Intelligent Medium Access Control Protocols for Wireless Sensor Networks

Yan Yan

Ph.D.

University of York

Electronics

May 2015

Abstract

The main contribution of this thesis is to present the design and evaluation of intelligent MAC protocols for Wireless Sensor Networks (WSNs). The objective of this research is to improve the channel utilisation of WSNs while providing flexibility and simplicity in channel access. As WSNs become an efficient tool for recognising and collecting various types of information from the physical world, sensor nodes are expected to be deployed in diverse geographical environments including volcanoes, jungles, and even rivers. Consequently, the requirements for the flexibility of deployment, the simplicity of maintenance, and system self-organisation are put into a higher level. A recently developed reinforcement learning-based MAC scheme referred as ALOHA-Q is adopted as the baseline MAC scheme in this thesis due to its intelligent collision avoidance feature, on-demand transmission strategy and relatively simple operation mechanism. Previous studies have shown that the reinforcement learning technique can considerably improve the system throughput and significantly reduce the probability of packet collisions. However, the implementation of reinforcement learning is based on assumptions about a number of critical network parameters. That impedes the usability of ALOHA-Q. To overcome the challenges in realistic scenarios, this thesis proposes numerous novel schemes and techniques. Two types of frame size evaluation schemes are designed to deal with the uncertainty of node population in single-hop systems, and the unpredictability of radio interference and node distribution in multi-hop systems. A slot swapping techniques is developed to solve the hidden node issue of multi-hop networks. Moreover, an intelligent frame adaptation scheme is introduced to assist sensor nodes to achieve collision-free scheduling in cross chain networks. The combination of these individual contributions forms state of the art MAC protocols, which offers a simple, intelligent and distributed solution to improving the channel utilisation and extend the lifetime of WSNs.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to dedicate this thesis to my parents for their love and support in my Ph.D. life. They have always encouraged me towards excellence.

My deepest gratitude is to Dr. Paul Daniel Mitchell for his patience, encouragement and persistent help throughout the past four years. As my first supervisor, Paul always give me immense knowledge and unique insights about my research work. Without his guidance, this thesis would not have been possible. I would also like to gratefully acknowledge the valuable help from my thesis advisor Dr. David Grace and my second supervisor Mr. Tim Clarke. They always give me inspiration and guidance, which is really beneficial to my research progress.

In addition, I want to extend my appreciation to my friends, colleagues and experts from the communications and signal processing group, their support is very important to my research as well. All in all, the merits of this thesis are due to the good efforts of contributors mentioned above, I wish to express my sincere gratitude to them.

Declaration

To the best knowledge of the author, all work in this thesis claimed as original is so. Any research not original is clearly specified. References and acknowledgements to other researchers have been given as appropriate. In addition, I certify that this thesis was never submitted in my name, for any other degree or diploma in any university or other tertiary institution.

Chapter 1

Introduction

Contents

1.1 Research Background

This thesis focuses on the techniques and challenges concerning the design of intelligent Medium Access Control Protocols for WSNs. The concept of WSNs originated in the 1950s. The earliest application was the Sound Surveillance System (SOSUS), designed by the United States Military for the purpose of detecting and tracking submarines [1]. This system contains a large number of acoustic sensors deployed in the Atlantic and Pacific oceans for gathering sound signals. Since then, sensing and wireless communication technologies have been constantly and gradually developed throughout the 1960s and 1970s. In the early 1980s, the Defence Advanced Research Projects Agency (DARPA) initiated the modern WSN programme called Distributed Sensor Networks (DSN) [4]. The creation of DSN led researchers to realise the potential benefits of WSNs in consumer markets. From the mid of 1990s to the present, WSN technology became an important focus in academia and civilian scientific research. As a multidisciplinary technology, WSNs merge a wide range of techniques of wireless communication and networking, distributed sensing, signal processing and pervasive computing. A wireless sensor network is a system consisting of densely distributed sensor nodes that gives people the capability to observe and react to events and phenomena within a specific sensing area [2]. Currently, WSNs play a key role in real-time information monitoring and collection due to their flexibility and efficiency. To some extent, WSNs represent the unique means of real-time remote information sensing and collecting.

WSNs have been developed for decades, but the market demands were mainly driven by the military and heavy industry. After entering the 21st century, corresponding research gained increasing attention due to breakthroughs in multiple key areas. As the standardisation of CMOS processing technologies for most semiconductor components begun from early 2000s. Network designer can use simplified hardware solutions such as wireless Micro-controllers (MCUs) that usually consist of a general-purpose MCU and an RF transceiver in a single chip. Therefore, the cost of high node count WSN applications finally reaches an affordable level. Besides, the development of battery and energy harvesting technologies enable longer operation of sensor nodes. Moreover, System-on-Chip (SoC) [3] integration technology has achieved unprecedented development during the past decade. Sensing units, processing units, memory and antennas can be integrated at a cubic millimetre sized scale. Owing to those technological advances, the production of small-sized, low-cost and multifunctional wireless sensor nodes becomes technically and economically feasible. The current trend of WSNs turns towards long term distributed sensing and intelligent self-maintenance [94]. Sensor nodes can be eventually deployed on any physical object and in any geographic area, perform various applications in diverse environments, including soil, natural habitats, oceans, volcanoes and the human body. In the near future, WSNs are expected to be integrated into many life-changing technologies such as the Internet of Things and Smart Cities [95]. Across a wide range of applications, sensor networks can help us to understand and manage an increasingly interconnected physical world.

Energy efficiency of wireless sensor nodes is usually treated as the paramount priority while designing a WSN [6]. It determines the operating period of WSNs due to the battery-driven nature of typical wireless sensor nodes. How to minimise the power consumption of nodes, on the condition that the reliability of the data transmission is guaranteed, is an important research topic for WSNs. To eliminate power inefficiencies, researchers have devoted substantial effort to extend the lifetime of sensor nodes from all aspects. Current studies about the energy consumption of sensor networks suggest that the power consumption of nodes is strongly dependent on their radio modes which are directly controlled by the Medium Access Control (MAC) protocols [13]. MAC protocols aims to regulate radio activities of nodes and coordinate channel sharing in order to avoid retransmissions, idle listening, overhearing and other energy waste activities. Consequently, the design of efficient MAC protocols plays a decisive role in ensuring system QoS and prolonging the operation time of sensor nodes. Compared to conventional wireless networks, the design of MAC protocols for WSNs needs to overcome some unique challenges about hardware constraints, power consumption, channel bandwidth, topology management, etc. The demand for higher channel utilisation and reduction in the coordination overheads intensifies the need for an intelligent channel sharing policy. The network uncertainty associated with environmental changes creates the demand for node self-organisation. In addition, to improve the reliability and efficiency of protocol implementation, a simpler operation mechanism is desired.

1.2 Overview of Medium Access Control Layer

WSNs rely on a group of protocols that are running concurrently to fulfil required applications. Those protocols are need to be properly managed to ensure the operation of individual sensor nodes and increase the overall efficiency of the network. Therefore, a protocol stack architecture is required to standardise and abstract the internal functions of sensor nodes. The protocol stack of WSNs is similar to Open Systems Interconnection (OSI) network model proposed by International Organization for Standardization (ISO). OSI model divides the communication functions of an open system into seven logical layers including physical layer, data link layer, network layer, transport layer, session layer, presentation layer and application layer [12]. Each layer deals with a particular problem. Compared to standard OSI model, the WSNs have a relative simplified protocol stack (see Fig. 1.1) due to their unique characteristics [13].

Amongst the layers demonstrated in Fig. 1.1, this thesis puts emphasis on the Media Access Control (MAC) layer which plays significant role in channel utilisation, system Quality of Service (QoS) and the lifetime of sensor nodes. The Media Access Control Layer is one of the two sublayers that form the Data Link Layer (DLL). The MAC sublayer is located above the physical layer and it is responsible for the following basic functions.

- Frame encapsulation and disassembling.
- Addressing of destination stations.
- Conveyance of source-station addressing information.
- Protection against errors, generally by means of generating and .checking frame check sequences.
- coordination of access to the physical communication medium.

The prime focus of MAC layer is the channel access control mechanisms which are also known as MAC protocols [15]. A MAC protocol is composed of a set of rules

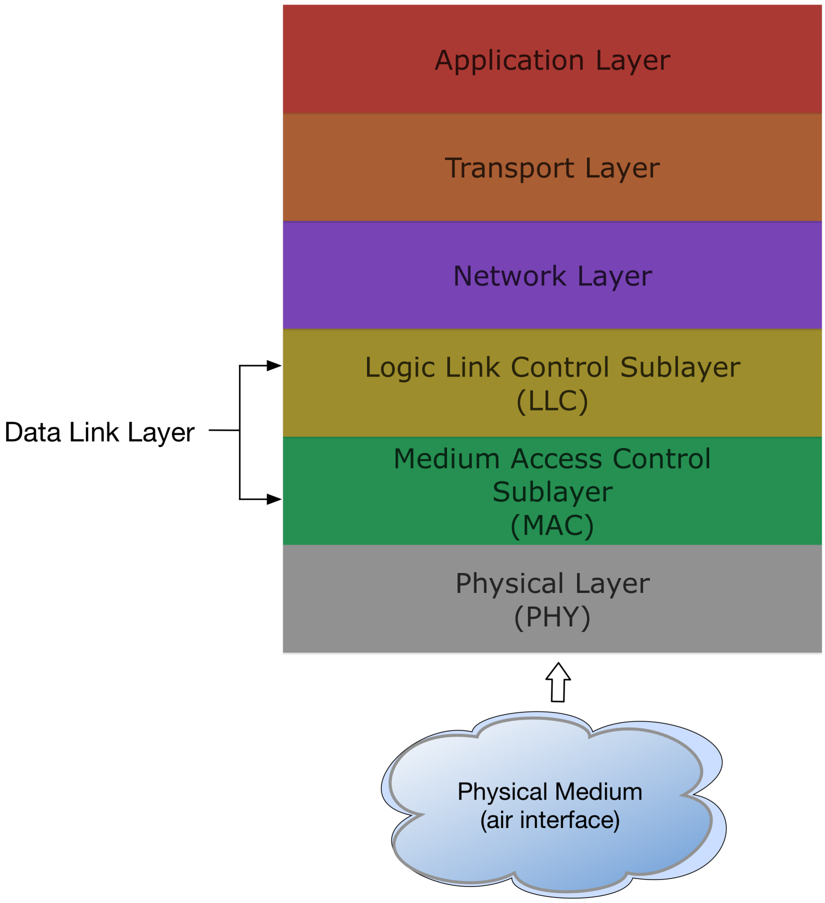


Figure 1.1: Protocol Stack of WSNs

allow multiple stations connected to the same physical medium to share it. Compared to wired networks such as ethernet. The data transmission of wireless networks can be easily disturbed due to channel noise, long range signal fading and environmental issues, so that the design of an efficient MAC protocol is vitally important for WSNs to successfully carry out their required operations.

1.3 Overview of Reinforcement Learning Technique

To improve the intelligence of sensor nodes for access channel, this research exploits reinforcement learning techniques, which are specific learning algorithms in the machine learning family. Reinforcement learning is expected to enable sensor nodes to form an optimal scheduling policy without consuming any significant control overheads [53]. In a general sense, machine learning is a subfield of artificial intelligence concerned with techniques that intelligent systems (i.e. computers, robots, sensor nodes.) simulating the learning behaviour of humans with the purpose of imbibing knowledge and skills from the external environment [89]. Machine learning allows learners to gain experience by doing tasks and then improve their performance on the same task or similar tasks in the domain. A learner is composed of a learning module, database and action module. These components help a learner capture useful information from the environment according to supervised or unsupervised approaches. The core idea of reinforcement learning is simple: when a learning process starts, the action module firstly executes random interaction between the learner and the environment [88]. The information from the environment will then be converted into the valuable experience by the learning module. These experience are always stored in the database module. From the next interaction, an improved action will be performed by analysing the database.

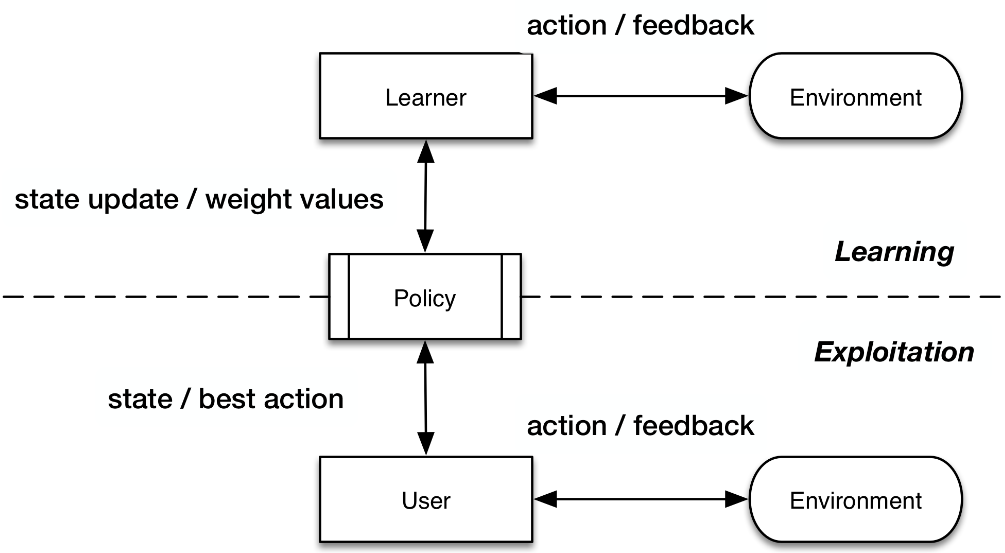


Figure 1.2: Reinforcement Learning Process

As an important machine learning category, reinforcement learning has been widely applied in many areas including intelligent robots, financial forecasting and analysis, etc [87]. It can be considered as a computational algorithm for transforming real-world situations into actions. As Fig. 1.2 shows, a typical reinforcement learning process can be divided into two sub-processes: learning process and exploitation process. During the learning process, a learner attempts to obtain an optimal policy for achieving its goal. Specifically, a learner performs a random action and environment will respond a feedback to the learner at the beginning of the learning process. Subsequently, the learner transforms the received feedback into reward or punishment value and then evaluate which is the optimal action based on a specific reward function. Any action performed by a learner also affects subsequent actions and rewards. In this manner, a learner continuously reinforces its actions by receiving feedback from the environment. After a period of time, the learner is expected to find an optimal action policy which always leads to maximum reward, it then enters an exploitation status where the learner does not need to gather more experience but makes best action according to current policy.

Compared to supervised learning algorithms, reinforcement learning does not assume that labelled or explicit patterns or examples are given to the learner for training. Moreover, reinforcement learning differs from unsupervised learning in that it introduces the error or reward signal to evaluate a potential solution. In reinforcement learning, a learner is not told directly what to do or which action to take. Instead, reinforcement learning eliminates examples and requires that the learners form and evaluate concepts based on action/reward functions. There are two important features of reinforcement learning: trial-and-error search and delayed reinforcement [88]. The learners are able to learn an optimal policy for accomplishing goals by trial, error and feedback. In order to optimise reward possibilities, the learner should not just do what it already knows but must explore further options. Generally, exploration must be made multiple times to gain reliable estimates of rewards. The learner that explores more has a higher probability to make the best selection in future, but learners who rarely explore or never explore can struggle to learn useful knowledge from the environment.

The environment is described as a set of states S which map to a set of corresponding actions A. In the Markov Decision Process (MDP), during every discrete time t, a learner will observe its current state s_t then take an optimal action a_t. The environment will respond to a_t with a reward $r_t = r(s_t, a_t)$, and generate a subsequent state $s_{t+1} = \sigma(s_t, a_t)$. In MDP, $\sigma(s_t, a_t)$ and $r_t = r(s_t, a_t)$ just relate to the current state and action pair of a learner rather than a historic state and action pair. The main task of the learner is to maximise the control policy $\pi : S \to A$ to accomplish its task. The control policy defines the learner's choices and methods of action at any given time, it helps the learner take the most appropriate action a_t based on the current state s_t. In order to make precise decisions, a learner keeps an accumulated value $V_\pi(s_t) = \sum_{i=0}^{\infty} \sigma_i r_{t+i}$ by obey the control policy π. The action a_t brings highest $V_\pi(s_t)$, so a learner has to estimate and then re-estimate from the successes and failures over time. In fact, the most important aspect of reinforcement learning is to create the accumulated value for efficiently determining actions.

1.3.1 Q-learning

Q-learning is a model-free reinforcement learning technique. It enables learner to learn an optimal policy from local experience without require established state to action ($S \to A$) policies [88]. Therefore, a learner can explore an optimal action policy in unknown environments. In Q-learning, the experience of a learner is represented by a Q-value ($Q[S,A]$), which is a function of state-action pairs to learned values. By executing different actions, the learner can move from state to state. Performing a particular action a offers the learner a reward r. The main objective of the learner is to maximum its total reward. To achieve this, a learner maintains a table of $Q[S,A]$, where S is the tuple of states and A is the tuple of actions. $Q[S,A]$ represents a learner's current estimate of Q-value. During the learning process, the learner always takes the action which is expected to result in the highest $Q[S,A]$.

The procedure of Q-learning algorithm can be summarised as follows:

- Initialise the table entry Q_{s_0,a_0} to zero when learning process starts.

- Observe the current state,s_i
- Choose an action a_i for that state based on current $Q_{s,a}$.
- Execute the action a_i.
- Receive immediate reward r_i.
- Update the table entry Q_{s_i,a_i} as follows:

$$Q_{s_{i+1},a_{i+1}} = (1-\alpha)Q_{s_i,a_i} + \alpha(r + maxQ_{s_i,a_i}). \tag{1.1}$$

Where α is called learning rate, and controls the extent to which the newly acquired experience will override the past experience. When $\alpha = 0$, the learner will not learn anything while $\alpha = 1$ will make the learner only learn the most recent experience. When solving a stochastic problem, a small constant learning rate (such as 0.1) is used so the learner can find the optimal policy quickly. γ is a discount factor, which decides the importance of expected rewards. When $\gamma = 0$, the learner only considers instant rewards while γ approaching 1 will make the learner pursue long-term high rewards. The introduction of learning rate and discount factor allows learner to decide its future actions according to instant rewards or accumulate rewards.

The Following example illustrates how Q-learning can be used to solve a practical problem. Consider there are 5 rooms (A, B, C, D and E) in a building. Rooms are interconnected by doors as shown in Fig. 1.3. We put a learner in room C and ask it to learn how to get out of this building in an optimal way. In other words, the learner has to reach F from its current location. Fig. 1.4 shows a simplified graph (of Fig. 1.3). Individual arrows represent mono-directional paths from one room to another, and each arrow is associated with a reward value. The paths that directly connect to F have an instant reward 100, the rest of paths which do not directly lead to F have zero rewards. In the context of Q-learning, each room can be considered as a *state* and the learner's movement refers to *action*.

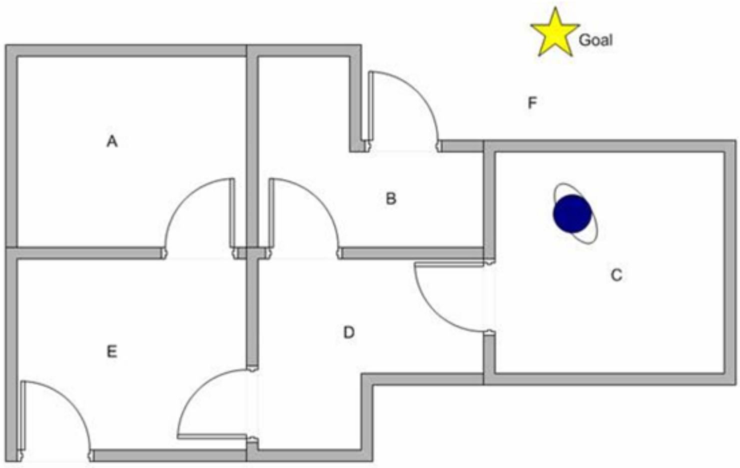


Figure 1.3: Example of Q-learning

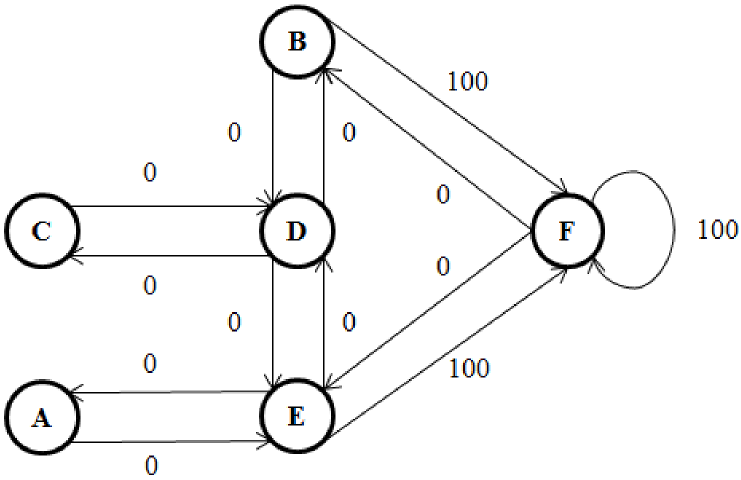


Figure 1.4: Simplified Environment

When the learner is in the initial state C, it can only go the state D since C is directly connected to D. From state D, the learner has three possible actions: either go to state B or E or back to state C. Once the learner reaches state E, the possible actions are go to state A, F or D. Moreover, if the learner is in state B, it can either go to state F or state D. From state A, the learner can only go back to state E. The state diagram and instant reward values can be represented by following matrix Q:

Each row of matrix Q represents a unique state, and each columns refers to a possible action leading to the next state. The matrix Q is initialised to zero prior to the beginning

$$
Q = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array}
\begin{array}{c}
\begin{array}{cccccc} A & B & C & D & E & F \end{array} \\
\begin{pmatrix}
- & - & - & - & 0 & - \\
- & - & - & 0 & - & 100 \\
- & - & - & 0 & - & - \\
- & 0 & 0 & - & 0 & - \\
0 & - & - & 0 & - & 100 \\
- & 0 & - & - & 0 & 100
\end{pmatrix}
\end{array}
$$

of learning process. When the learner stays in a certain state, it updates $Q_{s,a}$ by assigning specific reward value from Q based on (1.1). For example, let assume the discount value is 0.8 and learning rate is 0.1. Therefore, Q_{S_D,S_B}=0+0.1(0.8+max(0,100))=10.08. Accordingly, the Q-value tables can be continuously updated when the learner gains more and more experience through state-action pairs. Matrix Q will eventually reach a convergence point (where the leaner keeps updating its Q-value function but the value of $Q_{s,a}$ remains the same) such as:

$$
Q = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \end{array}
\begin{array}{c}
\begin{array}{cccccc} A & B & C & D & E & F \end{array} \\
\begin{pmatrix}
- & - & - & - & 90 & - \\
- & - & - & 74 & - & 100 \\
- & - & - & 74 & - & - \\
- & 90 & 71 & - & 90 & - \\
74 & - & - & 74 & - & 100 \\
- & 90 & - & - & 90 & 100
\end{pmatrix}
\end{array}
$$

Subsequently, an optimal path to F is obtained by the leaner, which is C→D→B→F.

1.4 The Scope of Thesis

This thesis presents the design and evaluation of intelligent MAC protocols for WSNs with particular focus on addressing the challenges faced by their practical operation.

The ultimate goal is to apply the proposed MAC protocols into broadly dispersed high-node-count applications such as target tracking and event monitoring applications. This thesis starts out by introducing basic knowledge of WSNs and MAC protocols so as to profoundly understand the various techniques and challenges among MAC design. Existing MAC schemes struggle to achieve a high level of channel utilisation with a relatively simple operating mechanism. CSMA/CA-based schemes have been shown to solve the hidden node problem ineffectively and consume significant control overheads [14]. Time-division schemes often require strict time synchronisation and continuous centralised control that may increase the cost of system maintenance [15]. To successfully complete required missions, nodes are supposed to self-organise their operations to deal with underlying challenges resulting from environmental changes.

Recently, a novel machine learning based MAC scheme called ALOHA-Q has been proposed [66]. ALOHA-Q utilises a frame-slotted structure and a Q-learning based slot-selection strategy. It combines the general merits of contention-free and contention-based protocols but eliminates their shortcomings. In ALOHA-Q, a node can manipulate its transmission history in order to form an unique scheduling strategy. The transmission behaviour of nodes starts in a contention-based manner but eventually ends up with contention-free based performance level. Nodes need to expand a certain amount of overhead at the beginning of network operation but they can enjoy the hugely improved throughput performance during the remaining lifetime. ALOHA-Q was originally applied to a single-hop network model and its performance has been examined in both software and hardware platforms. The simulation results from [66] suggested that the throughput of ALOHA-Q protocol is as good as TDMA under saturated traffic conditions, and the average convergence time of the whole system is predictable if the optimal number of slots per frame is selected. The study in [67] has applied the ALOHA-Q into real sensor test-beds, and corresponding practical results also showed the power of ALOHA-Q on improving throughput and delay performance. However, the implementation of ALOHA-Q relies on a known number of active nodes and relatively stable network environment. With a rapid increase of network scale and node populations, a multi-hop based data transmission strategy gains popularity in mainstream WSN applications. This trend brings extra challenges to MAC protocol design, such like the uncertainty of network size, unpredictable node interference range, dynamic network topologies, unstable traffic loads, etc. In this thesis, several novel schemes are presented in order to improve the functionalities of ALOHA-Q. The contribution of this research includes, but is not limited to:

- Designing a distributed and intelligent frame size adaptation scheme for maximising the performance of ALOHA-Q under single-hop and multi-hop conditions.
- Developing a slot swapping scheme for helping ALOHA-Q to overcome hidden nodes problem under multi-hop conditions.
- Introducing a sub-frame adaptation scheme which allows sensor nodes to assign unique scheduling policies to individual incoming traffic flows and proposal of a set of self-organisation functions to sensor nodes for the purpose of achieve optimal performance level in event-driven sensor networks.

1.5 Thesis Outline

The rest of thesis is divided into seven chapters and the structure is presented in Fig. 1.5

Chapter II begins with a brief introductory survey about the evolution of WSNs. The classification and the layered models of WSNs are presented, followed by an investigation into basic wireless sensor node technologies. Several representative WSN applications are introduced. Moreover, design challenges and technical hurdles of WSNs are summarised.

Chapter III places special focus on the Medium Access Control protocols of WSNs. This chapter begins with an exposition about the fundamental aspects of MAC protocols including performance trade-offs, energy efficiency design and MAC taxonomy.

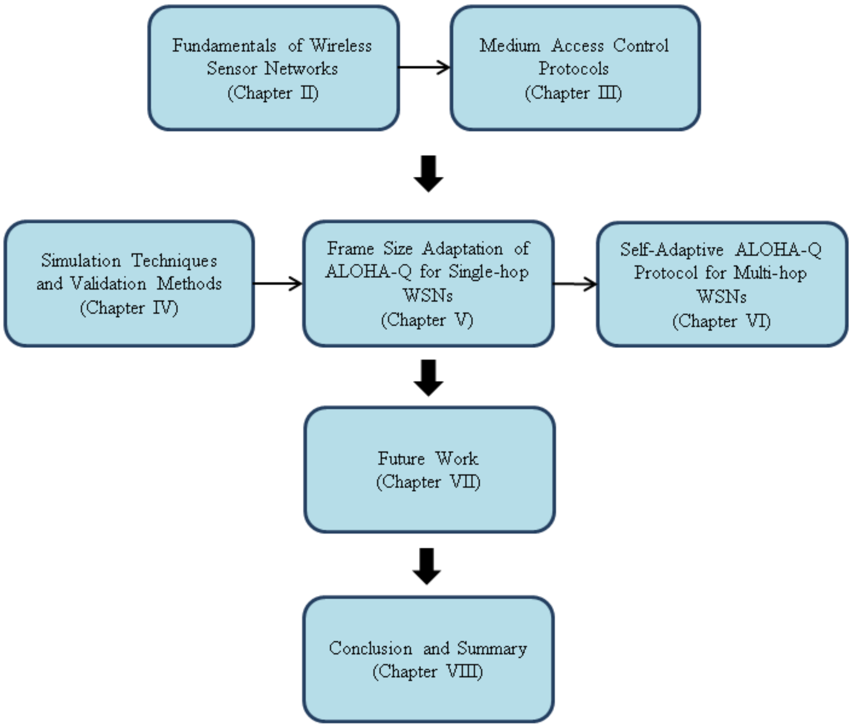


Figure 1.5: Thesis Outline

Several representative MAC protocols are introduced, include the discussion of their pros and cons. Subsequently, a machine learning MAC scheme named as ALOHA-Q is introduced.

Chapter IV covers the discussion of the simulation techniques and the validation methods. Software based simulation techniques and theoretical modelling methodologies are presented, followed by an introduction of the performance measures used to evaluate the proposed protocols.

Chapter V investigates the relationship between frame size selection and throughput performance of ALOHA-Q under single-hop conditions. A frame size adaptation scheme of ALOHA-Q for single-hop networks is proposed and analysed for the purpose of dealing with the uncertainty of node population.

Chapter VI investigates the challenges when implementing ALOHA-Q in multi-hop networks. The impact of the hidden node problem on throughput performance of ALOHA-Q in linear chain networks is analysed and simulated. A slot-swapping technique is proposed to resolve potential traffic congestion caused by the effect of hidden nodes. Besides, a frame size adaptation scheme specifically designed for linear chain networks is also introduced and analysed. The later sections of this chapter introduce a subframe adaptation scheme, which help sensor nodes to achieve collision-free scheduling in cross chain networks. Plentiful simulation results are illustrated to highlight unique features and evaluate the overall performance of proposed schemes and techniques.

Chapter VII provides potential further research to extend this thesis. The frame size adaptation schemes introduced in this thesis can be enhanced to achieve better channel utilisation for event-centric applications. Moreover, the introduction of intelligent duty cycle mechanism and Lifelong Machine Learning (LML) algorithm can further improve the adaptivity and energy efficiency of proposed MAC protocols.

Chapter VIII concludes the work in this thesis. The unique contributions are highlighted followed by corresponding publications developed as a result of the work undertaken.

Chapter 2

Fundamentals of Wireless Sensor Networks

Contents

2.1 A Brief History of Wireless Sensor Networks

A wireless sensor network (WSN) is an infrastructure comprising a group of spatially dispersed sensor nodes, which are responsible for measuring parameters from a specific geographic area [13]. WSNs have become an active research field and intensively studied over the past few decades. During the early stage, related applications were mainly designed to fulfil either military or scientific research purposes, and the corresponding research were limited in universities and institutes. In the early 1980s, the development of WSNs has gained the major impetus from the Distributed Sensor Network (DSN) program of the Defence Advanced Research Agency (DARPA). DSN initialised numerous novel techniques relate to distributed sensing, signal processing and wireless networking. It applied the newly established Transmission Control Protocol/Internet Protocol (TCP/IP) protocols and ARPnet's (the predecessor of the internet) technologies to the context of sensor networks. Based on the outcomes of the DSN program, some well-known corporations like Intel, Boeing, Motorola and Siemens and their academic partners such as Carnegie Mellon University and the Massachusetts Institute of Technology have greatly promoted the development of the sensor hardware, the application software, and the network protocols of WSNs [4]. As a consequence, WSNs have been applied to many fields for various purposes (i.e., environment monitoring, traffic management, healthcare, industrial production). These applications promise tremendous societal benefits, from disaster prevention to productivity increasing. From the late 1990s, advances in semiconductor, distributed computing, wireless communication, and energy storage result in a new generation of WSNs. Today's state-of-the-art WSNs employ inexpensive compact sensor nodes that can complete required tasks in a collaborative way and self-organise their operations [10]. Compared to traditional WSNs, the cost of deployment and the complexity of maintenance are significantly reduced. The evolved WSNs support a diverse range of connectivity technologies and network standards so that sensed information can easily be uploaded to the internet through wireless access points (as Fig. 2.1 shows). This offers a huge improvement in the flexibility of information access. WSNs are finding their way into countless applications in our homes, offices, factories, and beyond, provide an unprecedented way to help people to capture information about various real-world phenomena and convert these information into a form that can be further processed and shared.

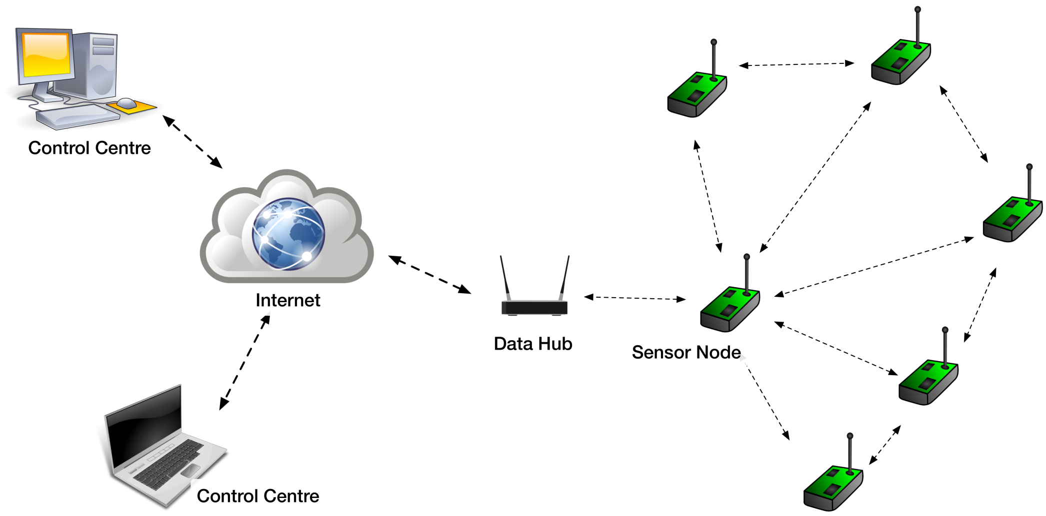


Figure 2.1: Modern Wireless Sensor Network Arrangement

2.2 System Classification

There has been extensive science and engineering progress on WSNs over the recent decade. The high-speed development of WSNs enables a wide range of applications for information sensing and controlling. Conventional WSN applications are distinct from each other in operating scenarios. From an applications perspective, WSN systems can be classified into the following two categories [4]:

- **Category 1 Wireless Sensor Networks (C1WSNs)**: they represent the mesh-based WSN system with multi-hop wireless connectivity between source nodes and forwarding nodes (see Fig. 2.2). Source nodes refer to the sensor nodes in the usual sense, which can capture, transmit and receive information. In practical scenarios, the distance between a source node and the back-end data sink may be too far to establish a direct radio link. Therefore, the messages generated from

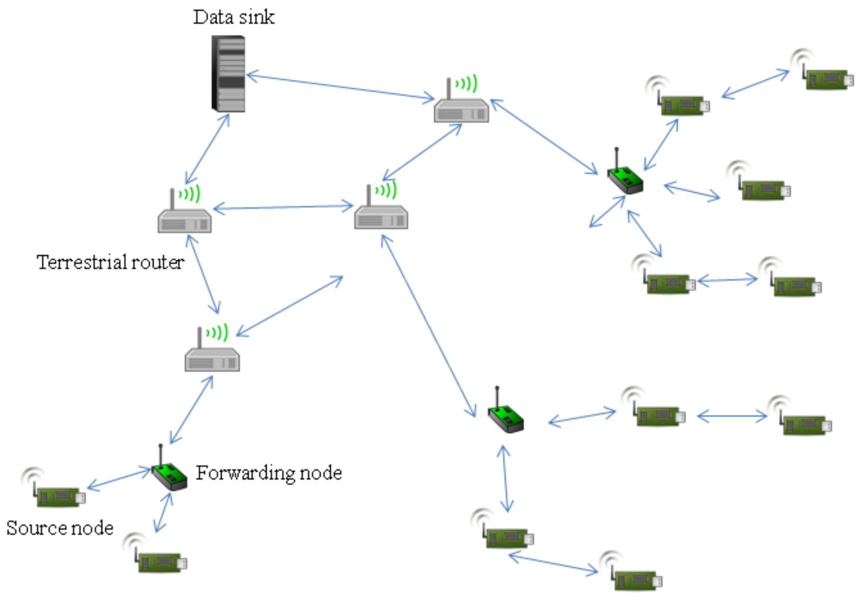


Figure 2.2: Category 1 WSNs

source nodes need to be relayed by intermediate devices called forwarding nodes that might be multiple link hops away from a source node. Forwarding nodes are interconnected via wireless links. Each of them acts as a wireless router that bridges the source nodes and the data sink. They can process and/or compress the messages from neighbouring source nodes, and send these messages to the data sink via an optimal route by using a dynamic routing technique.

