

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter aims to review the existing literature on network connectivity and clustering algorithm in Wireless Sensor Network (WSN). The literature focus on the characteristic of WSN, the behaviour of the sensor nodes and existing algorithm on clustering network. At the end of this chapter, the gap analysis will be discussed and the improvement from the current clustering algorithm also is suggested.

2.2 Wireless Sensor Network

WSN have been identified as a system that self-reconfigures in response to external conditions, as it consists of several different nodes that communicate with each other wirelessly, and contain smart and inexpensive sensors that monitor the surrounding changes and then collect and process them, and send them to observers. A wireless sensor networks are distinguished by several advantages, and those set of characteristics which explain in the next section.

2.2.1 Characteristic of WSNs

WSNs have some unique characteristics such as mobile node, ad hoc connectivity, low battery, dynamics network topology, node deployment, and communication capabilities. The characteristics of WSN are shown as following:

a) Ad Hoc Connectivity

An ad hoc connectivity is characterized by its dynamic network topological ups and downs, limited communication bandwidth, and low battery power of the neighborhood. The

network topology of a WSNs can change frequently and dramatically. One case is that nodes in a WSNs are capable of network clustering collectively or randomly. When one node is moved out of or in to the transmission range of another sensor node, the link between the nodes becomes down or up. Some other reason that causes the network topological random is the unstable wireless links, which might become up and down due to the sensor fading obstacles between of the nodes, interference from other sensors, or the changing of transfer power levels. Many cases the sensor nodes are not stable, the topology might change frequently because the sensor nodes of the network clustering are dynamic (Ejmaa & Subramaniam, 2016).

b) Low battery

The key to achieving a longer lifetime for WSN is to improve network connectivity for wireless sensor networks that minimize energy consumption of wireless sensor node, hence the name low battery. To cut down on overall energy consumption, low battery wireless sensor networks control the active node or awake node of the cluster such as transmission range and distance and limit the current draw when they are sleeping. These networks accomplish this by varying the power setting behaviours of the nodes, such as failure node, selfish node, malicious node and cooperative node.

Inactive mode, the node uses power to improve network connectivity and to manipulate data using a clustering algorithm, but the node does not have to do this constantly. When not in active mode, the distance can return to sleep mode until more sample correlated degree measurements are taken. Then, an establish network connectivity, nodes are required to be in a cooperative behaviour state to perform data transmission. If a node is not active, it will go back to sleep, conserving power usage. When the amount of time the transmission spends running is smartly managed and controlled, the overall amount of energy consumption is drastically reduced.

A sensor network which often becomes the network may become inaccessible because of the network cluster becoming even more challenged. Hence, the clustering algorithms should be considered for battery energy conservation of the network. In improver, the node use of the energy when transmits data is more than the energy of the sensor nodes implement computing consume. The power consumption of sensor node when transmitting a bit of the data information can be used to execute line calculation instructions. Therefore, energy is saved in the network processing and sensor of node does not become lost as to maximize wireless sensor network's lifetime (Mittal, N., Singh, U., & Sohi, 2016).

Clearly, a node behaviour deployment randomly of the networking cluster threshold to distinguish trustworthy nodes or not trustworthy nodes in a wireless sensor network. The node failure could be induced by the physical node failure and physical damage when a neighbouring node encounter different network conditions which cause packet drop without the transmission range due to buffer overflow.

c) Mobile node

The mobile node in wireless sensor networks, wherein numerous approach has been examined to control the active node when the node transmits data activity to another node to charge dynamism to achieve a network lifetime and awareness of multi-hop clustering algorithms (Chanak, P., Banerjee, I., & Sherratt., 2016). Many studies have been done to assess the possible impact on the accuracy and real-time properties of the sensor node by analyzing the likely advances in life battery that can be achieved by tuning easy-to-adjust parameters in a number of sensors. Those studies offer a quantitative foundation for future research in low battery-aware scheduling techniques for a number of the sensor.

That opportunities may exist for adapting transmission power levels, and scheduling number of sensor node which activates in a way that depends on the battery power efficiency on the network's lifetime. This is observed through an application that collects environmental

information in an area exhibiting large differences in temperature. Application duty-cycles and/or data precision could be held to receive the answer of the sensed data based on ambient temperature.

A mobile node is very similar to a mobile wireless sensor network, but resource constraint attributes of Mobile node demand performance analysis of effects caused by various node mobility patterns. A correlated degree is widely used in studies of the mobile node, which will give a true random node movement as expected in many mobile node applications. Mobility of the cluster heads (CHs) can be either stationary or mobile. The mobile CHs can move for a limited distance, although the topology management process of mobile CHs is more difficult than in a network with stationary CHs.

Although many research has been suggested in the clustering algorithms for ad-hoc wireless sensor networks, the objective was mainly to generate stable clusters in environments with mobile nodes. Many of the techniques focus on the node reachability and route stability, without degrading the degree about critical design goals of wireless sensor networks such as connectivity and coverage. In recent times, a number of clustering algorithms have been specifically designed for WSNs.

d) Dynamics Network topology

Mobility node would make clustering very challenging since the guest neighboring node will change dynamically which in turn may change clusters over time. On the other hand, the network depends on the sensory of the node either intermittent or continuous depending on the network clustering. In a target detection/tracking network, the event phenomenon is dynamic whereas node on network monitoring for early failure prevention is a network topology of intermittent issues. Monitoring intermittent events allow the network topology to operate in a reactive way, simply getting traffic when reporting. Continual events in the most

network require periodic reporting and consequently generating significant traffic to keep the lifeline of the network. (P. Singh & Satvir, 2017).

e) Heterogeneity of nodes

Heterogeneous means different nature of nodes, which are different from each other by the communication range, mobility, sensing parameters and work at the different cluster. Heterogeneous for WSN are collection of various different types of sensor nodes with have different features, follow different protocol, different computation capacity and different sensing and monitoring range. Deployment of a heterogeneous sensor network clustering more typical than homogeneous wireless sensor network clustering.

The heterogeneity is quite contextual in WSN. In the case of hardware terms, there might be a different kind of communication protocol, MAC layer, memory, computing architecture. In the abstract or higher level of thinking, you may consider different numbers of neighbors surrounding by a node might be different than others, in this way either of the nodes is heterogeneous to each other. Might be those nodes neighbors are pushing traffics on different rates or moving fast or slow. These contextual differences make WSN also heterogeneous (Farman, Javed, Ahmad, Jan, & Zeeshan, 2016).

There is a different number of sensors deployed in a hierarchical manner. In this case, the leaf nodes are typically energy-constraint, and that do not require much memory. Now, consider the nodes which are at the top level. These nodes should have a larger memory to store and process information from their child nodes. However, they can be charged. So, there are four types of nodes such as failure node, selfish node, malicious node and cooperative node with different properties which are energy and memory. Such sensor network can be treated as heterogeneity sensor networks.

