

# Analysis of Multi-hop Wireless Sensor Networks using Probability Propagation Models

Komgrit Jaksukam<sup>\*1</sup>, Teerawat Tongloy<sup>2</sup>  
Santad Chuwongin<sup>3</sup>

Center of Industrial Robot and Automation (CiRA)  
College of Advanced Manufacturing Innovation (AMI)  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand

Siridech Boonsang<sup>4</sup>

Department of Electrical Engineering  
Faculty of Engineering  
King Mongkut's Institute of Technology Ladkrabang  
Bangkok, Thailand

**Abstract**—This paper presents a formula for estimating the probability of collecting a given amount of data from a propagation model and multi-hop wireless sensor networks (WSNs) based on Monte Carlo simulation with cluster-tree topology. The probabilistic model is based on an analytical model of the IEEE 802.15.4 MAC protocol. The probability of successful node transmission is extended to the probabilities of successful collection at the cluster  $P(X=k)$  and sink node  $P(X \geq k)$ . A numerical example has been provided for comparing the probabilities. We propose a model to calculate the probability from the ratio of the collection rate to the total number of nodes and therefore provide the likeliness of complete data collection. Finally, the results from our analysis provide an estimation of the probability of achieving successful transmission in WSNs.

**Keywords**—Probabilistic modelling; wireless sensor network; multi-hop networks; data collection scheme; Monte Carlo simulation; probability propagation models; probabilistic analysis

## I. INTRODUCTION

Wireless network communication technology has been widely used owing to its infrastructure network that can link and exchange data based on an ad hoc network [1]. The ad hoc network uses distributed node computing without any control centres [2-5, 16]. In addition, it is a self-organization wireless network [6] having a significant number of sensor nodes that respond to several physical aspects, for example, agricultural parameter controlling, transportation planning, logistics planning, and structural testing [7, 24]. This can be achieved by physical data measurement and collection from sensor nodes; the data is then transferred to the coordinator nodes and finally to the sink nodes, based on data exchange of trade-off networks [3-4, 21].

General wireless sensor networks (WSNs) can support single-hop data transfer where the sink node is located at a far distance, which causes energy loss during data transfer. For multi-hop data transfer, the sensor nodes are linked together from the node source to the node sink. In the case of a large network, the data transfer becomes complicated and accumulates more data as the number of hops increase. This contributes to the weakness of multi-hop wireless communication, causing packet loss from the network traffic load, which in turn creates data overflow and network congestion [10, 11]. This then results in channel access failure. For this reason, it was found that general WSNs could not

guarantee packet transfer from the sensor to the sink nodes in multi-hop data transfer. Some probability models, based on the IEEE 802.15.4 standard of CSMA/CA protocol [6], have been proposed to solve the aforementioned problem. An example is the model proposed by Marco et al. [9]. This model focuses on the impact of the reliability and delay at the MAC layer on the routing algorithm's performance [2, 20]. This is done by balancing traffic distribution through the adjustment of three parameters: macMinBE, macMaxCSMABackoffs, and macMaxFrame-Retries. However, this model does not directly guarantee packet transfer to the sink nodes.

This paper presents a new model for calculating the probability of a successful collection. The proposed model was developed from the analytical model of the IEEE 802.15.4 MAC layer networks proposed by Buratti et al. [8]. The Buratti model for single-hop is expanded and applied to the probability framework of WSNs with multi-hop. Extending into a cluster-tree topology, the model for multi-hop networks is formulated by using a propagation model or by following a Monte Carlo simulation of the probability of successful reception for each cluster [12]. A simple numerical result is given to demonstrate the application of our proposed model to estimate the probability of a successful collection rate.

This paper is divided into six sections. In Section II, the statement of contribution is to analyse multi-hop wireless communication in the related works, the underlying analysis model for IEEE 802.15.4 networks is explained. This is followed by the development of multi-hop WSNs based on cluster-tree topology, in Section IV. Section V presents the numerical results that express the probability of successful reception in the propagation model and a probabilistic modelling based on Monte-Carlo simulation at the sink node of WSNs. Finally, the conclusions are provided in Section VI.

## II. RELATED WORKS

For applications based on IEEE 802.15.4 WSN, there are two types of simulation models to evaluate the performance of the data transfer as follows: (1) simulation model using a computer simulation program, (2) simulations models using the probability model for WSNs based on Markov chain's principle. In the first type of simulation model, a computer program, consisting of the OMNeT++ [16, 17], TinyOS [18] and the NS3 [20], is used to arrange the network in various

<sup>\*</sup>Corresponding Author.

kind of topologies by ways of random data to achieve different results. The weakness of this method is some limitations in adjusting some MAC-layer parameters.

For the second type of simulation model [8-9, 12, 19, 22], which is the simulation using the probability model for WSNs based on Markov chain's principle, this simulation described that WSN work mechanism includes two steps, that is, the transition state diagram and the process of every event with the transition probability. In this simulation type, the probability of successful data transfer can be easily calculated and adjusted by changing MAC-layer parameters. Due to these strong points, the topology designed [23] in this research was based on this type of this simulation model. The specific model used in this paper was the IEEE 802.15.4, which was non beacon-enabled model presented by Buratti et al. [8]. The model determines the trade-off between the probability success and energy consumption [22, 25] by adjusting three MAC parameters consists of number of nodes, number of time slots and size of packet. Subsequently, the MAC losses are reduced.

In view of the above, one of the probability models that have been proposed to analyse the performance of IEEE 802.15.4 is the one developed by single hop model [8]. On one hand, this model allows the calculation of the probability of success for a transmission packet. On the other hand, it is based on single-hop communication and can be operated only in a short range due to low-power radio. Some other probability models have been proposed based on multi-hop communication. However, in the case of an extensive network, data collection becomes complicated, and more data is accumulated as the number of hops increases. This contributes to the weakness of multi-hop wireless communication [5], causing packet loss from the network traffic load, which creates data overflow and network congestion [10]. Finally, this leads to channel access failure [11]. Therefore, there is no guarantee that the data packet will be transferred from the sensor to the sink nodes for multi-hop data collection.

