

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This chapter presents the methodology employed in the development of clustering algorithm for the correlated node behaviour. The algorithm mainly focuses on energy consumption for network connectivity. This is followed by discussion on the theoretical enhancement clustering algorithm in correlated node behaviour.

3.2 Research Process

The methodology comprised three stages, each fulfilling the research objectives in Chapter 1. The research procedures are illustrated in Figure 3.1.

3.2.1 Stage 1: Investigating on the Parameters of Correlated Nodes.

The identification of the key parameters in WSNs, such as energy consumption, distance, packet and security were important. This literature review aims to identify the parameters which could define network connectivity as the crucial method for clustering algorithm development to improve network connectivity in WSNs. Accordingly, the aforementioned parameters were identified through self-organization method. In this stage, the parameters of the clustering algorithm of correlated nodes were also identified. The contribution of this study is to find parameters that effect network connectivity, which could be implemented to improve network performance in wireless sensor network. Based on the

comprehensive review on clustering, the strengths and the weaknesses of each algorithm were determined.

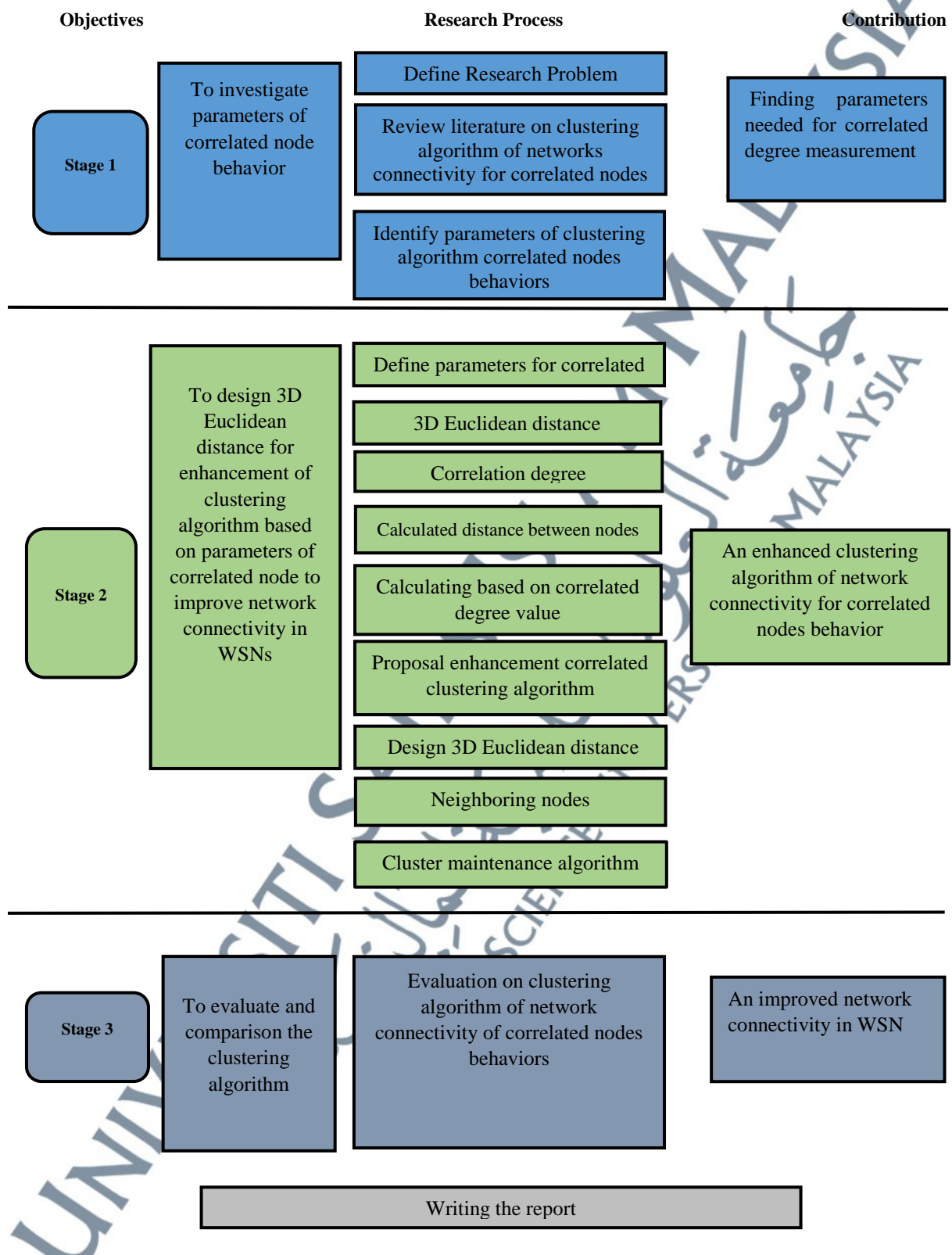


Figure 3.1: Research Process

3.2.2 Stage 2: The design 3D Euclidean Distance for Enhancement in Clustering Algorithm

In real network scenarios, misbehaving nodes will impact other nodes behaviour and lead to node failures. To illustrate, with constant failure of node neighbours, more traffic needs to be loaded to transmit the energy from the failure neighbours. Consequently, failure of this node may occur at a faster rate due to excessive energy consumption. Similarly, it is possible that when the malice of the node neighbour is higher, the possibility for the node to be compromised by its malicious neighbours is higher. Additionally, upon node failure, the network suffers from degraded performance due to the absence of the failure nodes. The subsequent impacts from this issue ranged from insignificant topological survivability to harmful network shutdown.

The establishment of network connectivity in WSN applications required a degree of dense deployment of sensor nodes, which led to a single event being recorded by several nodes. The sensor observation of these leads, which had a correlation degree or correlation nodes often identified similar values within a close proximity. Assuming that cluster head (CHs) were selected using three parameters, namely distance (D), packet (β), and energy (ℓ), all sensor nodes in 3D Euclidean distance were made up of $a_1, a_2, a_3, \dots, a_{100}$ and cluster head as anchor nodes. Furthermore, distance was labelled as (D), packet transfer as (β), and energy as (ℓ). Therefore, \mathcal{T}_{a_i} represented the threshold value which would either become the cluster head or vice versa. If node a_i has the right to become a Cluster head, a_j must fulfil the condition of $a_1 \geq 1.0$, then node a_i is entitled to become a Cluster head. However, when $a_j \geq 1.0$, node a_j is not entitled to become CH. The calculation of the cluster heads based on the 3D Euclidean distance took place after the establishment of the selected CHs using three criteria, namely distance (D), packet (β), and energy (ℓ). These criteria were essential for $(x_{a_i} - x_f)^2 + (y_{a_i} - y_f)^2 + (z_{a_i} - z_f)^2$ between one-hop

neighbor nodes by the 3D Euclidean Distance method. Meanwhile, x, y, z referred to the coordinates of the nodes in 3D Euclidean distance for node $a_1, a_2, a_3, \dots, a_{100}$.

Based on the aforementioned parameters, this study aims to formulate 3D Euclidean distances for the correlated degree of network clustering.

3.2.2.1 Parameters Used for Correlated Degree

There are three parameters used for correlated degree which are packet delivery ratio, distance and energy consumption.

a) Packet delivery ratio

Packet delivery is defined as the ratio of total number of packets, which have reached the destination node. It also defines the total number of packets originated at the source node and the effects of the number of nodes (Bhadra, Soni, & Vyas, 2015). In the Packet Delivery Ratio (PDR) of the wireless network, the packet relay nodes were involved to mitigate the network geometric deficiencies. This was followed by the use of the enhancement clustering algorithm to locate the optimal sink location with respect to the sensor nodes to solve the issues of network lifetime. In this case, appropriate input parameters from routing packets were essential to determine its behaviour node and network resilience. When the link was initiated by the node, a message broadcast would take place periodically to update link information to the node neighbours. Moreover, the message carried routing information regarding the number of packets, number of bytes sent and received, and its residual energy. Notably, every node possessed the same initial energy, which might put the packet forwarding functionality into a halt when the residual energy reverted to the initial energy. Packet delivery ratio can be measured as the ratio of number of packets delivered in total to the total number of packets sent from the sensor node to the neighbour node in the cluster head. It is desired that the maximum number of data packets has to be reached to the destination. PDR is the rate between the number

of packets received and the number of packets sent across the network as shown in. Equation.