For improvement, the communication range and the relative network clustering proximity to the sink have to be considered. The communication range is usually fixed sensors

in the CH which may not be able to receive at the sink. When node directly communicate with the sink, it is still better to pursue multichip directions. Thus, inter connectivity cluster heads becomes an important component that involves the network clustering (F. Wang, Wu, Wang, & Hu, 2016). On the other hand, heterogeneous sensor network may impose more constraints on the network clustering process since some nodes may be specified for particular tasks or empowered with distinct capabilities. It may then be examined to avoid such specific nodes to maintain their resources or limit the cluster heads selection.

f) Scalability to large scale of deployment

The sensor nodes are capable of aggregating data packets from the sensor nodes in their clusters and transmitting them to the sink node wireless multi-hop paths. The sensor node placement problem in large scale wireless sensor networks. The objective is to improve network connectivity with a number of nodes in the playing field of a network clustering such that each sensor node can communicate with at least one node and the network of nodes is connected. However, placement strategies realizing goals and do not provide any fault-tolerance as the network may lose functionality after the failure of some of the node. To incorporate fault-tolerance in such a network, that ensure that every sensor node can communicate with at least two sensor nodes and the induced network clustering. This strategy will ensure the clustering of the network in the event of a single fault, instead of higher node placement cost.

A sensor node may be added to the network and existing sensor nodes may become cluster head. This large scale and frequently changing network require scalable protocols and algorithms. Factors, such as energy depletion, harsh environmental conditions, and/or malicious attacks may result in node failures in a wireless sensor network. Therefore, survivability of sensor networks is a critical design goal. Moreover, energy is one of the most

precious resource in wireless sensor networks. Sensor nodes are normally powered by batteries and can only last for a fairly short period of time if operated at high transmission power levels. As a consequence, energy efficient design is needed for prolonging network lifetime (Chhaya, Sharma, Bhagwatikar, & Kumar, 2017).

2.2.2 Node Behaviour

There are four states of node behaviour in network which give different impact towards network performance: failure node, selfish node, malicious node and cooperative node. The following describes each node in details:

a) Failure node

The sensor nodes behaviour is the main issue in Wireless Sensor Network (WSN), especially network connectivity due to dynamic nature of the network. Achieving energy efficiency in WSNs is a challenge due to the direct transmission from sensor node to the sink node. This problem arises from the effect for each sensor node upon failure in node activation (W. Wang, 2006). To establish a network connectivity, nodes are required to be in a cooperative behaviour state to perform data transmission.

However, the changes of nodes' behaviours may occur from cooperative to misbehaving, which directly affects the network's connectivity. In respect of the correlated node behaviour model in a study by Azni et al., (2013), misbehaving nodes consists of failure nodes, malicious nodes, and selfish nodes. These nodes often have a correlated effect, leading to partition in the network. In respect of the failure node, it could be illustrated through node x_3 , as illustrated in Figure 2.1. When node u initiates a route discovery to another node v , the failure neighbour node x_3 will be unable to transmit the route discovery. With the failure of all node u neighbours, node u would no longer communication with other nodes. In this case, node

u is considered as a node correlated with its failure neighbours. The classification of failure node is set to the value of energy threshold < 0 .

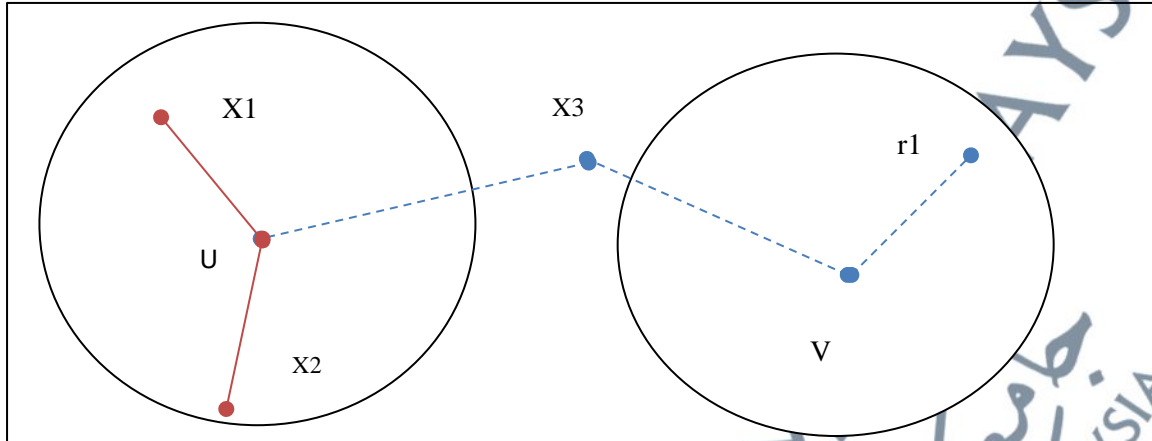


Figure 2.1: Failure node

The failure of the node neighbours leads to the occurrence of the correlated behaviour. As a result, more traffic, which is originally brought by these neighbours, needs to be loaded. This failure could occur at a more rapid rate due to excessive energy consumption. Consequently, the energy consumed through the transfer of the packet or data from the node-to-node destination could raise a critical issue in the development of a reasonable network connectivity for wireless sensor network. Another misbehaviour of the node, which causes a critical impact on connectivity is its transition from cooperative to selfish node.

b) Selfish node

Nodes that refuse to forward other's packets defined as selfish nodes. It thus maximizing their benefit at the expense of others. They are acted cooperatively and cheat only if it gives them an advantage. In addition, this kind of selfish node can accelerate the average end to end delay in which whenever the number of nodes become selfish increased, less feasible route will be available and results to longer delays.

Figure 2.2 illustrates the behaviour of selfish nodes, particularly node x_3 . When node u initiates a route discovery to another node v , difficulty would occur for the selfish neighbour node x_3 to broadcast the route request from node u . In this case, the behaviour of node x_3 is similar to that of a failure node. It is also possible for node x_3 to transmit the control packet. However, a more critical situation would occur as node u may select node x_3 as the next hop and transmit data to it. Consequently, all data transmitted through node x_3 may be discarded, which hinders the communication between node u and node v . With the selfish behaviour of all node u neighbours, node u is unable to establish any communications with other nodes at a distance farther than a hop. In this case, a node is isolated by its selfish neighbours and partition in the network. To determine if a node is selfish, the classification range is set to energy threshold value ≤ 0.7 .

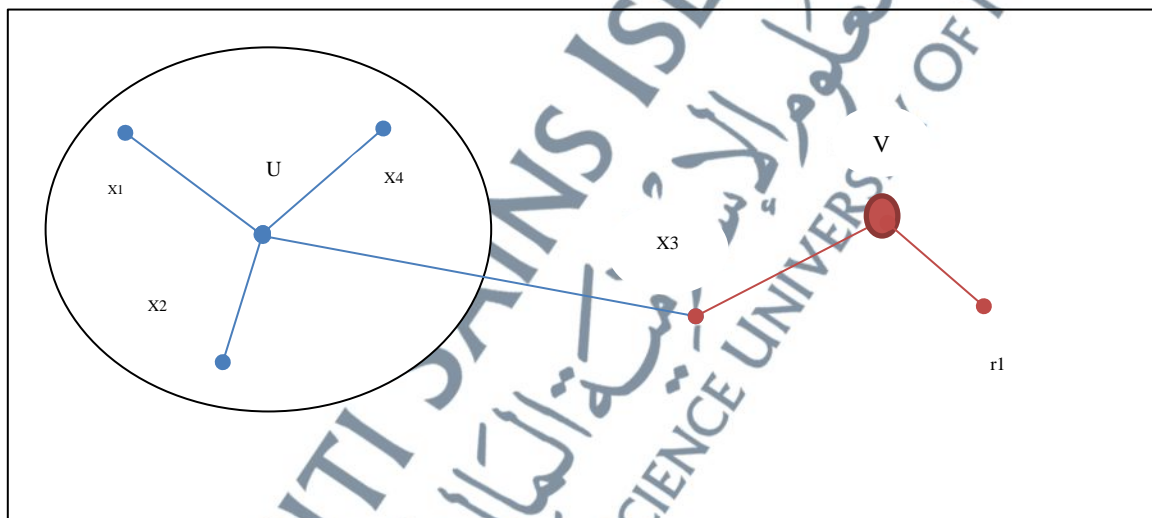


Figure 2.2: Selfish neighbours' nodes

c) Malicious node

In general, malicious node defined as node which active in both data discovery and launching attacks. Malicious event in somehow highly correlated in the anomaly event occurrence whereby intruder make the sensors to get the false reading (Rana, Jigish, Sangeeta

Vhatkar., 2015). In distributed network, the false reading yields to retransmission of packet data in alternate path hence energy is exhausted in large amount.