The proposed probabilistic model used to express the probability of successful collection at the sink node of WSNs is based on two topology parameters, namely cluster size and hop distance. The simulation was repeated using the Monte Carlo method. From the simulation results, a histogram was created to obtain the probability distribution of data collection. This study provides a reliable multi-hop WSN topology, which can guarantee packet transfer at an acceptable confidence percentage. The analytical multi-hop network has a uniform cluster-tree topology structure composed of small clusters. All clusters are connected to each other via collect and forward methods. In each cluster, the probability of data transfer is calculated based on the Buratti's model for single-hop data transfer. The overall probability of success is then calculated by probability propagation and classified using the number of nodes and hops. They also focus on the reception rate and the probability of success in data collection in WSNs.

### III. SINGLE HOP COMMUNICATION MODEL

The analytical model, the symmetrical topology design of multi-hop WSNs using probability propagation models in this study was based on this type of analytical model [23, 25], which was presented by single hop model [8]. These are the

hypotheses and assumptions for a multi-hop WSN in this study: (1) no hidden nodes, (2) no data retransmission, (3) no acknowledgement for node sensors, and (4) no propagation delay (no data collision). The WSN had a star topology structure and a query-based application. Spaces between each node were equal or symmetric, where each cluster had N nodes ( $N = 1, 2, \dots, N$ ). The data comprised the size of packet (D) which is an integer, the backoff time ( $d_b = 20T_s$ , where  $T_s = 16 \mu s$ ), the time slot (j), and the maximum data transmission time ( $T_{max} = 120$ ). The size of the data transmitted by each node was  $10 * D$  bytes. The time for sending a packet is  $D * d_b$ .

The IEEE 802.15.4 non-beacon-enabled model explains the behaviour of the nodes' operation in one cluster with single-hop communication, as shown in Fig. 1(a). Each single hop consists of 4 statuses: backoff, sensing, transmission, and idle. The procedure starts when the transmission node needs to transmit data. It will check the status of the channel, whether the node is busy or free, by using probability sensing of slot j or  $P\{S^j\}$ . If the status is busy, then the node will keep increasing the number of backoff slots each time until it finds a free channel. Once the data transmission is completed, the probability that a node transmits its packet in slot j for the final node,  $P\{T^j\}$  can be calculated. Then the subsequent transmission node, which is in idle status, will start the data transmission procedure once again, where the number of nodes, N, increases from N-1 to N until the data transmission is completed. As N increases, the number of slots j decreases. This mechanism is shown in Fig. 1(b).

From the data transmission mechanism of a single-hop cluster [8], once the probability  $P\{T^j\}$  is determined, the probability that a node transmits its packet with success in slot j,  $j \in [0, T_{max} + D - 1]$ , where  $P\{Z^j\}$  can be calculated and expressed as follows.

$$P\{Z^j\} = (1 - p_b^{j-D})P\{C^{j-D}\} \prod_{i=0}^{N_{Bmax}} (1 - P\{S^{j-D}\})^{N_c^{j-D}-1} \quad (1)$$

where the probability,  $P\{Z^j\}$  depends on other probabilities: the probability of a busy channel  $p_b^j$ , the probability of a node to sense a channel  $P\{C^j\}$ , and  $P\{S^j\}$ . Therefore, the probability success,  $p_s$ , of data transmission in every node in the cluster can also be determined by

$$p_s = \sum_{j=0}^{T_{max}+D-1} P\{Z^j\} \quad (2)$$

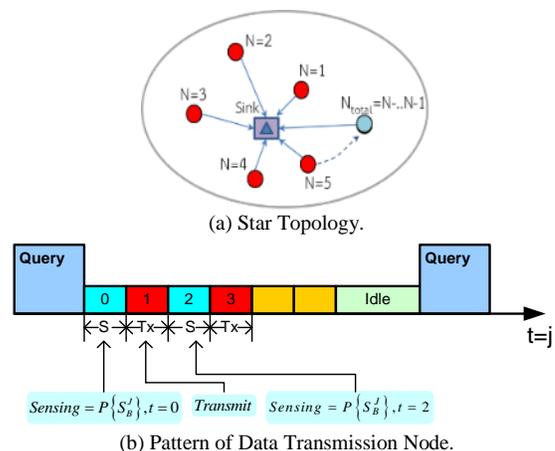


Fig. 1. Single-Hop Cluster using a Probability Model.

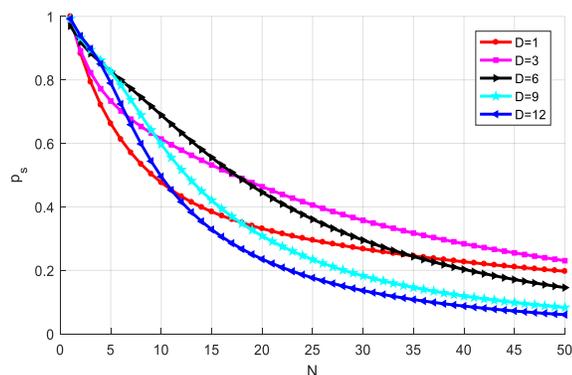
In addition, the function of accumulative probability  $P\{Z^j\}$  is affected by the number of nodes,  $N$ , and the size of the packet,  $D$ , from the following equation:

$$F\{Z^j\} = \sum_{v=0}^j P\{Z^v\} \quad (3)$$

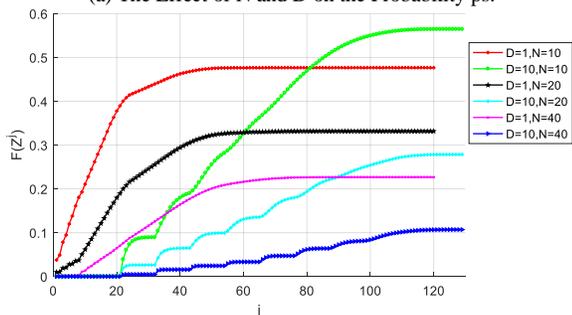
where  $F\{Z^j\}$  is a cumulative function for the probability that a node successfully transmits its packet in time slot  $j$ . Furthermore, the probability model is an algorithm with  $N \geq 1$ ,  $D \geq 1$ , and  $T_{max} \geq 120$  for single-hop, developed using a MATLAB script as follows: (1)

- 1) Initialize MAC parameters for  $j = 0$ .
- 2) Loop for  $j \leq (T_{max} + D - 1)$ 
  - a) Set  $N_c^j = N$ .
  - b) Compute probability  $p_b^j$ .
  - c) Compute probabilities  $P\{S_0^j\}$ ,  $P\{S_1^j\}$ ,  $P\{S_2^j\}$ ,  $P\{S_3^j\}$ , and  $P\{S_4^j\}$ .
  - d) Compute probabilities  $P\{C^j\}$ ,  $P\{Z^j\}$ ,  $p_s$ .
  - e) Compute probabilities  $P\{T^j\}$ ,  $P\{R^j\}$ .
  - f) Compute cumulative probability  $F\{Z^j\}$ .