(3.1)

$$PDR = \frac{\sum \text{total of received packet}}{\sum \text{total of send packet}} \quad (3.1)$$

Clearly when a total of receive drop packet more than two in all network nodes because it does not have a congestion control mechanism. Whilst, the PDR of the cluster head is between node independent with consideration of network connectivity value of sensor nodes. The amount of packet transmitted from sensor nodes to destination nodes, through the network at a specified. The packet delivery ratio changes due to varying of the network connectivity of both cooperating and misbehaviour nodes. Packet delivery ratio of cooperating nodes is greater than that of misbehaviour nodes. When there are value between range 0 to 0.69 as cooperative nodes in the network, the packet delivery ratio for all nodes is greater than > 1 based on a correlated degree is greater than threshold because a portion of the communication happens between nodes that are within each other's radio range. Usually, it is rated in terms of packets per message or bits per message.

b) Distance

In this research context, distance is referred to the proximity between two nodes without the needs for special-purpose chips or a redesign of the existing nodes. The measurement of distance is known as 'ranging'. To refine the node measurements, a radio runtime measurement was used to calculate the distance between nodes and present algorithms (Veremyev & Pasiliao, 2015). Most localization systems for WSNs implemented distance estimation techniques for the calculation of nodes position. Moreover, the distances could be collected in several ways, such as the estimation by received signal strength indicator (RSSI) or radio runtime measurement. As the RSSI value in indoor deployments was influenced by various

parameters, highly imprecise distance estimation based on RSSI value was obtained. An anchor node is a normal WSN node in the network with an a priori known position. To localise the WSN node in a network, identifying the distance of the WSN node to several anchor nodes was essential.

The objective function for distance is given by Equation 3.2 where x_i is distance between node x_j which belongs to cluster y_i and y_j is the cluster head node for network cluster.

$$d = \sqrt{(x_i - x_j)^2 (y_i - y_j)^2 (z_i - z_j)^2} \quad (3.2)$$

The distance between nodes and sink considered with energy threshold based on cluster head selection in that improves network connectivity. However, the node value is the cooperative node or misbehaviour node based on network connectivity then assign the node to cluster and check this process using the enhancement clustering algorithm using 3D Euclidean distance based on correlated degree.

c) Energy Consumption

Energy is the most important solution for the node network connectivity. The energy consumption in network connectivity was identified upon node transmission and packet collection. This subject was presented by Elrahim, Elsayed, Ramly and Ibrahim (2010), which involved the calculation of the energy consumption of a node. Specifically, it took place when the node transferred or received the packet networks and the node interface. In this section, it was assumed that the changing behaviour of the static node in the model occurred as per the definition. To illustrate, a cooperative node might turn into a selfish node when the node energy level was reduced below the minimum residual energy threshold (N. Kumar et al., 2010). Besides, a selfish node might be rehabilitated into a cooperative node by charging its energy. Provided if the packet delivery rate exceeded the threshold, a cooperative node or a selfish

node might be reduced into a failure node when it was placed outside of the radius of communication (R_c) (S. K. Gupta et al., 2015). This transition would also occur when the energy level of node was at its minimum (Aslam et al., 2014). However, a failure node might be rehabilitated into cooperative when the node was within the radius of communication R_c and responded to routing operations.

3.2.2.2 Enhancement of Clustering Algorithm

Clustering is the most compatible method with medium to large-scale WSNs. It is also a useful topology-management approach used to reduce communication overheads and exploit data aggregation in sensor networks (Rana, Jigish, Sangeeta Vhatkar, 2015). To maintain the connectivity of the sensor network, specific parameters were formulated in the proposed clustering algorithm for the improvement of this connectivity under correlated node behaviour. The parameters included packet, energy, and distance. Justification of choosing these parameters to reduce energy consumption to prolong node lifetime in the network clustering. There were three steps proposed for this algorithm, as seen in the Enhancement Clustering Algorithm in Figure 3.2.