- **Category 2 Wireless Sensor Networks (C2WSNs)**: they are networks in which source nodes are one link hop away from each forwarding node (see Fig. 2.3). The major characteristics include: there is no direct link between any two source nodes; Forwarding nodes support static routing technique instead of dynamic routing; Forwarding nodes cannot process and/or compress the sensed message on the behalf of source nodes.

The two categories of WSNs have different scopes. C1WSNs focus the large-scale highly distributed WSN systems. These systems usually cover a broad sensing area, and individual sensors are assumed to self-operate over a relatively long period. Corresponding applications include habitat monitoring and battlefield surveillance. C2WSNs

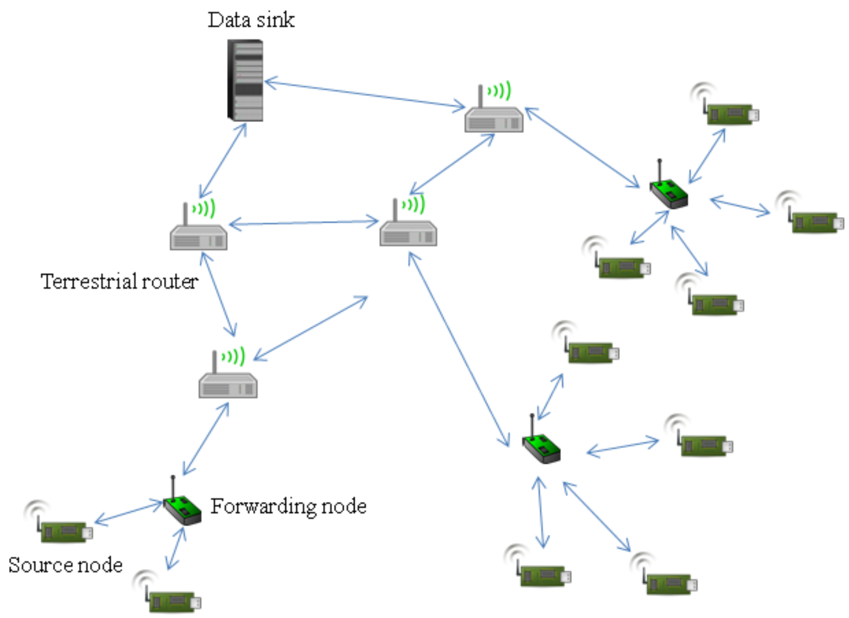


Figure 2.3: Category 2 WSNs

tend to handle the short-range single-hop systems with low traffic loads. Related applications include the Radio Frequency Identification (RFID), Indoor Intrusion Detection, and Wearable Human Health Monitoring.

2.3 Sensor Node Technology

2.3.1 Basic Functionalities

Wireless sensor nodes (also referred to motes) are the fundamental element of a WSN. They provide the functionalities of sensing, processing and communication by executing built-in communication protocols, data-processing algorithms and application programs. The design and implementation of sensor nodes determine the quality of information which can be accessed by the network. As technology-intensive research, the development of wireless sensor nodes were stymied by many technical issues during the early stages. Since the early 1990s, advances in Micro Electro Mechanical Systems (MEMs), SoC, Wireless Communications and Low Power Consumption Embedded technologies has enabled the great development of wireless sensor networks (WSNs). It has become technically and economically possible to manufacture powerful small-sized (According to the needs of the application, the size of a sensor node may vary from the size of a shoe box to a microscopically small particle) and low-cost wireless sensor nodes such as the example shown as Fig. 2.4.

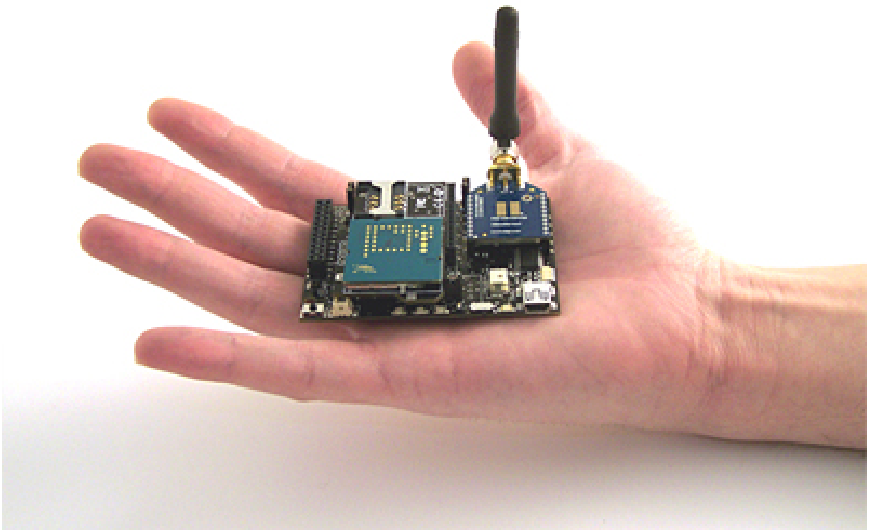


Figure 2.4: Typical Sensor Mote [117]

From a functional perspective, wireless sensor nodes combine physical sensing of parameters with computation and networking capabilities. The main responsibility of a sensor node is cover a geographic area in order to monitor certain parameters [13]. These parameters include but not limited to, mechanical, chemical, thermal, electrical, chromatographic, magnetic, biological, optical, ultrasonic information. Furthermore, sensor nodes have wireless communication capabilities and some logic for data sensing, signal processing and transmission handling. The basic functionalities of a wireless sensor node mainly depend on a specific application, but the following requirements are typical.

- **Determine the value of a parameter at a given locations**: for example, in a environment-oriented WSN, one may need to capture the temperature of the crater of a volcano, sense the atmospheric pressure of aircrafts, or obtain the value of humidity in a room.

- **Detect the occurrence of interesting events and estimate the parameters of the events**: for instance, in a traffic-oriented sensor network, one would like to detect a vehicle is moving through an intersection and estimate the speed and direction of the vehicle.

- **Classify an object that has been detected**: for example, whether a fire disaster occurs in sensing areas by comparing the current values of temperature and smoke particle density with the threshold values.

- **Track an object**: for example, in a military sensor network, track an enemy vehicle as it moves through the battlefield a covered by wireless sensor nodes.

Typical parameters that can be measured by sensor nodes include:

- **Physical measurement**: examples include speed and displacement (accelerometers measure acceleration); radiation levels (Ionisation detectors); acoustic wave's intensity (resonators) and location (Global Positioning System (GPS)).

- **Chemical and biological measurement**: examples include the air composition (infrared gas sensors); body acidity or alkalinity (PH sensors) and elements of drugs (electrochemical sensors).

Some of the design requirements that sensor nodes need take into account include the following:

- For military and security applications, sensor nodes need to support rapid deployment which must be supportable in an ad-hoc manner.

- Sensor nodes may be prone to failure, unattended, untethered, self-powered low-duty-cycle.

- The topology that the sensor nodes need to maintain may change very frequently. the cost of supporting high-capacity and long-range communications may be expensive.
- Sensor nodes may not have global addresses due to the potentially large number of nodes and overhead needed to support such global addresses.
- Sensor nodes usually require in-network processing, even while data are being routed. Typical processing jobs involve signal processing, data aggregation, data fusion, and data analysis.
- Sensor nodes may be deployed in a dense manner. The level of node local interference is not predictable.

2.3.2 Node Architecture

Wireless sensor nodes come in a variety of hardware components, and these components can be categorised into four subsystems: sensing, processing, communication and power storage [8]. Fig. 2.5 shows the architecture of a typical wireless sensor node. The power subsystem and its relationship with the other subsystems are not shown here. The sensing subsystem is the interface between the virtual data and physical world. It is responsible for the measurement and quantification of interested physical attributes. A sensing subsystem consists of single or multiple sensing units and/or analog to digital signal converters. It is through these components that the target information are converted into meaningful discrete digital signals. The processing subsystem can be thought of as a data hub that connects to the rest of subsystems and additional peripherals. This subsystem is used to handle processing and management tasks including data processing and manipulation, short-term data storage, digital modulation, self-organisation and communication control. To fulfil these tasks, the processing subsystem includes one of components from a micro-controller, Digital Signal Processor (DSP), Application-Specific Integrated Circuit (ASIC), or Field-Programmable Gate Array (FPGA) according to its needs [4]. The selection and implementation of each component of the processing subsystem are vital to the performance, cost, as well as the energy consumption of a sensor node. The communication subsystem is mainly composed of a radio transceiver. It is capable of establishing radio connectivity between sensor nodes and to transmit the encrypted data packets output from the processing subsystem. As a energy intensive subsystem, the operation of the communication subsystem is critical to the energy-efficiency of a sensor node as well as the overall efficiency of a network [8]. Finally, the power subsystem which usually contains a built-in battery or energy harvesting component provides sustained energy to all the other subsystems. Clearly, the architecture of sensor nodes varies with cost, application and operating environment. However, typical sensor nodes are comprised of some necessary components which are introduced as follows

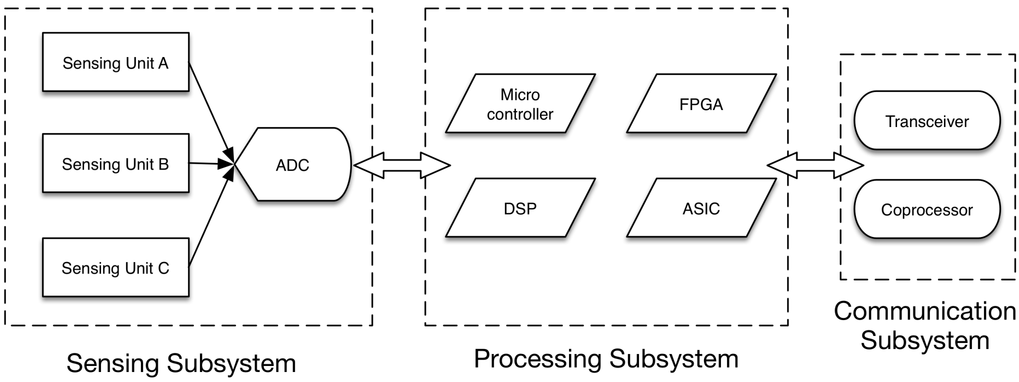


Figure 2.5: Architecture of Typical Wireless Sensor Nodes

- **Sensing units** are devices that produce a measurable response to a change in a physical condition. That converts a form of energy quantity (usually voltage) into analog signals. Typical sensing units can be classified as passive or active units. Passive sensing units tend to be low-energy devices as they just rely on environmental changes. Corresponding examples include the measurement of acoustic waves, humidity, temperature, and vibration. On the contrary, active sensing units transmit probing signals (e.g., microwaves, light, sound, infrared) to a target and then capture the interesting information by detecting the changes in the energy of the transmitted signals. Related examples include ultrasonic measuring and infrared measuring.

- **Micro-controllers** have been used widely by most embedded systems including home appliances, vending machines, elevators, etc. It is a miniaturised programmable computer that performs most functions of a typical personal computer. A micro-controller contains the following units: a Central Processing Unit (CPU) a Random Access Memory (RAM) unit, programmable input/output interfaces and a clock generator. Amongst these units, the CPU is the core element of a micro-controller. It is responsible for handling the locally collected data or the received data from other sensor nodes, and supervise different components of a sensor node. The most prominent advantages of micro-controllers is their programming flexibility, which enables developers to reduce the cost and time when trying to adapt new algorithms and applications to a given sensor node. Sometimes, micro-controllers can be replaced by DSP or FPGA if the applications required intensive computation or high speed data acquisition [13].

- **Radio transceiver** is the most necessary component of wireless sensor nodes. It is a single component but comprises both a transmitter and a receiver that are used to exchange information with other sensor nodes.Currently, frequencies used for wireless sensor systems include 315 MHz, 433 MHz, 868 MHz (Europe), 915 MHz (North America), and the 2.45-GHz Industrial-Scientific-Medical (ISM) band. When a sensor node is turned on, the radio transceiver has four operational states: transmitting, receiving, idle, and sleep [3]. A built-in coprocessor controls in which state the radio transceiver stays automatically. Compared to other onboard components, radio transceivers consume a significant amount of energy during operation time [45]. To reduce unnecessary energy consumption, the transceiver is expected to stay in the idle or the sleep state when a node has no message to send. Therefore, the regulation of the data transmissions is the most important factor that affects the energy conservation of sensor nodes.

- **Power supply** takes charge of providing DC currents to each component of a sensor node. For a typical sensor node, the primary power supply is a built-in battery that could support the operation of sensor nodes ranging from a few days to several years. In addition to relying on its built-in batteries, an energy-harvesting device can also provide energy to sensor nodes from external sources including sunlight, water stream, wind and etc [97]. Currently, energy harvesting has become a key research topic due to its advantages of perpetual power supply, simple device and good portability. However, the operation of energy-harvesting devices has specific requirements regarding node operating environments and would increase the cost of sensor nodes.

Small, low-cost, robust and reliable wireless sensor nodes are needed to enable the wide-scale deployment of practical and economical WSNs, which are expected to become ubiquitous in the future for improving the quality of human life. The current trend of wireless sensor nodes includes several aspects: to enhance the sensitivity, the speed, and the robustness of data acquisition; to increase the functions to respond to more physical phenomenon; To improve the ability to complete tasks in complex and dynamic environments. to address these needs, advanced research in broadband wireless communication, miniaturisation of embedded systems, pervasive computing and energy harvesting can be brought in the future stage [94].

2.4 Communication Protocol Stack

In addition to selecting superior sensor hardware, the most efficient approach to enhancing the performance of a WSN is to design a set of flexible, robust and interoperable networking protocols. A protocol refers to a rule that achieves a specific internal function of a node [96]. To characterise and standardise individual protocols of an open communication system, researchers introduced the concept of Open Systems Interconnection (OSI) reference model. This model divides the system architecture into seven layers: physical layer; data link layer; network layer; transport layer; session layer; presentation layer and application layer [12]. Each layer is composed of a group of protocols that deal with the same set of issues. The inspiration for the OSI reference model leads to a similar layered model for WSNs. Generally, Communication protocols of WSNs can be classified into five layers: physical layer, data link layer, network layer, transport layer and application layer [4]. The main descriptions of these layers are presented as follows.

- **Physical layer**: this layer consists of the basic transmission technologies of a node. It provides the means of generating raw bit streams and transmitting these streams over robust physical links between dispersed nodes. The functions performed by the physical layer protocols including medium and frequency selection, signal detection, carrier modulation, data encryption, etc.
- **Data Link layer**: this layer provides functional and procedural means to transmit data between nodes. The data link layer concerns some key issues like the utilisation of available radio resources and management of the transmission activity. As the interface between the physical layer and the upper layers, the data link layer is also responsible for detecting and correcting errors generated in the physical layer. The protocols of the data link layer are used to provide services related to medium access, error control, timing, and locality.
- **Network layer**: this layer is responsible for transmitting packets between a source node and a destination node while maintaining Quality of Service (QoS). The network layer protocols mainly focus on the services regard to routing, topology management, flow control, and congestion detection.
- **Transport layer**: the objective of the transport layer is to deliver data packets along a established transmission route. Transport layer protocols mainly concern the services related to data dissemination, caching, storage and congestion control. For WSNs, the design of transport layer protocols has unique focus. For example, how to achieve reliable transport in a WSN with dynamic topologies caused by node mobility, failure or power-down.

- **Application layer**: this layer directly connects to the application interface and mainly responsible for the execution of specific application services. As the highest layer, the application layer controls all the functions provided by lower layers. The protocols of the upper layer are mainly used to perform self-organisation tasks, including data processing and aggregation, time synchronisation, and clustering of nodes.

2.5 Current Applications

Through decades of development, envisaged sensor-network applications have been spread over the real world with commercial or military availability. The market scope of WSNs is expanding rapidly since the beginning of 21st century. Many academic institutes and companies have emerged as suppliers of the necessary hardware and software building blocks. Wireless sensor nodes have considerable deployment flexibility. They can be deployed in various sensing fields to detecting interesting events, to track moving targets, to support the operation of intelligent systems. The diversity of WSN applications can be remarkable, ranging from environment monitoring to human health care [10]. In this section, some representative applications are given to provide a better insight into the potential of WSNs.

2.5.1 Environmental Application: Habitat Monitoring

The evolution of wireless sensor networks has enabled new classes of applications that benefit the environmental monitoring domain. In recent decades, many research groups have proposed using WSNs for habitat monitoring. Traditional data loggers for habitat monitoring are typically large and expensive [11]. To obtain accurate results, probes need to be redeployed manually and constantly in different locations of the interest area. They also have to be connected to additional equipment for recording and analysing the collected raw data. Moreover, using probes often result in a ”shadow effect", a situation that occurs when an organism alters its behavioural patterns due to interference in their space or lifestyle. Instead, biologists argue for the small-sized devices that may be deployed on the surface, in burrows or trees. Since the interference is such a significant concern, the sensor must be inconspicuous. They should not disrupt the natural processes or behaviours of human or animal.

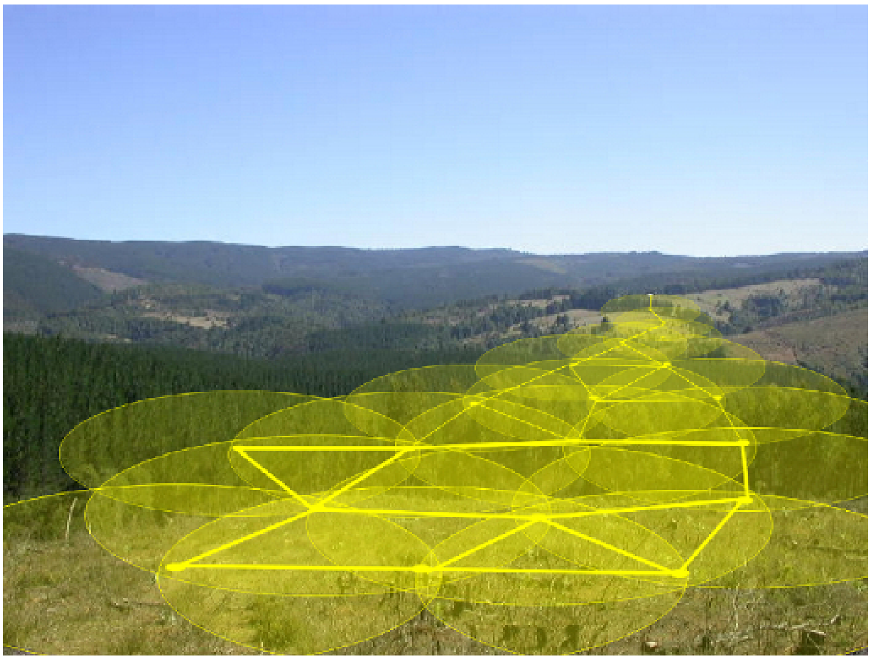


Figure 2.6: Habitat Monitoring Application [118]

WSNs represent a significant advance over the traditional methods of habitat monitoring [9]. Small nodes can be deployed prior to the sensitive period (e.g., breeding season for animals, plant dormancy period). Besides, sensor nodes can be deployed in small areas where it would be unsafe or unwise to frequently reached by the human. A key difference between WSNs and traditional probes or data loggers is that WSNs permit real-time data access without repeated visits to habitats [11]. Moreover, deploying sensor networks is a more economical and efficient method for conducting long-term studies than traditional methods such as data loggers, since sensor nodes can be deployed and left easily in habitat areas. WSNs may organise themselves and store data that may be later retrieved and notify that the network needs servicing. With the help of WSNs, researchers could access more information that often limited by concerns about disturbance or lack of easy access.

2.5.2 Military Application: Smart Dust

WSNs can be an integral part of military Command, Control, Communications, Computers, Intelligence, Surveillance, Reconnaissance, and Targeting (C4ISRT) systems. The clearest example is the Smart Dust project which was developed by UC Berkeley in 1998 [7]. Researchers want to use most advanced technologies from MEMs and SOC with the purpose of producing ultra-small, ultra-light wireless sensors which could float in the air like dust (see Fig. 2.7). Similar to the typical application of WSNs, the smart dust is a complete sensor system. Each dust contains a sensing unit, microprocessor, transceiver and power supply, and all of these components are embedded into a cubic-millimetre package [18].

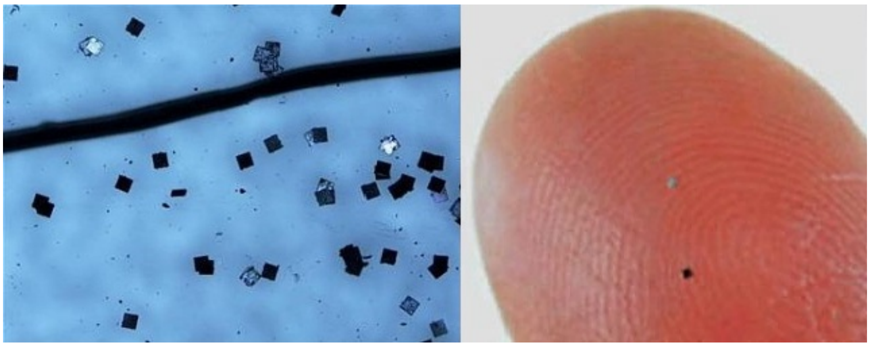


Figure 2.7: Smart Dust Application [119]

The US army believes that smart dust will play a crucial role in future battlefield monitoring systems, due to significant advantages in terms of reliability and fault-tolerance under harsh environmental conditions. The smart dust sensors can be deployed in a war zone by soldiers, vehicles or aircrafts. These sensors are usually randomly distributed and work in a self-organised manner. The performance of the whole system would not be severely affected even if the malfunction occurs in some nodes. If one node goes down another will immediately pick up its task. Commonly, smart dust sensors can be classified as source nodes, sink nodes and target nodes. The source nodes can collect real-time data of interest (i.e. position of enemies) and transmit perceived information to sink nodes through multi-hop communication routes. The sink nodes take charge of the analysis, processing and relaying of received data. Eventually, the data received by sink nodes will be further passed to data hubs out of the sensing area. Compared to traditional technologies, smart dust provides an invisible, low-cost and low-risk solution for gathering the useful data from a whole battlefield.

The prospect of the smart dust project is truly appealing. However, researchers still face a lot of technical issues that may restrict the development of smart dust for a relatively long period. With current technologies, it is difficult to integrate high-end electronic components into the expected cubic-millimetre package, so there will be a trade-off between performance and size. Moreover, power consumption is another challenge because the battery of the sensor is almost irreplaceable. Some of the researchers decide to investigate various low-power network protocols that are expected to help individual sensor nodes to reduce unnecessary power overheads.

2.5.3 Health Application: Wearable Medical Sensors

The use of WSNs in health care systems has yielded a tremendous effort in recent years [24]. Wearable health monitoring systems have gained great popularity during the past decade. These systems consist of multiple wearable medical sensors that can be easily placed in the human body to record their health status. The traditional health monitoring technologies are quite expensive and often feature unwieldy wires between the sensors and the control system. The activity and feeling of patients will be restricted and degraded by wires, so that affects the accuracy of the measured results. However, low weight and low-cost wearable medical sensors can provide a more convenient and affordable solution. Patients will benefit from continuous long-term, accurate, and real-time or near real-time monitoring. In addition, the wearable sensors could be connected with other telemedicine systems through the internet or some other wireless local networks [10]. Patients can make a medical enquiry at home, their physiological measures such as body temperature, heart rate, blood pressure will be transmitted to a computer viewed by remote doctors who then make an instantaneous diagnose and return useful feedback to the patients. The advantages of wearable medical sensors could significantly reduce the time and cost for patients and hospitals.

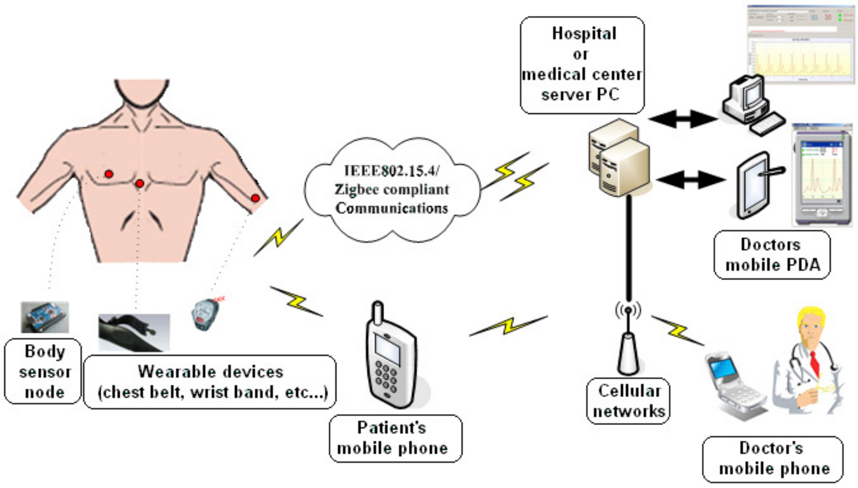


Figure 2.8: Wearable Medical Sensor Networks [120]

2.5.4 Summary

In this section, a set of existing and possible WSN applications were introduced. WSNs have very broad application spectrum. Wherever people want to observe and react to events and phenomena in a particular environment, they can use WSNs. The environment can be the physical world, a biological system, infrastructure or even IT framework. In terms of the market scope, WSNs are expected to see further expansion in the next decade. This expansion not only relates to science and engineering applications but also to upcoming new technologies from various emerging areas.

2.6 Design Challenges

WSNs offer powerful functionality of information sensing, data processing, and communication. They have brought about a plethora of applications and, in the meanwhile, bringing numerous technical challenges with respect to the unique traits of the sensor nodes and networks. The reliable and sustainable operation of WSNs is restricted without overcoming these challenges. The following section shows the major challenges that are shared among typical applications.

- **Energy efficiency**: energy efficiency is the paramount design consideration in WSNs. Wireless sensor nodes usually contain a limited built-in energy supply unit but are required to operate as long as possible in unpredictable environments [4]. To replace or recharge sensor batteries is a difficult effort, especially for large scale networks. In fact, some design goals of sensor networks are to build nodes that are cheap enough to be discarded rather than reused or that are efficient enough to operate only on ambient power sources. In all cases, prolonging the lifetime of sensor nodes is a critical issue. Using low-power units and energy saving circuit design is a popular approach for reducing node power consumption [6]. The study [121] suggests that the power rating of a typical wireless sensor node which is supported by two $2.7V$ *AA* batteries ranges from 10 mW to 1000 mW. Commonly, nodes consume a relatively small amount of energy when they sleep (processor off, sensor off, and radio off), but the energy consumption rises significantly when nodes switch from sleep state to active state (processor on, sensor on, radio on). The major energy consumption sources include communication, processing, status transient, sensor loggings and sensing [122]. These activities are controlled by different on-board units including micro-controller (MCU), radio transceiver, and sensing unit. The following table presents the current consumption for two representative sensor node platforms: IRIS and MICA2.

  Values shown in Table. 2.1 were collected from [123] and [124]. Each mote con-

Table 2.1: Current consumption for IRIS and MICA-2 motes

Mote Type	MCU Active	MCU Sleep	RF Tx	RF Rx	RF Sleep
IRIS	8 mA	8 μA	17 mA	16 mA	0 mA
MICA2	8 mA	<15 μA	27 mA	10 mA	<1 μA

sists of a micro-controller, a radio transceiver, and a sensor board that can be interfaced to specific sensing units according to applications. It can be clearly seen that the radio transceiver of both motes has higher current consumption than their MCUs if they are not in sleep mode. Since $Energy = Current * Voltage * Time$, radio transceiver can be considered more energy intensive when compared to MCU. However, it does not mean that the radio transceiver always the largest power consumer. The actual energy consumption of a transceiver is also determined by a number of factors including duty-cycle and data rate. Besides, some sophisticated sensing units such as infrared gas detectors may consume significant amount of energy if compared with MCU and radio transceiver. When comes to MAC layer, the most efficient way of improve energy efficiency is to properly manage the activities of radio transceivers [8]. To achieve this, the communication of nodes need to be operated efficiently. In some cases, sensor nodes can be powered by energy-harvesting devices such as solar panels or wind turbines. However, the cost of system raises and the deployment environments will be restricted.

- **Hardware performance**: a sensor node may need to fit into a tight module. Due to the restriction in size, the performance of wireless sensor nodes is significantly limited, especially in terms of processing capability (power consumption and cost of CPU), and the radio communication range (antenna size). In realistic scenarios, sensor nodes probably only have finite battery power but need to complete a large amount of tasks during the mission time or as long as possible [16]. Replacing the power sources of sensor nodes in a remote sensing field is usually not practicable. Sensor nodes also have to cooperate with adjacent nodes and adapt themselves to the environmental changes, which requires additional data computation and processing. To conserve energy, typical sensor nodes integrates a low-power and single-channel transceiver. Therefore, the communications between sensor nodes are more easily to be interrupted by noise signals.

- **Traffic pattern**: traffic pattern is an important consideration when designing WSNs. Modelling accurate traffic patterns brings great benefits to network optimisation. For example, better MAC protocols or routing strategies can be designed if the traffic burden among individual sensor nodes is better investigated [125]. According to [126], existing WSN applications can be classified as periodic or event-driven data generation. For periodic scenarios, the traffic arrival process follows either constant bit rate (CBR) where the bit rate [127] is always constant or Poisson distribution [128]. For event-driven scenarios, bursty traffic can arise from anywhere the sensing area when an event occurs. The burst phenomenon of the data generation can be simulated by ON/OFF model where the duration of ON/OFF periods follow the generalised Pareto distribution [129]. The WSNs simulated in this research are assumed to employ the Poisson based periodic traffic generation. The transmission data rate is set to 250kbps which fits the standard of IEEE 802.15.4 MAC layer. The ultimate goal is to achieve close-to-capacity throughput while fully eliminates packet collisions.

- **Operating environment**: sensor nodes may be deployed in harsh, hostile, or widely scattered environments. In such environments, nodes may die over time. The connectivity of network is easily affected by various environmental factors such as high electromagnetic interference, high humidity levels, vibrations, dirt and dust. Therefore, the reliability of a network may be threatened.

- **Topology management**: in some scenarios, sensor nodes are expected to be deployed in a random fashion so that the network topology is hardly to be predicted and managed [39]. For example, nodes could be dropped from a helicopter and the deployment location of each node is totally unknown. In more complex situations, the network topology may ensue, as the nodes may malfunction due to a lack of power or physical damage. To maintain reliable connections between nodes, an efficient mechanism is required to handle the changes of topology.

2.7 Summary

This chapter has introduced the basic concept of WSNs and supportive technologies. Typical WSNs are collections of miniaturised, inexpensive, self-operating computational nodes. These nodes cooperate with to each other and organise themselves into a single-hop or multi-hop network. The unique characteristics of WSNs such as connectivity models of the network, system operating environments, and hardware and software constraints of nodes have brought great challenges to researchers. Understanding these challenges is essential for the design and implementation of WSNs. Amongst all of these technical hurdles, the power consumption of nodes is the most important one since it directly affects the lifetime of entire network. In the following chapter, a technology called Medium Access Control (MAC) will be introduced, with detailed discussions of how MAC can be designed to improve the energy-efficiency of nodes.

Chapter 3

Medium Access Control Protocols

Contents

3.1 Fundamentals of MAC protocols

3.1.1 Introduction

In wireless networks, information are transmitted through various wireless communication media (e.g., radio, optical, sonic, electromagnetic induction, etc). Amongst these media, the radio channel is used by most of the high-speed and wide-coverage wireless networks (e.g., cellular networks, wireless broadband internet, wireless sensor networks and etc) [14]. To avoid communication interference, a single channel can not be simultaneously accessed by multiple nodes within a limited geographical area [85]. In practical sensing environments, the channel availability and capacity of WSNs are usually limited and variable due to the constraints of node hardware and environmental changes. Therefore, sensor nodes within a local area have to share a single channel (as Fig. 3.1 shows). To reduce the conflicts between nodes when they contend for limited communication channels, a suitable channel access scheme should be established. These schemes are called Medium Access Control (MAC) protocols. MAC protocols were originally developed for wired media and later adopted to the wireless media. Technically, they are responsible for the fair allocation of shared channels. More specifically, a MAC protocol determines when and how a node can access the shared medium and send data packets. If a collision occurs, the recipient could fail to receive any information and the sender need to spend extra time and energy on a retransmission. When designing MAC protocols, maximising channel utilisation and reducing unnecessary energy consumption activities will be the paramount considerations. The most effective approach for averting channel contention is to allow sensor nodes have unique scheduling policies which can help them to achieve a collision-free access over a time period [14]. Utilising the benefits of WSNs requires that sensor nodes operate in a distributed and self-organised manner instead of being centrally controlled by management infrastructure such as base stations.