The performance limitations of multi-hop networks were studied by evaluating the probability of success based on an assumption of star topology. We studied the effects of changing the size of packet ( $D$ ) on  $p_s$ , accumulated probability  $F\{Z^j\}$ . There were two cases studied of the probability for a data transfer in a single hop cluster: The first case studied for the effects of size of packet ( $D$ ) in the network to by fixing  $D$  are set 1, 3, 6, 9, 12 and the maximum data transmission time  $T_{max}$  at 120. For the effect of  $D$  on  $p_s$ ,  $N$  is varying due to 50 nodes while  $p_s$  for each is calculated and the result is show in Fig. 2(a).



(a) The Effect of  $N$  and  $D$  on the Probability  $p_s$ .



(b) The Effect of  $D$  and  $N$  on Accumulative Probability  $F\{Z^j\}$ .

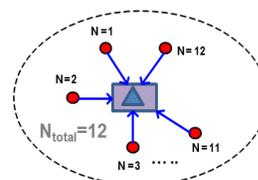
Fig. 2. Effect of Data Transfer in a Single hop Cluster.

The result show that the very low probability success as having a chance of occurrence of less than 0.25%, this caused  $N$  and  $D$  were increasingly. Second case studied for the effect  $D$  in the network on accumulated probability  $F\{Z^j\}$  by  $D$  was set 1, 10 and  $N$  was set to 10, 20, 40 nodes in each case shows that the node was transferring data in the network when time slot ( $j$ ) was increased due to steady-state, as show in Fig. 2(b). We are specifically interested in the channel distribution access in time slot  $j$  within the maximum data transmission time ( $T_{max}$ ).

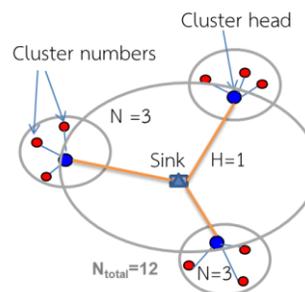
#### IV. MULTI-HOP WIRELESS SENSOR NETWORKS

##### A. Analysis Probability of Single-Hop Model

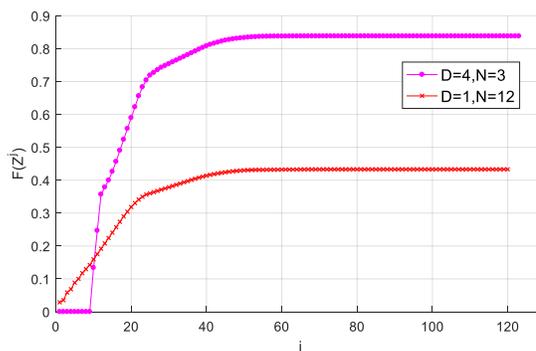
We studied the effects on the probability  $p_s$  of  $N$  and the number of hops ( $H$ ) in the network to learn which terms of data transfer were increasing with  $N$  per cluster and  $H$  per cluster. The varying parameters in the network were as follows: (1)  $N$  was set to 12 nodes per cluster and, for no data transfer,  $H$  was set to 0. The total for each node is equal to  $p_s(N=12, H=0)$  of  $N_{total}=12$ , as shown in Fig. 3(a). (2)  $N$  was set to 4 nodes per cluster with  $H$  set to 1, which makes the total per node equal to  $p_s(N=4, H=1)$  of  $N_{total}=12$ , as shown in Fig. 3(b). In Fig. 3(c), the comparison of probability  $p_s$  between  $N$  and  $H$  in the network shows that  $N$  has more impact than  $H$  on probability  $p_s$ . In case studies, we showed that the very low probability  $p_s$  could be improved by increasing the  $H$  of data transmission.



(a) Single-hop Transmission ( $H=0$ ) with  $N=12$ ,  $D=1$ .



(b) Two-hop Transmission ( $H=1$ ) with  $N=3$ ,  $D=4$ .



(c) A Comparison of the Probability  $p_s$  between  $N$ ,  $H$  in the Network.

Fig. 3. Effect of Data Transfer of  $N$ ,  $H$  Parameters in Network.

### B. Model of Multi-hop WSNs

As the previous model is for single-hop data transfer, it can be extended to support the multi-hop model. From the data transmission mechanism of a single-hop cluster, it was found that the effects of data collected in a single hop have a very low probability. The ideal goal would be to design a multi-hop uniform cluster-tree topology composed of small clusters, as shown in Fig. 4. According to the reasons mentioned above, we propose data transfer in-network, where small clusters are collected, and data is forwarded in each small cluster until it reaches the sink node. Each cluster is built by extending the single hop into the multi-hop network.

The model given in Section III is valid for WSNs with star topology. Considering its application to a multiple infrastructure network, there is a limitation due to the radius of low-power radio. Therefore, an extended model is proposed to multi-hop WSNs topology using a probability framework that can link and exchange data based on an ad hoc network. The cluster-tree formulation is one of the widely used topologies to expand the coverage area of WSNs to be larger than the distance of IEEE 802.15.4 radio. The multi-hop data transmission with cluster-tree formulation [12] with a simplifying model consists of clusters of single-hop models. The topology consists of clusters and a sink node in which each cluster is formed as a star network, which then forwards data to a next-hop router.