Based on Figure 2.3, nodes r_1 is a malicious node. Another effect of the malicious node can be illustrated through a scenario of one or more malicious nodes around the neighbourhood of node u . Provided if node r_1 is the next hope for node u , the communication between node u and the nodes u would be lost at least by two hops. With the selfish behaviour of the u neighbours, node u is isolated by malicious neighbours which have the same selfish behaviour as the u neighbours. To set the node behaviour as malicious, the value of packets delivery ration must be >1.0 due to excessive forwarding activity.

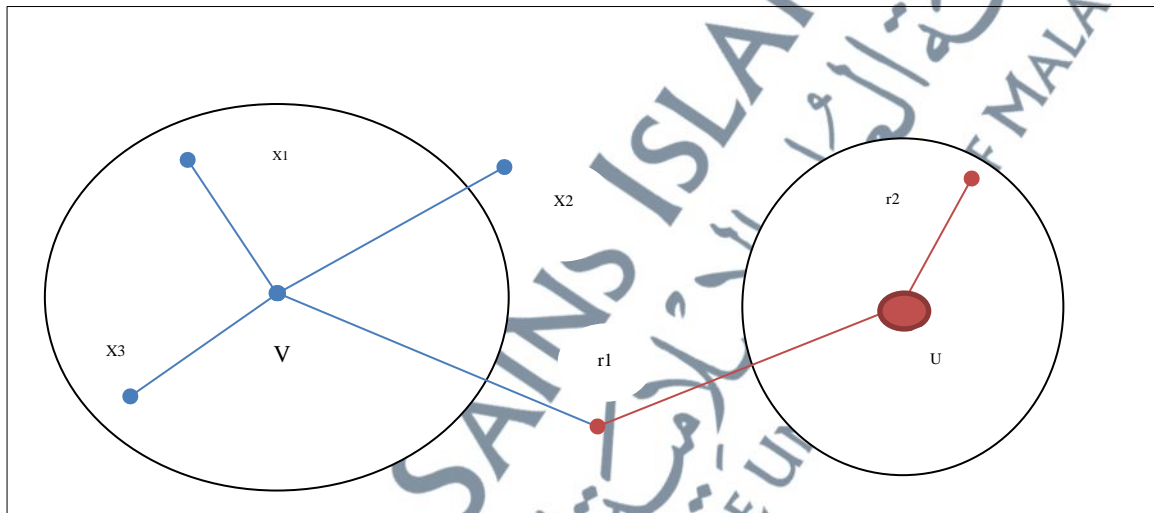


Figure 2.3: Malicious neighbours' nodes

d) Cooperative Node

The sensor nodes behaviour is usually randomly on the wireless sensor network, and then the action of the sensor nodes is formed to classify the behaviours nodes on the clustering algorithm on wireless sensor network on this configuration when nodes value from 0 to 0.69 are cooperative. Figure 2.4 illustrates the behaviour of cooperative node, when node v may become infected by its own behaviour such that at cooperative node k , node a is exposed to become failure node due to energy exhaustion, misconfiguration, and therefore it may change

cooperative node to selfish node, malicious node or failure node. Alternatively, a selfish node or cooperative node can become malicious node due to actuality bargained or failure node due to power depletion. A malicious node can become a failure node, but it will not be deemed to be cooperative node or selfish node somewhere if its disorder behaviours are intermittent only. That can a failure node become cooperative node again if it is recovered and responds to routing operations.

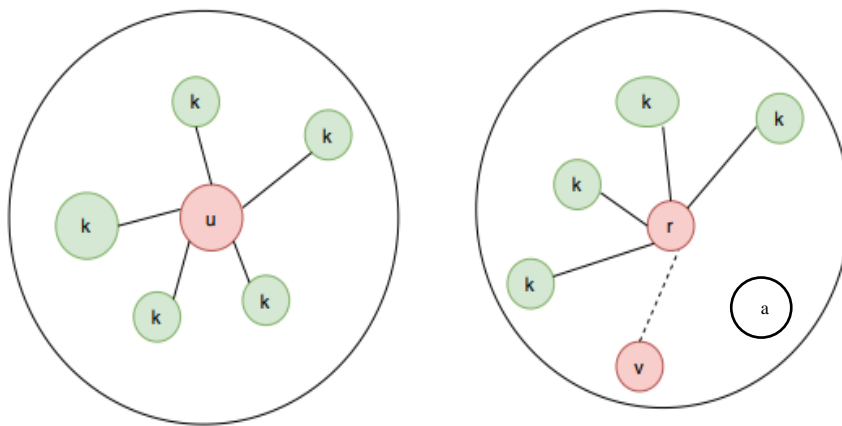


Figure 2.4 Cooperative node

2.3 Related Work

In recent years, wireless communication technology and the rapid development of miniaturization and low-cost sensor nodes have accelerated the development of the wireless sensor network. WSN comprises a large number of static sensor nodes, which form the wireless network through self-organization and multi-hop method. It aims to collaborate detection, process, and transmit the object for the monitoring of project in areas where the wired network is impossible. This literature review aims to present clustering algorithm method as the initial phase of a significant development of clustering approach in WSNs. Based on the selected

articles, ten methods of the algorithm are reviewed and analysed. The methods are listed in

Table 2.1.

Table 2.1: Summary of the literature work on related approaches and their parameters for the Clustering Algorithm Methods