MAC protocols became an active field of WSN research from the 1980s, and there exists considerable academic literature [15]. Previous studies indicate that the most

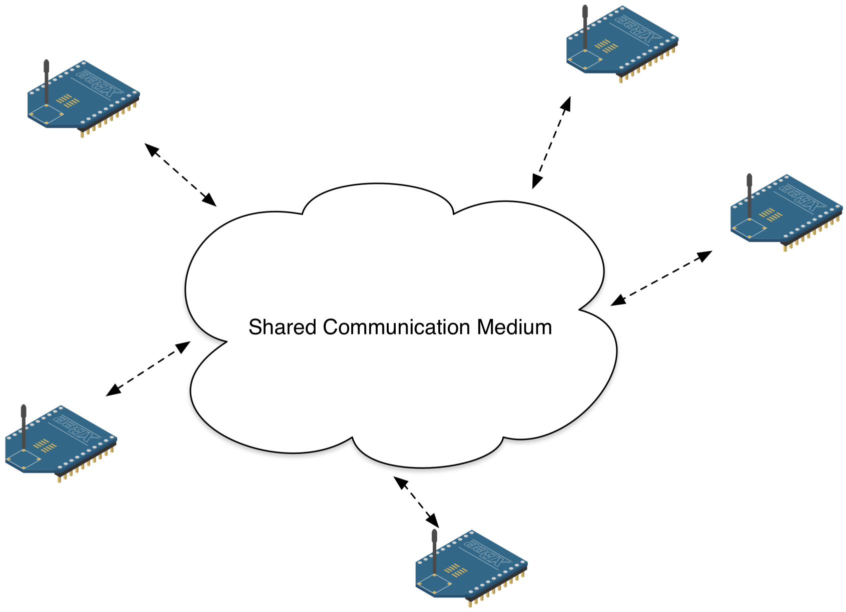


Figure 3.1: Medium Sharing Between Nodes

important source of low channel utilisation is message collisions that result from multiple adjacent nodes sending data at the same time and over the same channel [16]. Collisions can lead to incorrect decoding of received packets, result in additional retransmissions as well as energy consumption. Therefore, the operation time of nodes can be shortened. To reduce packet collisions, researchers have spent great effort and a number of MAC protocols have been proposed for WSNs. Existing channel access methods include centralised approaches such as coordinating transmission times of nodes by sending control messages from certain management nodes, and distributed approaches which allow individual nodes to exchange their schedule information periodically. However, none of these measures are ideal as they all introduce extra control overhead, which causes a degradation of channel utilisation and power efficiency [21]. To achieve a high level of network performance, MAC designers also need to deal with the challenges regarding the power supply, communication resources, computation and storage capacity of sensor nodes. An ideal MAC protocol must strike a balance in between these challenges. The objective of this chapter is to introduce the fundamental information of MAC design for WSNs. Some widely used MAC protocols are introduced, followed by the discussion of their advantages and disadvantages. A state of the art Q-learning based Slotted ALOHA (ALOHA-Q) is introduced. The Q-learning technique is expected to assist sensor nodes in learning a useful channel access policy from their transmission history. The combination of a conventional MAC scheme and reinforcement learning provides great improvement in throughput performance and the energy efficiency of nodes.

3.1.2 MAC Design Trade-offs

The research about MAC performance has been very broad, but there are a few important performance requirements which dominate MAC design. Such requirements usually include collision-avoidance, energy-efficiency, fairness, stability and scalability. They determine the QoS of the entire network and the energy consumption of individual nodes [15]. Due to the characteristics of WSNs, it is rare to see a MAC protocol achieve all of these requirements. Under common circumstances, researchers have to make trade-off decisions on MAC design. The following is a brief discussion of typical MAC performance requirements followed by a discussion of how their importance varies in the context of wireless sensor networks.

- **Collision avoidance** refers to the ability of a MAC protocol to eliminate packet collisions [24]. It is the major determining factor in MAC performance. The biggest difficulty in developing an effective collision avoidance scheme is the trade-off between channel efficiency and control overhead. To achieve collision-free channel sharing, sensor nodes need to either exchange coordinating information between each other or receive global scheduling information from centre management nodes. However, these approaches require nodes to transmit extra messages that do not contribute to the actual payload. Since sensor nodes have limited resources, nodes have to make quality decisions on channel access while reducing control overhead as much as they can. In common scenarios, collisions cannot always be fully avoided. Some MAC protocols introduce a few collisions during the transmission phase in order to achieve lower packet delay or better system throughput, but all MAC protocols are required to avoid frequent collisions.

- **Power consumption** is a central design consideration for MAC designers. Regardless of the improvements in low power consumption hardware design or energy harvesting solutions, an energy-aware MAC protocols is always needed. The cost and size of sensor nodes can significantly restrict the capacity and performance of their power units. Therefore, make efficient use of battery energy is critical to sensor nodes to fulfil their required applications. An efficient MAC protocols can manage the transceiver wisely, such as when to switching it on, make it listen, transmit, wait for or receive acknowledge signals, or re-transmit [16]. All of this is in accordance to MAC protocols been used. MAC designers have to consider all possible solutions to spending available energy in the most efficient way, without sacrificing performance (transmission rate, system throughput, latency, etc).

- **Scalability** is defined as the ability of a sensor network to maintain its performance metrics regardless the number of nodes in the network [26]. Scalability becomes an essential consideration for a network contains a relatively large number (thousands or even millions) of nodes. In typical WSNs, sensor nodes are deployed in an ad-hoc manner and often operate in an unpredictable harsh environment. Therefore, the network size are usually time-varying. Some nodes may die over time, some new nodes may join later and some nodes may move to a different location. To improve scalability, MAC designers should avoid using uniform global schedule. A common approach is to introduce the hierarchical structures that can group sensor nodes into many clusters of different levels (Such as Low Energy Adaptive Clustering Hierarchy (LEACH)). Communication between nodes is often localised within individual clusters. Once node population or distribution is changed, current channel access patterns can be efficiently scaled to adapt to a new environment.

- **Adaptability** refers to the ability of a MAC protocol to deal with fluctuating traffic loads over an extended period of time [15]. Typically, the traffic load of a WSN is time-varying. In some event-triggered monitoring or target-tracking applications, sensor nodes need to handle instantaneous traffic loads since the data generation time is unpredictable. To maintain system stability, the throughput should not decline as the offered load increases. Besides, the number of packets in the transmission queue should be limited. Handling bursty traffic while maintaining stable network performance is a challenging task, especially for large-scale WSNs. Achieving adaptability requires a MAC protocol to respond to the fluctuations of traffic carefully and promptly.

- **Fairness** reflects the ability of different sensor nodes to share the channel equally [21]. A MAC protocol is deemed to be *fair* if it allows competing nodes have an equal opportunity to share channel capacity. When all nodes have homogeneous resource demands, fairness aims to avoid situations where some nodes have excessive priority on channel access. When many sensor nodes cooperate for a single common task, they may have heterogeneous channel demands. That means some nodes may have more data to send than other nodes during a particular period of time. In this case, rather than treating each node equally, the performance of the application needs to be considered as a whole. The channel allocation just needs to be proportionally fair in order to achieve relative resource share of individual nodes.

The performance requirements mentioned above reflect the major challenges of MAC design. Existing MAC protocols struggle to achieve a great balance between all of these requirements. For example, centralised scheduling schemes can significantly reduce a great amount of energy consumption of nodes, but their scalability can be poor. Random access schemes bring an fairness to nodes on channel access but cannot achieve good energy efficiency. For wireless sensor networks, the most significant factors are effective collision avoidance in order to improve energy efficiency, and other attributes are normally secondary. The following section further discusses the characteristics of MAC energy efficiency.

3.1.3 Energy Efficiency of MAC Design

Energy efficiency is the most important issue in the design of MAC protocols. Methods to improve lifetime of sensor networks always attracts a lot of interest from researchers. One of the crucial aspects is to figure out the sources that contribute to energy inefficiency. According to previous studies, a significant amount of power is consumed by a node's on-board wireless transceiver module which is in charge of message transmission, reception, channel listening, etc [57]. The wireless transceiver module commonly consists of four modes: Transmission (TX), Reception (RX), idle listening and sleep. The energy consumption of these modes is in descending order [16]. In trying to maximise the use of energy, sensor nodes should make a reasonable choice about which mode to choose and eliminate potential energy waste sources. Based on a lot of experimental and theoretical analysis, it has been found that the energy waste of sensor nodes mainly includes the following aspects:

- **Message Collisions**: a transmitted packet cannot be decoded by the intended recipient correctly when it collides with other packets. A collision event does not waste energy but collided packets cannot usually be discarded immediately, and a follow on retransmission would consume extra energy plus the backoff time. Therefore, the more packets collide, the more energy is wasted.

- **Idle Listening**: sensor nodes are said to be idle listening if their radio is on while they are neither transmitting nor receiving [11]. A idle listening node can switch to receive mode if it hears transmission, and switch to transmit mode if it has packets to send. Commonly, sensor nodes cannot predict the specific times that they can start to send data packets. In order to receive possible data, nodes have to turn on their wireless transceiver module periodically for sensing the channel availability which may results extra power consumption.

- **Overhearing**: usually happens in the situation where multiple nodes are located relatively close to each other [15]. Some nodes may receive packets that are not destined to them because they are in the communication range of a sending node. In this case, the receiving and decoding RF signal process will lead to unnecessary energy waste.

- **Control-packet overheads**: control packets are used to coordinate the channel access of nodes [23]. Since these packets do not represent useful data, the more control packets are transmitted, the lower the energy efficiency achieved. For example, for some environment monitoring applications in which the system throughput is as low as few hundred kilobits per day, the control messages will represent a large proportion of the total energy consumption.

3.1.4 Taxonomy of MAC Protocols

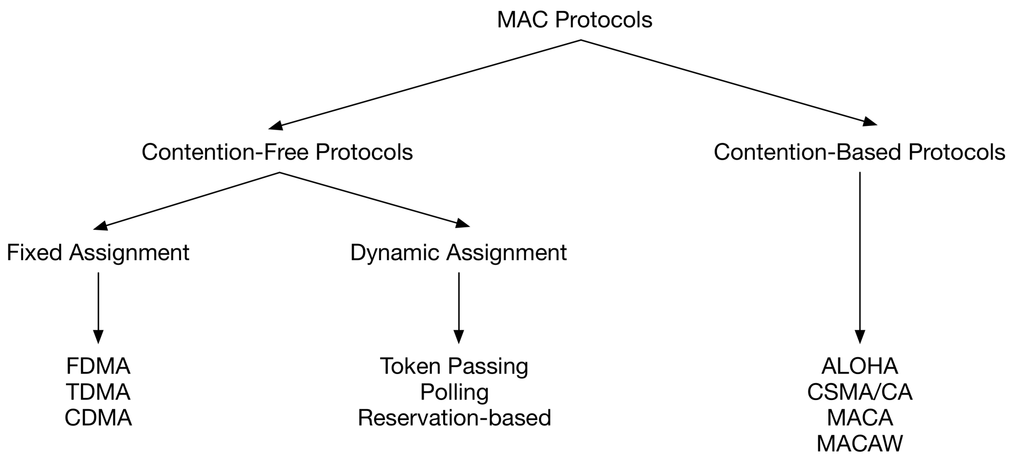


Figure 3.2: Classification of Existing MAC Protocols

Since the concept of MAC was created, researchers have devised and a large number of MAC protocols to solve the shared medium problem. These protocols, by various mechanisms, aim to strike a balance between achieving the efficient resource allocation decision and the overhead necessary to reach this decision. The most widely accepted classification strategy of MAC protocols is based on the resource allocation mechanism. From this angle, common MAC protocols can be defined as either contention-free protocols or contention-based protocols [15]. Contention-free protocols assume the existence of a schedule that regulates access to communication resource between nodes. The schedule will be allocated to individual nodes via centralised control methods. By such manner, a high level of channel utilisation can be achieved. Usually, contention-free protocols can be further divided into fixed assignment protocols and demand assignment protocols. contention-based protocols avoid the preallocation of communication resource to nodes. Instead, all nodes share a single radio channel and they can attempt to access this channel simultaneously. Contention-based protocols aim to provide on-demand channel access to nodes while minimising the occurrence of packet collisions. Fig. 3.2 demonstrates a clear taxonomy of MAC protocols.

Fixed assignment protocols

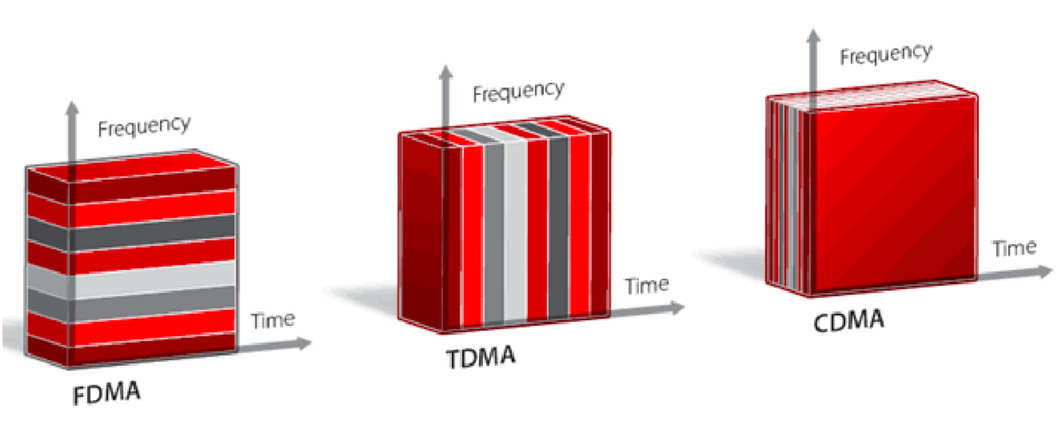


Figure 3.3: Fixed Assignment Multiple Access [121]

Fixed assignment protocols originate from traditional telecommunication systems. They allow nodes to share the available communication medium efficiently by dividing radio resources between nodes [98]. Each node can obtain a portion of the dedicated resource and use it without risk of collisions. Typical fixed assignment protocols include Frequency Division Multiple Access (FDMA), Time Division Multiple Access

(TDMA), andCode Division Multiple Access (CDMA). These access schemes adopt the same basic idea that is to avoid contention by scheduling the nodes in different sub-channels, separated in time, frequency or through the use of orthogonal codes so that each channel contender does not interfere with others.

- **FDMA**

  FDMA divides the channel into a number of sub-frequency bands of radio spectrum, and these are assigned to individual nodes [99]. All nodes can transmit packets simultaneously and nodes are separated by their exclusive sub-channels. Generally, frequency bands are assigned to nodes who request for service by a management station while other nodes cannot share the same frequency band during the period of a transmission. This scheme requires high performing narrowband filters in radio hardware and the ability of a receiver to tune to the channel used by a transmitter. FDMA can be used to transmit both of analog and digital signals. It does not require subscribers to synchronise scheduling since allocated frequency band is available for the entire period of communication. Compared to other fixed assignment schemes, FDMA cannot accommodate a large number of nodes due to its narrow channel bandwidth (usually from 30khz to 200khz). Another major shortcoming of FDMA is that user is impossible receive the data from more than one station at a single point of time. Besides, FDMA requires special filters to avoid interference between any narrow channel, which increases the operation cost.

- **TDMA**

  Amongst various scheduling based protocols, TDMA has attracted the most attention from researchers due to the great balance between channel capacity and system simplicity. TDMA divides the channel into N time slots. Only one user is allowed to transmit during a single slot [100]. The N slots comprise a frame, which repeats cyclically. Traditionally, TDMA is applied in cellular wireless networks such as GSM. A base station is used for allocating global information to each mobile user, and individual users communicate only with the base station. The main advantage of TDMA is the high energy efficiency resulting from the low duty cycle operation. However, TDMA also contains many drawbacks that limit its use in WSNs. In some cases, TDMA requires nodes to form clusters so that nodes in different clusters may share a same time slot. One of the nodes in a cluster is selected as cluster head to play the role of the base station. The cluster-based hierarchy consume a extra control overhead for exchanging scheduling knowledge between cluster heads and sensor nodes since it is difficult to establish peer-to-peer communications between inter-cluster nodes. Besides, TDMA restricts the scalability and adaptability to changes in the numbers of users. When a node population changes, the cluster head has to adjust the frame length or change the slot allocation scheme. The cluster head needs a fairly long time to handle potential changes in topology or density. Under a dynamic environment, the fixed slot allocation scheme can reduce the performance of the network. Furthermore, TDMA relies on global or in-cluster time synchronisation for slot allocation. However, perfect clocks matching over a long period can be a battery consuming task. This may affect nodes to carry out their required tasks.

- **CDMA**

  CDMA is originally used as the access scheme in mobile phone networks. It is more complex than other fixed assignment schemes due to its spread-spectrum technology and a special coding scheme. A CDMA system transmits data signals by combining the signal with a noise-like spreading signal [101]. The hybrid signal is a wide band signal that occupies a larger bandwidth than that required to transmit the original narrow band signal. The introduction of spread-spectrum techniques makes CDMA more secure than FDMA and TDMA as it is hard to capture the original signal from the wide band signal. Moreover, in CDMA systems, all nodes use the same frequency band to perform their transmissions. As FDMA, CDMA allows nodes send packets simultaneously. On the other hand, some nodes can share a same frequency band just like TDMA does. To eliminate the conflicts in share a radio resources, sensor nodes use different codes to separate their transmissions, and they spread signals over a much larger bandwidth than they needed. Signal encoding and decoding management is crucial to CDMA. A receiver has to know the code used by the transmitter. All parallel transmissions using other codes appear as noise.

Demand assignment protocols

In demand assignment strategies, the channel capacity is allocated exclusively to the sensor nodes that have communication needs on a real-time basis [102]. Compared to the fixed assignment protocols, where the capacity of the channel is assigned to every node regardless they are source nodes, relay nodes or sink nodes. Channel utilisation is greatly improved since public radio resources are only shared by the nodes that are ready to transmit instead of the idle or malfunctioned nodes [25]. Similar to TDMA, demand assignment protocols allow the sharing of the unique channel by multiple nodes at different time periods. However, in addition to the data channel, it also needs a control channel that arbitrates the requests for data channel access from active competing nodes. These protocols can be further divided into three major categories: polling, token passing and reservation-based schemes [59].

- **Polling**
  Polling is considered as a centralised demand assignment scheme. It allows a master control node to make queries to each slave nodes about whether it has a packet to send [103]. A slave node sends a request to the master node if it has data to transmit. Accordingly, the master node assigns an amount channel resources (usually through time slots) to the ready node, which can use it without encounter collisions. When the slave node is idle, it ignores the query sent from the master node. The polling process will be continued from the next node. By this means, the competing nodes can access the shared channel based on an equal opportunity. However, the polling scheme consumes a relatively large amount of

control overhead on transmitting the query messages. That reduces the channel utilisation and system throughput.

- **Token passing**

  The token-passing scheme is a distributed demand assignment scheme. To authorise the nodes to access the channel, a special type of data packet called a token is passed between nodes [104]. The token can be used by a single node to transmit data for a period of time. During the transmission phase, each node is continuously monitoring the passing token to determine whether it is the recipient of the token. If a node detects an empty token, it fills the token with data and the address of the destination node and then transmit the token. When a non-empty token is received by a node, it decodes the data, resets the token to empty, and passes it to the next node in the topology. The major advantage of the token-passing scheme is that packet collisions can be fully prevented. However, the channel bandwidth can not be fully utilised if traffic load is light since the nodes have packets to send need to waste an amount of time on waiting for the token.

- **Reservation-based**

  The core idea of reservation-based schemes is to employ small-sized time slots called mini slots for carrying reservation messages [105]. Similar to the TDMA-based schemes, the reservation-based scheme divides time into repeating frames. Each frame contains a fixed number of data slots and mini slots. If a node has packet to send, it requests a data slot by sending reservation message to the central management node via mini slots. In response, the management node assigns dedicated data slot in a frame to the requester. That means that packet collisions only occur when nodes compete in mini slots, each data slot will be allocated to a unique node. Generally, there are two approaches to contend for the mini slots of a frame: fixed-priority and dynamic-priority. In fixed-priority schemes, each contending node will be assign a unique mini slot. On the contrary, dynamic-priority schemes require individual nodes to contend the limited number of mini slots. Moreover, in reservation-based schemes, management nodes can schedule the data slots according to the priority levels obtained from the reservation requests. For example, the delay-sensitive data can be assigned an urgent time slots while ordinary data will be queued for a period of time. Therefore, the channel can be utilised more efficiently. Despite the enormous advantages in terms of channel utilisation, reservation-based protocols still face some issues due to their centralised control nature. This determines that scheduling based protocols are only suitable for small scale, many to one or cluster based sensor networks. In comparison,

Contention-based Protocols

Contention based protocols are widely adopted in most of the distributed WSNs applications due to several significant advantages [59]. Firstly, nodes can freely use the channel according to their need without obeying any command and protocol complexity will not raise as the network size increases. As a consequence, protocols can scale more easily to the fluctuations of node population and system traffic load. Besides, contention protocols can be more flexible as topologies change. They are inherently distributed and directly support peer to peer communication. Moreover, contention-based protocols do not require strict global or regional time synchronisation. By comparison with contention-free protocols, contention-based protocols have great advantages in applications which require relatively simple maintaining mechanism, low throughput, good scalability and adaptivity. Most existing contention-based protocols employ a Carrier Sense Multiple Access/ Collision Avoidance (CSMA/CA) mechanism for the purpose of reducing packet collisions [23]. The core idea of CSMA/CA is Listen Before Transmission (LBT). This idea assists a node to listen whether a current channel is occupied then decide on when to transmission [24]. Fig. 3.4 illustrates the workflow of CSMA/CA. Prior to any data transmission, a node senses the carrier to determine whether the public channel is currently busy. If a node detects an idle channel, it transmits incoming packets immediately or goes to a back-off stage and restarts carrier sensing with a certain probability [25]. Otherwise, if a current channel is occupied, the node waits a random time then continues to sense a channel.

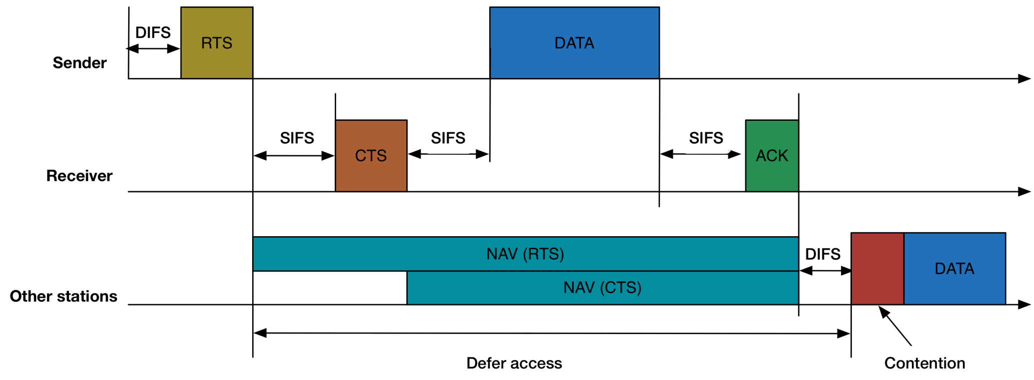


Figure 3.4: Contention Process of CSMA/CA

3.2 Examples of MAC protocols

This section introduces several representative contention-free and contention-based MAC protocols for WSNs, along with detailed discussions regard to their merits and drawbacks.

3.2.1 Pure ALOHA protocol

Pure ALOHA is the earliest random access wireless MAC protocol developed by University of Hawaii in the early 1970s [50]. This protocol offers complete random channel access. The data transmissions of nodes are independent of the current activity of the shared channel. The basic idea of the ALOHA protocol is very simple: when a node needs to transmit a message, it just transmits [49]. Once a transmission is completed, the node listens for a period of time called the guard band. If the message is successfully received, the node will receive an acknowledgement (ACK) message from the receiver within the guard-band of current time slot. In the absence of an ACK message, the node assumes the packet was lost due to errors caused by background noise or packet collision, and schedules a retransmission. If the number of retransmissions exceeds a threshold M, the node refrains from retransmission and reports an error. In simulations, M is set to seven which is defined according to the IEEE 802.11 specification.

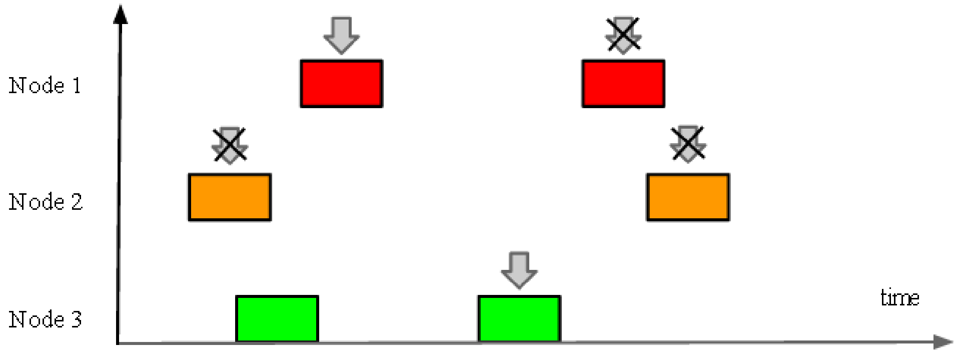


Figure 3.5: Contention process of Pure ALOHA

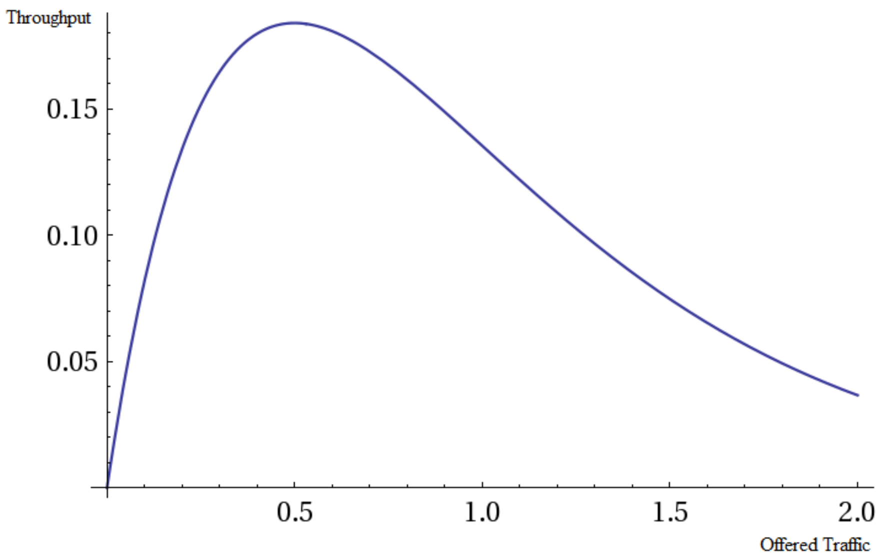


Figure 3.6: Throughput performance of Pure ALOHA

Fig. 3.5 and Fig. 3.6 illustrate the explicit workflow and throughput performance of Pure ALOHA respectively. Each node is allowed to transmit whenever it has a message to send. Under low traffic load conditions, nodes can access the channel within a relatively short period and throughput keeps raise as offered traffic increases. The maximum throughput is only 18.4% of the channel capacity, and it can be achieved when offered traffic reaches 0.5 Erlangs. Furthermore, there is no centralised control mechanism in Pure ALOHA, thereby offering great scalability since nodes can be added or removed easily. The main disadvantages of Pure ALOHA is that it does not check the availability of channel before a transmission. However, as traffic load increases, the probability of successful transmissions degrades severely due to the increasing probability of collisions.

3.2.2 Slotted ALOHA protocol

Slotted ALOHA is an enhanced version of Pure ALOHA [48]; it is widely used in passive WSNs such as RFID networks due to its simplicity and relatively good throughput performance under low traffic conditions. This protocol assumes all nodes are synchronised during the transmission phase [51]. The communication channel is divided into many equal-length discrete time slots whose duration is equal to the combination of the transmission time of a packet and a guard band. The channel access mechanism of slotted ALOHA is presented in Fig. 3.7. Contrary to Pure ALOHA, a node is only allowed to send messages at the beginning of each time slot. Therefore, data messages sent by different nodes will not partially overlap with each other, thereby a collision can occur only at the beginning of a time slot [51].

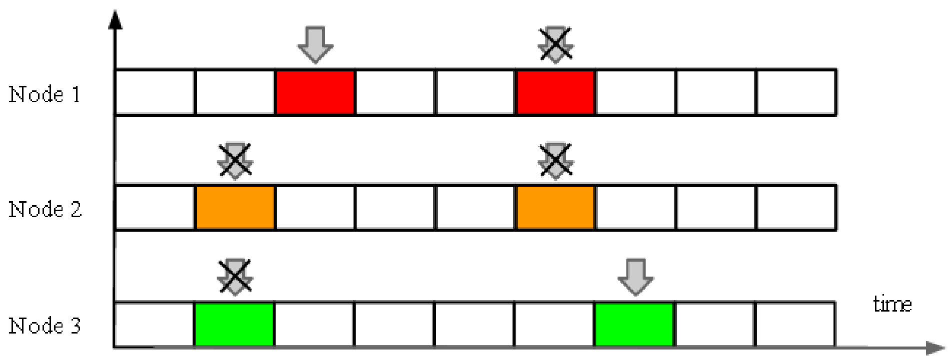


Figure 3.7: Contention process of Slotted ALOHA

Restricting data transmission to slot boundaries results in the significant decrease in

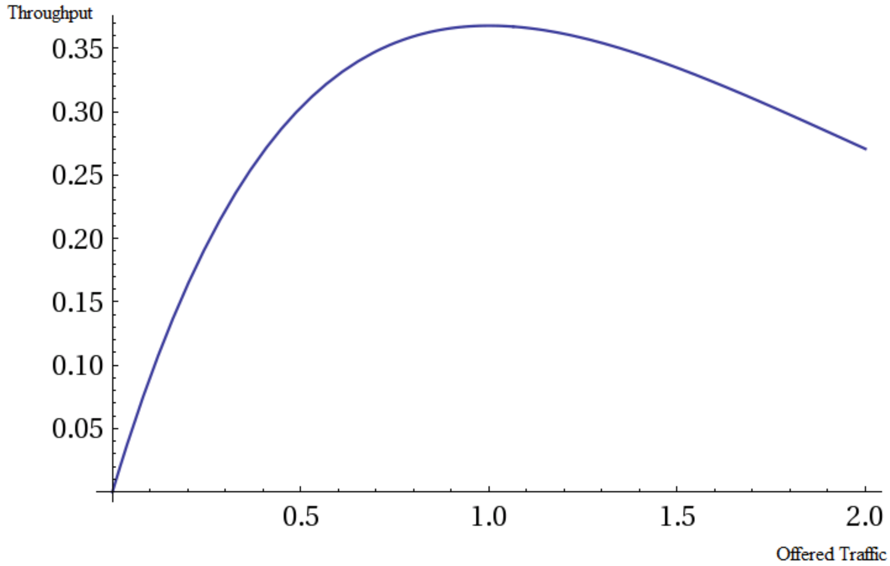


Figure 3.8: Throughput performance of Slotted ALOHA

the probability of collisions. Therefore, the average number of retransmissions can be decreased and the system throughput can be increased. Fig. 3.8 shows the throughput performance of Slotted ALOHA. Despite the improved performance, Slotted ALOHA remains inefficient under medium and heavy load conditions due to the absence of the ability to sense the channel availability. According to the theoretical analysis [52], the maximum throughput is 36.8% of the channel capacity, and it can only be achieved when offered traffic loads reaches one Erlang (represents the continuous use of the shared channel). The poor collision resolving capability of Slotted ALOHA limits its popularity in the applications which require high channel utilisation.

3.2.3 Sensor-MAC protocol: S-MAC

Sensor MAC protocol (S-MAC) was the first MAC protocol explicitly designed for WSNs. The objective of S-MAC is to reduce the energy consumption resulting from major energy wasting sources while obtaining a high level of channel utilisation and scalability [19]. This is achieved by allowing nodes to operate in a low-duty-cycle. Specifically, a periodic wake-up/sleep strategy is applied to each node (see Fig. 3.9).

The communication channel is divided into repeating frames, and each frame is formed by a complete sleep and active cycle. A node sets a wake-up timer and switches to the sleep mode for the specified period. During the sleep mode, the node turns off its wireless transceiver module in order to save energy. At the expiration of the timer, the node wakes up and checks if it has data traffic to deliver.

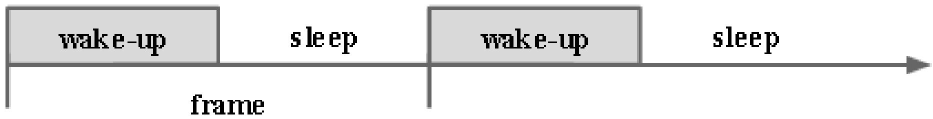


Figure 3.9: S-MAC wake-up and sleep modes of operation

To reduce control overheads, S-MAC requires neighbouring nodes to coordinate their wake-up/sleep schedules [17]. For example, a group of closely located nodes switch to sleep and active at same time. To achieve this, S-MAC introduces a synchronisation phase where nodes can broadcast their wake-up/sleep schedule to neighbours. Therefore, each node can obtain a schedule table that contains the schedules of its neighbours. Periodic exchange of schedules improves the energy efficiency of nodes. However, the average packet latency will be increased if nodes strictly follow the schedules of their neighbours, and data packets will be further delayed if nodes overhear the schedules that are not destined to themselves. To address this issue, S-MAC introduce the technique referred to adaptive listening. Upon transmitting a schedule, a node exchanges a Clear to Send (CTS) or Request to Send (RTS) packet to the neighbour. When an overhearing node in the next hop along the transmission route receives the CTS or RTS packet, it ignores the previously received schedule and schedules an extra listen period.

Similar to many other random access schemes, S-MAC uses a CSMA/CA-based mechanism to regulate channel access of contending nodes. In addition to the physical carrier sensing, a virtual carrier sensing procedure is introduced to nodes via the Network Allocation Vector (NAV), which represents the remaining time until the end of current data packet transmission. The NAV will be decremented as time passes and it finally

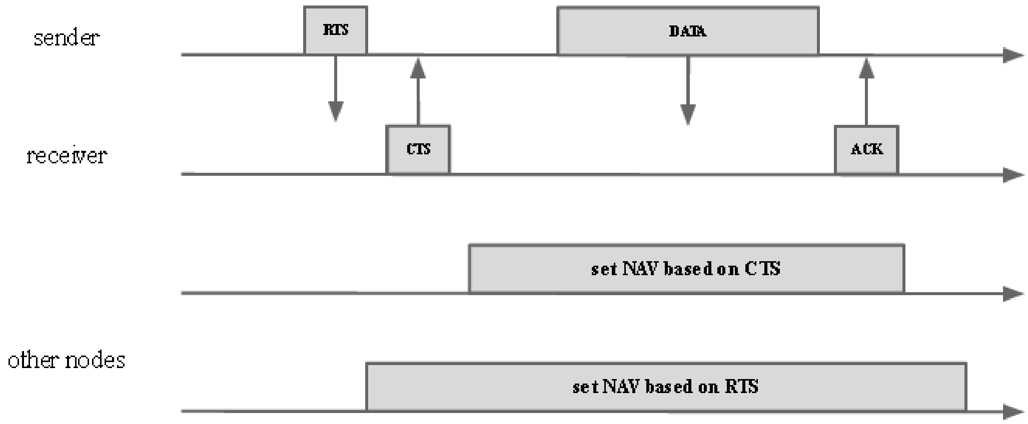


Figure 3.10: Contention process of S-MAC

reaches zero. Fig. 3.10 presents a simple sender and receiver communication of S-MAC. A node attempting to send a packet must first sense the channel. If the channel is idle, it issues an RTS packet and waits for the CTS packet from the receiver. Once a CTS is received, the node sends a data packet. During the handshaking process, The CTS and RTS packets will be broadcasted to the neighbouring nodes that are listening to the channel. These neighbours decode the NAV from a field of the CTS or the RTS packets then enter to sleep mode until the NAV reaches zero. The communication is complete when the sender gets an acknowledgement packet from the receiver. Compared to traditional MAC schemes, S-MAC employs many useful techniques improve the lifetime of nodes and overall performance of a WSN. However, the major shortcoming is that the duration of wake up/sleep time cannot be dynamically changed by nodes according to network traffic [18]. That degrades the efficiency of the channel access.