In addition, the model of multi-hop WSNs is based on the assumptions given in Section III. Symmetrical topology is assumed in the cluster-tree structure, as depicted in Fig. 3. In this study, we make four assumptions: (1) All nodes are identical. (2) All clusters have the same size. (3) Tree structure is uniform with respect to hop distance. (4) WSN communications use a collect-and-forward protocol, statistically independent in each layer. There were two types of distribution of the nodes in the network: the depth branch is increasing the H that data is transmitted over, and the breadth branch is increasing the number of child nodes in each cluster (N).

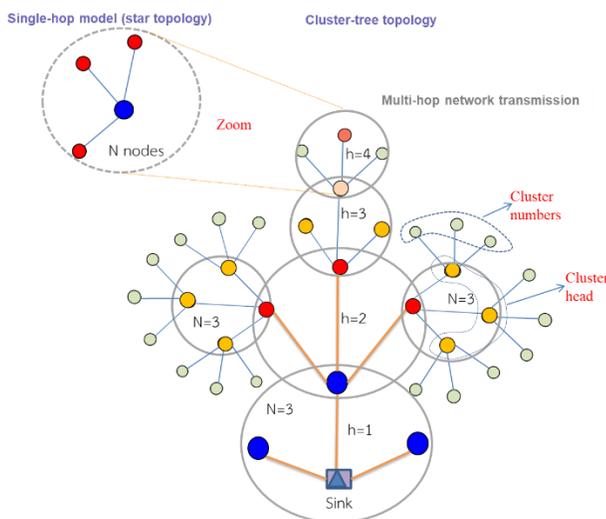


Fig. 4. Cluster-Tree Topology as Multi-Layer Star Topologies.

### C. Probabilistic Modelling based on Monte-Carlo Simulation

We can quickly solve this simplified model with a Monte Carlo simulation. A method for analysing the collected data in multi-hop WSNs with cluster-tree topology is to examine the probability of collected data in each hop independently. In this study, a probability propagation model was created using a pseudo-random Monte Carlo simulation [14, 15]. Such Monte Carlo simulation has been used to compute the concrete probability of successful collection rates to deploy node sensors in a network.

The probability distribution of internal communication networks with probability models was calculated using a pseudo-random method [15] with a uniform distribution [0, 1] of Monte Carlo simulation. Then the value  $p_s$  in each cluster was calculated. The value converges to the random number of the successful data transmitted ( $X=0, 1, 2, \dots, N$ ) by parameters  $N$  and  $H$  in the network. The expected value convergence is given by.

$$E[P(X = k)] = \frac{1}{L} \sum_{i=1}^L X_i \quad (4)$$

where  $E[P(X=k)]$  is the expected amount of collected data in each cluster, and  $L$  is the number of simulation loops. The estimated mean to be expected for collected data in a network can be realistically predicted by running enough simulations. The Monte Carlo model is computed further in loop simulations for the amount of data to be obtained using  $N$  and  $H$  parameters in WSNs that are closer to the behaviour of the data transmission within the networks.

The packet size per node ( $D$ ) in each cluster by  $H$  hops equals to  $(N+1) * D$  bytes is given by

$$D^h = ((N + 1) * D^{h-1}) + 1 \quad h=1, 2, 3, \dots, H \quad (5)$$

Then, the total number of nodes in the network ( $N_{total}$ ) is increasing by

$$N_{total} = \sum_{h=0}^{H+1} N^h \quad (6)$$

In this paper, multi-hop network transmission is considered to be a collect-and-forward protocol; thus, the probability of reception rate  $P(X = k)$  is statistically independent in each layer. In order to use the packet size to depend on the probability of data collected in each such cluster, nodes have to compute the probability of successful transmission ( $p_s$ ).

Therefore, the tree structure is created according to the parameters of WSNs, namely  $N$  and  $H$ . Then hop count =  $h$ ,  $p_s$  in each cluster is calculated recursively from the cluster at hop distance =  $H$ , and the loop ends at the sink node.

The algorithm to estimate the collected data of WSNs is calculated by following steps: (2)

- 1) Initialization for  $N, H$  parameters.
- 2) Loop for  $L \geq 10000$ 
  - a) Generate cluster-tree topologies ( $N, H$ )

- b) Compute probability  $p_s(N,D)$  in each cluster in WSN as H is decreasing.
- c) Compute probability  $p_s \in$  pseudo-random  $[0,1] \rightarrow X=0,1$  in each cluster.
- d) Compute path, counting the number of 1-valued nodes in the cluster; otherwise, it is equal to 0.
- e) Compute cumulative probability of reception rate  $P(X=k)$ .

This  $P(X=k)$  is the amount of data collection that will be used in the cluster that is closer to the sink node. The calculated value converges to the expected value when the number of simulation loops approaches infinity. The diagram of estimating collected data is outlined in Fig. 5.

In addition, to the algorithm to estimate the number of nodes of the data transmission, we calculate the  $p_s$  value for different N and D values by forming a tree topology. Each cycle of tree generation starts with N and H, which define the size of the network.

D. Probability propagation

This study analyses the performance limitations of data transmission from each node in the network. Then, we contribute to guidelines for a topology design [23, 25] with the optimal parameters  $\{N, H, N_{total}\}$ , based on their relationship to the probability  $p_s$ . The efficiency of data transfer can be calculated from the probability of success of collect-then-forward data passing to a cluster that is closer to the sink node.

Considering  $p_s$  as the probability for a successful collection from a child node, we can derive the probability of successful reception from k identical child nodes using the theory of binomial probability [26] for the data distribution by evaluating the success of data transfer in each H for all paths separately. The definitions of the variables are as follows: the number of independent nodes in each cluster is N+1, and the number of sample space (N(S)) has the sequence of the total probability as  $\{X_0, X_1, X_2, X_3, \dots, X_N\}$  where the number N(S) of nodes equals N, and k is the number of successful data transfers ( $k=0, 1, 2, \dots, N$ ), which means that N-k is the number of failed data transfer in each cluster. In addition, the determination of D and H in each cluster will be evaluated for the worst-case scenario, which equals to k+1.

