internet of Things Journal, 2018.
[23] Miead Tehrani Moayyed, Stefano Basagni, "Clustering Algorithms and Validation Indices for mmWave Radio Multipath Propagation", IEEE, 2019.