3.2.4 Timeout MAC protocol: T-MAC

Time-out MAC protocol (T-MAC) was proposed for addressing the fixed wake-up/sleep period of S-MAC [22]. T-MAC shares many mechanisms with the S-MAC, such as wake-up/sleep schedules, CSMA/CA-based channel access, RTS/CTS handshake, and frame structure. Fig. 3.11 illustrates the contention process of T-MAC. During active phase, a node stays in wake-up mode at the beginning of each frame and check if it has a task to do. When a node has a packet to send, it sends the packet and then goes to sleep until the beginning of the next frame. Besides, a node can also goes to sleep if no event has occurred for a predetermined period called TA. This design helps sensor nodes to dynamically adjust the duration of wake-up/sleep modes according to traffic fluctuations.

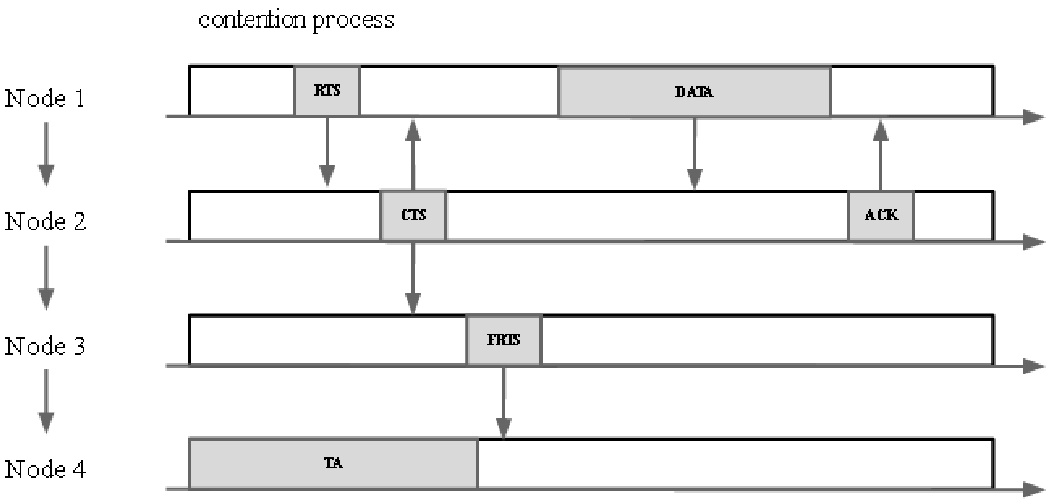


Figure 3.11: Contention Process of T-MAC

By allowing nodes to end their wake-up periods prematurely, T-MAC partially breaks the synchronisation among nodes, which results in the early sleep problem [20]. This problem occurs when a third hop node, expected to receive an ongoing packet from the second hop node, prematurely goes to sleep. T-MAC copes with this by using the FRTS (Future Request To Send) frames sent to the third hop node before its TA timer expires. Consequently, the third hop node remains active and then receives the next transmission instead of receiving it in the next active period. In variable traffic loads, T-MAC saves more energy than SMAC does. However, this is achieved at the cost of an increased latency and thus reduced throughput. Although T-MAC improves on S-MACs energy savings, it still suffers from the main problem of the extra cost of maintaining common wake-up/sleep schedules via the exchange of SYNC packets

[20].

3.2.5 Zebra MAC protocol: Z-MAC

Zebra MAC protocol (Z-MAC) is a MAC protocol that combines the strengths of TDMA and CSMA while offsetting their shortcomings. As a hybrid MAC protocol, it operates CSMA under low to medium traffic conditions and switches to TDMA under heavy traffic load conditions [58]. When a WSN is deployed, Z-MAC runs a Distributed Slot Allocation Algorithm (DRAND) to assign time slots. The objective of DRAND is to avoid the hidden node problem where a node is visible from its downstream receiver, but not from other nodes communicating with that receiver. DRAND ensures that no two nodes within the two-hop communication neighbourhood are assigned to the same slot. To further increase channel utilisation, a time frame rule is designed to adapt the local frame size to a nodes local neighbourhood size. To combine CSMA with TDMA, a node is allowed to contend for sending if a slot is not used by the owner. Consequently, Z-MAC performs as CSMA under low contention and it possesses high channel utilisation capability as TDMA under high contention.

However, the hybrid design faces challenges in a dense network. When some nodes have data to send, they have to contend for slots that are assigned to their neighbours who have no traffic and these contentions are synchronised in each slot. In every slot, a sender has to wait for a certain amount of time to ensure that the slot is abandoned by the owner. Each receiver also has to stay awake to check whether it is the target receiver. As a result, the slot stealing method introduces nontrivial additional energy consumption.

3.2.6 IEEE 802.15.4 MAC layer

IEEE 802.15.4 is an industrial standard defining the physical layer and the MAC layer for low-rate wireless personal area networks (LR-WPANs) [3]. The MAC layer of the IEEE 802.15.4 standard supports many home and industrial WSN applications which require low to medium data rates and moderate average delay. The MAC layer has a relatively simple operating mechanism that enables the power consumption of devices on a low level. To achieve flexible large-scale deployment and deal with the needs of application requirements, the IEEE 802.15.4 MAC layer embeds several unique features. In the following, a detailed description about these features is presented.

- ***Support for a vast number devices and various network topologies***

  The IEEE 802.15.4 MAC layer places physical devices into two categories according to their hardware complexity: Full Function Device (FFD) and the Reduced Function Device (RFD) [108]. These two types of devices differ from each other in function and embedded standards. An FFD is equipped with high-performance computing capability and adequate memory to deal with all the specifications required by the MAC layer standard. It serves three different roles including the network coordinator that chooses critical parameters of network configuration, the router that decides transmission route and forwards data between different devices across the networks, the end devices that capture intended information from environment and send the information to coordinators or routers. Compared to the FFD, the RFD is a device with reduced functionalities, and it only acts as the end device.

  According to the types of devices, the IEEE 802.15.4 MAC layer supports three general topologies: star-based, mesh-based, or cluster-tree-based [106]. In a star-based topology, one of the FFD devices will be selected as a coordinator whose responsible for initiating network parameters and managing other devices. The remaining devices in the network act as end devices that are only allowed to communicate with the coordinator. The mesh-based topology supports multi-hop connectivity between nodes. This topology divides devices into three types: coordinator, routers, and end devices. The routers can bridge end devices and the coordinator, and relay information from any source device to any destination device by using a table-driven routing algorithm. The cluster-tree-based topology is the combination of the other two topologies. It divides a network into up to 255 clusters of up to 254 nodes each. A single cluster is formed by star-based topology and neighbouring clusters are connected in a peer-to-peer manner. With such deployment, the cluster-tree-based networks can achieve a high level of fault tolerance and self-organisation.

- ***Embed a superframe structure for controlling the operational duty cycle of connection devices*** An optional superframe structure (see Fig. 3.12) is supported by the IEEE 802.15.4 MAC layer. A superframe contains 16 equal-length time slots. The first time slot of a superframe is used to transmit a beacon message that aims to synchronise the attached devices and coordinate communications. The rest of the slots can be accessed by contending nodes based on the CSMA/CA scheme during a Contention Access Period (CAP). The slots contention have to be started from a beginning of a beacon and finished by the end of current CAP.

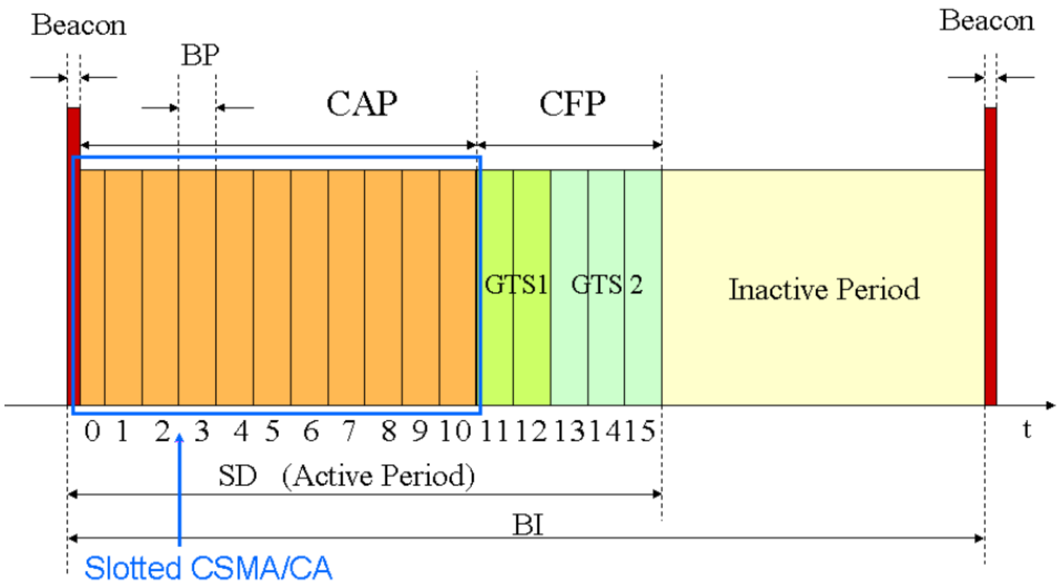


Figure 3.12: Superframe Structure of IEEE 802.15.4

- ***Support for switching between contention- and contention-free based channel access*** To balance the device population and bandwidth requirements of supported applications, the IEEE 802.15.4 MAC layer enables both of the contention-based and the contention-free channel access. During a Contention-free Period

(CFP), the network coordinator allocates at most seven contiguous time slots of a superframe to some active devices. These slots are identified as the Guaranteed Time Slots (GTSs). When a device has a packet to send, it sends a request to the coordinator for reserving GTSs. The request consists of number of desired GTSs, the data type, transmission or reception time, and some other information regard to data exchange. Once a request is received, the coordinator checks the availability of GTSs and issues a denial or a confirmation message. When the device gets a successful confirmation message, it sets up a specific timer and waits for the subsequent beacon message from the coordinator. When the number of allocated GTSs is not sufficient to cover the transmission requests of active devices, the coordinator automatically switches to the CAP mode. As mentioned above, network devices are allowed to contend for time slots during the CAP, using a non-persistent slotted CSMA mechanism. The selection of operation modes is mainly based on traffic types. The IEEE 802.15.4 MAC assumes three different patterns of traffic: periodic data which characterises the traffic generated by nodes when they switch between active and sleep modes, intermittent data which is often generated by stimulus of an application, and repetitive low-latency data which requires dedicated time slots to ensure low latency [107].

- ***Employ an energy-efficient sleep management scheme for a prolonged battery life*** To extend the lifetime of remote sensors that activate on a regular basis and report information when certain events occur such as motion detectors, the IEEE 802.15.4 MAC layer supports a beaconless mode. In this mode, the network coordinator does not allocate dedicated or random-access time slots to devices. Instead, the network devices can access the channel by using a non-persistent un-slotted CSMA mechanism. When a device needs to transmit data, it waits for a random back-off period and then senses channel availability. If the channel is busy, the device waits for another back-off period and increases the number of attempts by one. Once the number of attempt exceeds a limit, the device reports an error to the upper layers. The un-slotted access manner and absence of beacon message help the devices to operate in low-duty-cycle mode while maintain the connectivity to the network. Therefore, devices are expected to achieve extended battery life with minimum power consumption.

3.3 ALOHA-Q

The design of MAC protocols for WSNs has unique drivers. An ideal MAC protocol must first be energy efficient in order to prolong the lifetime of the network. In addition, the MAC protocol is expected to be scalable to handle the changes in system size, node density, topology and etc. Finally, the access fairness, packet transmission delay, throughput, are also essential attributes in the design of MAC protocols. Currently, most of the existing MAC protocols are unable to strike a balance between these attributes. Recently, a novel intelligent MAC scheme named ALOHA-Q (Q-learning based Slotted ALOHA) has been introduced. ALOHA-Q is a hybrid of Slotted ALOHA protocol and Q-learning technique. Instead of relying carrier sensing, ALOHA-Q offers an intelligent collision avoidance mechanism: Q-learning based slot-selection. The core idea of this mechanism is to allow individual sensor nodes to form a collision-free scheduling policy by evaluating their transmission history [66]. Compared with most existing MAC protocols, ALOHA-Q has advantages in simplicity, low control overhead and its on-demand transmission strategy. Thus ALOHA-Q is a new approach wich as potential to combine best features of contention-free protocols and contention-based protocols but eliminates their drawbacks. However, the implementation of this protocol relies on many assumptions about node population, interference model, and some other network parameters. This impede the application of ALOHA-Q in practical scenarios. The objective of ALOHA-Q is to achieve efficient channel allocation under a random access basis. In ALOHA-Q, the transmission behaviour of nodes starts in a random access manner but the channel performance eventually ends up with contention-free access. In other words, sensor nodes just need to consume a relatively small amount of overhead to learn a optimal transmission schedule during an initialisation stage, but they can benefit from the improved overall performance during the remaining operation time.

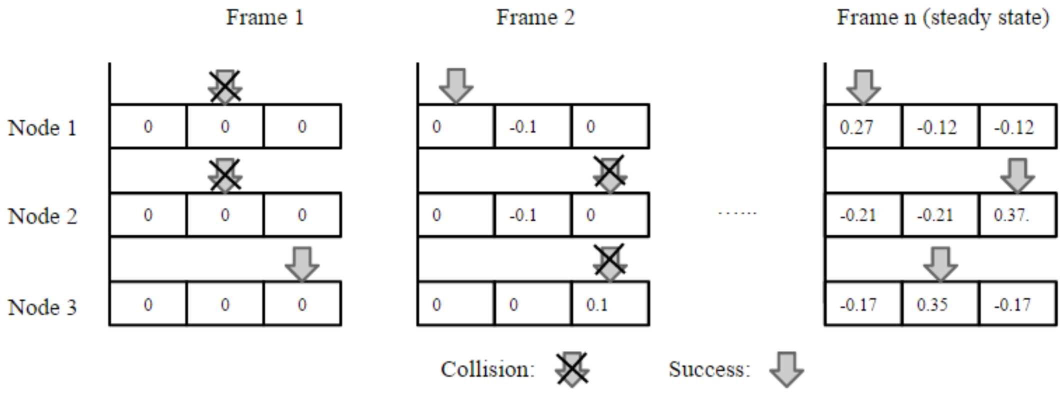


Figure 3.13: The Workflow of ALOHA-Q

ALOHA-Q adopts framed slotted ALOHA as the baseline MAC scheme. The channel is divided into repeating equal-length time frames and each time frame is composed of many equal-length time slots. A node selects a slot in a frame if it has packets to send and each node is allowed to transmit only once in each frame. After a successful transmission, a small sized acknowledgement packet will be sent to the sender. If the sender fails to transmit a packet, a retransmission will be made within the next frame. Compared to Slotted ALOHA, a significant advantage of ALOHA-Q is that the sensor nodes do not always transmit packets on a purely random basis. Instead, they infer transmission was success or failure based on acknowledgement packets sent by the receiver. These are converted into positive or negative rewards which are accumulated according to slot position using a Q-learning algorithm. Subsequently, the optimum slot is identified through this Q value-reward function. By this means, the transmission experience informs a slot selection policy which can be constantly reinforced during transmission stage. Eventually, in a steady state, all nodes are able to find dedicated slots which will not induce collisions. Fig. 3.13 illustrates the general workflow of ALOHA-Q.

The procedure for ALOHA-Q is as follows:

Each node assigns a weight value (Q) of zero to each slot prior to the first transmission attempt. After each transmission attempt, the weight value tables are updated according to receiver's acknowledgement.

A node sends a packet in a randomly chosen slot in the frame for the first transmission attempt. If it receives an acknowledgement packet, the corresponding Q is updated with a positive reward; otherwise a negative reward is used.

From the second transmission attempt, a node chooses the slot with highest Q, based on the current Q table. If multiple slots have the same highest Q, a random slot amongst them is chosen.

The Q value is updated as following equation:

$$Q = (1-\alpha)Q' \pm \alpha r; \tag{3.1}$$

Where α represents the learning rate, Q' refers to the iterated weight value and the r denotes the reward value. The weight Q value usually ranges from -r to r.

To some extent, a weight value (Q) represents a node's transmission experience in a specific slot. The higher the Q, the more reliable the slot. If a node sends a packet in the slot with the highest weight value, this packet will most likely to be successfully received. By this means, sensor nodes can continuously modify their slot selection policy. As the number of transmission attempts increases, system is expected to reach a steady state, where all active nodes are expected to find their unique contention-free slots. Once the system steady state is reached, nodes can exploit their dedicated slots and packet collision can never occurs.

3.4 Summary

This chapter has reviewed Medium Access Control technology for wireless sensor networks. The early chapters introduce several fundamental aspects of MAC protocol including trade-offs of MAC design, primary energy waste sources of nodes, and classification of MAC protocols. The following sections have presented several representative MAC protocols, discussed their benefits and shortcomings. The design of a MAC protocol should seriously consider specific application requirements such as operating environment, coverage area, traffic pattern, network topology, and etc. Contention-based protocols represent the future trend of MAC protocol for large-scale randomly deployed WSNs, but the performance of existing protocols still need to be further improved. In the final section, a machine-learning based the MAC protocol called ALOHA-Q protocol is introduced. ALOHA-Q provides a high level of channel utilisation while reducing the consumption of control overhead. In chapter V and VI, several important schemes and techniques will be introduced to improve the feasibility of ALOHA-Q in realistic scenarios.

Chapter 4

Simulation Techniques and Validation Methods

Contents

4.1 Introduction

Since the concept of WSNs was created, researchers and engineers made a good effort on developing advanced protocols and algorithms to optimise the overall performance of WSNs. With the enlargement of network scale and increasing application requirements, it has become very difficult to design and implement a MAC protocol purely based on hardware experiments and theoretical calculations [113]. Running experiments on real testbeds can increase the design cost. It also leads to great difficulty in data statistics and analysis since the expected theoretical performance can be affected by environmental factors such as weather, temperature, terrestrial obstacles, and background noise. On the other hand, theoretical calculations and analysis rely on many assumptions about system models and implementation environments so that the reliability and accuracy of results can be limited [109]. These issues and challenges have brought software-based simulation into the mainstream of designing, developing and evaluation of MAC protocols. Software-based simulation, by definition, is a method of running single instances of complex systems [109]. It has become an essential tool for researchers to test new applications, protocols, schemes and algorithms. Through software simulation, researchers can quickly establish WSN models and evaluate desired performance metrics. In addition, software-based simulation helps researchers to easily compare different schemes under the same network model or examine the performance of a single scheme under different network topologies by adjusting configuration parameters. Compared to traditional approaches, software simulation has its incomparable advantages in cost economy, efficiency and flexibility that make it to an most widely accepted way to conduct research into WSNs [110]. In this chapter, an overview of the simulation techniques and mathematical validation methods are given. Some critical performance measurements that are used to estimate the performance of protocols and schemes introduced in this thesis are also presented.

4.2 Discrete Event Simulation

Modelling WSNs in software can often be challenging. Many simulation paradigms have been developed to enable researchers to achieve parallel programming in simulations because of sensor nodes may execute different activities in progress simultaneously during their operating period [111]. Among these paradigms, Discrete Event Simulation (DES) paradigms have been widely used to simulating the behaviour of sensor nodes. DES can be considered as a process that codifies the operation of a complex system as a discrete sequence of well-defined events [112]. In briefly, DES defines a list of pending events that correspond to different activities of sensor nodes. The whole simulation period is divided into tiny time fragments. Once simulation starts, the system state will be updated based on a set of predefined events occurring in individual time fragments. A typical DES involves three different phases. In the first phase, system checks the chronological events list then jump to the nearest time fragment. The second phase is to execute all events that that unconditionally occur during current time fragment (refers to B-events) . The remaining conditional events (refers to C-events) will be executed during the third phase. The three phases will be repeated until simulation ends. By such means, DES can easily simulate concurrent jobs (such as transmission, reception, processing) running on different sensor nodes. In addition, DES provides dynamic memory management, which can add new events and drop old events in a same simulation scenario according to system requirements. To reduce the complexity of coding, debugger breakpoints are provided in DES, users thereby can check the code step by step without disrupting the programme operation.

4.3 OPNET Modeller

This section introduces a powerful simulation tool and analyser: OPNET Modeller. OPNET Modeller is an industry leading discrete-event network modelling and simulation software developed by MIL3 Company [28]. OPNET aims to be an efficient and powerful WSN simulator that is also easy to use. A variety of network devices or equipment can be simulated in OPNET. It has a good support with a graphical interface and object-oriented modelling. As a discrete event simulator, it supports interactive running and debugging tools, the graphical parser, and the dynamic observer [82]. All of these powerful features provide OPNET with the capability for accurate modelling, simulation and analysis of wired or wireless networks. Currently, OPNET is adopted by many leading communication equipment vendors, leading communications operators, military and institutions, universities and large corporations. Researchers can use it to simulate various protocols or applications of WSN. On the other hand, OPNET is large and complex. In order to use it efficiently, researchers may need spend much time to understand the plentiful tutorials, handouts and books.

The main features of OPNET include:

- OPNET is designed with object-oriented technologies. According to different requirements, each simulation attributes can be precisely and manually configured,
- OPNET provides a variety of the most common components and modules of communication networks and information systems.
- OPNET contains a powerful performance analyser, which can automatically collect corresponding simulation results in real-time.
- The simulation efficiency of OPNET is higher than other network simulators since the parallel simulation model is introduced.

The structure of network modelling can be divided into three parts: the network domain, the node domain, and the process domain [29]. Each domain uses a single paradigm to perform corresponding tasks that bring outstanding flexibility and scalability to the modelling process [80]. The following parts will introduce each domain in detail.

Network Domain

The Network Domain comprises objects that represent physical or logical parts of a network model. A typical network domain (see Fig. 4.1) is mainly composed of the following components: subnets, communication nodes, communication pipelines (invisible for wireless networks) [32]. A subnet can be imagined as a container that contains a group of nodes and communication pipelines. According to different application requirements, a large-scale network can also be divided into many subnets.

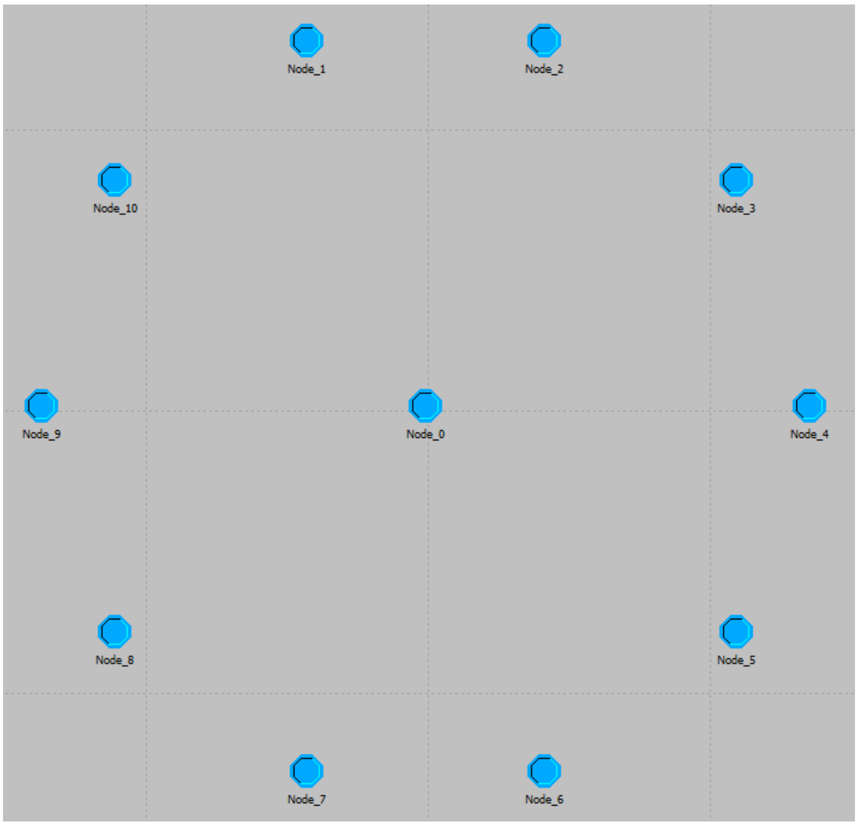


Figure 4.1: The Network Domain

A hierarchical structure of a network is formed when a subnet includes another or multiple subnets. This structure can be further expanded as needed. In addition to the objects, the primary attributes of subnets include the location, range and mobility. Based on these attributes, subnets can be classified as fixed subnets, mobile subnets or satellite subnets. For the simulations of wireless sensor networks, it is best to apply the fixed or mobile subnets since typical sensor nodes are commonly immobile or semi-mobile after deployment. Besides, the number of users, topology, and other macroscopic parameters about the WSNs could be defined in this domain.

Node Domain

As its name suggests, the node domain is used to define the inner structure of nodes [31]. From the functional perspective, sensor nodes can be represented as modular sys-

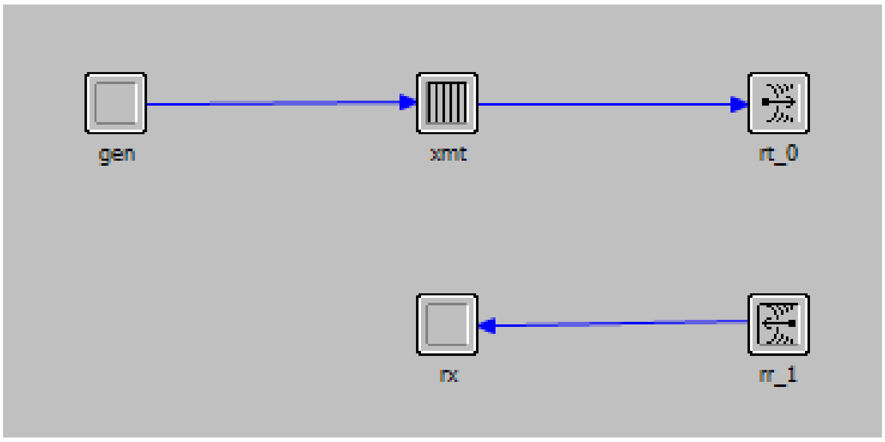


Figure 4.2: The Node Domain

tems. A node is considered as a combination of several modules, each module offers a specific function such as processing, transmission and reception. There are three types of modules in the node domain: processors, queues, and esys. These modules offer essentially the same capabilities with regard to their general behaviour and most of their physical resources. Queue modules provide special support for organised packet storage by enabling users to define internal sub queues in which packets can be inserted and sorted, and from which packets can be extracted according to a general, user-defined method. Esys modules provide queueing abilities and the ability to specify an ESD model for co-simulation. Modules can be linked with two types of pipelines. The first type is called packet stream that completes the transportation of packets from the source module to the destination module. Another type is called a statistic line, which transfers data values between individual modules. A module can also be regarded as a particular hardware or a layer of a protocol. As Fig. 4.2 indicates, a typical wireless sensor node can be divided into five modules: a transmitting processor module, a receiving processor module, a packet generator module, a wireless transmitter module, and a wireless receiver module. Data packets are created by a packet generator and sent to a transmitter process module through the packet stream. The transmitting processor takes charge of data transmission according to predefined algorithms. When a packet comes into the wireless receiver module, it will be forwarded to the receiving processor that can be used to process the incoming data packets and collect the final statistic data.

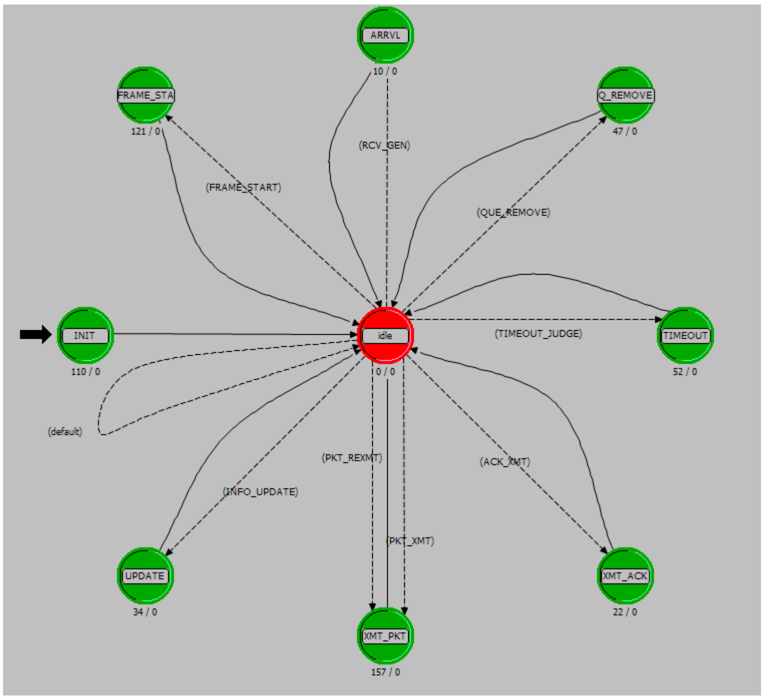


Figure 4.3: The Process Domain

Process Domain

The process domain is the lowest level of in the system design hierarchy. It is where the detailed protocol mechanisms are designed and implemented [30]. Fig. 4.3 illustrates a typical process domain. The process domain are defined by a computer language called PROTO-C, which provides a flexible programming ability to model a wide range of systems. A single process domain comprises of multiple interconnected states. These states represent a specific activity the node may perform. The evacuation of individual states follows a predefined sequence, and only one state is active during any particular instant. When a state is active, the corresponding codes are executed by the system until the end of current state. The process-switching is driven by events. When an event is actually delivered to a process, it is termed an interrupt. Once a process is interrupted, it will be invoked to allow it to take some action in response to the interrupt. OPNET manages the transfer of packets by implementing a series of computational procedures that form a complete radio transceiver pipeline. Each computational procedure defines particular aspects of link behaviour. The transceiver pipeline contains a fourteen-procedure radio link model (see Fig. 4.4) [34]. The sequence of the computational procedures and their interface are standardised for each type of link [33]. By this way, OPNET Modeller provides an open and modular architecture for implementing link behaviour.

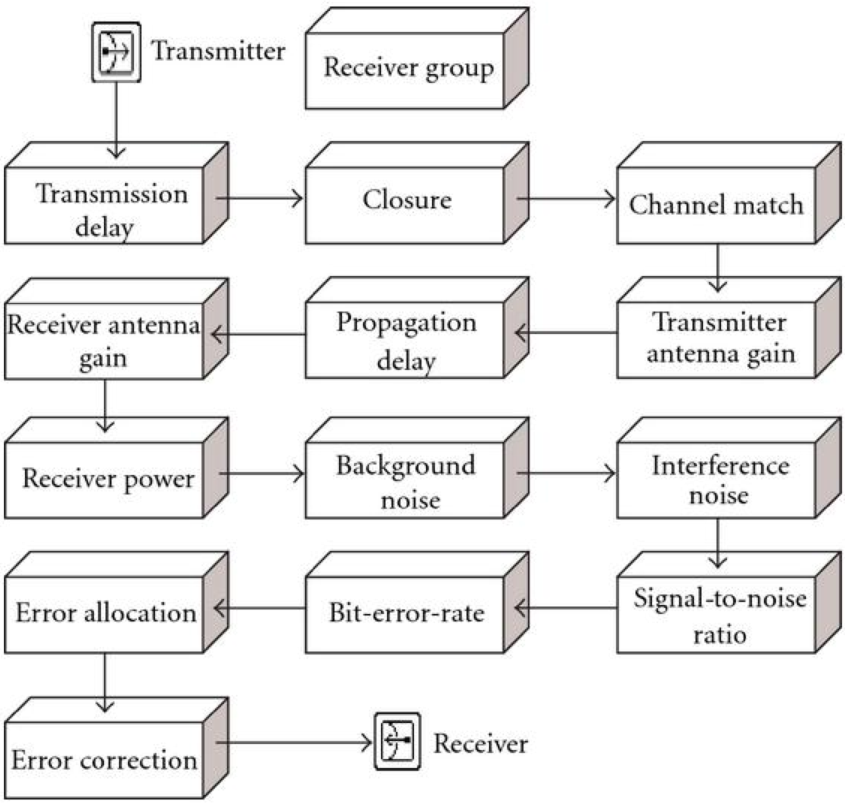


Figure 4.4: The Radio Transceiver Pipeline

4.4 Performance Measures and System Assumptions

The estimation of the simulation results generated from OPNET relies on a number of important performance metrics introduced at follows.

- **Offered Load** represents the average amount of data which trying to be sent by the network. Let G stand for the offered load and S refer to system throughput. G includes all the successfully transmitted packets and collided packets, Obviously, G is bigger than S in most cases and equals to S only if no collision occurs. It is noteworthy that that G could be much greater than 1. For example, $G \geqslant 1$ means each node may have multiple packets to transmit within a frame period.

- **Throughput** refers to the average numbers of packets delivered by the network over a predefined period of time. Let's use S to represent system throughput. When S=1, the channel is fully utilised, that means all the data packets can successful pass through the channel, and there have no interval between the sending of each packet. The relationship between offered load (G) and throughput (S) in the steady state is $S = G * P$, where P is the ratio of actually delivered packets to total transmission attempts. In chapter V and VI, we measure throughput values over different time periods, where each period starts at the beginning of simulation until different end points.

- **Channel Access Delay** is an important performance metric that represents the timeliness of packet delivery. It is the time duration since a packet is generated from a node until the packet is successfully transmitted.

- **End-to-end Delay** is the time duration since a packet is generated from a node until it is successfully received by the intended sink node.Node population, channel contention level, network density and many other interference related factors have huge impact on end-to-end delay performance.

- **Channel Efficiency** denotes the ratio of throughput to offered load. It represents the maximum amount of traffic can be successful transmitted on the channel. Commonly, channel efficiency relates to packet delivery rate. In this thesis, we define a same packet delivery rate for transmitter and receiver.

- **Channel Capacity** represents the maximum amount of data can be successfully transmitted through a channel. Since the transmitter and receiver usually have a same data rate, the channel capacity is equivalent to the maximum data rate of sensor nodes which is usually defined as 1 Erlang in this thesis.

- **System Convergence Time** represents the period of an initialisation process. It corresponds to the learning behaviours introduced in the following chapters. Simply speak, convergence time is the period which all nodes in a network can find their dedicated collision-free slots. Once the system is converged, the network will benefits from the improved QoS during the rest of operation time.