The probability  $p_s$  of data transfer in each cluster can be calculated to wait for data until  $T_{max}$  by using Equation (7):

$$P^l(x=k) = \binom{N}{k} p_s^k (1-p_s)^{N-k} \tag{7}$$

where  $0 \leq k \leq N$ , l is the layer number of a hop,  $P^l(x=k)$  is the probability of success for data transfers for each number of hops. Additionally, we studied the probability distribution that affects the probability  $p_s$  of data collection, and we can substitute  $p_s(T_{max}, D, N)$  with following the layers of data collection in the network. Thus, the probability as well as the sink node can be determined by the maximum amount of

data that can be collected at the sink. The total probability of success equals to

$$p_s^{topology} = \prod_{i=1}^l (p_s \rightarrow p'_s | p''_s \rightarrow p'''_s | p''''_s \rightarrow \dots \rightarrow p_s(\text{sink}) | p'_s \dots p'_l) \tag{8}$$

A method for analysing and classifying the propagation model in multi-hop WSNs with cluster-tree topology, where the relationship between N, D, j and H at  $T_{max}$  is assumed to be fixed. We studied the effect of generating the tree by probability of successful collection in which each cluster communicates independently from the others. For example, the application of the probability propagation [13] for successful data transmission into two nodes. The probability of success (X) of each cluster which has the probability distribution equals to N(S), and the factor of probability consists of  $\{V, \{V,A\}, \{V,B\}, \{V,C\}\}$ , as expressed in equation (9):

$$p_s = p_s^V p_s^{AV} p_s^{BV} p_s^{CV} \tag{9}$$

The expanding of the relationship of the joint probability distribution for  $\{V, A, B, C\}$  was shown in Fig. 6(a) by setting  $\{A,B,C\}$  as X and by letting A equal to 0, B equal to 1, and C equal to 2, as shown in Fig. 6(b).

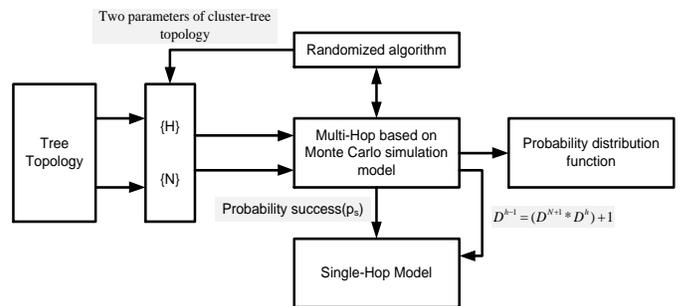
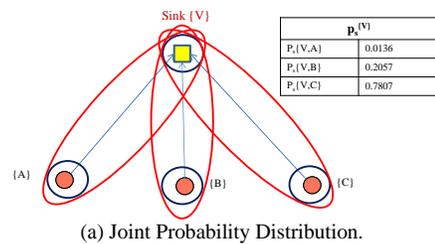
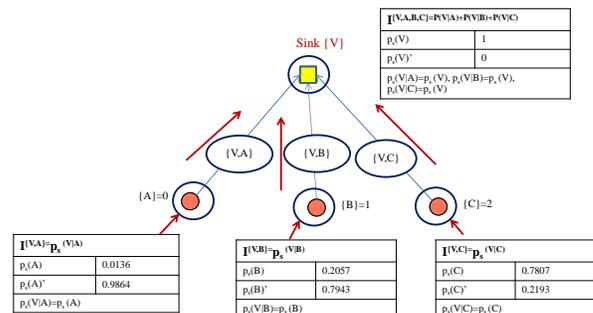


Fig. 5. Block Diagram of the Monte-Carlo Simulation Model.



(a) Joint Probability Distribution.



(b) Joint Probability Distribution of ps for 2 Nodes.

Fig. 6. Probability Propagation.

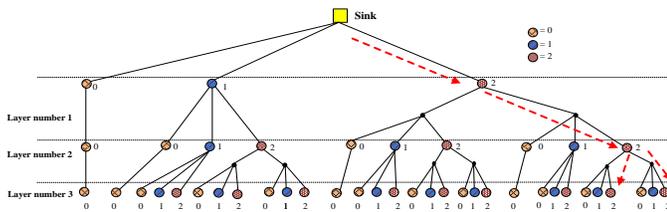


Fig. 7. Probability Propagation for WSNs with  $N = 2$  and  $H = 3$ .

Fig. 7 illustrates the backpropagation approach to calculate the probability of successful transmission from  $m$  nodes at the sink node. The amount of collected data can be computed by traversing down the tree where each cluster communicates independently from the others. For example, a set of arrows in Fig. 7 shows branches representing the scenario of which all 14 data can be collected in the networks. We can derive the probability of successful collection from  $m$  nodes at the sink node using the concept of probability propagation [13].

$$P(X = m) = \prod_{h=H}^l P_{h-1} \left( x = k_{h-1} \mid \prod_{l=1}^{k_{h-1}} P_h^{(l)} (x = k_h^{(l)}) \right) \quad (10)$$

where  $0 \leq m \leq (N_{\text{total}}-1)$ . Term  $P_h(x=k_h \mid \prod_{l=1}^{k_h} P_{h+1}(x=k_{h+1}^{(l)}))$  denotes the probability of successful collection from  $k_h$  nodes at hop  $h$  considering the probability of successful collection for each cluster at the next hop. Equation (10) leads to an approach for estimating the amount of collected data at hop  $h$  by back-propagating  $P(x=k)$  from the last hop.

For WSNs with star topology, the packet length  $D$  is an independent parameter and may be treated as a fixed value. However, the packet length of a node in WSNs with cluster-tree topology depends on both cluster size ( $N$ ) and hop distance ( $h$ ). Due to imperfections in transmission, a function  $D(N, h)$  is a random process with its upper bound given by.

$$D(N, h) \leq \sum_{p=0}^{H-h+1} N^p \quad (11)$$

To avoid random variables in the model expressed by (7) and (10), local maxima of  $D(N, h)$  may be used as an approximation for the computation of  $p_s$  at each hop.

## V. NUMERICAL RESULTS

For developing a model used to predict the behaviour of the internal data collection, the structure tree is assumed to have a uniform cluster-tree topology, and  $N$  and  $H$  are set to 2 and 3, respectively. This corresponds to the total number ( $N_{\text{total}} = 14$ ). The main results are as follows: There are two methods, namely propagation modelling and probabilistic modelling based on the Monte-Carlo simulation, which is used to calculate the probability of successful collection (see Section IV).