Author Year	Methods	Parameters	Weakness	Strengths
J.S.Lee &W.L. Cheng, (2012) & B.A. Attea & E. A. Khali, (2011). & Muthulakshmi, M. V. (2013).	LEACH	number of node, data transmission, Initial energy, round (frame), packet size Transmission Range, threshold	1-Inapplicable to time-constrained applications. 2-Inapplicable to large scale due to single-hop communication.	1-Localise the coordination and control of cluster set-up and operation. 2-Randomize rotation of cluster heads
Abidoeye, a.P.,et al (2011),& Technology, C, et al., (2014).	TL-LEACH	number of node, data transmission, Initial energy, Round (frame), packet size, threshold	1- Uneven distribution of CHs may lead to hot-spot issues in the network. 2- Increases delay time	1-Resolves reallocation time slot problem among the candidates and current CH nodes. 2-Distributes the energy load among the sensors in the dense networks
Gupta, S. K., et al., (2015), Transactions, W., (2016).	R-HEED	Number of nodes, residual energy , Transmission Range, Energy consumed, Distance between the node and the Base Station(sink), Data packet size, Threshold	1- Not suitable for the entire needs of WSN. 2- Minimum energy consumption in the intra-cluster communication is not achieved through the cluster topology.	1- The time in seconds elapsed until a re-clustering for the network. 2- Achieve power-balancing through residual energy and node degree.
Kumar, V., et al (2011), by Liu, X.-X. (2012). Driss, E., Abdelilah, M. (2013).& Zhang, M. (2013),	EEC	Number of nodes, Initial energy , Transmission Range, Energy consumed, Distance between the node and the Base Station (sink), Threshold, Broadcast packet size	1- The selection cluster head process of cluster construction lead to additional overheads. 2- Direct long-range transmissions from CHs to the BS could lead to high energy consumption.	1- Based on energy and distance, EEC develops the balancing point between intra-cluster energy consumption and inter-cluster communication. 2- The application to various environments to achieve enhanced lifetime, coverage.
Farahmand, E., and Mahani, A. (2016), M. Mirsadeghi, et al (2014).	Load Balancing	Nature, Overload Rejection, Reliability, Adaptability, Resource Utilization, Process Migration, Throughput, Turnaround Time, Response Time	1-distribution of sensor nodes among the clusters is usually considered for cluster construction. 2-energy loss for the subset of CHs	1-Load balancing is an essential element aiming towards prolonged network lifetime in WSNs. 2- Increased in balancing the energy between nodes.
Khan, R., et al., (2011), &Ali Peiravi, et al 2013, Rana, J et al., (2015), Gupta, S. K., et al., (2015)	PEGASIS	sensor of node, Initial energy, base station, Distance between the node, residual energy of sensor node, Data packet size, Threshold node	1-Requires dynamical topology adjustment. 2-Requires multi-hop data transmission from the sensor to the sink, leading to a packet delay problem.	1-Avoids excessive clustering. 2-Removes the clustering overhead
R. K. Vishwkarma, et al., (2013), Wang, Y.,et al., (2013).	GROUP	Number of node Initial energy, location of the base station Distribution of Node, Energy consumption, packets data, transmitting and receiving	1-Transmission data which isn't dependent on the location of the sensing node. 2-Transmitted from the sink in the network it is away from CH	1-Efficient for large amount of data. 2- Strong to outliers.

Li, Q.-Q., et al. (2011). W. C. Y. Lee, 2012, Karimi, A., et al (2014), Afsharfarnia, A., & Karimi, A., (2014).	Weighted Clustering Algorithm (WCA)	Number of node, throughput, node degree, mobility, energy, transmission range,	1- The number of nodes that a cluster head can handle is ideally constrained. 2- Where the compute the degree difference for every node to ensure that cluster heads are not over-loaded.	1-This algorithm is fully distributed. 2-A the received signal strength from its neighbours of each node.
Virmani, D., et al., (2015),& Khoozani, P. A., (2013)	Fault-Tolerance	sensor node, Initial energy , distance, residual energy, routing overhead, transmission range	1- A sensor nodes affected by energy depletion, transmission errors. 2- The source node becomes a bottleneck if many replicators for data updating at the same time.	1- CHs is usually required in this kind of applications. 2- Effective fault-tolerant approaches must be implemented in WSNs
Gupta, S. et al., (2015).	Highest-Connectivity Cluster Algorithm	number of nodes, network size, case of static nodes, the average number of neighbors of a node, transmission range, degree of a node	1- A node which is unconnected to the neighbor nodes. 2- Among the non-covered neighbours, cluster-heads are selected as the nodes with the lowest ID number.	1- Based on the lowest ID number, the connectivity of this node is considered. 2- The node connected to the highest number of nodes is elected as the cluster head.
Baker, D., & Ephremides, (1981), Ephremides, Baker, & Wieselthier, (1987), Kumarawadu, Dechene, Luccini, & Sauer, (2008), Lakshmaiah, K., Krishna, S. M., & Reddy, (2020).	Linked Clustering Algorithm	The number of nodes, transmission range, the communication range, a seed number for generating random node positions, and the radius of the circle within which all nodes must lie, that is the maximum spatial size of the network.	1-not suitable for non-convex data, relatively sensitive to the outliers, easily drawn into local optimal, the number of clusters needed to be preset, and the clustering result sensitive to the number of clusters.	1-The Linked clustering algorithm helps to prefer which clustering mechanism should be utilized to obtain a coherent and a mixed data significant mechanism for specific character deployment. 2-The optimization criteria during cluster formation for assisting clustering process toward better and efficient character interpretable technique

2.3.1 Low-Energy Adaptive Clustering Hierarchical (LEACH)

Low Energy Adaptive Clustering Hierarchy (LEACH) protocol is one of the most popular hierarchical routing algorithms for WSNs energy efficiency. It is the basis for the development of the next algorithms after it. Further studies on LEACH algorithm were conducted by Lee et al., (2012), Khalil and Attea., (2011). The operation of LEACH protocol is divided into rounds, with each round comprising the set-up and steady-state phases. During the set-up phase, sensor nodes are organised into different clusters based on the received signal strength. This is followed by the selection of cluster heads for each cluster as the routers to the base station. In the cluster forming phase, information will be transferred from the cluster head

nodes to other nodes after they are selected. Meanwhile, the ordinary nodes can automatically choose the particular clusters to be linked with based on their distance from the head nodes. Overall, the election of cluster head depends on Equation 2.1 on the selected threshold $T(m)$, which is as follows:

$$T(m) = \begin{cases} \frac{P}{1 - P * [r \bmod (1/P)]}, & m \in F \\ 0, & \text{else} \end{cases} \quad (2.1)$$

P represents the percentage of cluster head nodes in all nodes, while m represents the token of the node. Meanwhile, r refers to the number of rounds for the election. $r \bmod (1/p)$ represents the number of nodes elected as cluster heads in a cycle, while f refers to the set of nodes opted out from the election. Each node is elected as the cluster head with the same probability. Threshold $T(m)$ of the node is elected as the cluster head in each circle. Setting the cluster head to 0 leads to an increase in the probability of the remaining nodes. This process could prove that the node with equal probability is elected as the cluster head.

Notably, the LEACH algorithm only emphasises on the probability for each node to become the cluster head. However, it has been proven in many experiments that the specific geographical location of nodes is omitted, which leads to heterogeneity of the election of cluster heads. Additionally, LEACH allows the development of single-hop cluster, contributing in a high number of clusters to be formed. The formation of this amount of clusters leads to significant energy consumption for communications.

Due to the significant development of Wireless Sensor Networks in the past few decades, the needs for scalable and energy-efficient routing and data aggregation protocol for large scale deployments have increased. However, the protocol's performance could be disrupted by the simultaneous presence of extremely large and small clusters in the network.

This situation leads to the decrease in WSNs lifespan. Nevertheless, LEACH is a hierarchical clustering protocol which offers an effective solution for this situation.

An improved FZ-LEACH is a work carried out to analyse new energy efficient in clustering protocol, which solves the aforementioned protocol issues through Far-Zone formation. Far-Zone is a group of sensor nodes which are placed in locations where their energy level is lower than the threshold levels. Furthermore, the communication between nodes and sink is based on energy consumption and the minimum distance between nodes. Notably, the communicating nodes remain in active mode, with the remaining nodes constantly being in the sleep mode. Muthulakshmi., (2013) highlighted that improved FZ-LEACH algorithm outperformed LEACH in terms of energy consumption and network lifetime. In addition, it also outperformed LEACH protocol in terms of energy dissipation rate and network lifetime (R. Kumar & Rajpoot, 2018). Specifically, it was found that FZ-LEACH protocol saved around 30% of sensor network energy compared to LEACH.

