CHAPTER 1

INTRODUCTION

1.1 Introduction

Wireless sensor network (WSN) refers to a group of spatially dispersed and dedicated sensors for monitoring and recording the physical conditions of the environment and organizing the collected data at a central location. A WSNs measure environmental conditions like temperature, sound, pollution levels, humidity, wind, and to cooperatively pass their data through the system to the primary location or sink where the data can be observed and analyzed. Network connectivity is one fundamental issue in wireless sensor networks, which reflects the best way for a network to be connected with each other through sensors. It also reflects how the sink sensor node receives the sensed data for further processing. These functions indicate the importance of network connectivity. In this case, various network connectivity is present as the continuous movement of the nodes is due to their dynamic nature. Furthermore, as each node plays the role of the router in a network, the movements of one or more nodes from one point to another possible result in network partitioning. Due to the infrastructure-less nature of the network topology and the dynamic behaviour of the nodes, maintaining network connectivity is a challenging process (Sajadian, Ibrahim, Freitas, and Larsson, 2011). As a result, network connectivity is among the main problems in wireless sensor network. Similar to other previous studies, this research aims to improve network connectivity clustering.

The factors of network failure lead to the partitioning in the event of correlated node behaviour. These behaviours are divided into four categories, which cooperative node, failure node, malicious node, and selfish node (Taheri, Yaser, Hossein Gharaee Garakani, 2016). In the correlated node behaviours network partitioning, two disjoint sets of notes are created, which could either be failure nodes or selfish nodes. Specifically, failure nodes are inactive in route discovery. In contrast, although the selfish nodes are active in route discovery, they are inactive in packet forwarding. They tend to omit other data packets to save energy, as this energy could be used to transmit more of their own packets and reduce the packet latency. Therefore, this research aims to determine the capability of these nodes to conserve energy in critical operations under the aforementioned conditions and extend network lifetime through clustering algorithm (AC).

This research is based on the correlated node behaviour proposed by A. H. Azni, Ahmad & Noh (2013) to improve the network connectivity in wireless sensor networks (WSNs). In this research, a correlated behaviour was proposed as an extension to node behaviour, which clusters the static nodes with high correlation based on a semi Markov process in a continuous duration. The use of the semi Markov model determines the correlation between the nodes based on the probabilistic parameters, including the probability of selfish behaviour, forwarding, injection, loss, and average recovery. Another proposal of this research was that a node in this model was viewed in several conditions, such as forward, cooperative, failure, and selfish nodes. In this approach, node behaviour transitions were modelled based on the correlated transition probability parameters and the transition time distribution parameters.

The categorization of the nodes was performed through the clustering mechanism defined in this model based on their statuses and behaviour, leading to reduced computational complexity. This model also determined the state of each and every node through a functional mapping between state behaviour and transition probabilities. As a result, a decline in cooperation levels occurred with the increase in the number of nodes in the topology. In contrast, Rani, Malhotra and Talwar (2013) mentioned that although a large number of nodes in the topology of the network increased further, they became unstable due to the possibility of the initial energy level of nodes the range 0.5. Moreover, this study also proved the prevalence of Denial of Services (DOS), which occurred at a fast rate. Following that, a rapid decrease in the network took place, illustrating the possibility for a network failure to be critical in a wireless sensor network. Li, Gong, Liu, Yang & Zheng (2011) highlighted how the network was directly affected when the behaviour of node changed from cooperative to selfish. In addition, the selfish behaviour of static nodes not only affects the individual node, but it also affects multiple nodes involved in the routing process.