### E. Analysis of Multi-hop WSNs based on Monte Carlo Simulation

For numerical comparison, we used a model for estimating the probability of collecting a given amount of data from multi-hop WSNs based on Monte Carlo simulation with cluster-tree

topology. Based on the analytical model of IEEE 802.15.4 MAC, the probability of successful node transmission  $p_s$  is extended to the probabilities of successful collection at the cluster  $P(X=k)$  and sink node  $P(X \geq k)$ .

Fig. 8 shows the simulation results, demonstrating the effectiveness of our proposed model for the probability of data collection at the sink node, for  $N=2$  and  $H=3$ . The cluster-tree structure has a uniform topology with  $N = 2$  and  $H = 3$ , and  $N_{\text{total}}=14$ .

### F. Analysis of Multi-hop WSNs based on Propagation Probability

The probability propagation model has been used to calculate the probability distribution of internal communication in WSNs. We considered a multi-hop WSN with  $N = 2$  and  $H = 3$ . We name each node by concatenating the characters 'A' and 'B' to denote the hop distance from the sink node marked by 'V'. That is, a cluster "VA" consists of "VAA" and "VAB" nodes. Similarly, the "VAA" node also forms a cluster by collecting data from "VAAA" and "VAAB" nodes. There are eight nodes belonging to four clusters at the last hop ( $h = 2$ ), namely "VAAA", "VAAB", "VABA", "VABB", "VBAA", "VBAB", "VBBA", and "VBBB". For these clusters, the probability of successful collection from  $k$  nodes is deterministic and can be calculated from (3) using  $N = 2$  and  $D = 1$ . For a cluster with  $h=1$ , e.g. "VA", the probability of successful collection can be approximated conservatively from (4) by using  $D = 3$ .

Fig. 9 illustrates the backpropagation approach to calculate the probability of successful collection, if the number of successful transmissions at each cluster is provided. The completed collection in Fig. 10, it shows that the calculation of all nodes of the network can be successfully transferred 14 data sets by considering the reduced hop distance and a growing size of data will be increased according to reduced hop distances located on near the sink node. Also, the probability of successful transmission of a data will be decreased according to a growing size of data by collecting data at the sink node ( $X= 14$ ), and the probability of successful transmission of a data at 14 ( $P(X=14)$  was 0.3600.

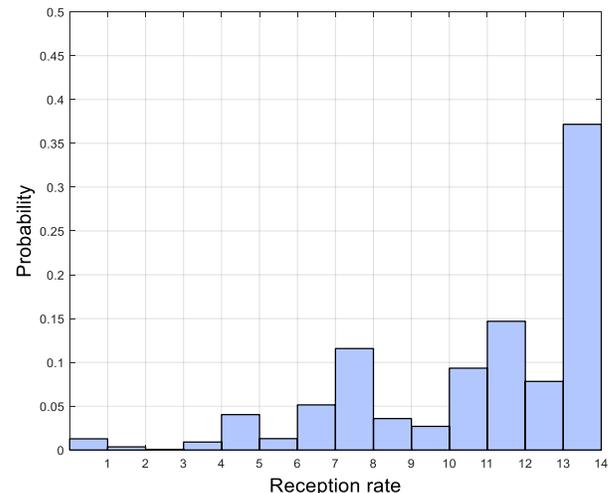


Fig. 8. Histogram of the Reception Rate with  $N = 2$ ,  $H = 3$ .

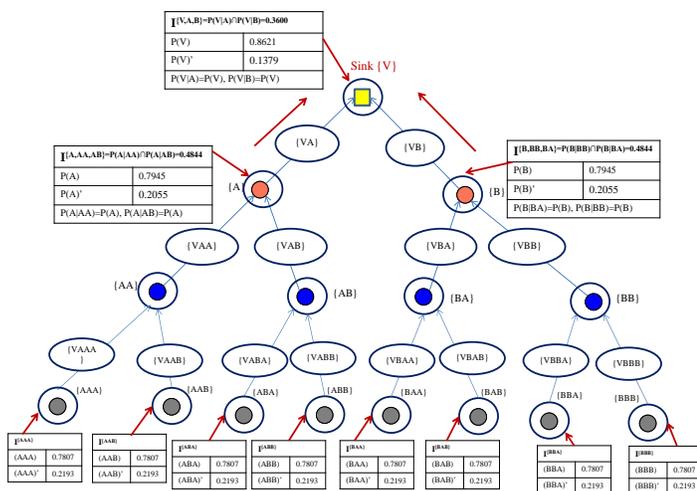


Fig. 9. Probability Propagation for WSNs with N=2, H=3.

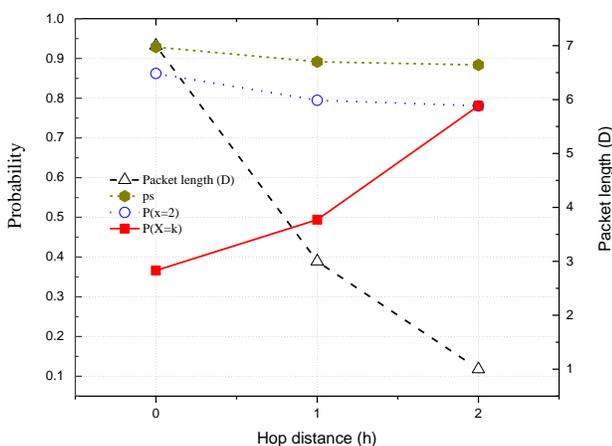


Fig. 10. Probability of Successful Collection with P(X=14).

The calculation of the probability of data collection viewed as an imperfect transmission in WSNs was at X=11 of N=14. The consideration of the probability of data collection to gather data of N=2, H=3 can be successfully transmitted in some sink nodes considered as imperfect transmissions in WSNs (X=11) consisting of the probability of propagation of imperfect transmission {7, 4}, {6, 5}, {5, 6} and {4, 7}.

This leads to the current researcher selects this model in order to explain the imperfect transmitting data to sink node. To test the probability propagation in imperfect transmission, the calculation of the probability for data collection was to gather transferring data at X=11 according to the formulation presented  $m_i=(7,4)$ ,  $(6,5)$ ,  $(5,6)$  and  $(4,7)$  as mentioned term  $P(X = 11) = \sum_{m_i} \prod_{h=H}^l P_{h-1} (X = k_{h-1} | \prod_{l=1}^{k_{h-1}} P_{h-1}^{(l)} (X = k_h^l))$  as presented in Table 1.