Stage 1: this is an initial setup of the network value for all node (Tam, N. T., & Hai, 2018). Then, setup of the parameters is to implement to the algorithm than calculate CH based on 3D Euclidean Distance.

Stage 2: cluster formation: compute each degree in the neighbouring nodes cluster. Then, calculate neighbouring node in cluster. The degree is calculated based on 3D Euclidean distance where cluster head will be selected if the value of ζ for correlated degree is greater than the energy threshold. The energy threshold can be selected randomly ≥ 1.0 . If the number is less than a threshold $T(a)$, then it is selected as a CH for that threshold value varies in each round randomly based on the 3D Euclidean Distance. Then, every sensor the energy threshold

value $T(a) \geq 1.0$, and compares the correlated degree of energy thresholds $T(a)$. If the correlated degree is greater than energy threshold value $T(a)$ the sensor node will periodically broadcast message to its neighboring nodes to inform that it will be the cluster head.

Stage 3: Clustering maintenance used is to transform data after clustering formation is presented. If several nodes in the cluster have the highest degree of correlated degree, the node must be added to the start cluster head for all its neighbours to see if any of them contains the required transmission data and if not, the search is forwarded to the neighbour with the highest degree. The clustering algorithm is a correlated degree based on 3D Euclidean distance that can identify clusters of arbitrary shape. The clustering maintenances are recycled by every 5–10 seconds.