In this research, we do not measure the energy consumption of nodes directly since the main objective of proposed MAC protocols is to achieve optimal scheduling. However, these protocols contains a set of novel schemes that are designed for eliminating potential energy waste sources including packet collisions, over hearing, hidden nodes problem, and etc. Consequently, the enhanced energy efficiency can be reflected by the improvements on some other performance metrics such as increased throughput, reduced system convergence time and channel access delay. When conducting simulations in OPNET, the hardware features sensor nodes of a network are assumed the same, and nodes are supposed to always maintain uniform transmission power so as to achieve the same physical transmission, reception, and interference range. To test the maximum theoretical performance of proposed protocols and eliminate the potential influence of environmental factors, sensor networks scenario considered in this thesis obeys following assumptions:

- The wireless channel is ideal without errors.
- Time clocks are perfectly synchronised across the entire network. During the transmission phase, each node only relies on their internal clock.
- The optimal transmission route is obtained by each sensor node before it transmit a packet. The selection of route is based on the minimum number of hops between source and sink.

- Some characteristics including internal node processing, propagation delay, frequency dependent path loss, signal fading, collision result in background noise are not considered in this thesis.
- The packet arrival process is treated as *M/M/c* (*c* stands for the number of source nodes) queue model [114].

4.5 Validation Methods

In order to examine the correctness of simulation results, some analytical models are carried out in this thesis. In this research, the validating tasks are performed by Matlab software that is widely used in science and industry areas. Matlab is designed by MATHWORKS Inc, it provides powerful mathematical and professional toolboxes, Researchers can use them to design and test the performance of various protocols and algorithms. MATLAB can be used to obtain theoretical results regard to specific performance metrics. These results will be compared with the corresponding results generated from simulations. The comparison of outputs from OPNET and Matlab can help researchers to identify the potential issues and problems of protocol design.

4.6 Summary

The attributes of WSNs and the complexity of hardware implementations make the evaluation of MAC performance very challenging. In this chapter, an exposition of software-based simulation approaches are presented, and followed by the introduction of the Discrete Event Simulation paradigm. The OPNET Modeller which is used for examine all the performance of MAC protocols proposed in this thesis was introduced, including its unique features and the general workflow of system modelling. Multiple performance metrics used to evaluate proposed MAC protocols are described. The validation methods are discussed at final part.

Chapter 5

Frame Size Adaptation of ALOHA-Q for Single-hop WSNs

Contents

5.1 Introduction

The previous chapter introduced several representative MAC protocols for WSNs. By comparing their merits and drawbacks, it can be found that these protocols either struggle to strike a balance between channel utilisation and network scalability (Pure ALOHA, Slotted ALOHA, Z-MAC) or require an exchange of control messages causing significant overhead (Bluetooth, Zigbee, S-MAC, T-MAC). To simultaneously achieve high level of channel utilisation, flexible channel access, a simpler operating mechanism, a reinforcement learning based ALOHA-Q MAC scheme has been proposed. The main advantage for ALOHA-Q comes from the Q-learning based slot selection algorithm that allows sensor nodes to independently and intelligently learn a collision-free channel access policy. [66] suggest that the throughput performance of ALOHA-Q under single-hop conditions is extremely close to the maximum channel capacity. In this respect, ALOHA-Q outperforms most existing MAC protocols.

ALOHA-Q adopts Framed ALOHA as the baseline MAC scheme it thereby retains the original advantages of random access protocols. Based on a Q-learning algorithm, sensor nodes can reinforce their slot selection behaviours according to the transmission feedback (acknowledgement messages) from intended receivers. Once the learning process is finished, the system will stay in a steady state where each node has obtained a dedicated time slot in a frame. As a consequence, the energy waste caused by packet collisions are eliminated. Compared with conventional ”non-intelligent” MAC protocols, ALOHA-Q resolves channel contention by fully utilising the local information of nodes, this can intensively reduce the system control overhead. However, due to the frame-based transmission mechanism, the number of time slots per frame (frame size) needs to be manually preconfigured when implementing ALOHA-Q. To some extent, the performance of ALOHA-Q relies on the selection of an appropriate frame size [68].

In this chapter, we model the ALOHA-Q protocol in single-hop networks and then investigate the impact of frame size on several important system performance metrics including throughput, end-to-end delay, and system convergence time. To realise the maximum performance of ALOHA-Q, a distributed frame size adaptation scheme is introduced. The objective of this scheme is to assist individual nodes to autonomously learn the optimal value of frame size. Corresponding theoretical analysis and simulation results are presented in this chapter for the purpose of demonstrating the effect of the proposed scheme.

5.2 Modelling of ALOHA-Q in Single-hop Network

5.2.1 Scenarios and Parameters

In this section, the single-hop based system models and network parameters used to evaluate the ALOHA-Q will be introduced. Single-hop network topology can be found mainly in the category 2 WSNs (see Chapter I). This topology usually represents a one-hop subnet or a single cluster for hierarchical WSNs. We consider a single-hop network in which all source nodes transmit packets to a single sink node. Perfect time synchronisation will be maintained across the network before transmission starts, and then each node relies on its internal clock to count the beginning of each frame and time slot during the transmission phase. Moreover, the single-hop network discussed in this chapter follows four important assumptions:

- Each source node is within the communication range of the sink.
- The sink node can successfully receive one packet during each time slot.
- Compared to the time that packets wait in sub-queues, the signal propagation delay will be insignificant.
- Internal process time of nodes is not considered.

Table. 5.1 shows a full set of network parameters.

Parameter	Value
Time Slot Length	0.0044 sec
Data Packet Size	1060 bits
ACK Packet Size	36 bits
Transmission Data Rate	250 kbits/sec
Receiver Data Rate	250 kbits/sec
Maximum Retry Limit	6

Table 5.1: Simulation Parameters I

The parameters listed in this table remain the same for all the simulations demonstrated in this chapter and Chapter 6. The simulation period is set to long enough in order to cover the period of the slot leaning process. To show the optimal performance of ALOHA-Q, most performance metrics presented in this chapter are calculated during the system steady state after converge. In some cases, we measure performance metrics throughout the entire simulation period. Figure. 5.1 demonstrates the structure of a single time slot used in simulations. The largest ratio of a slot is occupied by the payload which is the cargo of a data transmission. The payload is used to carry of a complete data packet. Once a packet transmission is finished, the time reaches the guard-band area that is used for receiving the acknowledgement (ACK) packets from the intended receiver. A guard band occupies a relatively small portion of a entire time slot. During the period of a guard band, each sender hears the response from the receiver. Therefore, the duration of guard band should not be less than the duration to receive an ACK packet. At the start of each guard band, a timer will be started for checking whether the original packet is successfully delivered to the destination node, which is named the timeout checkpoint. If the time reaches this checkpoint, the node needs to check if it received an ACK packet. If not, the previous packet might have collided with other packets, and hence the node will retransmit the failed packet. For the simulations carried out in this research, each data packet contains a 1048 bits of payload and 8bits of source address, 4 bits of destination address, and 4 bits of packet type. Each ACK packet contains 4bits of source address, 8 bits of destination address, and 4 bits of packet type.

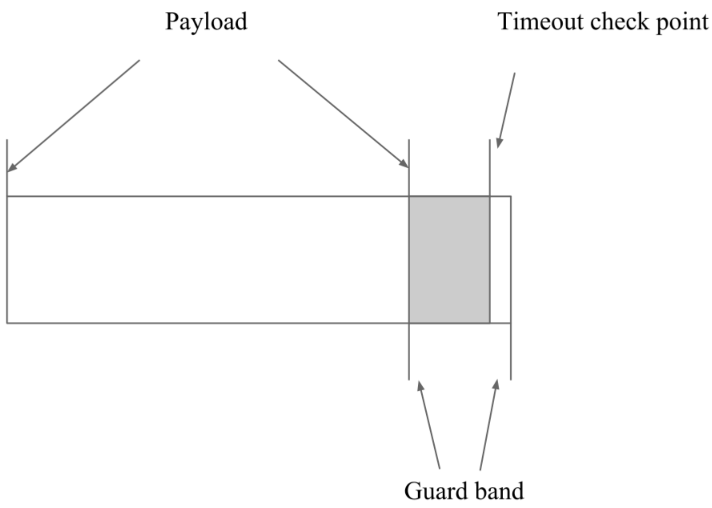


Figure 5.1: The Time Slot Structure

5.2.2 Traffic Model

In the simulations carried out in this thesis, the number of incoming packets per frame follow an exponential distribution. In the context of single-hop networks, all nodes in a network are supposed to have the same mean packet inter-arrival time. This value is determined by multiple factors including packet size, traffic load, node population and transmission data rate. The calculation of the mean packet inter-arrival time can be expressed by:

$$mean = \frac{packet\ size * number\ of\ nodes}{traffic\ load * transmission\ datarate} \tag{5.1}$$

The above equation clearly notes that the higher traffic load, the higher the packet generation rate. If the number of nodes in the network is relatively large, the data traffic can be considered as uniformly contributed by each transmitting node. Moreover, there are two traffic load conditions implemented in all the simulation works: saturated traffic load and practical traffic load. Saturated load conditions represents that each node always has packets to send at the beginning of each transmission round (frame). It is used to examine the performance of the learning process during maximum contention. A saturated load can also boost learning speed because of the higher the traffic load, the more the iteration of Q values and sooner the converge. The practical traffic load approximates to the ordinary traffic loads generated during runtime, thereby it is usually unsaturated.

The system offered traffic load and the system throughput are measured in Erlangs which is a standard unit to represents the traffic density in a system. One Erlang traffic is the equivalent of the continuous use of the shared channel over a period of time. In the context of ALOHA-Q, the period over which the average is calculated is often one frame. Let N denotes the number of active nodes, thus one Erlang offered load represents an average N concurrent transmissions during each frame. Once a data packet is generated, it will be stored in a sub-queue of unlimited size. A node always picks packets from the sub-queue based on a first come first serve basis.

5.3 Impact of Frame Size Selection on Performance of ALOHA-Q

The core design philosophy of ALOHA-Q is to allow individual nodes to learn an optimum slot selection policy within a relatively short initialisation stage. Subsequently, in a steady state, nodes can benefit from improved throughput and energy efficiency during the remaining period of network operation. ALOHA-Q assumes that nodes can exploit their dedicated time slots once the learning process is finished [66]. In other words, channel contention cannot be completely resolved until the owner of each occupied slot becomes unique. To achieve the maximum channel utilisation, a frame should contain an appropriate number of time slots to ensure that every node can eventually obtain a dedicated time slot and no empty slots will be left during the steady state. This conclusion indicates that the value of frame size has an considerable impact on the outcomes of the learning process.

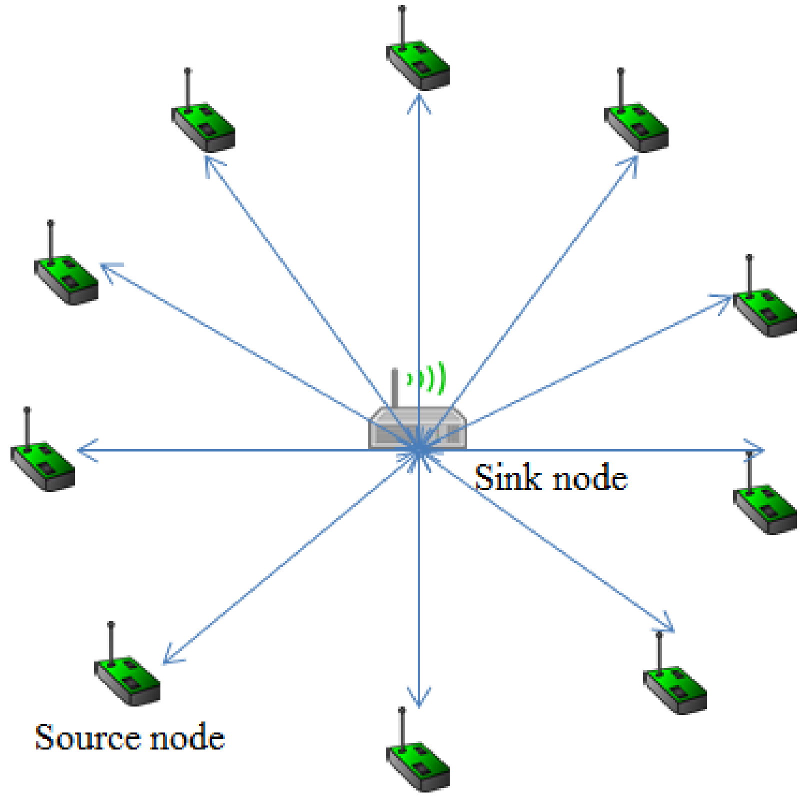


Figure 5.2: Typical Single-Hop Based WSN

Consider a single-hop WSN illustrated in the Fig. 5.2, where ten source nodes send packets directly to a sink node via a shared channel. According to the *pigeon hole* theory [115], sensor nodes can obtain their dedicated time slots only if the frame size is greater than the number of active nodes in the network. On the other hand, an overlarge frame size can lead to redundant slots once learning ends. In this case, the maximum system throughput will be degraded and the average end-to-end delay of packets will be raised owing to inefficient slot use. Therefore, we can safely assume that optimal channel utilisation of ALOHA-Q can be achieved if the value of the frame size equals to the number of active source nodes. To prove this hypothesis, we explore the performance of ALOHA-Q under different frame sizes by using OPNET. The simulated single-hop network (see Fig. 5.3) is deployed in a 100m x 100m square area. There are 100 source nodes which form a circular (with 50m radius) and a sink node locates in the centre point of the sensing area. The corresponding frame size is set to 50, 100, and 120 respectively.

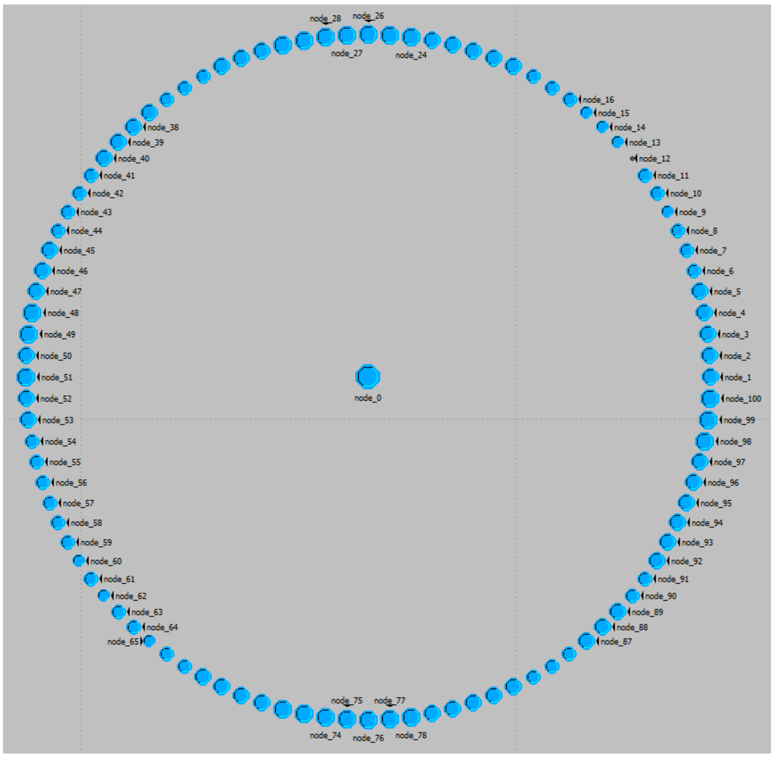


Figure 5.3: Topology of the Simulated Single-hop Network

5.3.1 Throughput Performance

Fig. 5.4 compares the throughput performance obtained during steady state for the three different frame sizes. The duration of the simulation period is set to 50000 slots which ensures the 100 nodes can their obtain dedicated time slots before the simulation ends. The throughput values for 100 slots and the 120 slots were collected during the system steady state, and the throughput values for 80 slots were collected from the beginning of the simulation since the system steady state can never be achieved. To improve the accuracy of the results, each presented values was calculated by taking the average values from 200 simulations. The offered traffic load ranges from 0.1 to

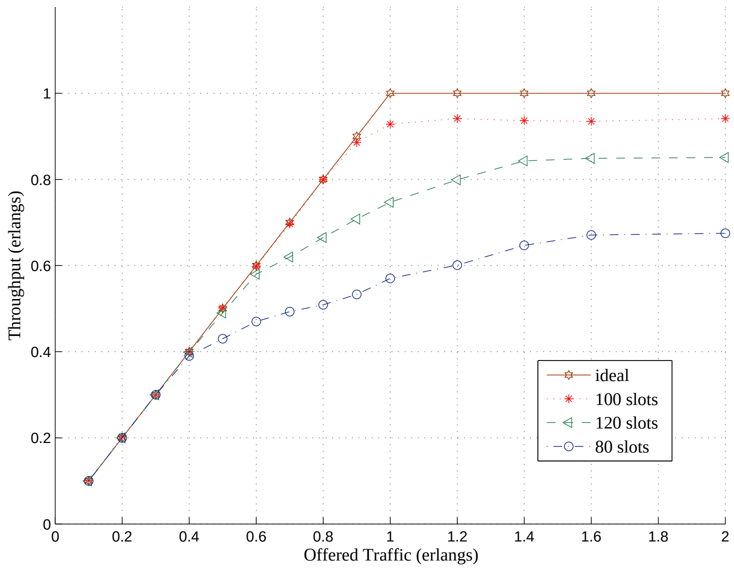


Figure 5.4: Comparison of Throughput Performance

2 Erlangs. Commonly, the higher the Erlang, the more packets will be sent per frame. An Erlang offered traffic means every node has exactly one packet to send per frame. When offered traffic is greater than 1 Erlangs, the traffic level is considered as saturate. In an ideal circumstance, the system throughput is supposed to increase linearly before the offered load become saturate, and the throughput values fully correspond to the values of offered traffic. Subsequently, the throughput maintains at one Erlang which represents all the generated traffic successfully transmitted through the shared channel when contention level becomes full (every node sends at least one packet per frame). However, the ideal throughput cannot be achieved in practical scenarios due to many factors including the effects of the duty cycle of payload. According to the simulation parameters, the maximum throughput can be calculated as:

$$Throughput = \frac{1044(packetsize)}{1100(slotlength)} \tag{5.2}$$

The simulation result clearly shows that the maximum throughput is achieved when a frame contains 100 slots. In this case, the throughput value at saturated traffic load can reach to 94.9% of channel capacity. The effect of guard-bands reflects the missing

5.1% throughput. When the frame size is either smaller (80) or larger than (120) the given number of nodes, the maximum throughput is adversely affected because smaller frame size leads to permanent slot hopping for some nodes and the overcapacity of frame results in empty slots during steady state.

5.3.2 Delay Performance

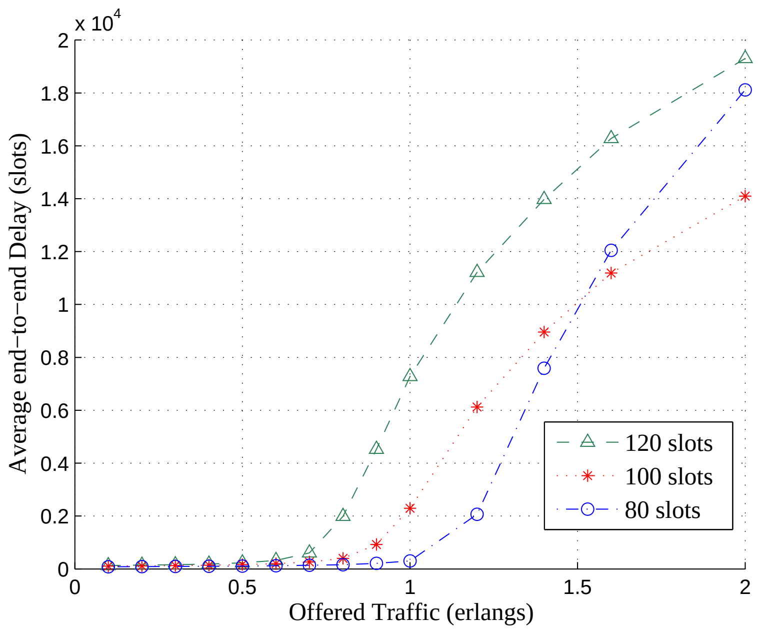


Figure 5.5: Comparison of Average End to End Delay Performance

Fig. 5.5 presents a comparison of the average end to end delay of received data packets. Similarly, the values for 100 slots and the 120 slots were collected during the system steady state, and the values for 80 slots were collected from the beginning of the simulation. It is clearly observed that the lower the frame size, the shorter the average delay that can be achieved when offered load is less than 1.5 erlangs. That is because the frame size is directly proportional to the queuing time of packets. However, the reduced delay is achieved at the cost of degraded throughput performance. Under the ultra heavily load conditions (1.0 - 2.0 erlangs), the delay of 80 slots increase rapidly as the offered load raises, and finally approaches the same level of 120 slots when offered load is 2.0 erlangs. This is caused by the insufficient number of time slots within a frame. When the frame size is less than the node population, some nodes must always share slots with others thereby the system steady state can never be achieved. Consequently, the packets generated by unsteady nodes will be backlogged in the subqueue. The more packets are generated, the longer the queuing time.

5.3.3 System Convergence Time

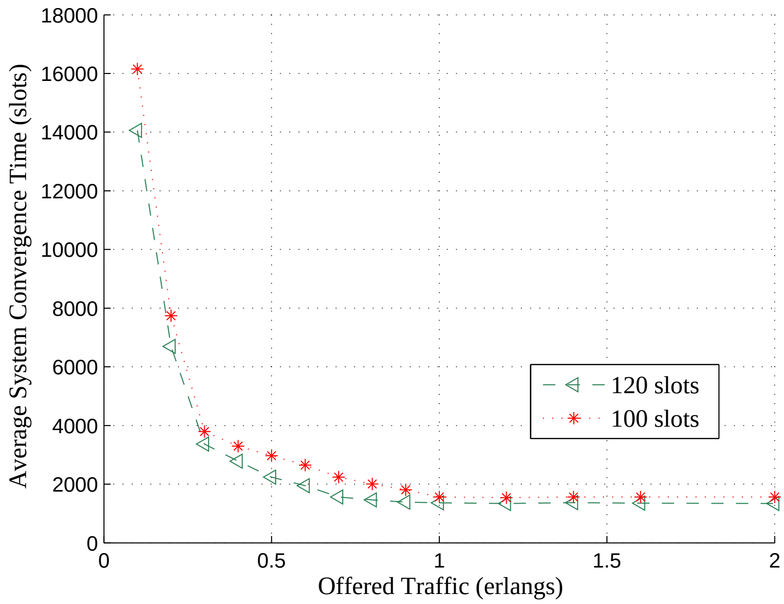


Figure 5.6: Comparison of Average System Convergence Time

A comparison of system convergence time is illustarted in Fig. 5.6. It can be seen that system convergence time for 120 slots is slightly shorter than the time for 100 slots. In ALOHA-Q, nodes randomly select slots at the beginning of the learning process. Before they have gained enough transmission experience, the probability of a collision is inversely proportional to the frame size. The greater the frame size, the sooner the nodes can find their dedicated slots. However, the system always can reach steady state whether a frame contains extra slots or not. Based on the overall evaluation, the system can strike a balance between throughput performance, delay performance, and system convergence when a frame is composed of 100 time slots. Therefore, one may draw a conclusion that the optimal value of frame size under single-hop conditions is equal to the node population. This conclusion implies that the node population must be accurately predicted for the purpose of determining the optimum frame size. Otherwise, ALOHA-Q cannot be optimally implemented.

5.4 Distributed Frame Size Adaptation

5.4.1 Basic Principle

Since the performance of ALOHA-Q is heavily dependent on the selection of frame size, the question of how to set up the optimal frame size becomes a primary challenge for implementing ALOHA-Q. Existing approaches for setting an ideal frame size relies on artificial operation, where the optimum frame size is ascertained and manually configured on each node. However, accurate prediction of the number of active nodes can be very difficult for large-scale networks that may contain hundreds or thousands of nodes. To solve this problem, a distributed frame size selection algorithm becomes appealing.

As previously mentioned, the general Q-learning algorithm assumes that if a learner keeps reinforcing its actions according to the feedback from an optimum environment, it is certain to form an optimum policy of action. This assumption indicates that the policy of action is related to the environment. We can reverse this assumption: an optimum environment is a prerequisite for a learners optimum policy of action. In the context of ALOHA-Q, if the environment is defined as the frame size then one can draw a conclusion that every node can find a dedicated slot if a frame contains adequate slots. Conversely, if the number of slots per frame is less than the number of nodes, there must be some nodes in a perpetually non-steady state (slot searching).

Based on the fundamental assumptions of the Q-Learning algorithm, we propose a Distributed Frame size Adaptation (DFA) scheme for helping sensor nodes to select an ideal frame size. The key idea of this scheme is to allow individual nodes to learn an optimal value of frame size by continuously examining channel performance. To this end, DFA divides time into repeating frame size evaluation windows. A window consists of an integer number of frames. During each window, nodes are required to transmit packets on a frame-by-frame basis based on the Q-learning based slot selection strategy. A failed packet is always retransmitted in next frame until the number of attempts reaches the threshold M (usually 6 according to the IEEE 802.15 standard). The transmission experience of nodes gained from the slot selection process within current window is utilised as a basis to select the frame size for the next window. Specifically, if a network can reach steady state at the end of a frame, the current frame size is thereby not less than the number of nodes. Otherwise, the frame size is too small to accommodate contending nodes.

The length of the frame size evaluation window is an important parameter in determining the performance of the DFA scheme. Commonly, it is determined convergence time of the Q-learning based slot selection. To ensure sensor nodes have adequate time to estimate channel performance, the length of each window should not be less than the system convergence time. In some cases, the window length needs to be dynamic as depend on the changing number of nodes. [86] presents the approach for estimating system convergence time through MDP. Let state i represent that a total of i nodes have found their dedicated slots and N denotes the assumed frame size in a window. Obviously, the system has N states during a slot learning process and the state N refers to the steady state. The state transitions of slot learning take place in every slot, and the process can move forward or backward or stay in the same state after each slot. Let $p_{i,j}$ refer to the state transition probability from state i (i=0,1,2,...,N) to state j (j=0,1,2,...,N), P denotes the state transaction probability matrix which has the elements of $p_{i,j}$. Multiply P by itself, the elements in P^2 are:

$$p_{i,j}^2 = \sum_{m=0}^{N} (p_{i,m} p_{m,j}) \tag{5.3}$$

Which is the probability that the slot learning process reaches state j via state m by starting with state i. According to (5.3), we can consider P^n as the matrix of state transition probabilities after n transactions (slots), and $p^n_{i,j}$ is the probability that the learning process reaches state j after n transitions starting in state i. According to the principle of Q-learning based slot selection, the system will always stay in steady state once all nodes have found their dedicated slots. In other words, the steady state is an absorbing state in MDP. Therefore, the system can reach convergence state by starting with state i if

$$\lim_{n \to \infty} p^n_{i,N} = 1 (i = 0, 1, 2, ..., N) \tag{5.4}$$

To obtain the time before convergence, it is needed to calculate the expected time that the slot learning process stays in all states except state N. Consequently, the expected system convergence time from the beginning of the learning process can be obtained by calculating

$$\sum_{n=1}^{\infty} \sum_{j=0}^{N-1} p^n_{0,j} \tag{5.5}$$

5.4.2 Frame Size Adaptation Process

Fig. 5.7 depicts a general workflow of the DFA. In this example, there are three nodes contending for the shared channel. An initial frame size of two slots is assigned to a node. When the adaptation process starts, the node keeps sending packets and updateing Q values table. Since the contention level is too high, it cannot find a dedicated slot at the end of the current window. Accordingly, the node considered the current value of the frame size as the lower bound of the ideal value and decided to increase the frame size from the beginning of the second window. After many transmission rounds, the node has finally obtained a contention-free slot so that the current frame size can be identified as the upper bound. Hence, the nodes decreased current frame size at the beginning of the third window. By this way, the bounds will be gradually narrowed. An optimal frame size can be eventually selected.

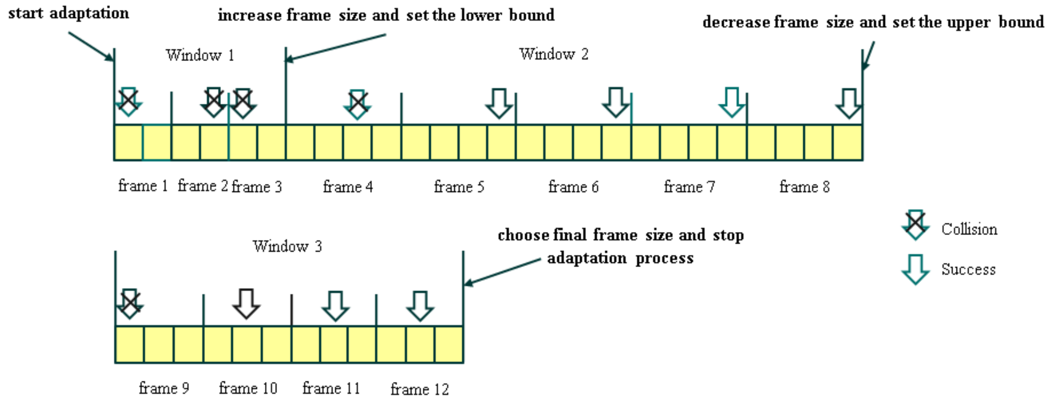


Figure 5.7: General Workflow of DFA

To speed up the frame size adaptation process, the magnitude of the frame size adjustment has to be properly considered. In DFA, nodes follow a binary search strategy to adjust their frame size. Consider the example illustrated in the Fig. 5.8. Let assume there are 10 source nodes in a single-hop network. In this particular case, each node is assigned an initial frame size of two slots. At end of each frame size evaluation window, nodes make a decision to increase or decrease current frame size according to the channel performance. Let S_i denotes the frame size during i-th window, S_{max} and S_{min} stand for the frame size upper bound and lower bound respectively. The most recent successful frame size is deemed as a frame size upper bound S_{max} and the most recent unsuccessful frame size becomes the lower bound S_{min}. Prior to a node finding the first S_{max}, the new frame size (S_{i+1}) is always twice as large as the previous frame size (S_i). If the first S_{max} has been found, nodes always choose the middle value between S_{max} and S_{min} as the new frame size. The constant updating of S_{max} and S_{min} helps nodes to quickly lock the range of the ideal frame size. When S_{max} and S_{min} differ by 1, S_{max} is the final frame size, and then the adaptation process automatically stops. Based the binary search strategy, the number of windows before the optimum frame size is found can be reduced to the minimum value.

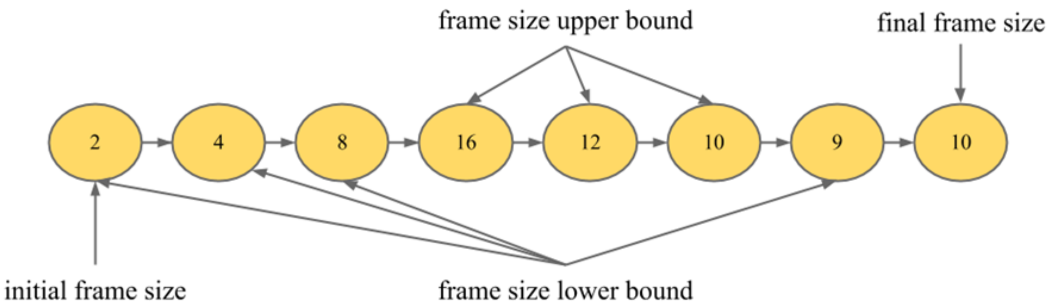


Figure 5.8: Frame Size Adjusting Process for 10 Nodes

5.4.3 Consistency of Frame Size Adjustment

The detail algorithm of DFA is presented as follows:

if The beginning of the i-th frame size adjustment window **then**
  send one packet per frame based on ALOHA-Q.
end if
if The end of the penultimate frame of the i-th frame size evaluation window **then**
  if Node in steady state **then**
    send a packet at the preferred slot during last frame of current window.
  else if Node not in steady state **then**
    send packets at each slot during last frame of current window.
  end if
end if
if The end of last frame of the i-th frame size evaluation window **then**
  if received an ACK packet **then**
    $S_{max} = S_i$, choose $\frac{(S_{max}-S_{min})}{2}$ as the new frame size from the next window.
  else
    if $S_{max} > 0$ **then**
      $S_{min} = S_i$, choose $\frac{(S_{max}-S_{min})}{2}$ as the new frame size from the next window.
    else if $S_{max} = 0$ **then**
      $S_{min} = S_i$, choose $2 * S_i$ as the new frame size from the next window.
    end if

    end if
    if $S_{max} - S_{min} = 1$ **then**
        stop frame size evaluation process, choose S_{max} as final frame size.
    end if
end if

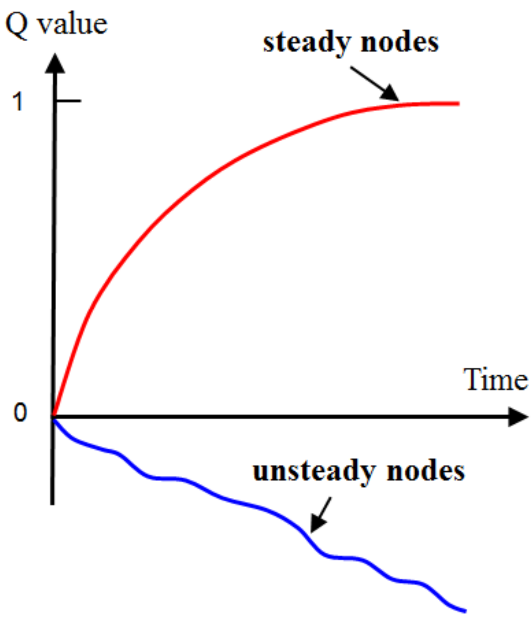


Figure 5.9: The Changing Trend of Highest Q value for Steady and Unsteady Nodes

From the beginning of the i-th window, nodes keep exploring the best slots until the end of the penultimate frame. At this point, nodes determine their state (steady or non-steady) by checking the highest Q value (Q_{max}), which is considered as the basis of frame size adjustment. During the Q-learning based slot selection process, the changing of the (Q_{max}) has two trends that are shown in Fig. 5.9. An ever-increasing Q_{max} indicates that a node has found a dedicated slot (steady node) and the current frame size is likely to be greater than the node population. On the contrary, an ever-decreasing (Q_{max}) represents that a node is still likely to seek the best slot (unsteady node) and the current frame size may less than the node population.

Apparently, when the current frame size is very close to but still less than node population, some nodes can become steady after many transmission rounds but some nodes can never find their dedicated slots. Therefore, the ranges of Q_{max} for each node may vary. Since Q_{max} is the only consideration for changing the frame size, nodes may have different beliefs about the new frame size for next window. For instance, for steady nodes, it is appropriate to maintain the current frame size, lower it from the next window onwards or stop the frame size learning process. However, for the unsteady nodes, it is appropriate to increase the frame size from the next window because there are no empty slots left. If that happens, the frame starting time of each will no longer be synchronised. Consequently, the system will never reach steady state again. Therefore, to avoid the partial overlap of frames, the consistency of frame size adjustment must be guaranteed.