Fig. 11 shows the probability of successful collection, which represents the reception rate, for X = 2, 7, 11, and 14. Both propagation and probabilistic modelling based on the Monte-Carlo simulation results are reported: the P(X=k) value according to (4) and the probabilistic modelling process in Section IV.

TABLE. I THE CALCULATION OF THE PROBABILITY FOR DATA COLLECTION AT X=11 IN WSNs

Probability of successful collection P(X=11) in WSNs		
Pattern Probability propagation	P(X=k)	$\sum_{m_i}^{m_k} \{P(X = 11)\}_{m_i}$
$m_1=7,4$	0.01418256	0.0722
$m_2=6,5$	0.0076	
$m_3=5,6$	0.0076	
$m_4=4,7$	0.01418256	

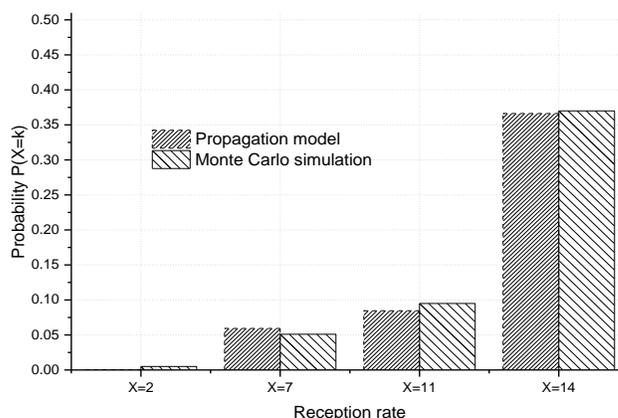


Fig. 11. Comparison of Probability Propagation vs Monte Carlo Simulation with N = 2 and H = 3.

As we can see, the two proposed models provide approximately the same results, and neither presents relevant differences with respect to simulations; therefore, the model is validated. The comparison of results obtained from the propagation and simulation models confirms that the proposed model used for probabilistic analysis of multi-hop WSNs leads to changing the N and H parameters in the topological structure of large-scale WSNs. From the case study of the multi-hop model, we considered the pattern of the probability distribution. The uniform topology structure for perfect transmission of the total packet has H set to 3,  $T_{max}$  set to 120, and D set to 1. In addition, we categorized the calculated probabilities of success for a perfect transmission packet (X=14 nodes) into three hops, as presented in Table 2.

### G. Design Topology of Multi-hop WSNs

The designed multi-hop topology has a uniform cluster-tree topology which is composed of small clusters. All clusters are connected to each other via a collect-and-forward method. From the design of the uniform cluster-tree topology, which the number of nodes was set to 14 nodes, the probability of success was evaluated using the probability propagation. Additionally, the result of perfect transmission of total packet at 14 nodes and imperfect transmission of total packet at 10 nodes, it was found that the  $p_s$  was equal to 20.22% and 0.89%, respectively. Furthermore, in Fig. 12, there were other parameters in the network such as N, D, j and H, and all of these parameters contributed to the reliability of data collection and forwarding for the multi-hop WSNs.

TABLE. II. THE PROBABILITY SUCCESS VALUE FOR ATTRIBUTION

Number of received packet (X)	Number of hop: 3		Number of hop: 2		Number of hop: 1		Topology (%)
	Parameters /cluster	$p_s^{[3]}$ (%)	Parameters/cluster	$p_s^{[2]}$ (%)	Parameters /cluster	$p_s^{[1]}$ (%)	
X=10 (5-5)	N=2,D=1, X=1	20.57	N=2,D=2, X=2	77.93	N=2,D=5, X=2	82.7	0.89
	N=2,D=1, X=1	20.57	N=2,D=2, X=2	77.93	N=2,D=5, X=2	82.7	
X=10 (5-5)	N=2,D=1, X=0	1.354	N=2,D=3, X=2	79.45	N=2,D=5, X=2	82.68	0.005
	N=2,D=1, X=2	78.07	N=2,D=3, X=2	79.45	N=2,D=5, X=2	82.68	
:	:	:	:	:	:	:	:
X=14	N=2,D=1, X=2	78.07	N=2,D=3, X=2	79.46	N=2,D=7, X=2	86.21	20.22
	N=2,D=1, X=2	78.07	N=2,D=3, X=2	79.46	N=2,D=7, X=2	86.21	

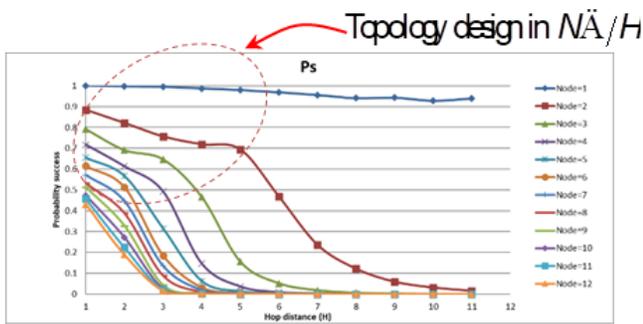


Fig. 12. WSNs Topology for Optimal Data Collection from N and H with  $P(X=k)$ .

As future work, we plan to design an algorithm for multi-hop WSN by optimising a technique that increases the data reception ratio (PRR) and is suitable for the usage of maximum topology with deep learning approaches.

### VI. CONCLUSION

This study presents an analysis model for multi-hop WSNs based on the Monte Carlo simulation model for the probability of data collection in cluster-tree WSNs using the IEEE 802.15.4, MAC protocol with the probability of success in data transmission ( $p_s$ ) of each node in the cluster by data transmission network expansion with a probability of obtaining data on each cluster  $P(X = k)$  and at the sink node  $P(X \geq k)$ , respectively.

The model for multi-hop WSNs with simulation data collection within WSNs for data received at the sink node is applied to calculate the probability of success for a maximum topology of a network model. The objective of this study was to create a model with an explanation of the conditions for collecting/forwarding data over a network with adjustable parameters, N and H, which have an effect on data collection, hop distance, and the number of nodes in each cluster. Results from the analysis of the collected data were applied in a probabilistic analysis of multi-hop WSNs, comparing probability propagation to probabilistic modelling based on Monte-Carlo simulation as an important indicator of WSNs data collection abilities.