2.3.2 Two-Level Low-Energy Adaptive Clustering Hierarchy (TL-LEACH)

TL-LEACH is an extension of the work in the LEACH algorithm. In addition to other sensing nodes, it involves two levels of cluster-heads, namely primary and secondary. Communication takes place between the primary cluster head for each cluster of TL-LEACH and the nodes in their sub-cluster. The arrangement of communication between nodes within a cluster is still performed with Time-Division-Multiple-Access (TDMA). TDMA is adapted for the cluster head with nodes in a fixed time slot per a packet to each node. Therefore, it is decided whether the node is a primary or secondary cluster-head or not in each round. This method of determining the nodes is similar to LEACH, where the probability of a node to be classified as a primary CH is lower than the probability for the secondary node. Furthermore, data-fusion could be performed in LEACH. Specifically, secondary cluster heads collect the

data from the other nodes. Following that, fusion of data is performed before it is transmitted to the primary CH. Then, further collection of data is performed by the primary cluster heads prior to the fusion of the received data and its transmission to the base station. However, energy efficiency may not be present in the algorithm when CH is situated within a certain range from the sink. There is also a high probability for the overhead to increase during the selection of both cluster heads, resulting in an increased energy consumption (Abidoeye, Azeez, Adesina, & Agbele, 2011). Nonetheless, TL-LEACH algorithm reduces the amount of cluster head which needs to be transmitted to the sink, leading to effective reduction of total energy usage (A. Gupta & Nayyar, 2014), and (Manzoor, Jokhio, Khanzada, & Jokhio, 2019).

2.3.3 Rotated Hybrid Energy-Efficient Distributed Clustering (R-HEED)

The Hybrid, Energy-Efficient, and Distributed (HEED) is a multi-hop clustering algorithm for wireless sensor networks, with an aim to distribute energy consumption to prolong network lifetime, minimize energy during the cluster-head pick period, and minimize the control cluster-head of the network by S. K. Gupta et al., (2015). This clustering approach is considered one of the most effective methods to prolong the network lifetime in WSNs. Furthermore, it is perceived as one of the most energy-efficient clustering algorithms which elect the Cluster Heads (CHs) through intra-communication cost and residual energy. This study aims to examine different inter-cluster routing protocols over HEED and evaluate their performance. Selection method of CH in HEED was determined based on two important parameters, namely the residual energy of each node. Meanwhile, Intra-Cluster Communication Cost aims to determine the node to join the cluster. The setup phase was conducted with an enhanced version to HEED, namely Rotated HEED (RHEED), based on certain rules and schedule. Simultaneously, HEED conducted this step at the beginning of several rounds. At the beginning of every round, the CHs were held on for a pre-defined period

of time to receive a re-clustering message from the sink. If the re-clustering message was not received by the CHs, the rotation of the cluster head would continue within the same cluster. As a result, the RHEED outperformed the HEED protocol by more than 20% in terms of network lifetime and residual energy (Transactions, 2016), and (Ullah, 2020).

2.3.4 Energy Efficient Clustering (EEC)

Velmani and Kaarthick., (2015) presented a-clustering algorithm, where competition between the cluster candidates took place in terms of their ability to elevate to cluster-head in a given round. In this competition, the candidates broadcasted their residual energy to the neighbouring candidates. If the given node did not identify the node with more residual energy, it would be classified as CH.

This cluster is different from the LEACH cluster formation by X.-X. Liu., (2012), where clusters are formed based on the minimum distance of nodes to the corresponding CH. In this algorithm LEACH, EEC extends this dynamic by sizing the clusters based on their distance from the base station (Ye, Li, Chen, & Wu, 2005). Therefore, EEC algorithm resolves the problem of clusters placed within a further distance from the base station. These clusters require more energy for transmission compared to those within a nearer distance. Subsequently, the distribution network of energy throughout improved, resulting in an operation involving a better resource usage in a larger area network lifetime (Driss & Abdelilah, 2013). (M. Zhang, 2013) implemented a new energy efficient clustering approach (EEC) for single-hop wireless networks for periodical data gathering applications. EECS improves LEACH protocol through the dynamic size of clusters, which were based on the distance of the clusters from the sink. However, EECS was different from LEACH in the cluster head selection, where it takes place in a localized competition to generate CH. Besides, there is no iteration property which distinguishes it from HEED (D. Lin, Member, & Wang, 2019).

2.3.5 Load Balancing

Farahmand and Mahani., (2016) suggested that the distribution of sensors among the clusters is normally an objective for setups, where data processing or significant intra-cluster management duties are performed by the CHs. Furthermore, the CHs have a role of balancing the load among them in order to achieve their target performance. However, load balancing is a more pressing issue in WSNs, where CHs are selected from the available sensors ((O. Younis & Fahmy, 2004). In this case, setting equal-sized clusters is crucial to extend the network lifetime. This action could prevent rapid and premature energy loss for the subset of CHs which could render them dysfunctional. Besides, data delay could be leveraged through even distribution of sensors.

When data aggregation is carried out through CHs, a similar number of nodes in the clusters is crucial. With this similarity, the combined data report would be ready almost at the same time for further processing at the base station or the next tier in the network. A new methodology, LEGAC, was also introduced. It involves the use of two-stage GA to determine the optimal set of cluster heads and their members. Specifically, the first stage involves the selection of optimal cluster heads, while the optimal assignment of cluster members to the selected cluster heads is performed in the second stage. The GA methodology aims to minimise energy consumption, minimise the number of orphan nodes, and maximise coverage. As a result, network lifetime was increased in balancing the energy utilization between neighbour nodes. Using a benchmark case study, comparison was made between the performance of this technique and similar methods, namely UCFIA, GCA, SCP, and the method proposed by Mirsadeghi, Mahani, and Shojaee., (2014). It was proven that the LEGAC technique exhibited a better performance compared to other available techniques in terms of orphan nodes, network lifetime, and energy efficiency (Ennaciri, Erritali, & Bengourram, 2019).

2.3.6 Power-Efficient Gathering in Sensor Information Systems (PEGASIS)

PEGASIS is an enhanced form of the LEACH protocol. However, PEGASIS forms CHs instead of clusters. After the transmission between neighbours' nodes, only one node is selected for communication to occur between it and the sink. In PEGASIS, data aggregation is carried out at every node, with the nodes located at the two ends of the chain as the exceptions. Furthermore, the chain is built to connect the nodes altogether through the greedy algorithm. This is followed by transmission of data from one node to another, the aggregation of it at each node, and its transmission to the sink.

The transmission mechanism in PEGASIS is triggered by the use of a control token issued by the leader and is passed on along the chain. It implements a greedy method to form a cluster head (Khan, Madani, Hayat, & Khan, 2011). One of the sensor nodes will vary in each node round, followed by the transmission of sensor node data to their node neighbours through the CH toward the node. In data gathering, each sensor node fuses the data which has been sent to and received from the neighbour node, with the other nodes in the cluster head as the exceptions. In fact, node usually receives data from both its neighbours and the data transmitted to the sink. Therefore, sensor nodes are only connected to their neighbours in the CH, leading to reduced and balanced energy consumption.