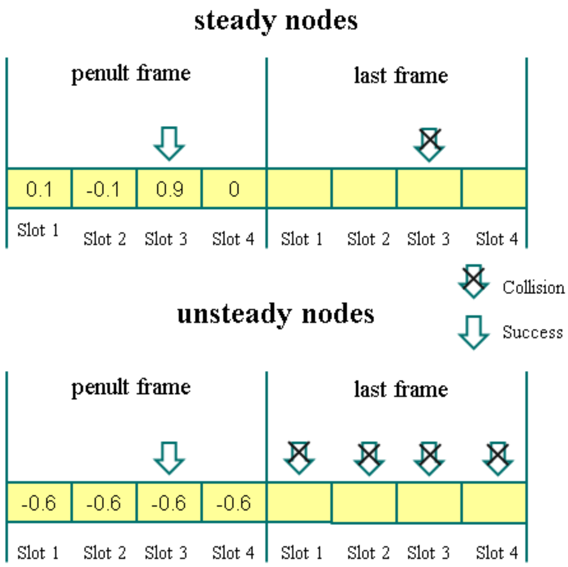


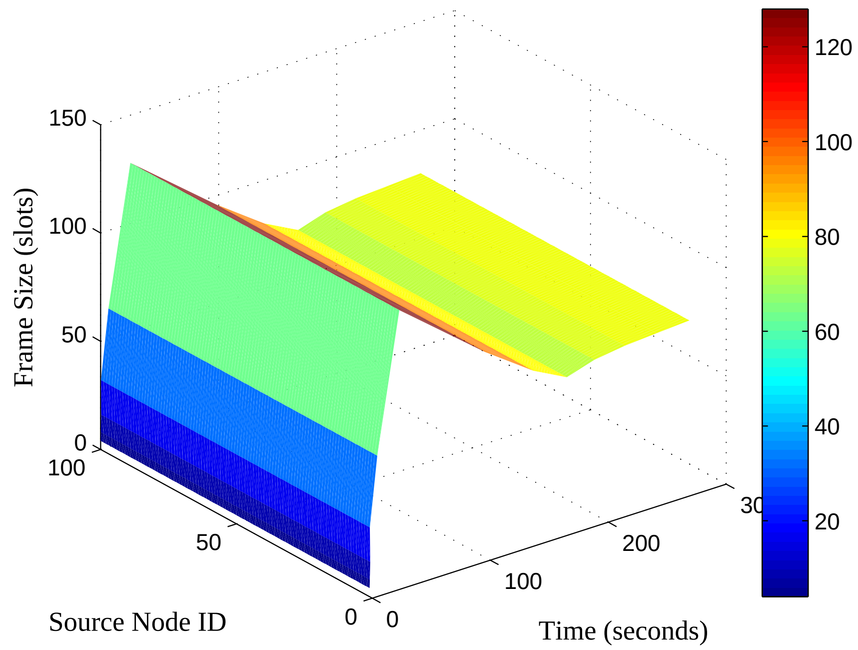
Figure 5.10: Proactive Jamming Strategy

The conventional single-hop networks assumed in this chapter do not support any direct communications between source nodes. However, to achieve the consistency of frame size adaptation, individual nodes need to exchange their beliefs about the current frame size via indirect ways. To this end, DFA introduces a proactive jamming strategy (see Fig. 5.10). To be specific, during the i-th frame size evaluation window, steady nodes are required to choose their preferred slots in the last frame and unsteady are needed to occupy all slots in the last frame. As a consequence, none of nodes can make a successful transmission during the last frame if one or more nodes remain in the unsteady state, indicating that they have not found a unique slot and wish for the frame size to be increased. At the end of each frame size evaluation window, nodes that have failed to transmit a packet during the last frame will increase their frame size for the next window. If nodes have successfully sent a packet during the last frame, they will decrease their current frame size for the next window. To put it simply, the frame size is reduced when all nodes have found their dedicated slots before the end of the last frame of each window. Otherwise, the frame size is increased. Based on this strategy, the consistency of frame size adaptation can be ensured.

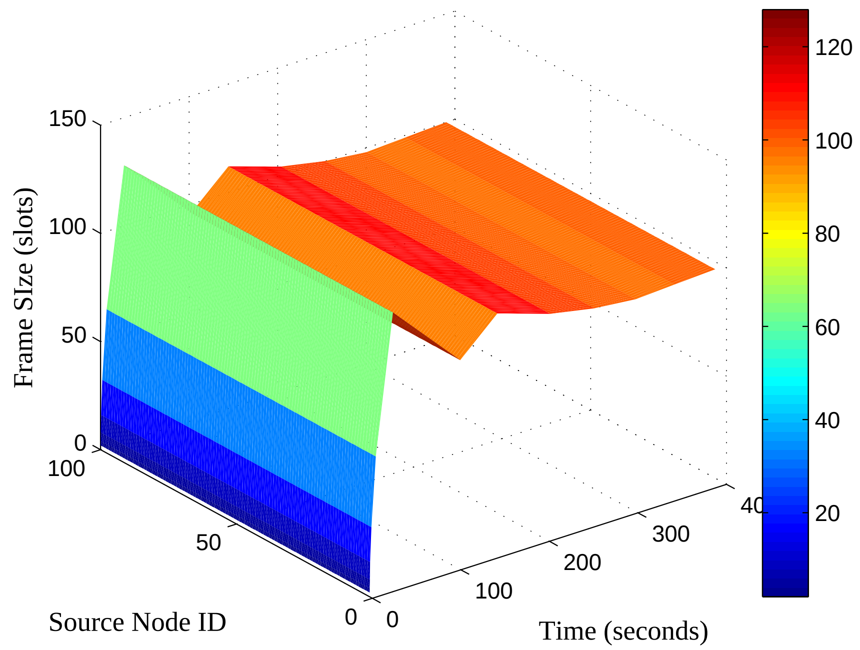
5.5 Performance Analysis

To examine the performance of the DFA, a set of simulations are carried out via OPNET. We model three single-hop networks which contain 80 nodes, 100 nodes and 150 nodes respectively. The initial frame size of each node is set to four slots and learning rate is set to 0.1. The same network parameters listed in Table.5.1 are selected.

Fig. 5.11 presents the individual frame size adaptation processes for different node populations. Consider the results shown in Fig. 5.11 (b), there are 100 source nodes in the simulated network. The value of frame size for each node starts from just four slots since the adaptation process begins. After a very short period to evaluate the channel performance, nodes have found that the current frame size does not enable every node become steady. Therefore, the frame size is increased to eight slots from the second window. When frame size is increased to 128, all nodes can obtain their dedicated slot during that window and the system can attain the steady state, so nodes consider 128 as the upper bound of the optimal frame size and select a lower frame size in next window. At this moment, the optimum value of frame size is bounded between 64 and 128. Hence, the nodes continuously and periodically evaluate and adjust the frame size and narrow the range of the optimal value. Eventually, the frame size settles at 100, corresponding exactly to the number of source nodes. Subsequently, the frame size learning process is automatically stopped, and the optimum frame size is attained by each node. The same frame size adaptation process is repeated when nodes population is changed to 80 and 150. No matter how many nodes are deployed in the network, an optimal frame size can be leant by nodes by implementing the DFA within an initialisation stage.

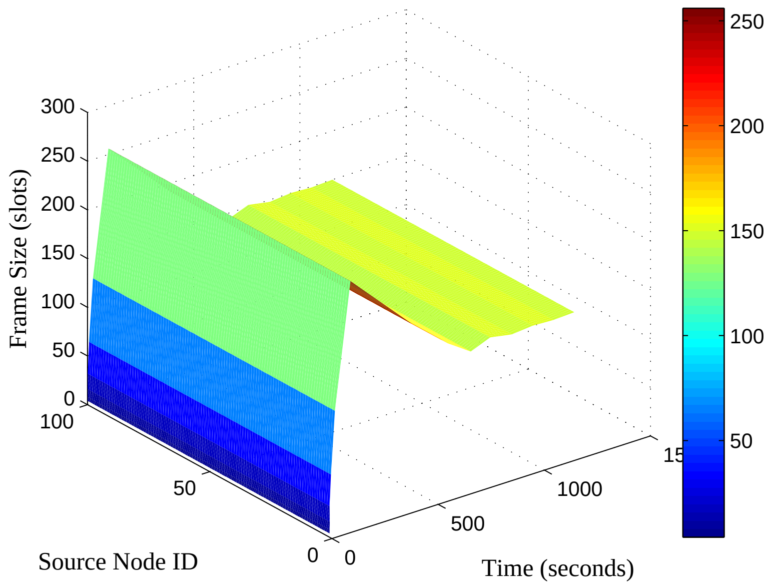


(a) 80 nodes



(b) 100 nodes

Figure 5.11: Frame Size Adaptation Process for Individual Nodes



(c) 150 nodes

Figure 5.11: Frame Size Adaptation Process for Individual Nodes

Fig. 5.12 and Fig. 5.13 demonstrate the results regarding the steady nodes population and the system running throughput as a function of time respectively. The simulations end up with 100th frame once frame size adaptation process is finished. These results can reflect the tendency of channel performance during adaptation process. At the beginning of the adaptation process, the frame size is too small compared to the number of nodes. A large number of nodes have to contend for a small number of slots. Under this circumstance, a node can hardly to find an empty slot unless other nodes stay in a back-off stage. The number of steady nodes and system running throughput thereby are very close to zero. As frame size increases, more empty slots can be occupied by nodes. Once the first frame size upper bound is found, the whole system reaches the steady state for the first time, and the value of optimum frame size can be bounded. When frame size is greater than node population, the running throughput will increase rapidly and all nodes can be identified as steady nodes within the current evaluation window. The range of optimal frame size is narrowed through the continuous frame size adjustment. At the end of frame size adaptation process, all nodes become steady and contention are fully reduced. As the consequence, the system running throughput of different node population tend to approach the maximum channel capacity. These results prove the effectiveness of the distributed frame-size adaptation scheme since all packets are transmitted without collision when adaptation process is finished. Therefore, the ALOHA-Q protocol can help sensor nodes achieve collision-free channel access and the throughput of the system can reach the full capacity of the channel under saturated traffic load conditions.

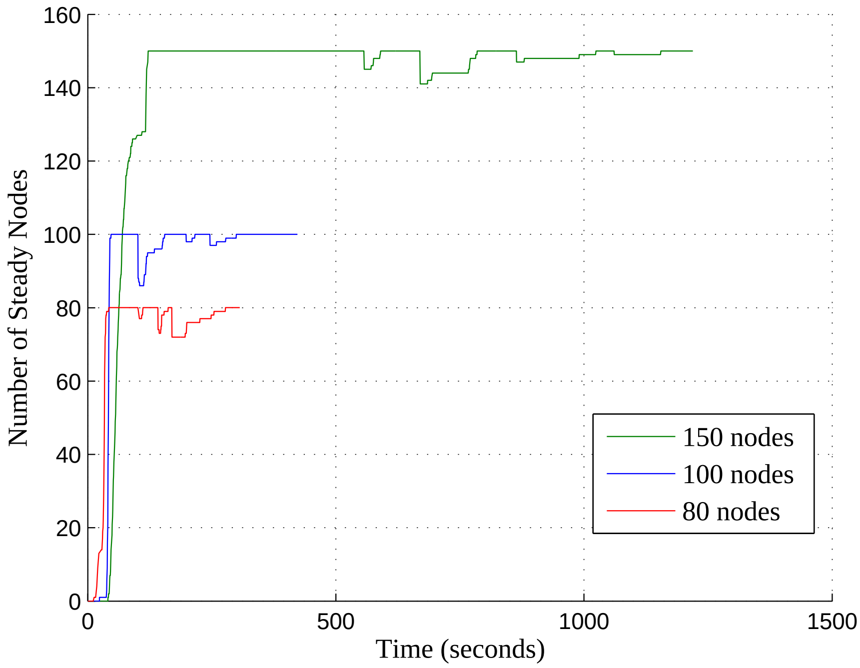


Figure 5.12: Real Time Number of Steady Nodes

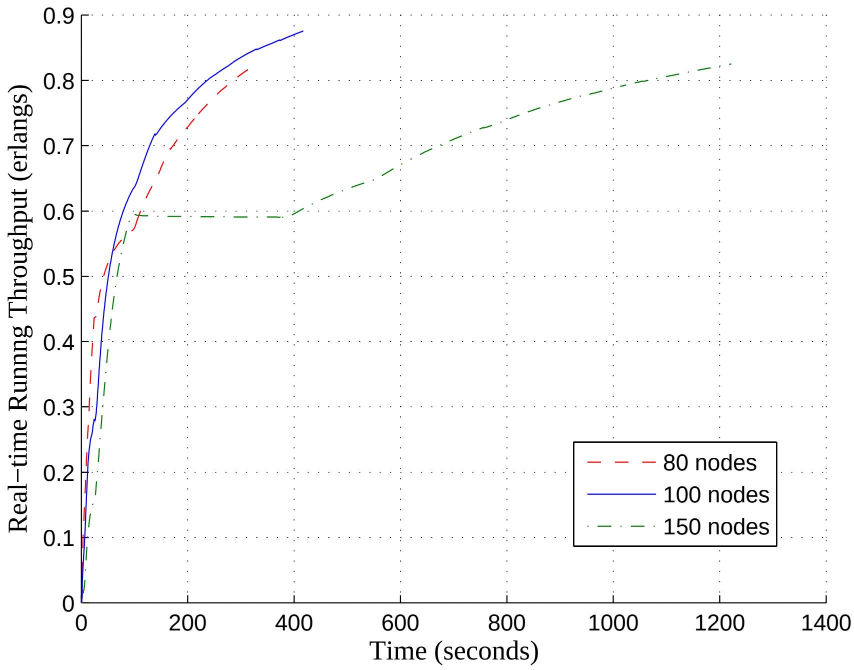


Figure 5.13: Real Time Running Throughput

5.6 Summary

Due to the issue arising from the setting of an optimal frame size, it is not feasible to implement the ALOHA-Q protocol to large-scale random-deployed single-hop WSNs. This chapter started out by investigating the relationship between the performance of ALOHA-Q and the selection of frame sizes. Results show that the maximum performance of ALOHA-Q in single-hop networks can be achieved if the optimal frame size is equal to the node population. Motivated by this problem, this chapter proposes a distributed frame size adaptation (DFA) scheme to help sensor nodes automatically learn an optimal value of frame size. Simulations in this chapter evaluated the proposed scheme under different conditions. Corresponding results bring confidence that DFA can effectively support ALOHA-Q in achieving collision-free scheduling without precise assumption the node population. In the following chapter, a similar scheme is designed for multi-hop WSNs, and more techniques are studied to maximise the benefits of ALOHA-Q.

Chapter 6

A Self-adaptive ALOHA-Q Protocol for Multi-hop WSNs

Contents

6.1 Introduction

The previous chapter presented a frame size adaptation scheme to achieve maximum performance of the ALOHA-Q protocol under single-hop conditions. Single-hop WSNs typically only support short-range communication between source nodes and sink nodes, and do not employ advanced features such as collaborative in-network processing [91]. Therefore, the scope of related applications is very limited. In recent decades, the development of multi-hop WSNs has been greatly advocated because of their broad application prospects. Compared to single-hop WSNs, multi-hop WSNs are more suitable for large scale WSNs due to multi-hop data forwarding. Moreover, multi-hop networks provide more robust data delivery since source nodes can select multiple transmission routes. These advantages enable multi-hop WSNs to become the mainstream in many professional applications including military purposes for monitoring, tracking and surveillance of moving objects, intelligent transportation systems, etc [62]. In this chapter, we present a self-adaptive ALOHA-Q protocol. Its objective is to adapt the Q-learning based slot-selection strategy of ALOHA-Q to multi-hop WSNs. Since ALOHA-Q was originally designed for single-hop WSNs, the differences in network topologies, transmission patterns, and radio interference model increase the complexity when developing the new protocol. To fulfil the optimal performance of ALOHA-Q under multi-hop conditions, a set of novel schemes and techniques are introduced for the purpose of overcoming the underlying practical challenges.

6.2 Network Topologies

In this chapter, the proposed protocol is applied to two representative multi-hop topologies: linear chain topology and cross chain topology. The detailed scenarios and assumptions of these topologies are described as follows.

6.2.1 Linear Chain Topology

The alignment of sensor nodes in a linear form can often be found in many applications such as monitoring of oil, gas, and water pipelines, railways/subway monitoring, highway driver-alert networks, etc. Fig.6.1 shows a simple linear network, where seven nodes (one source, one sink and five relay nodes) are deployed as a linear chain. Packets generated from the source node are forwarded by each relay node in sequence. R_r represents the node reception range of each node and R_i denotes the node interference range. To characterise the radio interference effect of sensor nodes, the simulations carried out in this chapter adopt a protocol interference model. The protocol model assumes that a node can successfully receive a packet if it is within the distance R_r of its intended transmitter and falls outside the R_i of other non-intended transmitters [90]. Realistically, a packet can be successfully received if the Signal-to-interference-plus-noise Ratio (SINR) at the intended receiver exceeds a threshold so that the transmitted signal can be decoded with an acceptable bit error probability. This indicates that a node's interference range usually exceeds its reception range. Since WSNs are usually homogeneous, it is a reasonable assumption that every wireless sensor node in the network has the same hardware features. For the simulations carried out in this chapter, sensor nodes are assumed to be deployed in a barrier-free area and they always maintain uniform transmission power so as to achieve same communication range and interference range.

As Fig.6.1 shows, each node communicates with the nodes within a one-hop distance and may interfere with the nodes within a two-hop distance. To better understand the slot selection of ALOHA-Q in this linear chain network, the nodes can be separated into two groups from the middle of node 4 and 5. Nodes in the left group have initially chosen four different time slots (slot 1, 2, 3 and 4) in order to avoid packet collisions caused by channel contention and radio interference. In the right group, time slot 1 has

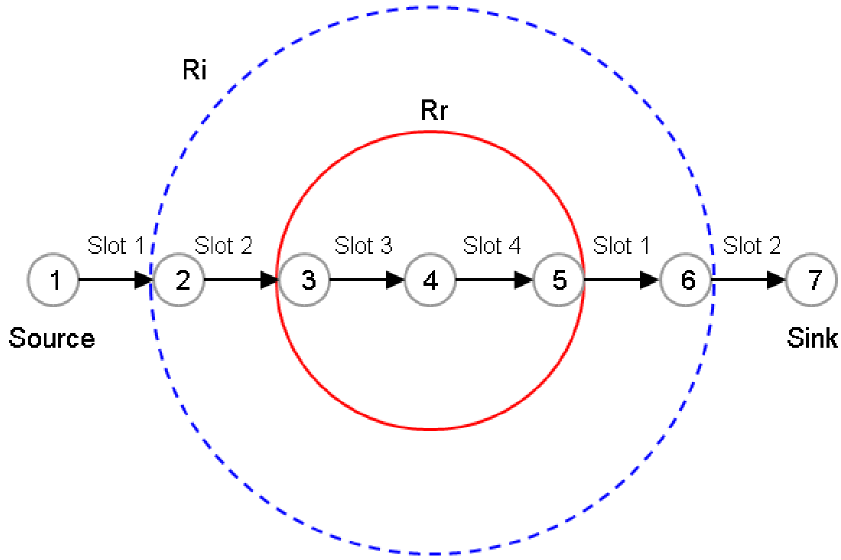
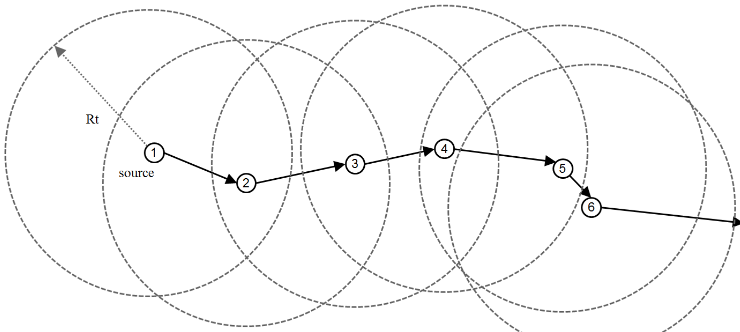


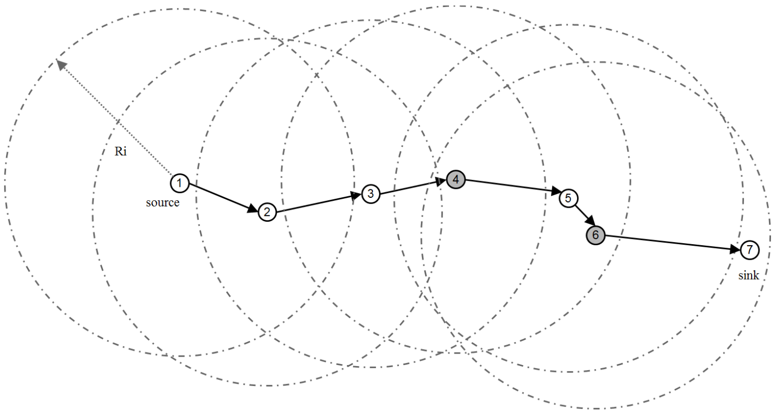
Figure 6.1: Example of Linear Chain Network

been reselected by node 5 since the link between node 5 and 6 cannot be disturbed by interference signals from node 1. So node 1 and 5 can safely share a same time slot. In the same way, node 6 has shared time slot 2 with node 2. Therefore, to achieve optimal channel utilisation, nodes within four-hop distance have to choose separate slots and any two nodes within a three hop range away can share the same slot in a frame. if the chain further extends, each time slot formerly occupied by the left group can be repeatedly reused. To achieve the most efficient slot utilisation, the groups of sensor nodes divided by spatial location have to maintain the same sequence of slots. Once the sequence of slot selection of the first group is fixed, the remaining nodes should follow this sequence to avoid packet collisions.

The linear topology illustrated in Fig.6.1 represents an ideal circumstance, where nodes are equidistantly distributed so that they have uniform reception range and interference range in hops. However, in realistic scenarios, nodes may be randomly deployed. As a result, the local interference level of nodes are not identical. Other effects such as shadowing from obstacles, background electromagnetic radiation may lead to many uncertainties associated with the node interference range. Fig. 6.2 shows a randomly deployed linear chain network. Nodes are assumed to have uniform coverage areas (in



(a) Node Reception Range



(b) Node Interference Range

Figure 6.2: Randomly Deployed Linear Chain Multi-hop Network

square meters) of reception and interference. The reception range of each node covers one adjacent neighbour on its each side and the interference range of most nodes covers one-hop distance on its each side. However, some nodes (4 and 6) can disturb the transmission of other nodes which locate more than a two-hop distance away because of the uneven density of nodes. In this case, node 4 and 6 become bottlenecks as they disturb more nodes than others. Therefore the interference range (in hops) of node 4 and 6 can be considered as the critical interference range (R_{imax}) for this chain. To achieve optimal channel utilisation, nodes within a four-hop distance have to choose separate slots and every one in four nodes along the chain can share the same slot in a frame. This conclusion implies that the optimal slot assignment policy of linear networks depends on node neighbourhood interference that can not be affected by the total number of nodes along the chain.

6.2.2 Cross Chain Topology

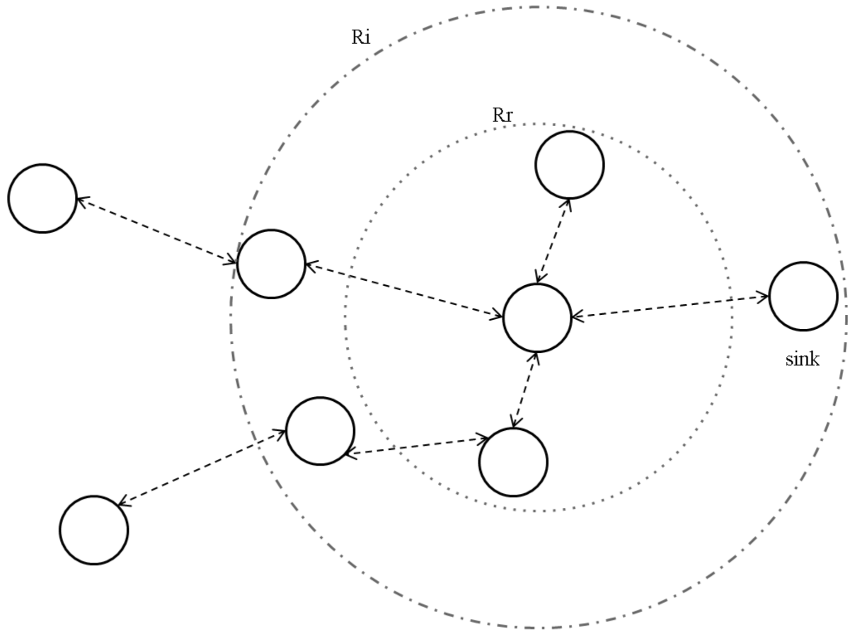


Figure 6.3: Example of Cross Chain Multi-hop Network

Cross chain topologies are very common in multi-hop wireless networks. In such networks, nodes are interconnected in an ad-hoc manner. During operation time, the role of a node can be switched between source, relay, and sink according to the application requirements. It is possible that multiple source nodes simultaneously transmit data packets to single or multiple sinks during runtime. Therefore, independent traffic flows may overlap with each other or converge at the same sink so that some nodes need to relay multiple packets generated from different source nodes. Fig. 6.3 illustrates an example of cross chain WSN. R_r and R_i represent node reception range and interference range respectively. For multi-hop networks mentioned in this chapter, nodes are assumed decide their transmission route based on the Most Forwarding progress within Radius (MFR) scheme since it is the most energy efficiency routing model. Specifically, a node selects a neighbour within R_r with the shortest geometric distance to the sink as the next hop node to send packets.

6.3 Adaptation of ALOHA-Q to Linear Chain Networks

6.3.1 Maximum Throughput and Optimal Frame Size Estimation

To implement the ALOHA-Q protocol in a linear network, it is necessary to determine the optimal frame size, which is determined by many specific network parameters. In this section, the maximum attainable throughput and optimal frame size for linear networks are deduced using analytical models. Consider a linear network (see Fig. 6.4) consisting one sink node and n sensor nodes denoted as N_i (i=1,2,3...,n). Let T stand for the duration of a slot, S_i be the number of source nodes between N_0 and N_i, and R_{imax} represent the critical interference range (hops) of the network sine nodes are equally spaced. Therefore, it is able to have the following theorems.

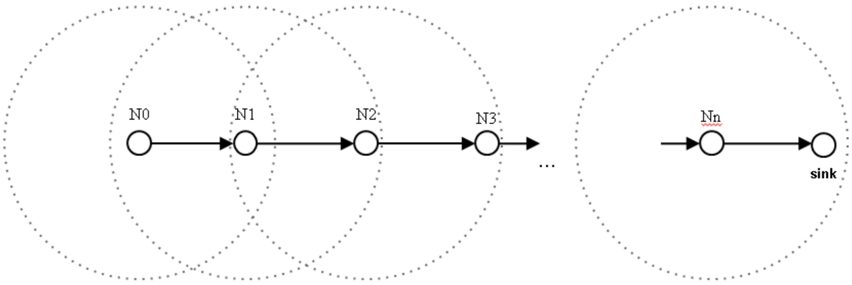


Figure 6.4: Linear Chain Network

Theorem 1. *The frame duration for each node D obeys:*

$$D \geqslant D_{min} = \begin{cases} (S_n + S_{n-1} + \sum_{i=0}^{R_{imax}-1} S_{n-2-i})T, & n \geqslant 2. \\ T, & n=1. \end{cases} \quad (6.1)$$

Proof. **For** $n \geqslant 2$: Let t represent the time period that the sink node has successfully received one packet from each source node along transmission route. At any instant, a sensor node can only be in one of three possible states (transmission, reception, idle). Thus, t is the combination of transmission time (t_T), reception time (t_R) and idle time (t_I). Thus, $t = t_T + t_R + t_I$ ($t_R = 0$ for source node and t_T=0 for sink node). During the time period t, the sink has received at least S_n packets from the N_n. Hence, N_n has made at least n transmission attempts, which includes delivery of $n-1$ relayed packets and one original packet. So the reception time of the sink node follows:

$$t_R \geqslant S_n T \tag{6.2}$$

Since the sink cannot receive packets when N_n stays either in reception mode or idle mode, the idle time of the sink node is not less than the sum of reception time and idle time of N_n. In order for N_n to receive S_{n-1} packets from node N_{n-1}, N_n needs to listen at least $(n-1)$ slots, during which time the sink node must be idle. In addition, the time period that N_n cannot transmits is depends on the critical interference range along transmission route R_{imax}. When N_{n-1} transmits, N_n cannot transmit since there is only a one-hop distance between both nodes. Furthermore, N_{n-2} needs to transmit (S_{n-2}) packets to N_{n-1}, during which time N_n cannot transmit either (i.e., N_n's transmission will disturb the packet reception of node $(n-1)$). Based on the above analysis, we can safely deduce to the idle time of the sink node:

$$t_I \geqslant S_{n-1}T + \sum_{i=0}^{R_{imax}-1} S_{n-2-i}T \tag{6.3}$$

Therefore,

$$t = t_R + t_I \geqslant (S_n + S_{n-1} + \sum_{i=0}^{R_{imax}-1} S_{n-2-i})T \tag{6.4}$$

For $n = 1$: obviously, $S_n = 1$ and $D = T$.

For $n \geqslant 2$ **and** $S_n = 1$: If a linear network has only one source, each sensor node transmits one packet to its downstream neighbour. The optimal frame size depends on the bottleneck interference range. $D_{min} = 3T$ if each node only can disturb its nearest neighbour on both sides. Therefore, the corresponding frame size can be 3 slots per frame. If some nodes have a large interference range than others like the model of Fig. 6.2, the minimum frame size D_{min} will be increased. □

Theorem 2. *The generalised system throughput (Erlangs) S can be expressed as:*

$$S \leqslant S_{max} = \begin{cases} \dfrac{S_n}{S_n + S_{n-1} + \sum\limits_{i=0}^{R_{imax}-1} S_{n-2-i}}, & n \geqslant 2. \\ \\ 1, & n = 1. \end{cases} \tag{6.5}$$

Proof. **For** $n \geqslant 2$: The generalised system throughput S represents the data received by the sink node over the period of time since the transmission started. Accordingly, there is a inverse relationship between D and S. The maximum S is obtained when the minimum value of D is achieved. We can derive (6.1) for the calculating S. During a frame duration D, there are S_n packets can be successfully received by sink under the fair-access criterion. Since we can minimise D to achieve maximum S, therefore:

$$S_{max} = \frac{S_n T}{D_{min}} \leqslant \frac{S_n T}{(S_n + S_{n-1} + \sum\limits_{i=0}^{R_{imax}-1} S_{n-2-i})T} \tag{6.6}$$

For $n = 1$: when a source node is locate one hop away from the sink, then the throughput is always the same as the traffic load, therefore S_{max} =1 Erlang.

For $n \geqslant 2$ **and** $S_n = 1$: If a linear chain has one source, the sink is supposed to receive one packet during t. Under a fair-access criterion, the upper bound of throughput only relies on the critical interference range of network. If each node can interfere at most one node on their each side, the S_{max}= 0.33 Erlangs, then one in every three nodes can transmit simultaneously. This means that no two nodes share a same time slot within three-hop distance. If the node interference range is not identical, following the Liebig's law of the minimum[116], the channel utilisation depends on the critical interference range R_{imax}. The higher the R_{imax} the lower the S. For example, the S for the model of Fig. 6.2 is 0.25 Erlangs since node 6 is a bottleneck.

□

6.3.2 Actual Throughput Analysis

To validate the theoretical analysis about frame size and throughput presented in the previous section, we simulate a network shown in Fig. 6.2 and basic network parameters as listed in Table.5.1. Node 1 is chosen as the only source node along the chain. According to (6.1) and (6.5), the optimal frame size for this model is four slots and the maximum system throughput will be 0.25 Erlangs. To demonstrate the impact of frame size on throughput performance, we collected the throughput values under different frame sizes (three slots, four slots, and five slots).

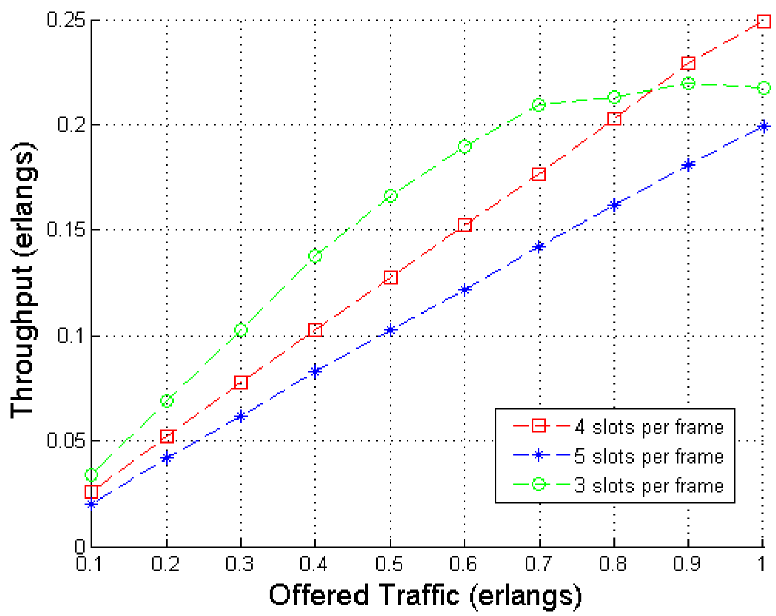


Figure 6.5: Comparison of Throughput Performance

According to the simulation results presented in Fig. 6.5, the throughput performance with three-slots per frame outperforms the throughput of four-slot and five-slot per frame when traffic load is less than 0.85 Erlang. However, the throughput with four-slots outperforms the other results in the range of 0.85 to 1 Erlang. When the traffic load is saturated (1.0 Erlang), the throughput with four-slot per frame reaches the highest value which is 0.25 Erlangs. By comparing throughput results, it can be seen that the optimal frame size cannot be determined by only considering throughput performance since none of these three results is the highest across the whole range of traffic loads. This observation is different from the throughput performance comparison of single-hop networks presented in Fig. 5.4. In multi-hop networks, system offered traffic is only contributed by source nodes. Therefore, the lower the traffic load, the lower the node neighbourhood contention level. In other words, bottleneck nodes may not receive enough packets from source nodes under low load conditions so that they can not disturb other nodes frequently. This states that the lower frame size may achieve higher throughputs under low load conditions. However, to ensure that each node can exploit a unique contention-free slot when the neighbourhood contention level is maximum (all source and relay nodes have packets to send during each frame) only the frame size which achieves highest throughput under saturated load conditions can be considered as the optimal frame size. By strike a balance between throughput and delay, four-slots is the optimal frame size for this scenario even if it does not alway achieve the highest throughput. This conclusion fully corresponds to the previous theoretical analysis.