### REFERENCES

- [1] I.F. Akyildiz and M.C. Vuran, *Wireless Sensor Networks*. Wiley, 2010.
- [2] Kiruthika.A 1, Sugandhi.N, "The Efficient Energy Aware Routing and Data Aggregation in Multi-Hop Wireless Sensor Network", *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 4, Issue 1, January 2015.
- [3] N. Patwari, J.N. Ash, S. Kyperountas, A.O. Hero, R.L. Moses, and N.S. Correal, "Locating the Nodes: Cooperative Localization in Wireless Sensor Networks" *IEEE Signal Processing Magazine*, vol. 22, pp. 54 - 69, July 2005.
- [4] PRABHA. N and ASHWIN.M, M.E," Mobile Data Collection in Wireless Sensor Networks Using Bounded Relay Hop", *International Journal of Innovative Research in Science, Engineering and Technology*, Volume 3, Special Issue 1, February 2014.
- [5] Senthil.T1, Bifrin Samuel.Y, "Energy Efficient MAC Protocol for Wireless Sensor Networks Using Multi hop Optimization", *International Journal of Innovative Research in Science, Engineering and Technology*, Volume 3, Special Issue 3, March 2014.
- [6] (2006). "IEEE Standard for Information technology--Local and metropolitan area networks-- Specific requirements-- Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPANs)." *IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4-2003)*: 1-320.
- [7] E. Callaway, P. Gorday, L. Hester, J.A. Gutierrez, M. Naeve, B. Heile, and V. Bahl, "Home Networking with IEEE 802.15.4: a Developing Standard for Low-rate Wireless Personal Area Networks", *IEEE Communications Magazine*, vol. 40, pp. 70-77, August 2002.
- [8] Buratti and R. Verdone, "Performance Analysis of IEEE 802.15.4 Non Beacon-Enabled Mode", *IEEE Trans. on Vehicular Technology*, vol.58, pp. 3480-3493, September 2009.
- [9] P. Di Marco, P. Park, C. Fischione, K. H. Johansson, "Analytical Modeling of Multi-hop IEEE 802.15.4 Networks", *IEEE Transactions on Vehicular Technology*, Vol. 61, No. 7, pp. 3191--3208, September 2012.
- [10] B. Hull, K. Jamieson and H. Balakrishnan , "Mitigating congestion in wireless sensor networks "Proceedings of the 2nd international conference on Embedded networked sensor systems, 2004, Baltimore, MD, USA.
- [11] Z. Fu "The Impact of Multihop Wireless Channel on TCP Throughput and Loss", *IEEE INFOCOM 2003*, 2003.
- [12] K.Jaksukam and S.Vorapojpisut, "A Study on Data Collection Behaviors for a Class of Multi-Hop Wireless Sensor Networks Based on Probability Model", *International Review on Computers and Software* 10(6):635, June 2015.
- [13] G. Shafer and P.P. Shenoy, "Probability Propagation", *Annals of Mathematics and Artificial Intelligence*, vol. 2, pp. 327-351, 1990.

- [14] R. Ricardo, H. G. Xiong, Q. Gao, A. Magallanes and F. Candilio, "Monte Carlo Analysis of nodes deployment for large-scale Wireless Sensor Network using range-free location methods", 2010 IEEE 2nd Symposium on Pages: 484-489, 16-17 Aug. 2010.
- [15] Rubinstein RY. Simulation and the Monte Carlo method. John Wiley & Sons, Inc.; 1981.
- [16] Fatima Laassiri, Mohamed Moughit and Nouredine Idboufker, "Improvement of the Vertical Handover Decision and Quality of Service in Heterogeneous Wireless Networks using Software Defined Network" International Journal of Advanced Computer Science and Applications (IJACSA), 9(12), 2018. <http://dx.doi.org/10.14569/IJACSA.2018.091210>
- [17] F. Chen and Falko Dressler, : A Simulation Model of IEEE 802.15.4 in OMNeT++, in Proceedings of 6. GI/ITG KuVS Fachgespräch Drahtlose Sensornetze, Poster Session, Aachen, Germany, July (2007), pp. 35-38.
- [18] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss and P. Levis, Collection tree protocol, Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, (2009).
- [19] Jurcik, P., A. Koubaa, et al.: A Simulation Model for the IEEE 802.15.4 protocol: Delay/Throughput Evaluation of the GTS Mechanism. Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, MASCOTS (2007).
- [20] Nouredine Seddiki and Bassou Abedsalem, "Study of the Performance of Multi-hop Routing Protocols in Wireless Sensor Networks" International Journal of Advanced Computer Science and Applications (ijacsa), 8(2), 2017. <http://dx.doi.org/10.14569/IJACSA.2017.080248>.
- [21] Omar Fouad Mohammed, Burairah Hussin and Abd Samad Hasan Basari, "On the Probability of Detection Ability in Observing Dynamic Environmental Phenomena using Wireless Sensor Networks" International Journal of Advanced Computer Science and Applications (ijacsa), 8(5), 2017.
- [22] Hui, J. and H. Aida,: An analytical approach to optimization of throughput for IEEE 802.15.4 slotted CSMA/CA networks., CCNC 2011, (2011).
- [23] Buratti, C. and R. Verdone : Tree-based Topology Design for Multi-Sink Wireless Sensor Networks. Personal, Proceedings of the IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2007, 3-7 September 2007.
- [24] Chen, F., N. Wang, et al.: Simulation study of IEEE 802.15.4 LR-WPAN for industrial applications. Wireless Communications and Mobile Computing vol 10(5), (2010), p.609-621.
- [25] Martalo, M., C. Buratti, et al. : Optimum Topology in Clustered IEEE 802.15.4 Sensor Networks with Decentralized Detection, Vehicular Technology Conference (VTC Spring), (2011).
- [26] Jianguo Ding: Probabilistic Inferences in Bayesian Networks, Bayesian Network. Ahmed Rebai (Ed.), Teck, (2010), p. 39-52.