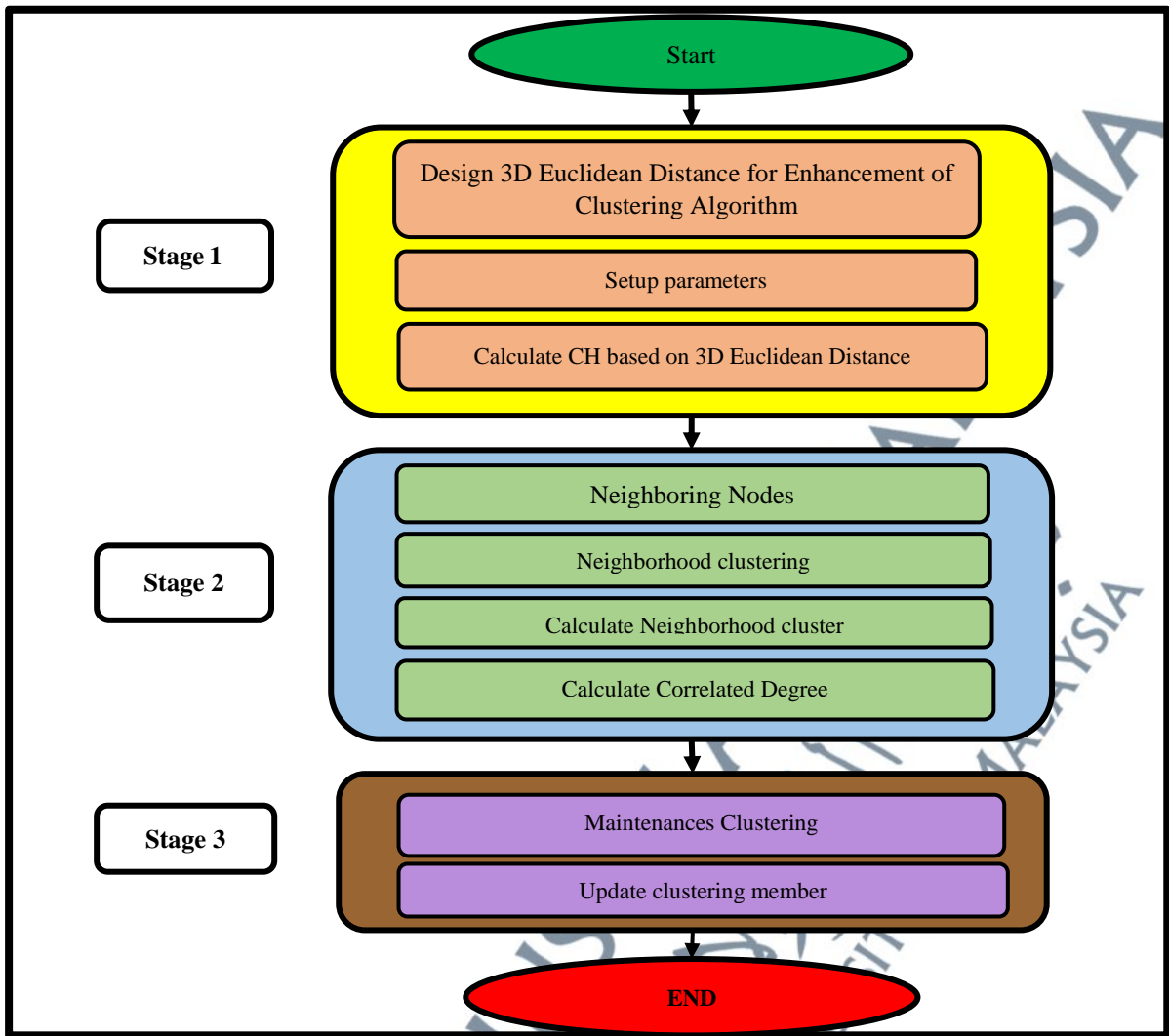


Figure 3.2: Enhancement Clustering Algorithm

a) Original Clustering Algorithm

Figure 3.3 shows the clustering algorithm proposed by Yang et al. (2017) in wireless sensor network where the sensor nodes consume energy efficiently by communicating directly with the nodes without a network infrastructure. However, this algorithm required regular re-establishment of clusters, which leads to unnecessary energy consumption. Moreover, there is a large energy consumption because a cluster head that is placed far apart from a sink node directly transmits data to the sink nodes. The clustering algorithm, which has been used frequently thus far, resolves energy imbalance between nodes and increases the entire network's lifetime through effective cluster establishment and efficient merging of data.

However, because the clustering algorithm selects cluster heads probabilistically and re-establishes clusters regularly, unnecessary energy is consumed.

Therefore, to resolve this issue, an enhancement of clustering algorithm for Efficient Energy Management in WSN proposed by Yang et al., (2017) is proposed to resolve the problem of unnecessary energy consumption in the previous clustering algorithm. Unnecessary energy consumption was reduced by re-establishing a cluster only if the remaining energy is below the set threshold while omitting re-establishment if it is above the threshold. Moreover, to resolve a single-hop transmission method's problems, a multi-hop method was used, which allows for efficient energy management. Based on the multi-hop method, that original a clustering algorithm can maximize the power efficiency of data transmission by selecting the minimal and optimal route based on the threshold.