Peiravi, A., Mashhadi, H. R. & Hamed Javadi (2013) evaluated the difference between LEACH and PEGASIS. Compared to LEACH, PEGASIS involves multi-hop routing through the formation of CH and selection of one node instead of multiple nodes for data transmission to the sink. Multi-hop transmission is adopted by CHs for data transmission to the sink node. All CH are chosen based on the weighed score for energy and distance. Furthermore, the performance of the cluster is assessed, followed by the acquirement of improvements in the lifetime, residual energy, and delay in WSNs. Therefore, compared to LEACH where each CH is involved in the node distance between from the sink, PEGASIS is an energy saving

algorithm. In addition, PEGASIS involves minimum energy consumption as only CH will be included in data aggregation and data fusion. Moreover, the distance of the node transmission is minimum in PEGASIS compared to the CH. However, the data received by the cluster head from the neighbours are not present in LEACH (Rana, Jigish, Sangeeta Vhatkar, 2015). The algorithm PEGASIS is in a data-gathering, which establishes a concept that energy conservation could prevent the direct formation of clusters from the nodes. Data fusion occurs at every node in the network for all relevant information which spreads throughout the network (S. K. Gupta et al., 2015, J. Wang et al., 2018). Although it may not be functional due to the nature of the application, the incorporation of BS inside the network increases network lifetime. Therefore, both cases will be discussed separately in this study.

2.3.7 A Grid-Clustering Routing Protocol (GROUP)

The Group algorithm, which is a grid-based clustering algorithm, was proposed by Rohit kumar ishwkarma and Arvind kumar (2013). It defines the base stations where one of the primary sink will be staged dynamically and randomly build the cluster-grid. The cluster-heads are arranged in a grid-like manner. The propagation of data query forwarding from the sink to source node takes place from the Grid Seed (GS) to its cluster-heads (CH). GS is a node located within a given radius from the primary sink. In the case of CH selection, the selection of GS by the primary sink is based on residual energy. As the GS is selected, the selection of CH takes place along the corners of the grid within R range. Each new CH will then select more CHs along the grid the selection of all CHs is complete.

Data transmission in GROUP depends on the type of data being collected. Meanwhile, location-unaware data query refers to data which is not dependent on the location of the sensing node. In this case, the query is transmitted from the central most sink in the network to its nearest CH. This is followed by a message broadcast by the CH to neighbouring CHs. If

awareness of location is present in the data, the requests would be transferred to the chain of CHs towards the specified region using unicast packets (Kaur, 2020). For both data queries, data is transmitted upstream through the chain of cluster-heads established during cluster formation. As a result, energy conservation is achieved through narrower transmission distance for upstream data, leading to prolonged network lifetime. Network grid is performed by Y. Wang et al. (2013), followed by the combination of the grids based on a user-defined value. Notably, three advantages of this protocol are that 1) sensors will be allocated by high density in the same cluster regardless the conditions, 2) the cluster head rotation mechanism is added, leading to equal allocation of workload to every node, and 3) the routing protocol could conserve more energy and prolong the network lifetime. Based on the comparisons made between GROUP and CH, the method proposed in this study conserved more power and prolonged network lifetime.

2.3.8 Weighted Clustering Algorithm (WCA)

The methods based on weighted clustering are demand-based non-periodic and cluster head selection procedure. It attempts to identify the most stable configuration in its first set-up. Whenever a sensor node loses its connectivity, a new configuration in the network would take place, leading to the formation of new topology. This algorithm involves several system parameters, such as the ideal node degree, transmission power, mobility, and the remaining energy of the nodes as the inputs. It also involves a selection of cluster head and topology. Any or all of the parameters could be used as such based on application requirement. It is fully distributed so that the entire node performs the same role when they function as the CH selection procedure displayed in Equation 2.2. The selection process involves a global parameter known as the combined weight.

$$w_v = w_1\Delta_v + w_2D_v + w_3M_v + w_4P_v \quad (2.2)$$

$w_1, w_2, w_3,$ and w_4 are the weighing factors for the corresponding system parameters.

In weighting factors, the nodes can be chosen based on the cluster-head. After the combination of it, weight is calculated by each node and transmitted across the entire neighbours. Therefore, the node with smallest w_v is chosen as the CH. The first essential, $w_1\Delta_v$, assists in efficient medium-access control protocol (MAC) functionality. However, this essential is constantly preferable for a cluster-head and the maximum number of nodes in a cluster. Meanwhile, the second essential, D_v , is the average range from the neighbours, which is strictly related to energy consumption. Besides, it is known that more power is required for a transmission through a long distance. The third essential takes place due to the mobility of the nodes. Due to the preference for a slow transfer of cluster-head, the stability of the cluster head architecture is maintained as long as possible as moving nodes improve the cluster-head. The last essential and energy at each node are directly related to the load carried by each node (Q.-Q. Li et al., 2011). Moreover, if a node was already a CH, the weighing factors ($w_1, w_2, w_3,$ and w_4) may be selected according to the system needs. For instance, due to high importance of power control is in Code-Division Multiple Access (CDMA) networks, the weight of the corresponding parameter may be increased (Lee et al., 2012). The flexibility of changing the weight factors assists in the application of this algorithm for different implementations.

Low stability of clusters may lead to rapid failure of clusters, high energy consumption for re-clustering, and decrease in the overall network stability in static ad hoc network. To improve the stability of clusters, weight-based clustering algorithms are utilised. However, these algorithms involve the use of limited features of the nodes. Therefore, the weight accuracy is decreased to determine node's competency, leading to incorrect selection of cluster heads. Furthermore, the new weight-based algorithm presented in their paper did not only

determine the node's weight based on its features, but it also emphasised the direct effects network of the adjacent node features. This algorithm also determined the weight of virtual links between the nodes and the effect of the weights on the identification of the node's final weight. Through this strategy, the heaviest weight was selected as the cluster head, which increased the accuracy of nodes selection (Arafat & Moh, 2019), and (Karimi, Afsharfarnia & Zarafshan, 2014). WCA provides one weight-based algorithm, where each node's weight is determined through the node and neighbour node information. In this case, the links' weight between nodes which provide connection between them is also determined. In this method, the ideal options could be classified as the cluster-head (Afsharfarnia & Karimi, 2014).

2.3.9 Fault-Tolerance

In general, clustering routing algorithms, may be compatible with the constraints and challenges of WSNs. In many applications, WSNs will be operational in harsh environments. Therefore, nodes are usually exposed to increased risk of malfunction and physical damage (Virmani, Kaur, & Jain., 2015). In these applications, tolerating the failure of CHs is normally necessary for the prevention from the loss of important sensors data. In this case, re-clustering the network is the most effective method of recovering from CH failure. However, it is not only a resource which burdens the nodes, but it is also highly disruptive to the on-going operation. Therefore, contemporary fault-tolerance techniques would be more appropriate in this case.

Assigning backups for CHs is the most popular scheme highlighted in the literature about CH failure recovery. The selection of a backup and the role of CH backups would vary in normal network operation. With a long radio range in CHs, neighbouring CHs could adapt the sensors to the failing cluster (M. Younis, Youssef, & Arisha., 2003) (S. K. Gupta, Ranjan, & Sharma., 2015). As these networks could be deployed in harsh environments, the possibility

for node failure and link failure is high (Jassbi & Moridi, 2019). Therefore, continuous function of the network is essential during the failure of node or link. Specifically, when node failure occurs, their roles would be played by others nodes, such as backup or other active nodes. Meanwhile, when link failure occurs, nodes must establish and maintain communications between one another (Khoozani., 2013).