Network clustering is highly vulnerable to random failures and selfishness due to their characteristics, including dynamic network topology, limited availability of energy, and error prone communication link (Thakkar & Kotecha, 2014). Therefore, the primary goal for high network connectivity is to establish and maintain a cooperative environment between the static nodes of the network during a reliable dissemination of data. In overcoming the issue of vulnerable network connectivity, clustering is perceived as the ideal method. It is used in WSNs for effective data communication and energy efficiency, as proposed by (Zhu, Zheng, Shu and, Han., 2012). In each cluster, sensor nodes have different roles, such as cluster head, ordinary member node, or gateway node. Specifically, a cluster head (CH) is the group leader in each cluster which collects the sensed data from member nodes. This is followed by the aggregation and transmission of data to the next CH or the base station, as documented by (Li, Gong, Liu, Yang and, Zheng., 2011 and Y. Wang, Lin, and Chang., 2013). Meanwhile, the role of ordinary member node is to sense the data from the environment the data are deployed in. Gateway nodes are the nodes belonging under more than one clusters, with a role to transmit data between the two clusters. Notably, several studies have proposed various traditional clustering algorithms for wireless networks, including the studies by (Lee, Memberm and, Cheng., 2012, and Peiravi, A., Mashhadi, H. R. and, Hamed Javadi., 2013, and X.-X. Liu., 2012). However, these algorithms are not compatible for sensor networks as connectivity and energy efficiency is the primary factor of wireless networks.

1.2 Problem Statement

In wireless sensor networks, nodes which operate under dynamic topology are often correlated with their behaviour. Correlated behaviour may have a critical impact on network connectivity. To illustrate, this direct critical impact would occur when the behaviour of a node transitions from cooperative to misbehaving nodes. These nodes are specified into failure, malicious, and selfish nodes. Furthermore, they often have a correlated effect, leading to partition in the network (D. Zhang, Liu, Zhang and, Liang, 2017).

Although many studies have explored the mobility of sensors node to improve the network connectivity, little attention has been paid to the correlation of sensors movement, which often consumes the majority of the limited energy of sensors and thus shortens the network lifetime significantly (Al-kiyumi, Foh, Vural, Chatzimisios, & Tafazolli, 2018). This address the challenges of the mobile sensor node problem and there is a need to investigate how to deploy mobile sensors with minimum movement to form a cluster in WSN that provide network connectivity. The mobile sensor node problem is decomposed into the problem of the network connectivity. The first such clustering algorithms (initially developed for wired sensor networks) was the Linked Cluster Algorithm (LCA) (G. Li et al., 2018). LCA was a distributed ID- based, one-hop, static clustering algorithm which trying to maximize network connectivity. However, LCA only cater static sensor node and not suitable to be implemented in mobile sensor node. For a mobile sensor node, an efficient solution based on the correlation minimum cluster with constrained edge length is 3D Euclidean distance is proposed in this research to solve network connectivity in WSN.

Previous research worked by J. Wang, Gao, Yin, Li and, Kim (2018) also showed that the method used in PEGASIS do not take into consideration on the network partitioning in which resulted in network degradation under the event of node misbehaviour. The PEGASIS method used chain-based clustering algorithm with the implementation of a Genetic Algorithm to build the chain of network (Somauroo & Bassoo, 2019). The chain based clustering algorithm has delayed the node to join a clustering network due to rotation policy imposed by the chain algorithm. Thus, the PEGASUS is not solving the partition problem which always occurred in WSN.

The second issue related to partitioning in WSN is to measure the degree of correlation between nodes (Muriira, L. M., Zhao, Z., & Min, 2018). According to Saleem et al. (2017), the higher the partition shows more misbehaviour nodes in the network cluster. Thus, a formulation to measure the degree of correlation is needed in order to improve network connectivity by enhancing the existing clustering algorithm. Nodes in the network need to scan the neighboring node thoroughly before cluster head can be elected. In most network clustering highlighted in WSNs literature, the clustering algorithm was only taken into account within a 2D Euclidean distance (Baghouri, Hajraoui, & Chakkor, 2015). Besides, in various recent cases, 3D Euclidean distance in WSN has more significance compared to the 2D Euclidean distance in WSN (Boufares, N., Khoufi, I., Minet, P., Saidane, L., and Saied., 2015). In this study, 3D Euclidean distance is introduced in wireless sensor network structure based on new parameters which are the distance, packet and total energy consumption. The incorporation of static nodes in the formulation of correlated degree through 3D Euclidean distance in wireless sensor network architecture is also considered. Notably, the key idea behind the 3D Euclidean distance in network clustering is to identify a set of cluster network based on three-point distance within a correlation region to identify the level of correlation between the neighboring nodes.