To further investigate the throughput performance of ALOHA-Q in linear chain networks, we simulate multiple linear chain networks with different route lengths (L) and node interference ranges (R_i). In the simulated networks, nodes are equidistantly distributed and they have a uniform reception range and interference range. We consider the throughput at the sink node as the system throughput. Some default network parameters all listed in following Table.

Parameter	Value
Number of Source Nodes	1
Number of Relay Nodes	2, 6, 14
Transmission Route Length	4 hops, 8 hops, 16 hops
Node Interference Range	1, 2, 3 hops on each side
Offered Traffic Load	0.1-1.0 Erlang

Table 6.1: Simulation Parameters II

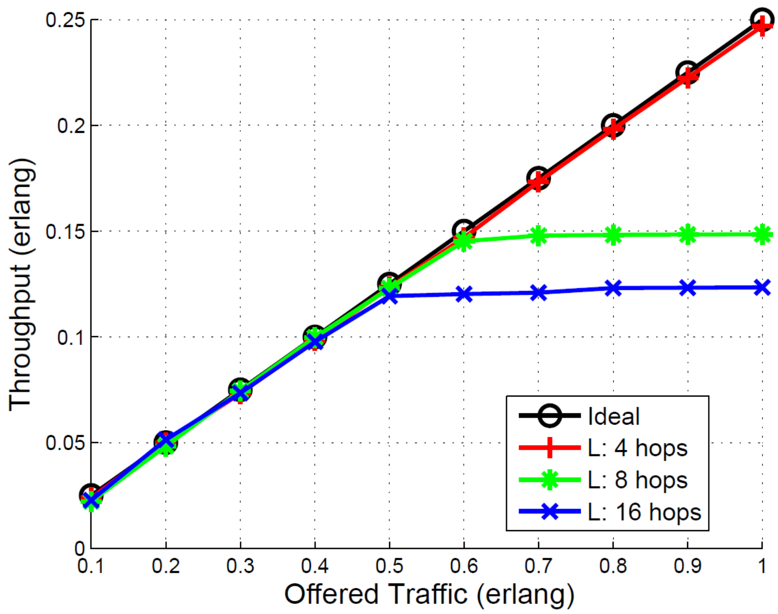


Figure 6.6: Comparison of Throughput Performance

Fig. 6.6 shows a comparison of system throughput during the steady state. In this scenario, R_i and R_r covers a two-hop and one-hop distance respectively. Based on (6.1), the corresponding frame size is set to four slots (optimal value) to ensure the optimum schedule to be feasible. It is clear to see that the maximum throughput performance is reached when the route length L is four hops. This is consistent with the maximum throughput calculated by (6.5). In this case, the throughput increases linearly as the offered traffic level rises. Although the effect of the guard band slightly degrades the throughput performance, the level of maximum throughput is still close to the theoretically attainable throughput: 0.25 Erlangs. This indicates that nearly all the traffic offered by source node can be received by the sink node, and no traffic will be congested in the route. However, when the route length is increased to eight hops, the throughput goes up linearly from 0.1 Erlangs to 0.6 Erlangs. Subsequently, the throughput value is maintained at a level around 0.15 Erlangs. The throughput performance becomes even worse when the route length is further increased to sixteen hops, where the maximum throughput can only reach 0.125 Erlangs. These results indicate that the longer the route, the lower the maximum throughput level can be obtained. This does not match the previous analysis which shows that the maximum throughput is only affected by R_{imax} and S_n

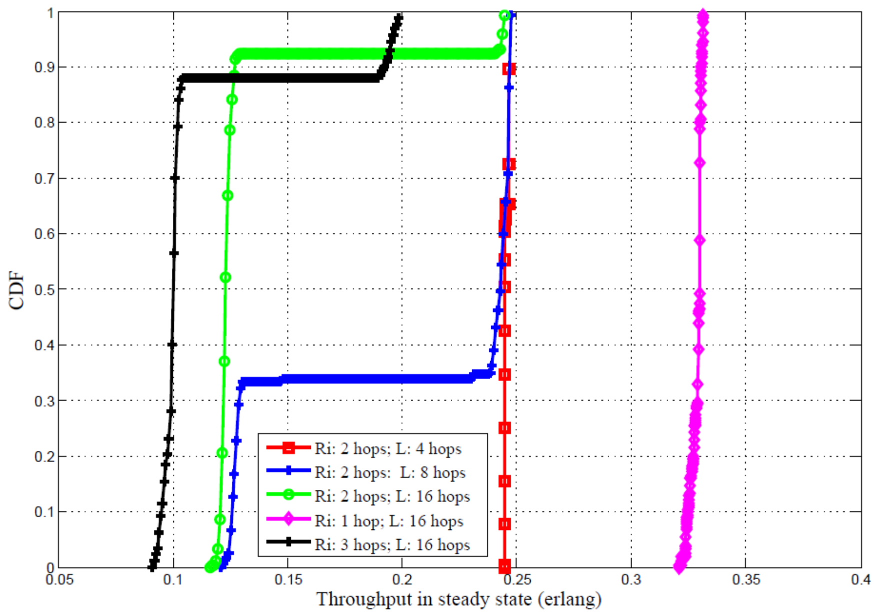


Figure 6.7: The CDF of Throughput Performance

Fig. 6.7 demonstrates the distribution of system throughput performance when L and R_i are varied. The frame size is set to the optimum value according to (6.1). To investigate the performance of ALOHA-Q under heavy load conditions, the offered traffic load is fixed at the saturated level which means the source node always has a packet to send during each transmission round. According to (6.5), maximum throughput when R_i covers nodes in a one hop, two hops and three hops distance will be 0.33 Erlangs, 0.25 Erlangs and 0.2 Erlangs respectively. As shown in Fig. 6.7, when R_i is set to two hops and L is set to sixteen hops, the throughput values focus on the region of 0.333 Erlangs. That fully corresponds to the steady state where a unique contention-free slot is obtained by every node. When L is increased but R_i remains the same, there is a decreasing proportion of the throughput values focused on the region of maximum throughput and a increasing proportion of throughput values converge to the region between 0.12 Erlangs to 0.125 Erlangs which is about half of the maximum throughput. In addition, the throughput values are distributed in the ideal region when R_i covers node within one hop distance and the L is sixteen hops. However, when L

stays same but R_i is increased, the throughput values will no longer be distributed in the region of maximum throughput. It can be found that the higher the R_i, a lower proportion of throughput values converge to the maximum throughput.

6.3.3 Hidden Node Problem

The simulation results presented in the Fig. 6.6 and the Fig. 6.7 indicate an important conclusion: increasing the transmission route length or the node interference range can negatively affect the throughput performance of ALOHA-Q. This conclusion contradicts with the assumptions of ALOHA-Q. The reason can by found by investigating the effect of the hidden node problem.

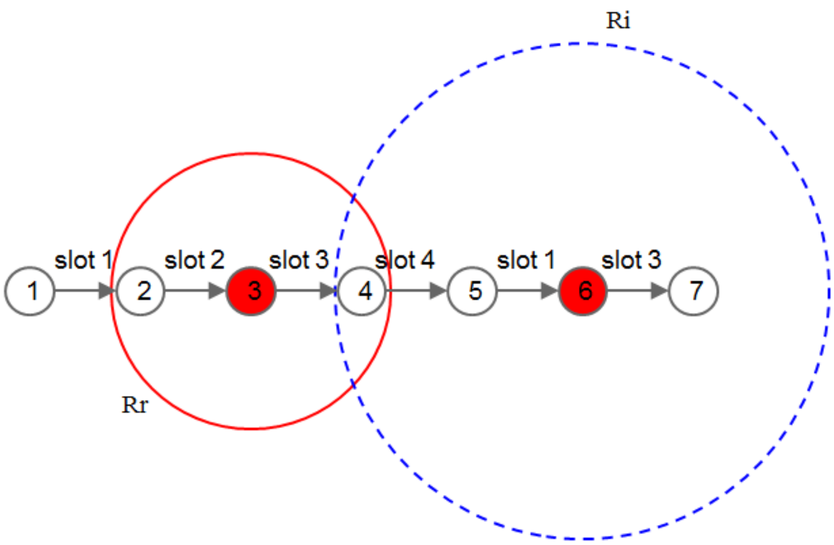


Figure 6.8: Hidden Node Problem

Fig. 6.8 illustrates the hidden node problem. In this scenario, nodes 3 and 6 share time slot 3. Node 3 does not interrupt the connectivity between node 6 and 7 because node 6 is out of the interference range of node 3. In another respect, since node 4 is within the interference range of node 6, the connectivity between node 3 and 4 could be impacted by node 6. Therefore, node 6 can be considered as a hidden node to node 3. When node 3 and 6 send a packet simultaneously during time slot 3, node 3 will fail to deliver a packet to node 4 while the packet transmitted from node 6 will be successfully received by node 7. Consequently, node 3 is expected to pick a new time slot if it experiences successive interference from node 6. However, in multi-hop networks, the incoming traffic of each relay node is provided by the nodes in the previous hop. Thus, node 6 can only disturb packet reception at node 4 when it has packets to send. After a few transmission rounds, there are no packets in the sub-queue of node 6 because newly generated packets have been backlogged at node 3. Once node 6 is run of packets, then node 3 can successfully deliver data packets to node 4 until node 6 receives a new packet. In this process, node 3 experiences periodic failure and successful transmissions as node 6 periodically receives data packets. Accordingly, the probability of successful transmission of node 3 is 50% and half of data traffic provided by the source node will be congested at node 3.

Nodes suffering interference from hidden nodes can be regarded as bottleneck nodes. To further explain the slot selection behaviour of bottleneck nodes when the hidden node problem occurs, it is necessary to examine the iteration of Q values. Let Q_0 represent a bottleneck node weight value of its preferred slot just prior to experiencing interference from a hidden node, and Q_i (i $\in$ 1,2,3...) be the weight value of its preferred slot after the i-th transmission attempt since interference starts. Therefore, Q_1 will be:

$$Q_1 = (1-\alpha)Q_0 - \alpha r; \tag{6.7}$$

For the simulations presented in this thesis, the reward value r was set to one, so (6.8) only involves the learning rate α and the initial weight value Q_0. Let N $\in$[R_r,R_i] refer to the number of hops between a bottleneck node and a hidden node. According to the previous discussion, a bottleneck node suffers N successive failures and successful transmissions in turn due to periodic interference from a hidden node. So Q_N can be expressed as:

$$Q_N = (1-\alpha)^N (Q_0 + 1) - 1; \tag{6.8}$$

After 2N transmission attempts, a node has experienced a period of N successive failures and N successful transmissions, so Q_{2N} can be calculated as:

$$Q_{2N} = (1-\alpha)^N (Q_N - 1) + 1; \tag{6.9}$$

Using (6) and (7), Q_{2N} can be rewritten as:

$$Q_{2N} = (1-\alpha)^{2N} (Q_0 + 1) - 2(1-\alpha)^N + 1; \tag{6.10}$$

Letting β = 1- α, (6.11) can be rewritten as:

$$Q_{2N} = \beta^{2N} Q_0 + (\beta^N - 1)^2; \tag{6.11}$$

After $m \in$ (1,2,3...) pairs of failure & successful cycles, the bottleneck node weight value for its preferred slot will be:

$$Q_{2mN} = \beta^{2mN}(1 - Q_0) + \frac{(1-\beta^{2mN})(1-\beta^N)}{1+\beta^N}; \tag{6.12}$$

The partial derivative of function Q_{2mN} with respect to m is:

$$\frac{dQ_{2mN}}{\mathrm{d}m} = \ln \beta^{2N(1+m)} \left(\frac{2\beta^N}{1+\beta^N} - Q_0\right); \tag{6.13}$$

Since $\beta \in (0,1)$, $\frac{dQ_{2mN}}{\mathrm{d}m} < 0$. This implies that a node's highest weight value maintains a downward trend when facing the hidden node problem. However, since a bottleneck node makes a great number of transmission attempts in its preferred slot, the corresponding weight value is close in the limit to the minimum critical value, which can be expressed as:

$$\lim_{m \to \infty} Q_{2mN} = \frac{1-\beta^N}{1+\beta^N} \tag{6.14}$$

Let A_{max} and A denote the range of the highest weight value and all possible weight values of bottleneck nodes respectively. Hence,

$$A_{max} = (\frac{1-\beta^N}{1+\beta^N}, r); A = (-r, r); \tag{6.15}$$

It is easy to figure out that $A_{max} \subset A_{rest}$. Therefore, a bottleneck node may never consider reselecting a new slot because the lower bound of its highest weight value may always be bigger than the remaining weight values. When that happens, the system will remain in a pseudo steady state where all nodes retain their preferred slots but packet collisions still exist. To illustrate, consider the interference model demonstrated in Fig. 6.8. If node 6 chooses the same slot as node 3, both of them will share a same slot. As a result, a considerable amount of traffic will be congested at node 3 and the end-to-end throughput will be significantly reduced. According to simulation results, the effect of hidden nodes happens more frequently if a multi-hop model has a relatively longer route length or a higher local interference level.

6.3.4 Slot Swapping Technique

To overcome the hidden node problem, a bottleneck node has to give up its preferred slot in this pseudo steady state because hidden nodes can not observe that they disturb the transmission of hidden nodes. This can be achieved by increasing the punishment value. However, forcing bottleneck nodes to reselect new time slots during the pseudo steady state could disturb the transmission of other nodes that are not suffering the hidden node problems thereby leading to extra packet collisions. This section proposes a slot swapping technique allowing two adjacent nodes to exchange their preferred slots with the aim of smoothly extending the distance between any two nodes that select same time slot.

Fig. 6.9 illustrates a clear workflow of the slot swapping technique. In this case, R_r and R_i are chosen as one-hops and two-hops respectively. In round one, node 3 is experiencing interference from node 5 because they are using time slot 3. By checking

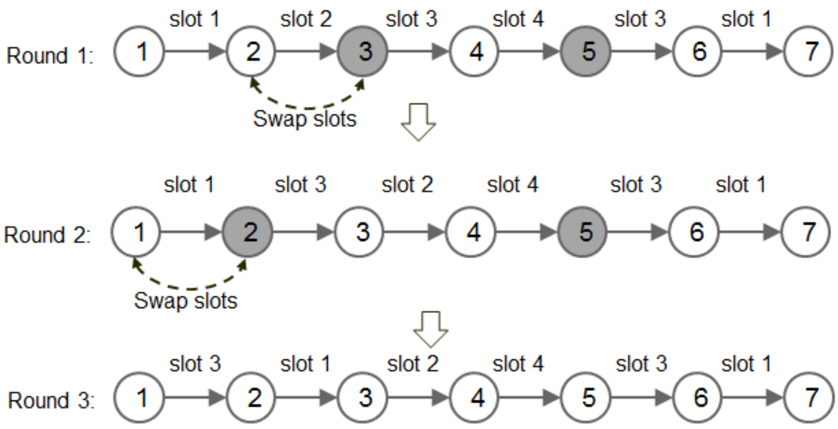


Figure 6.9: The Workflow of Slot Swapping Technique

its transmission history, node 3 observes that it is suffering periodic interference if it experiences periodic successful and failure transmissions. Subsequently, node 3 decides to exchange the preferred slot with node 2 through ACK packets. During the second round, node 2 and 5 share time slot 3 and node 2 is suffering periodic interference from node 5. Therefore, the slot swapping process will occur between node 2 and 1. The distance between the nodes sharing slot 3 has been extended to four hops from the third round. Hence, signals from node 5 will no longer interfere with the transmissions of other nodes. Therefore, all nodes have entered into the real steady state. Furthermore, if there are more bottleneck nodes in the transmission route, the slot swapping process will be repeatedly executed until there is no contention over the channel. In addition, the source nodes will not swap slots with other nodes. Instead, the source nodes are required to immediately reselect another time slot if they suffer periodic interference. The detailed algorithm for slot swapping is shown as follows:

while Node is in active mode **do**
    if Node periodically experiences success and failure transmission **then**
        if Node is a source node **then**
            Reselect a new slot as a transmission slot.
        else if Node is a relay node **then**
            Add the information about the preferred transmission slot to an ACK packet

and send this packet to the the upstream node.
if ACK packet has been sent **then**
Select the current reception slot as the new transmission slot.
end if
end if
else if Node receives an ACK packet **then**
Decode the ACK packet.
if Packet contains information about preferred transmission slot of the downstream node **then**
Select this new slot as the transmission slot.
end if
end if
Finish current slot swapping round.
end while

The effect of this slot swapping technique can be observed in the Fig. 6.10 and the Fig. 6.11. Through exchanging preferred slots, the route length and the node interference range can no longer affect the throughput performance in the steady state and the hidden node problem will not be the barrier to preventing individual nodes finding their contention-free slots. With the slot swapping technique, the best performance of the ALOHA-Q protocol can always be achieved once the optimum frame size is found.

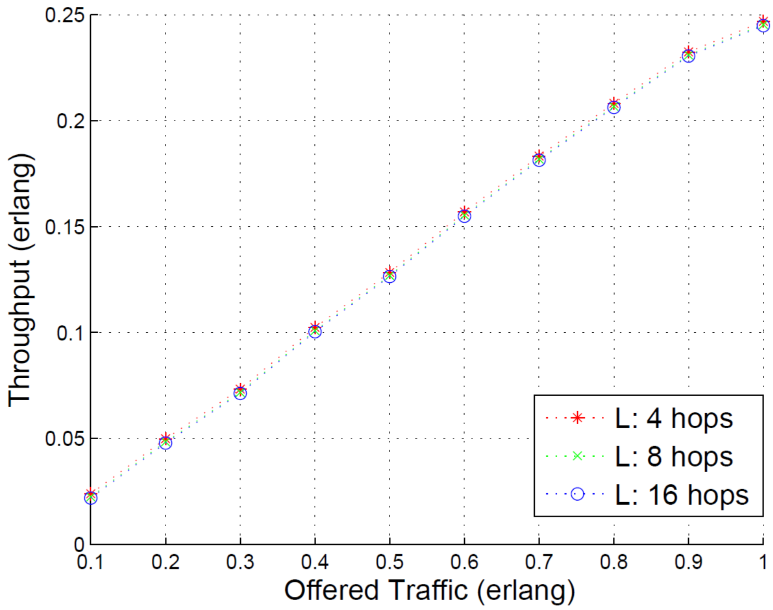


Figure 6.10: The Comparison of Average End-to-end Throughput Performance with Slot Swapping

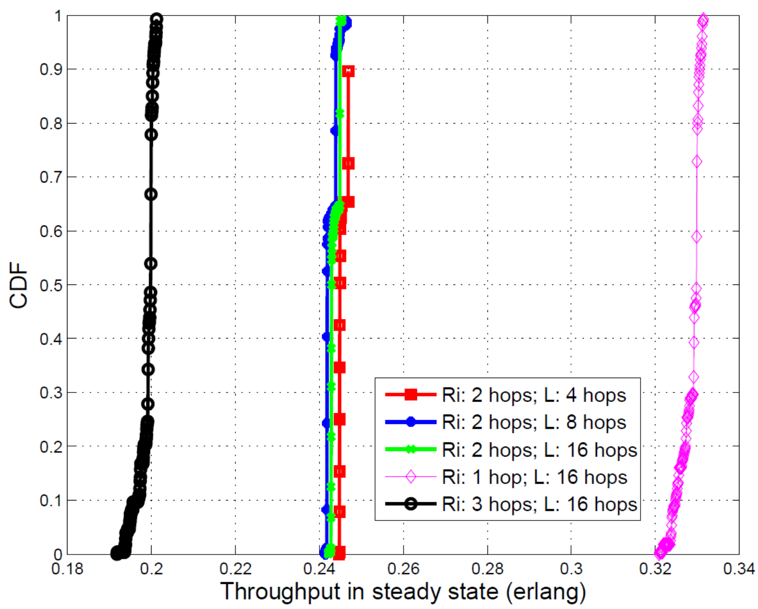


Figure 6.11: The CDF of End-to-end Throughput Performance with Slot Swapping

6.3.5 Frame Size Adaptation for Linear Chain Networks

The main contribution of the slot swapping technique is to allow sensor nodes to always achieve the maximum attainable throughput when the frame size is optimal. Consequently, we can surely draw a conclusion that the selection of optimal frame size leads to the maximum value of throughput. In realistic scenarios, the ideal frame size can only be manually set-up at each node if the distribution of nodes is explicit. However, if the number of sensor nodes and the local interference range (in hops) of nodes are uncertain, these parameters can not be manually calculated. Even if the optimal frame size is obtained before implementation, the network topology and node interference range could be affected by hardware problems or environmental changes during system run time. This potentially increases the complexity of selecting an ideal frame size. To solve this problem, we propose an intelligent frame size adaptation scheme for linear chain networks.

The key idea of the proposed scheme is similar to the DFA scheme introduced in Chapter V, in which sensor nodes periodically examine their local channel performance under different frame sizes obtained from the source, but their opinion about the current frame size will be back-propagated to the top end source node through ACK packets. To this end, sensor nodes are allowed to determine whether the frame size is optimum or not by monitoring their local transmission history. The workflow of the frame size adaptation process is shown in Fig. 6.12, and detailed algorithm of frame size selection is as follows:

if Source node == TRUE **then**
  if The beginning of the first frame size adjusting window **then**
    Each node selects an initial frame size S of two slots.
  else if The end of the i-th frame size evaluation window **then**
    if Source node has received a frame size adjustment request **then**
      Set the current frame size value as a new lower bound S_{min}.
      if $S_{max} == 0$ **then**
        Choose $2S_{min}$ as the new frame size from next the window.
      else if $S_{max} > 0$ **then**
        Choose $\frac{(S_{max}-S_{min})}{2}$ as the new frame size from the next window.
      end if

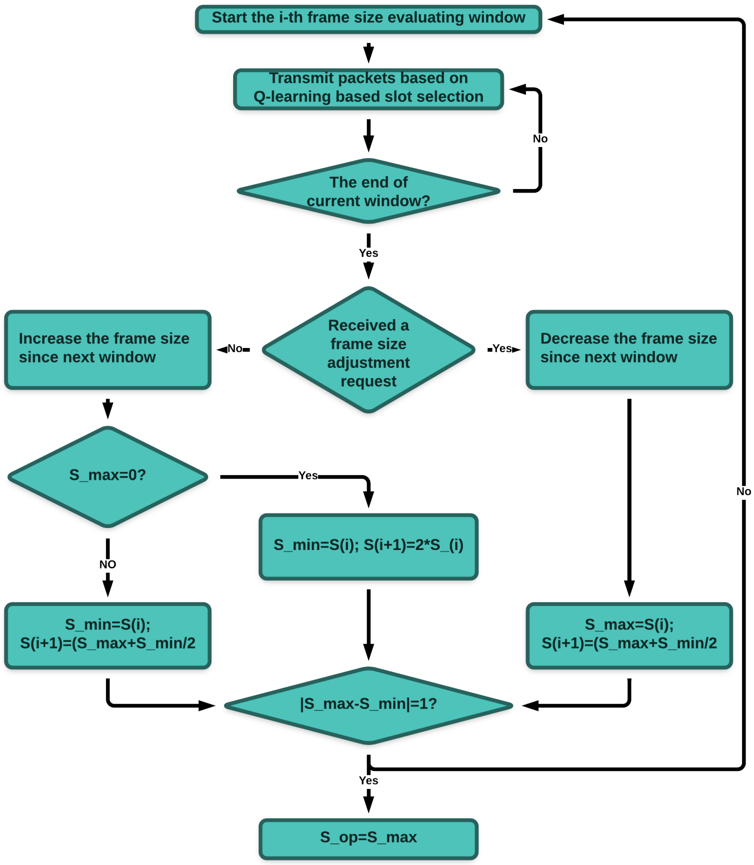


Figure 6.12: The Workflow of Frame Size Evaluating Process

else if source node has not received any frame size adjustment request **then**

  Set current frame size value as a new upper bound S_{max}.

  if $(S_{max} - S_{min})$==1 **then**

    Choose S_{max} as the final frame size.

  else if $(S_{max} - S_{min}) > 1$ **then**

    Choose $\frac{(S_{max}-S_{min})}{2}$ as the new frame size from the next window.

  end if

end if

Spread current frame size to the rest of nodes along current transmission route via ping packets.

end if

else if Source node == TRUE **then**

Spread current frame size to the rest of nodes along current transmission route via ping packets.

end if

In the proposed frame size adaptation scheme, the decision of frame size adjustment can only be made by the top end source node. Once a source node decides to adjust the current frame size, it will embed its decision into a small-sized ping packet and then spread the ping packet to the rest of the nodes along the transmission route. The duration of the frame size adaptation process is divided into repeating frame size evaluation windows. Each window is composed of a fixed number of frames. Before starting the first window, each node selects an initial frame size of two slots, which corresponds to an assumption that the current transmission route contains at least one source node and one relay node. During each window, nodes evaluate the current frame size by monitoring its subqueue size. If the value of the preliminary frame size is not large enough, steady state can never be reached. As a result, packets will be backlogged within the transmission route. An ever-increasing subqueue size indicates that a node is still likely to seek a collision-free slot, and the current frame size may less than the optimal value. On the other hand, if the number of packets in a subqueue is unchanged over a period of time, the node may has found a collision-free slot and the current frame size is likely to be greater than the optimal value. If the subqueue size of a certain relay node exceeds a threshold N, it will send a request to the top end source node to increase the current frame size. The request will be embedded in an ACK packet, which ensures that the request can be quickly sent back to the source node after a number of frames. The magnitude of frame size adjustment follows the same binary search strategy mentioned in Chapter V. If a source node receives a frame size adjustment request or its subqueue size exceeds N, it will set the current frame size at a new lower bound S_{min} and then select $2S_{min}$ as the new frame size from the next window. If the source node has not received a frame size adjustment request and its subqueue size is less than N until the last frame of each window, it will consider the current frame size as a new upper bound S_{max} and then choose $\frac{S_{max} - S_{min}}{2}$ as the new frame size from the next window. By periodically adjusting the frame size, the range of optimal frame sizes is kept narrow. When there is just one slot difference between the upper and lower bound, the upper bound is selected as the final frame size.

6.3.6 Performance Analysis

To evaluate the effect of the frame size adaptation scheme, a set of simulations were conducted via OPNET. The network topology modelled in simulations is shown in Fig. 6.13. A group of ten nodes form a single-dimensional chain topology. The nodes along the chain are unequally spaced, every node maintains communication with its downstream and/or upstream neighbour. The distance between any two nodes are D is a random number between 0 and R_r. The two nodes located at both ends (node 1 and node 10) are chosen as a source node or a sink node. Other basic simulation parameters are listed in following Table 6.2

Parameter	Value
Source Node ID	1, 10
Sink Node ID	10, 1
Node Interference Range	25 m
Node Interference Range	30 m
Offered Traffic Load	1.0 Erlang
Initial Frame Size	2 Slots
Frame Size Evaluation Window Size	200 frames
Subqueue Size Threshold N	100

Table 6.2: Simulation Parameters III

In the simulated network, we define two different data transmission routes: route A (from left to right) and route B (from right to left). It can be seen that the densest slot contention group in this network is composed of node 1, 2 and 3. Under saturated load conditions, node 1 can be disturbed by its two downstream neighbours (2 and 3). In contrary, node 10 can only be interfered by node 9. Therefore, route A refers to the situation that the source node is considered as a bottleneck node while the route B represents the scenario that bottleneck nodes are closes to the sink.

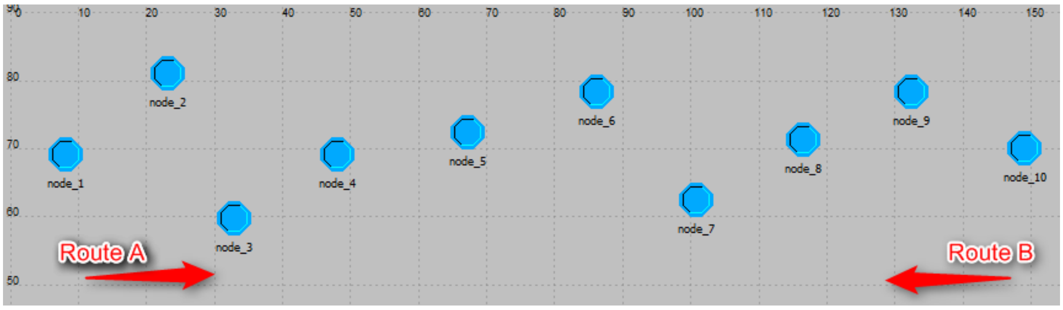


Figure 6.13: Linear Chain Network

Fig. 6.14 demonstrates the frame size adaptation process of individual nodes during the whole simulation period. This result applies to both of route A and B since the adaptation process is not affected by the direction of traffic flow. Based on network parameters and the distribution of nodes, it can be observed that the critical interference range R_{imax} for this network is a two hop distance. Accordingly, we can figure out the optimal frame size based on (6.1), which is four slots. From the beginning of the adaptation process, the frame size of individual nodes starts from a initial value of two. As previously mentioned, the adjustment of frame size is controlled by the source node. Once the subqueue size of source node exceeds the threshold N or the source node receives a frame size adjustment request during a frame size evaluation window, it will spread a ping packet contains a new frame size to rest of nodes along transmission route. In such a way, the consistency of the frame size adjustments can be guaranteed. By periodically examining the channel performance, the frame size eventually reach to four slots, which is exactly corresponds to the optimal value.

Fig. 6.15 and Fig. 6.16 present the comparisons of the real-time system offered traffic

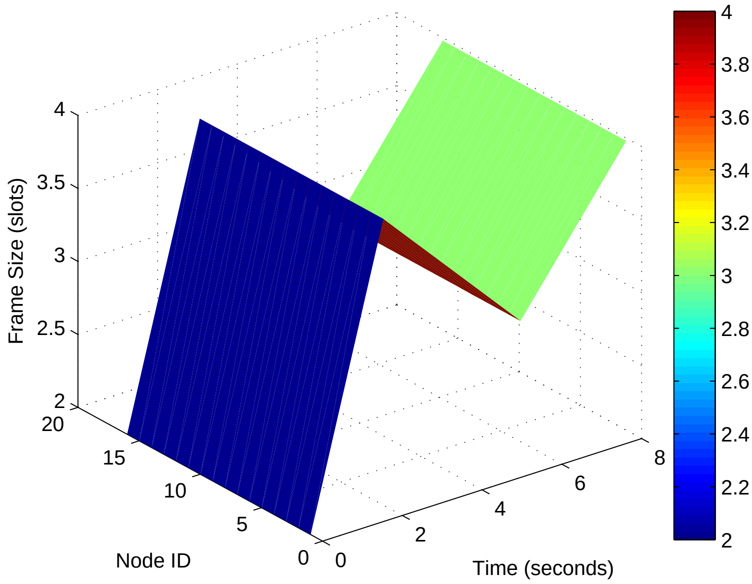


Figure 6.14: Frame Size Adaptation Process for Individual Nodes

and system throughput of two transmission routes. Although the two results were collected from a single simulation, the corresponding results remain the same in all cases when repeatedly run simulations. It can be clearly observed that the offered traffic load for both routes starts at 0.5 Erlangs, which corresponds to the initial frame size of 2 slots. Due to differences in neighbourhood interference, the offered traffic of source node 1 drops rapidly meanwhile the offered traffic of source node 10 is maintained around the maximum level during the first evaluation window. Accordingly, the two source nodes may have different opinions about current channel performance result in different beliefs in adjusting frame size. However, based on the proposed adaptation algorithm, node 10 can still increase current frame size once it receives a frame size adjustment request from other downstream nodes. When the adaptation process ends, the offered traffic of source nodes and sink nodes for both routes approach a same level when frame size is either equal or greater to the optimal value. Otherwise, the throughput of sink node always less than the throughput actually contributed by the source node. That is because of the proportion of traffic backlogged in relay nodes due to the contention in slots.

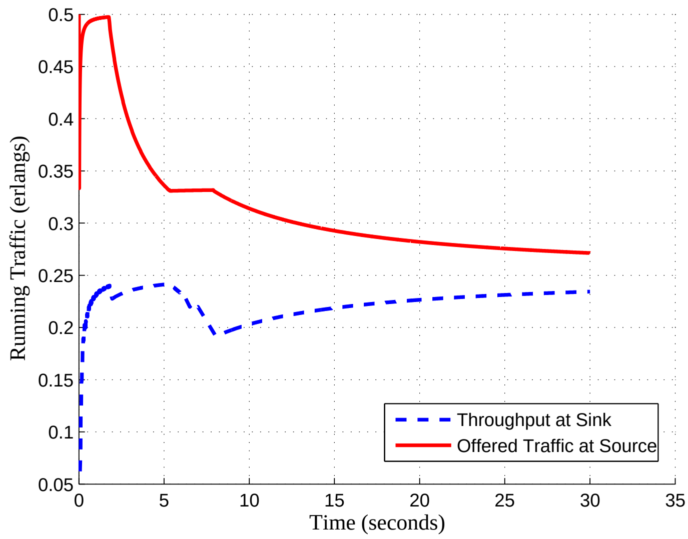


Figure 6.15: Comparison of Real-time Offered Traffic and Throughput for Route A

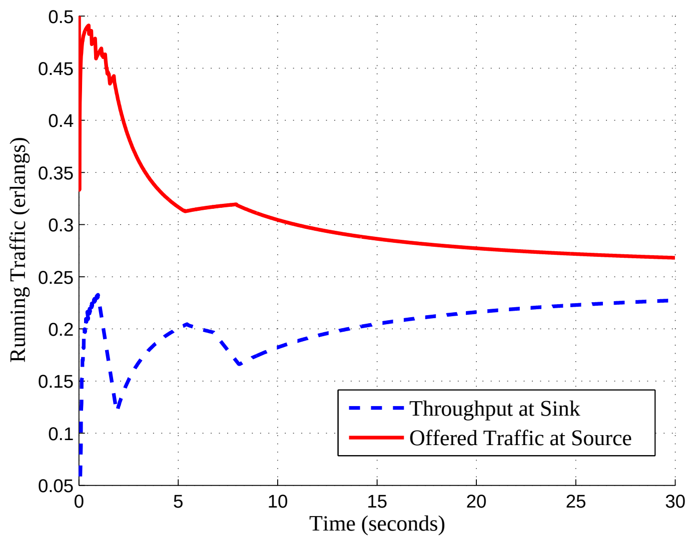


Figure 6.16: Comparison of Real-time Offered Traffic and Throughput for Route B

6.4 Adaptation of ALOHA-Q to Cross Chain Networks

6.4.1 Subframe Adaptation Scheme for Cross Chain Networks

The previous section presented the adaptation of ALOHA-Q to a linear chain topology. Simulation results imply that the maximum throughput of a linear chain network can be guaranteed if an optimal frame size is selected by each node along the transmission route. In this section, the adaptation of ALOHA-Q to two dimensional cross chain networks is presented. Consider the example in Fig. 6.17. In this network, two transmission routes are overlap with each other and meet at the node $N_{1,j}$. Let's assume the source node $N_{1,1}$ detects an event at time instant T and the source node $N_{2,1}$ detects another event at time instant T', where $T' > T$. Obviously, the traffic flow generated by $N_{1,1}$ will be disturbed after T' because the relay node $N_{1,j}$ starts to forward extra packets generated by $N_{2,1}$. This leads to difficulty in achieving frame size adaptation for this network because the frame starting times of the two groups of nodes are not synchronised. Therefore, the consistency of frame size adaptation cannot be ensured so that system steady state can never be reached.