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

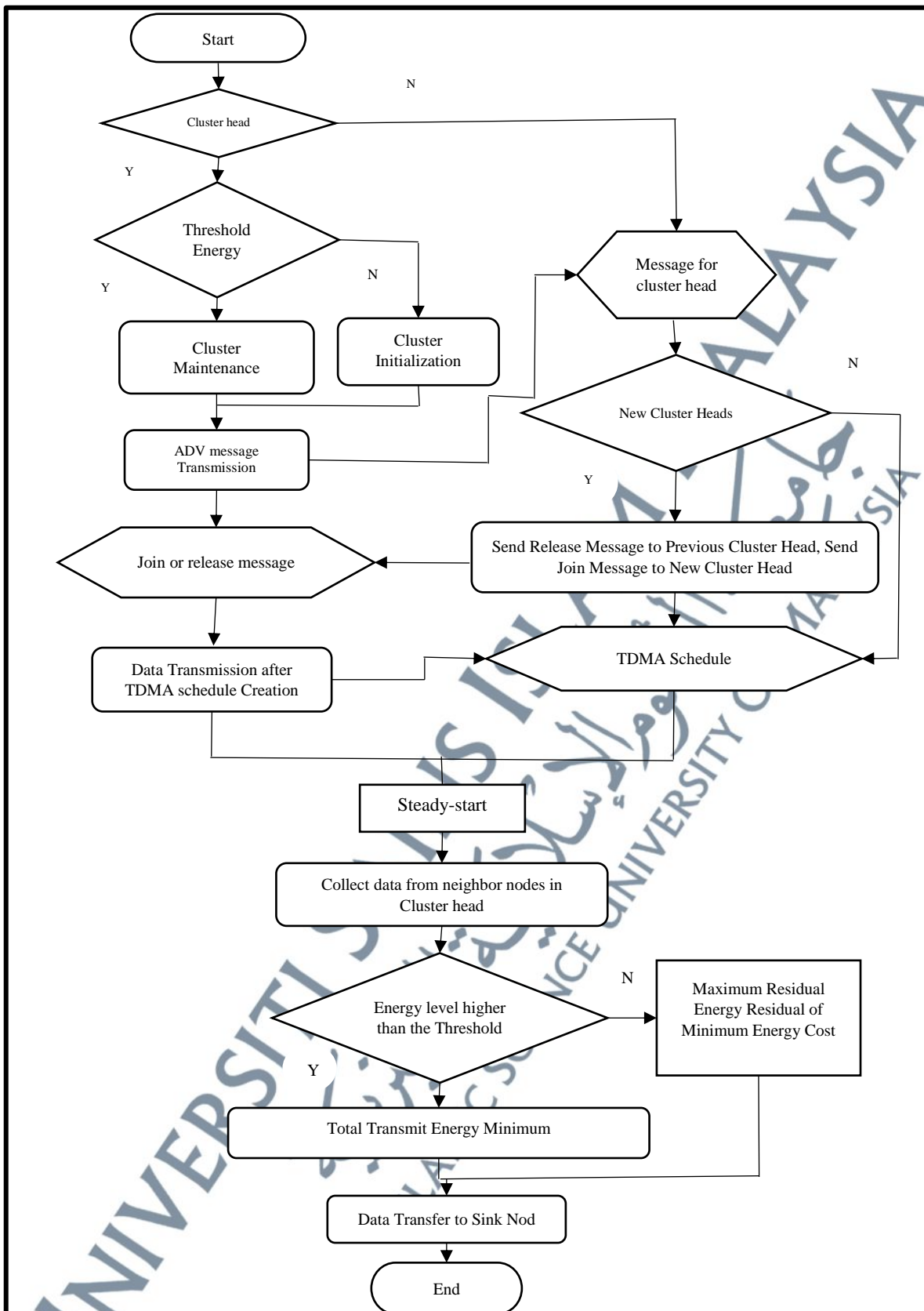


Figure 3.3: The Original Clustering Algorithm

3.2.3 Stage 3: To evaluate and compare the Clustering Algorithm

An experiment was carried out to test the proposed algorithm using object-oriented programming languages, namely NS-2 and C++. Specifically, NS-2 is a discrete event and object-oriented simulator, which is written in C++ and OTcl (an object-oriented version of Tcl). Meanwhile, OTcl was used as a script language to develop the simulation. To meet the aim and objectives of the research, the impacts of network connectivity based on distance, energy, and packet and the node behaviour factors of correlated node behaviour in this network were evaluated. In this phase, the performance of the suggested algorithm was evaluated. The testing of the enhancement clustering algorithm was based on the quality of the output, network connectivity, distance, energy, and packet of the simulation. Accordingly, the enhancement clustering algorithm was highlighted when the network connectivity in WSNs was improved through output simulation for network performance and energy consumption.

3.2.3.1 Experiment Setup

Experimental-based simulation of enhancement clustering algorithm was used, including NS-2 and C++ programming. NS-2 is the most popular and widely used general-purpose network simulator. It supports simulations for various kinds of wired and wireless networks, such as an IP-based network protocol and an 802.11 standard based wireless network. It is also a discrete event and object-oriented simulator, written in C++ and OTcl (an object-oriented version of Tcl). Meanwhile, OTcl was used as a script language to develop the simulation. Specifically, the trace file used for data storage was generated during the simulation. Subsequently, the results in the trace file could be graphically observed and analyzed through the Network Animator (NAM).