2.3.10 Highest-Connectivity Cluster Algorithm

This algorithm is similar to Linked Cluster Algorithm (LCA). In this scheme, the number of node neighbours is broadcasted to the surrounding nodes. As a result, the connectivity of a node is taken into account rather than the ID number. The node with the highest connectivity to the highest number of nodes is selected as the cluster-head. However, in the case of a tie, the node with the lowest ID is elected as such (S. K. Gupta et al., 2015) (Amine, 2019).

2.3.11 Linked Clustering Algorithms

Linked Clustering Algorithm (LCA) is designed to find a number of clusters on network clustering to produce results for several sensor nodes Baker, D., & Ephremides, (1981), Ephremides, Baker, & Wieselthier, (1987), Kumarawadu, Dechene, Luccini, & Sauer, (2008), and Lakshmaiah, K., Krishna, S. M., & Reddy, (2020). The LCA assists to identify which clustering mechanism should be applied to achieve a coherent and transfer data significant method for a network connectivity. Moreover, the method is used to optimize criteria during cluster designing for helping the clustering process since some nodes may be specified for particular tasks or empowered with distinct capabilities approach by S. E. Khediri, A. Thaljaoui, (2018) and Yang, Shim, & Jang, (2017),

The number cluster is among the early ones on clustering of wireless networks Chu-hang, Xiao-li, Huang-shui, You-jia, & Mei-qin, (2020), and J. Wang, Gao, Liu, & Sangaiah,

(2019). The focus is mainly on forming an efficient network topology that can handle the mobility of nodes. By clustering, CHs are hoped to form a backbone network to which cluster members can connect while on the move. The Objective of the proposed distributed algorithm is to form clusters such that a CH is directly connected to all nodes in its cluster. LCA is thus geared for maximizing network connectivity.

The algorithm assumes synchronized nodes and time-based medium access. A node is assigned the slot in the frame that matches its ID. First, each node broadcasts its ID and listens to transmission of other nodes. In the next round, a node broadcast the set of neighbors that it heard from and thus every node will eventually know its 1-hop and 2-hop neighbors. A node a_i becomes a CH if it has the highest ID among its neighbors or does not have the highest ID in its 1-hop neighborhood, but there exists at least one neighboring node a_j such that a_i is the highest ID node in a_j 's 1-hop neighborhood. Since LCA is found to yield excessive number of clusters, the approach is refined in Z. Wang, Zheng, Li, & Wang, (2019). The idea is to pick a node a_i at random as the first CH and assign its neighbors to such first cluster. Then the node a_j with the lowest ID in the cluster is nominated as a CH. The neighbors of a_j that are not reachable to a_i would join the second cluster. The procedure is repeated for the third cluster and so on.

The linked cluster algorithm to optimally control the cluster size by balancing the interest in the spatial reuse of node, which is increased by having small clusters, and data delivery delay, which gets reduced by avoiding inter-cluster routing, large cluster sizes. Time-Division-Multiple-Access (TDMA) is used for intra-cluster communication. However, every cluster would use a distinct code resulting is simplified implementation and great potential for meeting the QoS requirements often found in the network connectivity for wireless sensor network. Hence, the correlated degree is to be estimated according to similarity measurement

of the node objects, and then 3D Euclidean distance will be designed. Next, the constraint-based number of sensor nodes need to be cluster with the different cluster heads. The LCA method estimates the correlated degree adjacent relationship among the nodes along with the relevant information of the network clustering. The linked clustering algorithm the nodes which are very closest to each other.

The LCA is a clustering algorithm that avoids communication collisions among nodes and uses TDMA frames for inter-node communication, with each frame having a slot for each node in the network for communication. Proposing cluster formation and CH election algorithms, many researchers focus on single-hop clustering and guarantees that no node will be more than one hop away from leader. In LCA, every node requires $2n$ time slots, where n is the number of nodes in the network, to have knowledge of all nodes in its neighborhood. If a node a_i has the highest identity among all nodes within one wireless hop of it or does not have the highest identity in its one hop neighborhood, but there exists at least one neighboring node y such that a_i is the highest identity node in a_{j_1} one hop neighborhood, it becomes a cluster-head. Basically, the LCA approach was designed to be used in the networks with less than 100 nodes. In such small networks, the delay between the node transmissions is minor and may be accepted. Each node has two ways of becoming a cluster head and each node is assigned a unique ID number. The one way is if the node has the highest ID number in the network including all neighbor nodes and the node itself. Other way, assuming none of its neighbors are cluster heads, then it became a cluster-head.

Hence, the cluster head is represented on behalf of the link clustering. CH is expressed based on the following information like the total of the sensor node. It contains correlated degree for the cluster with the number of nodes of the correlation relations in the network clustering algorithm.

2.4 Discussion

This section will discuss the issues in the research area that contributes to the WSN clustering algorithm. The research shows the lack of a clustering algorithm approach to avoid partitioning in WSN. Partitioning as mentioned in Chapter 1 will degrade the network performance. Furthermore; the method for clustering algorithm focus on individual nodes which causes the noncommutative measurement for connectivity and inaccurate network performance (Adamou, Ari, Omer, Labraoui, & Damakoa, 2016).

From this point of view, each one of these methods are only suitable for a specific application, rather than any arbitrary wireless network. In this research, the clustering algorithm is proposed to cater partitioning issues by taking factors and parameters defined in the selected papers. Finally, misbehaved node leads to network partitioning hence degrading the network's performance caused by the unavailability of the nodes. Thus, finding the best clustering algorithm to be implemented for nodes partitioning is the key to solve this problem.

In wireless sensor networks, the energy limitations of nodes play a crucial role in designing any clustering algorithm for implementation. In addition, Quality of Service parameters such as node degree, distance, transmission range, mobility energy consumption, bandwidth, packet loss, link lifetime, and failure rate will affect the performance of wireless sensor networks. Furthermore, wireless network performance will be degraded when a network is under attacked or there is a node or link failure. Fundamentally, network connectivity also determines QoS based on throughput, energy consumption, bandwidth, and packet loss, and network lifetime. These important parameters are often opposed, as one often has a negative impact on the other. Of the limitation enhance the existing clustering algorithm to improve network connectivity of correlated node behaviour. To optimal clustering in terms of energy efficiency should eliminate all overhead associated not only with the cluster-head selection process but also with node association to their respective cluster heads in (Kareliya, 2013).

Thus, this research proposed promising improvement over conventional enhancement clustering algorithm. Many energy improvements thus far have focused with minimization of energy associated in the cluster-head, in which will examine some future research problems and draw conclusions on the current state of wireless network clustering. The summary of the compared methods is listed in Table 2.2.