1.3 Research Objectives

The objectives of the research are as follow:

1. To investigate the parameters of correlated nodes behaviour in clustering algorithm for network connectivity.
2. To design 3D Euclidean distance for enhancement of clustering algorithm based on parameters of correlated nodes to improve networks connectivity in WSN.
3. To evaluate and compare the enhancement of clustering algorithm based on network performance and energy consumption.

1.4 Research Questions (RQ)

Research questions will help to find an idea on how, what and where to improve the clustering algorithm. It also will help to achieve the objectives as well as solving the problem.

Table 1.1: Research Questions

Research Objectives	**Research Questions**
To investigate the parameters of correlated nodes behaviour in clustering algorithm for network connectivity.	What are the parameters involved in clustering algorithm? What are the methods used for clustering? What are the factors that impact the network connectivity? What factors influence the network connectivity? How node effect network connectivity?
To design 3D Euclidean distance for enhancement of clustering algorithm based on parameters of correlated nodes to improve network connectivity in WSN.	How to formulate the clustering algorithm? How you will improve the network connectivity in clustering algorithm? How to design clustering algorithm for correlated node behaviour? How you want to design clustering algorithm? What method wants to design? What does different factors in impact of the network?
To evaluate and compare the enhancement of clustering algorithm based on network performance and energy consumption.	How to evaluate the energy consumptions in clustering algorithm to improve network connectivity? What you want to evaluate of the clustering? How you want to do evaluate of clustering algorithm in network connectivity of WSN? What you may achieve?

1.5 Scope of the Research

In this section, the scope of the research is only based on the clustering algorithm to improve network connectivity for WSNs. This focusing on the correlated degree. The algorithm is tested based on distance, energy consumption, and packet delivery ratio. To achieve network connectivity in WSNs. Then, the algorithm will measure based on network performance and the energy consumption is to improve network connectivity for WSNs.

1.6 Thesis Organisation

This thesis contains total of seven chapters including the current chapter. Chapter 1 is the introduction that covers the research background and motivation, research problem and overview of the research idea.

Chapter 2 reviews the available literature review in clustering algorithm to improve network connectivity in wireless sensor networks (WSN) problem in general, and then concentrates upon reviews and analyses the current published researches especially with respect of correlated node behaviour. The available tested methods are also presented together with best-known result in the literature.

Chapter 3 demonstrates the research methodology used in this thesis. It consists of three main phases starting from the identification of the problem domain, identification of the possible solution approaches, and finally the performance and evaluation phases.

Chapter 4 presents the formula of correlated degree which is derived based on 3D Euclidean distance and the correlation coefficient between two nodes which involves computing the square root of the sum of the squares of the variances between the corresponding values in the cluster. Three parameters which are distance, packets and energy are chosen to be used to measure correlated degree due to its high flexibility and ability to support long-range, large-scale, and highly distributed applications. Using the calculation, it is shown that the correlations have a drastic influence on the topological properties of networks. Assortative networks tend to form highly connected groups of nodes with correlated degree which value in an increase of the average path length and correlation coefficient.

Chapter 5 presents the design of the clustering algorithm for correlated nodes behaviour. If the correlated degree value falls below the energy threshold, the node is discarded, and new clustering is reconstructed. The algorithm focuses only in energy consumption to improve network connectivity. Correlation clustering provides a method for clustering as a set of objects into the optimum number of clusters without specifying that number in advance. To keep the connectivity of the sensor network, the parameters formulated in Chapter 4 is applied in the proposed clustering algorithm to enhance connectivity under correlated node behaviour.

Chapter 6 presents an analysis and evaluation based on the results obtained from the best performed algorithm in this work in comparison with outer available approaches in the literature with respect to the experimental.

Finally, the overall conclusions of the work and research direction for future work in this area are presented in Chapter 7.