Due to NS-2's excellent extensibility, it was maintained by a large number of researchers, with many third-party add-ons being adopted to solve new problems. However,

NS-2 did not possess appropriate application models, packet formats, sensing models, protocols, and a configurable GUI. Besides, requirement of expert skills is especially unsuitable for beginners. Also, the exponential time and execution overhead resulted in a limited NS-2 performance in terms of scalability for large sensor networks. However, the energy consumption model in NS-2 was simple. This could be seen from the transfer of an initial amount of energy to each node. Following that, corresponding energy was subtracted from the total energy after packet transmission and reception until the value was reduced to zero. Two important and frequently mentioned NS-2 were based on WSN simulation environments.

3.2.3.2 Data Collection

The data collection area was randomly divided into nodes. Misbehaviour node was used to transmit data from the sensor nodes to the cluster head. Because, the sensor nodes behaviour is the main issue in WSNs, especially network connectivity due to dynamic nature of the network. The changes of nodes behaviours may occur from cooperative to misbehaving, which directly affects the network's connectivity. Misbehaving nodes consists of failure nodes, malicious nodes, and selfish nodes. As the data was transmitted to the selected misbehaviour node, it was recorded according to the configuration table and the correlated degree values. Specific parameters including packet, distance, and energy were provided. As supported by the clustering algorithm, the received data could be seen in failure node value, selfish node value, malicious node value, current value, and correlated degree when the.

Other specific functions could also be efficiently extended to enhancement clustering algorithm based on user requirements. Furthermore, all the displayed correlated degree values were printed out into a trace file used for data storage. This trace file was generated in the simulation, and its results could be graphically observed. The convenience of subsequent data processing and evaluation was also present. Therefore, the result of the correlated degree in the

clustering algorithm would contribute to the connectivity of the neighbouring nodes. Overall, this research might improve network connectivity, network performance, and energy consumption in wireless sensor networks.

3.2.3.3 Performance Evaluation and Comparison

The overall performance evaluation of the enhancement clustering algorithm scenario was conducted on the testing. In this case, the enhancement clustering algorithm was mentioned upon when the output simulation process for network performance was conducted. In respect of energy consumption, numerous objectives needed to be considered for the improvement in the network connectivity in WSN.

a) Network Connectivity

The sensor nodes u were uniformly distributed within a 3D square on a rectangular field $(a_1, a_2, a_3, \dots, a_n)$. The base station was positioned at the center of the square region (Somauroo & Bassoo, 2019). All nodes were deployed randomly, with the number of sensor nodes a to be deployed being based on the simulation. Furthermore, each sensor could identify the environment in the 3D sphere of radius r . All sensor nodes in 3D Euclidean distance consisted of $(a_1, a_2, a_3, \dots, a_n)$ and cluster head (CH) due to the homogeneity of the anchor nodes and sensors, where both had the same capacities. Distance was denoted as (D) , with packet transfer being labelled as (β) and energy as (ℓ) . Moreover, all the sensor nodes comprised a particular identifier allocated to them. Each cluster head coordinates the MAC and routing of packets were placed with their clusters. The measurement of the correlated degree from CH component with a set of sensor node parameters, namely distance, packet, and energy reflected the corresponding value of correlated degree based on distance D , packet β , and energy ℓ .

Besides, $(D_{a_i} - D_{a_{i,f}})^2 + (\beta_{a_i} - \beta_{a_{i,f}})^2 + (\ell_{a_i} - \ell_{a_{i,f}})^2$ equation was needed in one-hop neighbor nodes by the 3D Euclidean Distance method. Meanwhile, $(a_1, a_2, a_3, \dots, a_n)$ represented the coordinates of the nodes in 3D Euclidean distance ji and $(a_1, a_2, a_3, \dots, a_n)$. Based on Equation 3.3, (a) denoted the number of nodes, while D represented the distance or communication radius.

$$\zeta_{a_{i,j},f} = \frac{\sum (D_{a_i} - D_{a_{i,f}})^2 (\beta_{a_i} - \beta_{a_{i,f}})^2 (\ell_{a_i} - \ell_{a_{i,f}})^2}{\sqrt{\sum (D_{a_i} - D_{a_{i,f}})^2 (\beta_{a_i} - \beta_{a_{i,f}})^2 (\ell_{a_i} - \ell_{a_{i,f}})^2}} \quad (3.3)$$

b) Energy Consumption

Provided the diverse distances of information transmission between nodes to cluster head, a free 3D Euclidean distance and the multiple path fading cluster were used (Abdullah, Saudi, & Marsal, 2017). Energy consumption in a WSN has direct implications on the lifetime of the network and thus poor energy management techniques can result in the network exhausting their energy resource quickly.