Table 2.2: Comparison of the clustering algorithm's method

Clustering algorithms	LEACH	TL-LEACH	EECs	L-B	F-T	WCA	R-HEED	PEGASIS	GROUP
Load balancing	medium	bad	medium	bad	good	moderate	medium	moderate	good
Energy efficiency	very low	low	moderate	good	good	very low	medium	low	good
Mobility	very low stationary	low stationary	stationary	stationary	stationary	stationary	stationary	Fixed BS	stationary
Distance	100	150	90	70	90	100	100	60	100
Energy Consumption	very low	low	moderate	low	low	moderate	moderate	low	very low
Number of node	50	100	100	40	10	50	80	10	10
Packet Forwarding	high	low	medium	low	high	high	medium	high	high
Delivery Delay	very small	small	good	bad	good	good	medium	very large	very large
Linked Clustering algorithm	moderate	good	stationary	moderate	good	good	stationary	good	good

2.4.1 Factors

Network connectivity is influenced by mobility prediction of the node, which estimates the future locations. The definition of location is based on the link of wireless networks. In infrastructure networks, location refers to the access points to which static terminal is connected. In infrastructure-less networks, it refers to geographical coordinates (Jain & Chand, 2016). Furthermore, its main advantage is its prediction of the link expiration time for node connectivity and routing performance improvement. Although various location prediction methods have been discussed by Kaaniche and Kamoun (2010), the prediction of future locations based on realistic scenarios is yet to be studied. Nevertheless, realistic prediction of

future locations is one of the crucial challenges in network connectivity. To illustrate, the mobility of the nodes could affect the outcome of the wireless network. In the group mobility wireless network, when the nodes move from cluster head groups, the mobility of the nodes in the same cluster head group depends on the mobility of the cluster head groups.

Static node may be the most ideal solution in connecting two clusters in a highly dynamic environment. To deploy static nodes in place of critical nodes, the cost of static node deployment must be considered. When the nodes are placed, maximised network connectivity and minimum cost must be prioritised. Meanwhile, achieving connectivity efficiency and cost objectivity is an issue in deploying the critical points is an NP-complete. Although this initiative was performed by Romoozi and Babaei (2011), no investigation on critical links and nodes has considered the factors of cost objectivity and connectivity efficiency. As these factors are the issues in NP-complete, more optimised technique is required to achieve this goal. Due to long distance data transmission, the nodes could consume more energy to neighbours. Besides, the battery resources of sensor nodes are limited in WSNs. Therefore, an extended network lifetime should be achieved by the sensor nodes via balanced energy consumption.

Energy consumption is the key challenge in energy conservation for a prolonged network lifetime due to the network connectivity constraints. Provided that energy is the most valuable resource in network connectivity, its condition should constantly be monitored after network deployment (N. Kumar, Kumar, & Patel, 2010). The random waypoint mobility model was found to be an unrealistic model based on previous studies. It was assumed that the required energy for data transmission of the node is proportional to the square of transmission distance. Therefore, the best clustering algorithm for minimum energy consumption for data transmission should minimise the total square of data transmission distance in a network (X. Liu & Wu, 2019).

2.4.2 Parameters

Several methods highlighted in primary studies defined and quantified the parameters to evaluate the clustering algorithms for an improved network connectivity. To design a wireless network, identifying the key parameters in WSNs is essential, such as network lifetime, data gathering, and security. Furthermore, the parameters affecting the performance of wireless networks include node degree, distance, transmission range, mobility energy consumption, bandwidth, packet loss, link lifetime, and failure rate. Moreover, wireless network performance will degrade when the network is under attacked or node or link failure occurs. Essentially, network connectivity also determines QoS based on throughput, energy consumption, bandwidth, and packet loss.

To achieve network connectivity, a wireless multi-hop path is essential from each static node to other static node. Therefore, network connectivity depends on the number of nodes on the area of the node density and the transmission range area. Furthermore, each static node contributes to the connectivity of the entire network. The bandwidth (throughput) measures the speed of data transmission when the data starts to flow. To illustrate, it represents the amount of data sent over a particular duration or the duration required to transmit a certain amount of data from a static node to another static node. Meanwhile, in wireless network, packet loss refers to the failure of packets during data transfer to each node due to weak signal strength, interference by nature or human, noise, hardware, and software failure. It is an prominent performance issue which would affect the wireless network (Gebreyohannes, 2014). Studies were conducted on the security aspect of sensor wireless network as network connectivity parameters, such as confidentiality, integrity, and authenticity. These parameters were used in qualitative and usually immeasurable, as shown in Table 2.3.

Table 2.3: Index and parameters

Index	Parameters
Network Connectivity	Degree of a node Distance radius Transmission range Portability
Network Performance	Throughput Energy consumption Bandwidth Packet loss
Fault-tolerance	Link lifetime Failure rate Transmission range
Security	Confidentiality Integrity Authenticity

2.4.3 Limitation

This section will discuss the limitations in this study which led to the WSN clustering algorithm. It could be seen from the literature review that there was insufficient clustering algorithm approach used to prevent partitioning in WSN. The justify network clustering into non equal-sized partitions and each partition into some unequal clusters. As mentioned earlier in the previous unequal approaches, partition energy balanced and efficient clustering have smaller clusters that are closer to the sink, and larger clusters in the farther areas from the sink. The Cluster Heads selection in a partition of network clustering is performed by a clustering algorithm which includes the node degree and location of the nodes as its main criteria for Cluster Heads selection (Jiang et al., 2019).

As mentioned in Section 2.1, essentially, connectivity would not be successful when partitioning occurs. As a result, poor network performance is created. The method for

clustering algorithm emphasises on individual nodes, which led to noncommutative connectivity and inaccurate performance (Kosunalp, 2015). Furthermore, rather than any arbitrary wireless network, each of these methods was only suitable for a specific application. In this research, clustering algorithm was proposed to cater partitioning issues by taking the factors and parameters defined in the selected articles into account. Moreover, misbehaved node led to network partitioning, therefore degrading the network's performance caused by the absence of the nodes. Therefore, identifying the best clustering algorithm to be implemented for nodes partitioning is the solution to this issue.

In wireless sensor networks, the energy limitation of nodes plays a crucial role in the design of any clustering algorithm for implementation. In addition, Quality of Service parameters, such as node degree, distance, transmission range, mobility energy consumption, bandwidth, packet loss, link lifetime, and failure rate would affect the performance of wireless sensor networks. Fundamentally, network connectivity also determines QoS based on throughput, energy consumption, bandwidth, packet loss, and network lifetime. However, these parameters were often opposed due to their negative impacts on the other.

The limitations in the improvement in network connectivity, node may change its behaviour from normal node to misbehaved node leading to correlated node behaviour are also highlighted. To achieve an optimal clustering in terms of energy efficiency, all overheads should be eliminated. They were not only associated with the cluster-head selection process, but also with node association to their respective cluster heads (Kareliya, 2013).

This research proposed promising improvement network performance in enhancement clustering algorithm. Various improvement methods for energy were highlighted, with minimisation of energy associated in the cluster-head. Accordingly, several future research issues will be discussed, followed by conclusions made on the current state of wireless network clustering.

2.5 Summary

Literature review was conducted in this study to gain an understanding of the issues related to clustering algorithm for an improved network connectivity in the WSNs of correlated node behaviour. This review result highlighted the parameters for performance measurement and the factors of network critical condition. It was found that network connectivity practices were still focused on theoretical works, with significant number of studies ensuring that connectivity of nodes maintain network performance. However, the studies conducted on enhanced network connectivity were scarce, despite its notable impact on the clustering algorithm in WSNs. Moreover, the constant development of wireless technology keeps growing lead to more challenges to be addressed. Therefore, the key points in resolving this issue are the reduction of network connectivity effect, advanced method to improve network connectivity, and the reliable transmission of field information.