Furthermore, aggressive energy consumption can yield poor performance by nodes oversleeping. The main cause of energy loss by a node is due to sending and receiving messages. Loss of energy due to amount of messages being sent can be simulated using a formula, such as the one in Equation 3.4 which has been derived from the nodes distance and the loss of packet delivery per message being sent. Since each of the sensors in the network cluster is energy-constrained and each component in a sensor node consumes a certain amount of energy, power supply becomes important to ensure proper operation of the entire system as 10 the number of sensors deployed in a network clustering.

$$\ell = \ell - \frac{D * \beta pm}{10} \quad (3.4)$$

where ℓ = energy,

D = distance,

βpm = packet per message,

10 = the number of sensors nodes

Energy loss is greater when sending a message, however just receiving a message causes energy to be used. Equation 3.5 demonstrates a possible formula to use in the loss of energy due to a message being received. In the equation, only a small fraction of the amount of the packet delivery per message is being lost.

Other forms of energy loss come in the form of computation, and node overhead. To account for this type of loss, a new equation is used. Equation 3.6 demonstrates how much energy is being lost due to overhead per time frame.

$$\ell = \ell - \frac{\beta pm}{10} \quad (3.5)$$

where ℓ = energy,

βpm = packet per message

$$\ell = \ell + \ell ps \quad (\text{Eq3.6})$$

where ℓ = energy,

ℓps = energy per step,

βpm = packet per message.

Whenever the energy level of a node drops beyond a certain threshold, the node will enter a sleep state. The node will remain in the sleep state till it regains enough energy to become functional again. Energy in a node is produced primarily through the use of the

connectivity network. The amount of energy that is being produced is limited and is dependent on the node's environment. By using an average node energy production as formulated in Equation 3.7, one can estimate the amount of energy produced at every time frame.

$$e = e + eP \quad (3.7)$$

where e = energy,

eP = energy produced.

The node will lose energy from messages being sent and received as well as from node overhead. While the amount of energy that the node is producing might not be enough when their initial energy level was below 0.5, unless the energy level was above 0.5 to permit the node to run constantly appropriate energy management done allow the node to prolong its life allowing the network to continue functioning.

3.4 Writing Report

The enhancement clustering algorithm for network connectivity in wireless sensor networks was used for the correlated node behaviour framework in wireless sensor network. This was followed by the development of a correlated node behaviour model which could dynamically affect network connectivity. In Chapter 7, this research proposal will be included, which consists of chapter 1, 2, and 3 of this thesis. Following that, Chapter 4 and Chapter 5 will formulate the algorithm, which will be tested by the experiment developed in this study. Chapter 6 will evaluate results of the performance of simulations. Finally, chapter 7 will elaborate suggestions for future studies and conclude this research.

3.5 Summary

WSNs energy is the most valuable resource of network connectivity. The wireless sensor network and the nodes with different energy levels could prolong the network lifetime and its reliability. Therefore, this study mainly focuses on the clustering algorithm in terms of energy requirements. Numerous studies had highlighted on this algorithm and offered a potential of improvement in the conventional algorithm. However, further studies still need to be conducted. Furthermore, optimal clustering in terms of energy efficiency should eliminate all the overheads of cluster head selection process and cluster member selection process. Efficient clustering is important to improve the network lifetime. Moreover, load balancing could also be used to extend the lifetime of a node by reducing energy consumption. This method could also increase network scalability with the inclusion of clustering algorithm.

Wireless sensor networks (WSNs) and the clustering algorithm in the scenarios of WSN network connectivity have garnered significant attention over the past few years. Accordingly, the literature review of this study highlighted the main contributions from each previous research including the year when the review was conducted. However, this study was different from other studies due to several distinctive parameters, which shall be elaborated in future studies. Last but not least, this study will focus on the enhancement clustering algorithm in network connectivity through correlated node behaviours in these networks, such as limited bandwidth and energy consumption.