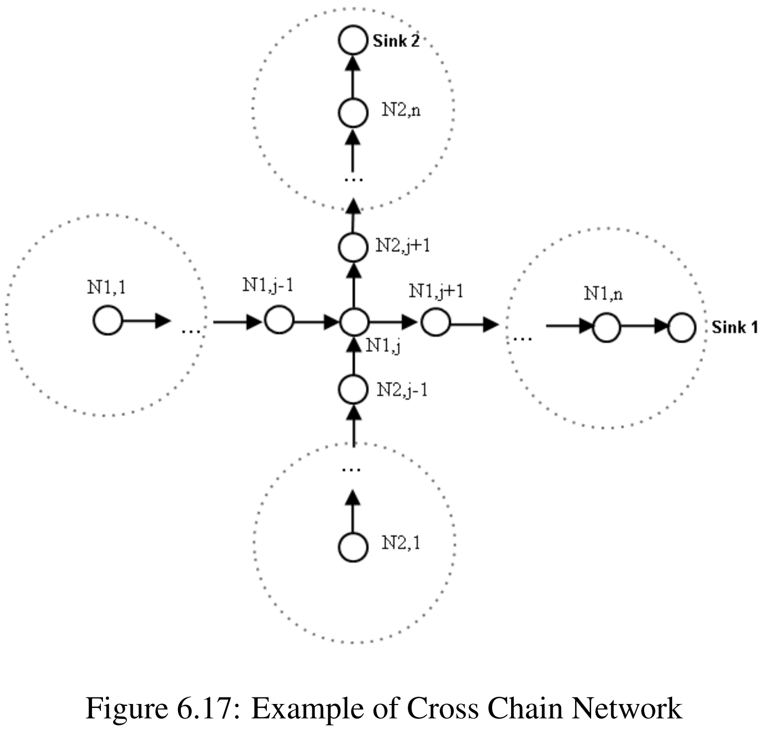


Figure 6.17: Example of Cross Chain Network

In this section, a subframe adaptation scheme is proposed to assist sensor nodes to deal with the unpredictable disturbance caused by extra incoming traffic flow. The core design idea of the proposed scheme is to assign unique subframes to the nodes along different transmission routes in order to avoid channel contention. Consider the model in Fig. 6.17, let's assume the optimal frame sizes for horizontal chain and vertical chain are S and S' respectively. Since all nodes can obtain their optimal frame size by conducting frame size adaptation, a superframe can be created which contains two subframes. Each subframe is dedicated to the nodes in the same transmission route. As Fig. 6.18 illustrates, the nodes in the horizontal chain can use subframe A which contains S slots, and a subframe B contains S' slots is assigned to the nodes in vertical chain. By this way, nodes along different transmission routes can exploit their best slots within a unique subframe. As a consequence, channel contention can be eliminated. This example serves to demonstrate the adaptability of the system in response to significant environmental change after initial convergence.

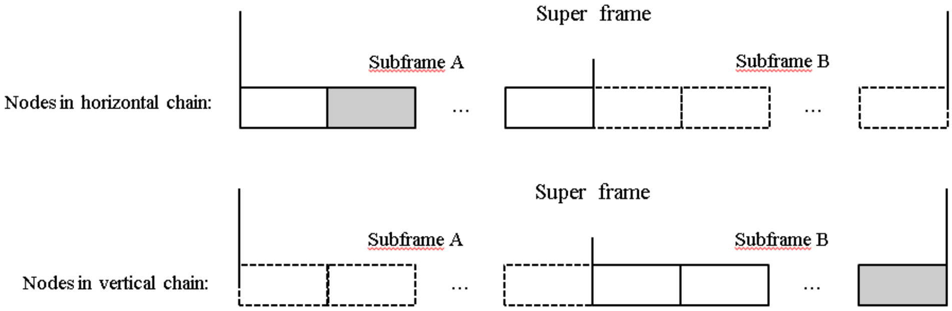


Figure 6.18: Subframe Structure

The biggest challenge for creating the superframe structure is to obtain the optimal subframe sizes of individual routes. In the proposed subframe adaptation scheme, the optimal frame size for a single transmission route can be obtained by conducting frame size adaptation within its individual sub-routes. Consider the model in Fig. 6.17, when $N_{1,j}$ receives a packet generated from $N_{2,1}$, it considers itself as a multi-reception node and then sends a frame size adaptation request to $N_{2,1}$ and $N_{2,j+1}$ via ACK packet and ping packet. When the request is received, $N_{2,1}$ will start to learn the optimal frame size S_{opt1} for the sub-route between $N_{2,1}$ to $N_{1,j}$ and $N_{2,j+1}$ start to learn the optimal frame size S_{opt2} for the sub-route between $N_{1,j}$ to sink 2. Meanwhile, $N_{1,j}$ keeps forwarding the packets generated from $N_{1,1}$. When S_{opt1} and S_{opt2} are obtained, a S_{opt}=max(S_{opt1},S_{opt2}) will be selected as the optimal frame size for the vertical chain. Subsequently, the two source nodes can exchange their optimal frame size through

$N_{1,j}$. Therefore, the superframe can be created.

The detailed algorithm for subframe adaptation scheme is shown as follows:

while Node is in active mode **do**
if Relay node == TRUE **then**
if Receive a packet from a new source node **then**
Multi-reception flag = TRUE.
Decode the packet and obtain the new routing list.
Add its downstream node address and a frame size adaptation request to an ACK packet, then back propagate this packet to the new source node.
Add the destination address and a frame size adaptation request to a ping packet, then send this packet to its upstream node.
end if
if Receive the optimal frame sizes for individual sub-routes **then**
Create a superframe and send the information about superframe size to individual source nodes via ACK packets.
end if
end if
if Source node == TURE **then**
if Sink is reached **then**
To learn optimal frame size for the entire transmission route.
if Optimal frame size is obtained **then**
Transmit packet at its preferred time slot according to ALOHA-Q.
end if
end if
else if sink is not reached **then**
if Receive a frame size adaptation request **then**
To learn the optimal frame size for the sub-route .
if Optimal frame size is obtained **then**
send optimal frame size for the sub-route to multi-reception node via

ping packets.

end if

end if

if Receive a superframe size from multi-reception node **then**

To create a superframe structure, and send information about the superframe size to the rest of nodes along current route via ping packets.

end if

end if

end while

6.4.2 Performance Analysis

To evaluate the proposed subframe adaptation scheme, a set of simulations were conducted via OPNET. We consider a cross chain network (as Fig. 6.19) consisting of twenty one nodes deployed in a 100m X 200m square sensing area. Eleven nodes equidistantly distributed in the horizontal chain and the vertical chain respectively. We assume that the generation of two traffic flows are triggered by two different events. The starting times of individual events follow a predefined sequence and with an interval of 20 seconds. The transmission begins with the assumption that the node neighbourhood interference level and the spatial distribution of nodes are unknown to each node, and each node will select two slots per frame as the initial frame size. In addition, some default simulation parameters are listed in Table.6.3

Parameter	Value
Source Node ID	1, 12
Sink Node ID	11, 21
Starting Time Sequence	1, 12
Starting Time Interval	20 seconds
Node Reception Range	15 m
Node Interference Range	25 m
Node Spacing Distance (horizontal chain)	10 m
Node Spacing Distance (vertical chain)	15 m
Offered Traffic Load	1.0 Erlang
Initial Frame Size	2 Slots
Simulation period	60 seconds
Frame Size Evaluation Window	200 frames

Table 6.3: Simulation Parameters IV

The frame size adaptation process of two source nodes is presented in Fig. 6.20. According to simulation parameters, each node in the horizontal chain has uniform interference range (hops) which is a two-hop distance, and the interference range (hops) of each node in the vertical chain is one-hop distance. The frame size shown in Fig. 6.20 refer to the total frame size of each node. Based on (6.1), the optimal frame sizes for the horizontal chain and the vertical chain are four slots and three slots respectively. It is observed that source node 1 has an initial frame size of two slots at the beginning of its transmission phase. After a short period, the first frame size upper bound (four slots) is obtained. By periodically evaluating channel performance, the frame size of node 1 is finally reaches three slots which is corresponds to the optimal value. On the other hand, source node 12 starts to transmit packets from twentieth second. However, the packets generated from node 12 cannot be directly forwarded to sink node 21 since node 6 is dominated by the first traffic flow. According to the subframe adaptation algorithm, node 12 is required to learn an optimal frame size for the sub-route between

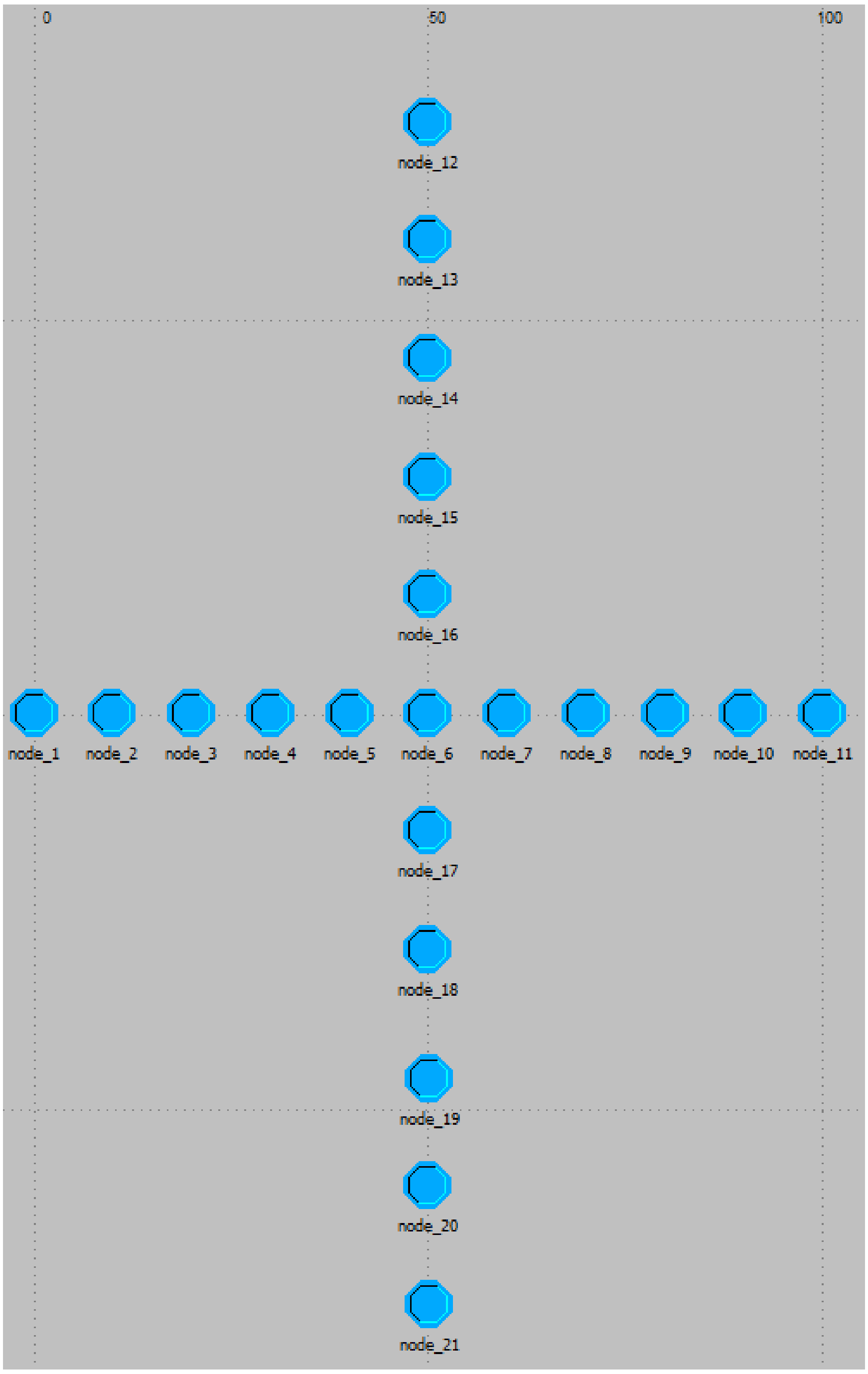


Figure 6.19: Cross Chain Network

node 12 and node 6. As Fig. 6.20 shows, the frame size of node 12 starts from two slots and ends up with three slots. This is also corresponds to the optimal value. Once the two source nodes have found their optimal frame sizes, they can exchange their current frame size through the multi-reception node 6. Eventually, a superframe contains

7 slots is created by each source node.

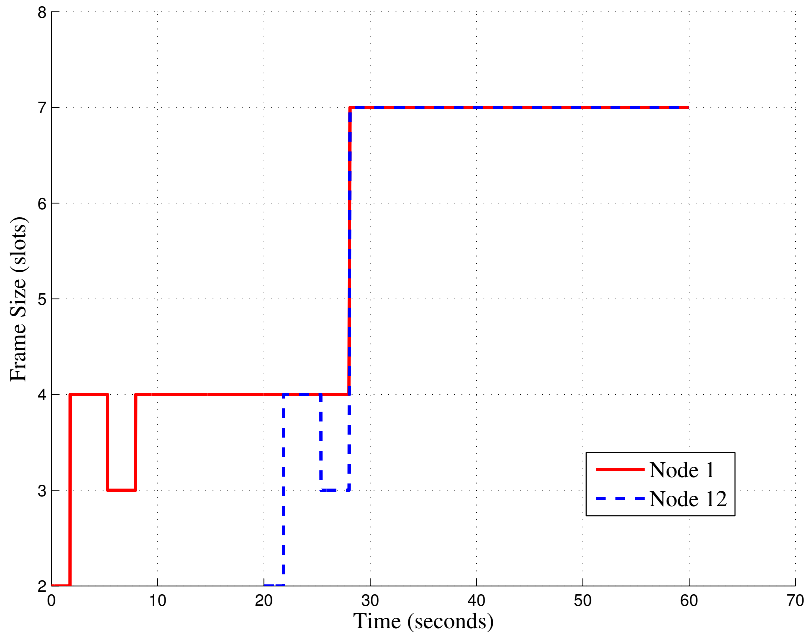


Figure 6.20: Subframe Adaptation Process

Fig. 6.21 demonstrates the probability of successful transmission over time for the two source nodes. Let P represents the probability of successful transmission, N_{ack} and N_{apt} refer to the number of received ACK packets and the number of total transmission attempts during a frame size evaluation window respectively. So P can be expressed by:$P = \frac{N_{ack}}{N_{apt}}$. By comparing Fig. 6.21 with Fig. 6.20, We can observe that P is maintained at a level very close to 1 only if current frame size is greater than or equal to the optimal value, which means the packet collision can be entirely eliminated during corresponding frame size evaluation windows. Besides, the P of the source node 1 keeps approaching 1 once it has found the optimal frame size. The generation of the second traffic flows does not affect the data delivery of the horizontal chain and the vertical chain adapts its frame size to achieve convergence itself, with P=1. This indicates that the proposed subframe adaptation scheme can effectively help nodes to accommodate the disturbance caused by overlapping of traffic flows without introducing packet collisions.

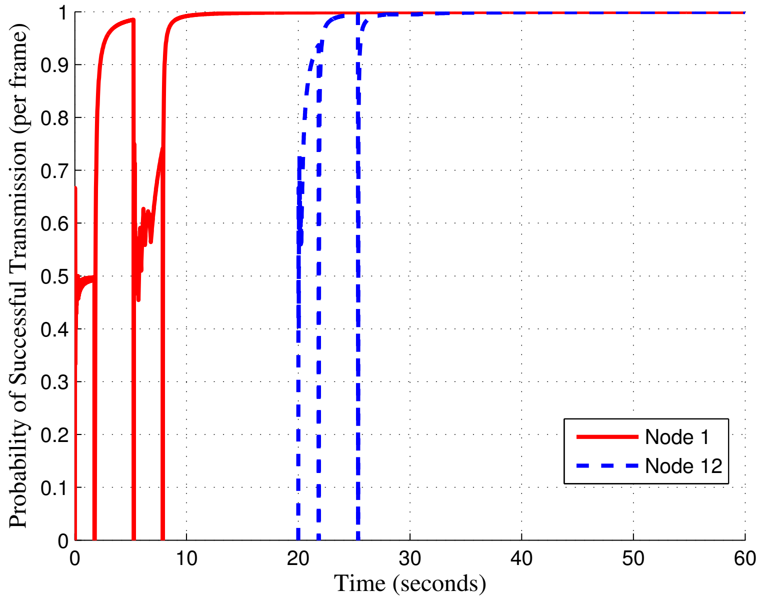


Figure 6.21: Real-time Probability of Successful Transmission

To investigate the channel performance throughout whole simulation period, we can compare the real-time running offered traffic load of each source node with the real-time running throughput of each sink node and multi-reception node (see Fig. 6.22). It can be found that the running offered traffic load of source node 1 and the running throughput of sink node 11 are close to 0.25 Erlangs after 7.9 seconds in which source node 1 has found its optimal frame size. This fully corresponds to the frame size adaptation process presented in Fig. 6.20. Similarly, the running offered traffic load of source node 12 converges to 0.33 Erlangs after 27.9 seconds in which it has found an optimal frame size for the sub-route between node 12 and node 16. When the subframe adaptation process ends, each node is assigned a superframe that contains seven slots. Therefore, each source node is allowed to transmit at most one packet during a superframe so that the maximum offered traffic load of each source node is 1/7 Erlangs. Moreover, the multi-reception node can use at most 2 slots per superframe since it has to relay all the packets generated from node 1 and node 12. Accordingly, the offered traffic load of each source node and the throughput of each sink node reduces rapidly after 27.9 seconds and will finally converge to 1/7 Erlangs. Meanwhile, the throughput of the multi-reception node 6 is going to reach 2/7 Erlangs.

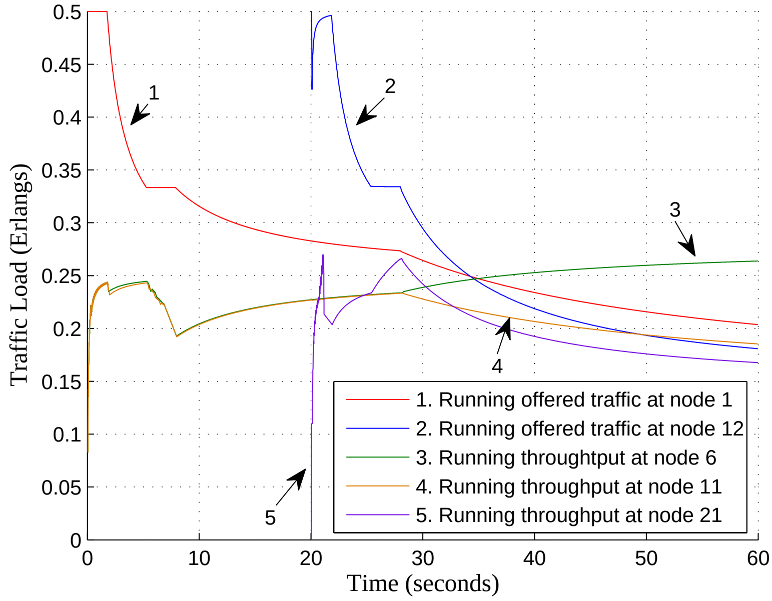


Figure 6.22: Comparison of Real-time Running Offered Traffic and Throughput

6.5 Summary

In this chapter, the design and evaluation of a self-adaptive ALOHA-Q protocol is presented. The proposed protocol adopts ALOHA-Q as a baseline MAC scheme and contains a set of novel features for the purpose of assisting sensor nodes to achieve collision-free scheduling under multi-hop conditions. The early sections demonstrate the adaptation of ALOHA-Q to a linear chain topology. Corresponding simulation results suggest that the throughput performance of ALOHA-Q can be negatively affected by increasing the transmission route length or the node interference range due to the effect of hidden nodes, leading to traffic congestion in the transmission route. A slot swapping technique has been proposed to help sensor nodes to smoothly exchange their scheduling policy to avoid interference from hidden nodes. To overcome the challenge in setting up optimal frame size, a frame size adaptation scheme specially designed for multi-hop networks is introduced. The proposed new frame size adaptation scheme allows sensor nodes to learning the optimal frame size of current transmission route. The subsequent section introduced a subframe adaptation scheme to adapt ALOHA-Q to cross chain topology. Its core idea is that data packets generated from different source nodes can be transmitted within unique subframes. Simulation results shown that the subframe adaptation scheme provides effectively solution to help nodes to achieve guaranteed channel utilisation and collision-free scheduling when they suffer the disturbance caused by extra incoming traffic flows.

Chapter 7

Future Work

Contents

This chapter discusses future directions based on the research work in this thesis to investigate potential enhancements about the proposed MAC protocols.

7.1 Frame Size Adaptation to Event-based Traffic Delivery

The frame size adaptation schemes proposed in this thesis are aimed to assist ALOHA-Q to achieve maximum channel utilisation under saturated traffic load conditions. The traffic model used in corresponding simulations assumes data packets are continuously generated by a large number of independent source nodes. However, many WSN applications are designed for monitoring specific events so that not all data captured by sensor nodes need to be transmitted and analysed. For such event-centric applications, traffic generation only corresponds to the occurrence of events. It is not efficient to allocate a dedicated slot to each node when channel contention level is low. To improved the efficiency of channel access, the proposed frame size adaptation schemes can be further enhanced to respond to traffic fluctuation. For example, to periodically estimate the channel contention level and ensure only active nodes have their dedicated slot. As a consequence, the system throughput can be increased, and the average packet latency can be reduced.

7.2 Intelligent Duty Cycling

Duty Cycling is one of the most effective approaches to conserving the power consumption of nodes. A number of MAC protocols support adaptive duty cycling strategies (such as T-MAC, U-MAC, and LPL-MAC). However, the MAC protocols proposed in this thesis do not introduce a wake-up/sleep mechanism as they focus on the initialisation of an optimal channel access policy. To extend network lifetime, it is desirable to develop a duty cycle scheme which can dynamically adjust the ratio of active and sleep periods of nodes under network condition changes. Besides, the learning experience of nodes can also be used to control their duty cycle. For example, nodes can switch to sleep during the slots which they have experienced a lot of failed transmissions, and increase the duration of the listening period when traffic load risese or interesting events occur. By increasing the intelligence of duty cycling, nodes can strike a balance between throughput performance and energy efficiency.

7.3 Lifelong Machine Learning

The concept of ALOHA-Q allows individual nodes to learn optimal scheduling policies within a set-up stage. Once system steady state is reached, the optimal scheduling policies will be exploited by nodes during their remaining operation time. However, the outcome of the initial learning process cannot be applied to a dynamic environment where node population, locations and some other network parameters are frequently changed. To handle various environmental changes, nodes have to relearn their best scheduling policies periodically. This leads to extra computational overhead and power consumption of nodes. Lifelong machine learning (LML) provides a potential solution to help sensor nodes quickly adapt to a constant changing environment. For example, nodes can retain the historical knowledge about slot/frame size selection in long-term memory. When the environment changes, they can select the most related prior knowledge to improve efficiency and accuracy of the learning of a new task. The use of previous experience by nodes can reduce the learning time and cost for developing a new scheduling policy in a new environment compared to using only the available knowledge.

Chapter 8

Summary and Conclusions

Contents

This thesis has presented the research work undertaking during Ph.D. study from 2011 to 2014 at the University of York. The early chapters have introduced the research motivation, comprehensive background knowledge, and useful methodologies. The later chapters demonstrate the main contributions of the thesis, which mainly focuses on developing intelligent MAC protocols to improve channel utilisation while reduce control overhead.

The choice of MAC protocols is one of the primary determining factors for wireless sensor nodes to successfully fulfil their missions. An efficient design of MAC protocol must be energy efficient to prolong the lifetime of the network. To balance system QoS and network scalability, an efficient MAC protocol should also provide an effective anti-collision solution and a flexible transmission mechanism. Moreover, system throughput, packet latency, bandwidth utilisation are also significant attributes in the design of MAC protocols. The recently proposed ALOHA-Q protocol is considered as a potential replacement to conventional signalling based or carrier-sensing based MAC protocols due to its simplicity and intelligent collision resolving features. However, the implementation of ALOHA-Q relies on many assumptions about node population, network topology, radio interference level, etc. These parameters cannot be easily obtained in realistic scenarios. The research work in this thesis emphasises on the frame adaptation of ALOHA-Q through the design of effective and efficient self-adaptive approaches to deal with foregoing practical problems.

The basic concept of WSNs is introduced in Chapter I. A detailed discussion about the history, design challenges, system classification, communication protocol stack, applications, and some other supportive technologies of WSNs are presented in this chapter. As an emerging technology, WSNs have a broad spectrum of potential applications. The special design and features of sensors make WSNs different from traditional data gathering tools. These characteristics pose great challenges for architecture and protocol design, performance modelling, and implementation.

The fundamental background and a literature review of MAC protocols for WSNs are presented in Chapter II. The attributes of WSNs and the characteristics of operating environment of sensor nodes make the design of MAC protocol very challenging. Contention-free protocols have appealing advantages in reducing collisions and ensuring fairness, but require strict time synchronisation. Contention-based protocols provide flexible transmission mechanism and good scalability. However, they are unable to achieve high throughput under heavy load conditions. The ALOHA-Q protocol introduced in the final part of this chapter is expected to keep all the advantages of contention-free and contention-based protocols but eliminate their drawbacks.

To achieve the maximum performance of ALOHA-Q under single-hop conditions, a frame size adaptation scheme is introduced in Chapter V. Nodes can learn an optimal value of frame size by periodically evaluating the channel performance with different frame sizes. A proactive jamming strategy is proposed to ensure the consistency of the frame size adaptation process. Results show that the optimal frame size can be selected by individual nodes once the adaptation process ends. As the consequence, a collision-free time slot can be exploited by each node during the remaining operation time.

In Chapter VI, ALOHA-Q is applied to two representative multi-hop topologies: linear chain topology and cross chain topology. The simulation results of the throughput performance of linear chain networks revealed the effect of hidden node problem. To overcome this issue, a slot swapping technique is proposed which helps ALOHA-Q to achieve maximum attainable throughput. This chapter also introduced a frame size adaptation scheme specially designed for multi-hop networks. The proposed new frame size adaptation scheme allows sensor nodes along a single transmission route to learn an optimal frame size. To allow nodes accommodate unpredictable incoming traffic flows, a subframe adaptation scheme is introduced. By this approach, data packets generated from different source nodes can be transmitted within individual subframes. Therefore, channel contention can be entirely avoided.

8.1 Novel Contributions

The main research contributions are summarised as follows:

8.1.1 Frame Size Adaptation for Single-hop Networks

The ALOHA-Q protocol provides a novel solution to reducing packet collisions and improve energy-efficiency of WSNs. Nodes can reinforce their slot selection policies by conducting trial-and-error interactions based on a Q-learning algorithm. When the learning process ends, the system reaches a steady state where perfect scheduling can be achieved. In single-hop WSNs, the performance of ALOHA-Q is determined by the number of slots per frame, which is difficult to determined in an unpredictable environment. This thesis presents a distributed frame size adaptation (DFA) approach to solving this issue. According to the principle of DFA, sensor nodes can independently adapt their frame size to the number of nodes by continuously evaluating the contention level of the public channel and periodically adjust frame size based on their needs. Simulation results have demonstrated that the DFA algorithm can effectively support the ALOHA-Q protocol in achieving maximum performance within a predictable period. Once the best frame size is obtained, packet collisions can be fully eliminated and system throughput can be maximised.

8.1.2 Adaptation of ALOHA-Q to Linear Chain Networks

The early sections of Chapter VI introduced the implementation of ALOHA-Q in linear chain networks and examined corresponding throughput performance through software simulations. Results suggest that the throughput performance of ALOHA-Q is not stable under multi-hop conditions due to the effect of hidden nodes, leading to traffic congestion in the transmission route. A slot swapping technique has been proposed to help sensor nodes to smoothly exchange their scheduling policy for the purpose of avoiding interference from hidden nodes. Besides, a frame size adaptation scheme for linear chain network is proposed to help sensor nodes automatically figure out an optimal value of frame size by periodically evaluating their transmission histories. Simulation results have demonstrated that the combination of slot swapping technique and frame size adaptation scheme can effectively support the ALOHA-Q protocol in achieving perfect scheduling in linear chain networks without accurate prediction of node interference range and network topology.

8.1.3 Subframe Adaptation

In multi-hop WSNs, individual traffic flows may overlap with each other and relay nodes may need to forward packets from multiple source nodes. Compared with single-hop or linear chain networks, the channel contention level is hugely increased. This makes the implementation of ALOHA-Q in cross chain networks complex and inefficient. In the later sections of Chapter VI, a novel subframe adaptation scheme is proposed. The proposed scheme introduces a superframe structure which allow nodes along individual transmission routes to have their dedicated subframes. The slot selection algorithm of ALOHA-Q protocol can be used by nodes to choose their preferred time slots within corresponding subframes. Consequently, channel contention between different transmission routes can be fully resolved.

8.2 Publications

Yan, Yan; Mitchell, Paul; Clarke, Tim; Grace, David, ”Adaptation of the ALOHA-Q protocol to Multi-Hop Wireless Sensor Networks”, European Wireless 2014; 20th European Wireless Conference; Proceedings of, May 2014.

Yan, Yan; Mitchell, Paul; Clarke, Tim; Grace, David, ”Distributed Frame Size Selection for a Q learning based Slotted ALOHA Protocol”, Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on, Aug, 2013.

Glossary

ACK acknowledgement. 51

ALOHA-Q Q-learning based Slotted ALOHA. 40

ASIC Application-Specific Integrated Circuit. 23

C1WSNs Category 1 Wireless Sensor Networks. 18

C2WSNs Category 2 Wireless Sensor Networks. 19

C4ISRT Command, Control, Communications, Computers, Intelligence, Surveillance, Reconnaissance, and Targeting. 30

CAP Contention Access Period. 60

CDMA Code Division Multiple Access. 46

CFP Contention-free Period. 60

CPU Central Processing Unit. 25

CSMA/CA Carrier Sense Multiple Access/ Collision Avoidance. 50

CTS Clear to Send. 55

DARPA Defence Advanced Research Projects Agency. 2

DES Discrete Event Simulation. 67

DFA Distributed Frame size Adaptation. 88

DLL Data Link Layer. 4

DRAND Distributed Slot Allocation Algorithm. 58

DSN Distributed Sensor Networks. 2

DSP Digital Signal Processor. 23

FDMA Frequency Division Multiple Access. 45

FFD Full Function Device. 59

FPGA Field-Programmable Gate Array. 23

GPS Global Positioning System. 22

GTSs Guaranteed Time Slots. 61

ISO International Organization for Standardization. 4

LBT Listen Before Transmission. 50

LEACH Low Energy Adaptive Clustering Hierarchy. 41

MAC Medium Access Control. 3

MDP Markov Decision Process. 8, 89

MEMs Micro Electro Mechanical Systems. 20

MFR Most Forwarding progress within Radius. 106

NAV Network Allocation Vector. 55

OSI Open Systems Interconnection. 4

QoS Quality of Service. 4

RAM Random Access Memory. 25

RFD Reduced Function Device. 59

RFID Radio Frequency Identification. 20

RTS Request to Send. 55

RX Reception. 43

S-MAC Sensor MAC protocol. 54

SINR Signal-to-interference-plus-noise Ratio. 103

SoC System-on-Chip. 2, 20

SOSUS Sound Surveillance System. 1

T-MAC Time-out MAC protocol. 56

TCP/IP Transmission Control Protocol/Internet Protocol. 17

TDMA Time Division Multiple Access. 45

TX Transmission. 43

WSNs Wireless Sensor Networks. ii, 1

Z-MAC Zebra MAC protocol. 58

Bibliography

[1] C.Raghavendra, K.M.Sivalingam and Taieb Znati, *"Wirelss Sensor Networks"*, Springer, Sept 2006.

[2] F. Akyildiz, Weilian Su, Yogesh Sankarasubramaniam and Erdal Cayirci, *"A Survey on Sensor Networks"*, Wiley, 2nd Edition, Aug 2002.

[3] Kazem Sohraby, Daniel Minoli, and Taieb Znati, *"Wireless Sensor Networks: Technology, Protocols, and Applications"*, IEEE Communications Magazine, Volume:40, Issue:8, May 2007.

[4] Nirupama Bulusu, and Sanjay Jha, *"Wireless Sensor Networks A Systems Perspective"*, Artech Hous 1st Edition, July 2005.

[5] Mohammad Alex, Simon Carlsen and Stig Petersen, *"Applications of Wireless Sensor Networks in the Oil, Gas and Resources Industries"* 2010 International Conference on Advanced Information Networking and Applications, pp 941-948, 2010.

[6] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. Wang and A. Chandrakasan, *"Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks"*, ACM MobiCom'01, Rome, Italy, pp. 272-286, 2001.

[7] J. M. Kahn, R. H. Katz and K. S. J. Pister, *"Next century challenges: Mobile networking for smart dust"*, ACM MobiCom'99, Washington, USA, pp. 271-278, 1999.

[8] A.S. Porret, T. Melly, C. C. Enz and E. A. Vittoz, *"A low-power low-voltage transceiver architecture suitable for wireless distributed sensor network"*, ISCAS 2000, Geneva, Switzerland, pp. 56-59, 2000.

[9] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler and J. Anderson, *"Wireless sensor networks for habitat monitoring"*, WSNA'02, Atlanta, USA, pp. 88-97, 2002.

[10] C.Y. Chong and S. P. Kumar, *"Sensor networks: Evolution, opportunities and challenges"*, Proceedings of the IEEE, vol.91, pp. 1247-1256, 2003.

[11] J. Polastre, R. Szewczyk, A. Mainwaring, D. Culler, and J. Anderson *"ANALYSIS OF WIRELESS SENSOR NETWORKS FOR HABITAT MONITORING"*, Wireless sensor networks, Kluwer Academic Publishers Norwell, pp. 399 - 423, 2004.

[12] J.Day and H.Zimmermann, *"The OSI Reference Model"*, Proceedings of the IEEE, volume:71, Number: 12, Dec 1983.

[13] Waltenegus Dargie, and Christian Poellabauer, *"Fundamentals of Wireless Sensor Networks: Theory and Practice"*, Wiley, 2nd Edition, 2010.

[14] Andrew Myers, and Stefano Basagni, *"Wireless Media Access Control"*, PDF report (http://www.ece.neu.edu/fac-ece/basagni/Classes/F016V81.503/WMAC.pdf), University of Texas at Dallas.

[15] Ilker Demirkol, Cem Ersoy, and Fatih Alagz, *"MAC protocols for wireless sensor networks: a survey"*, IEEE Communications Magazine, Volume:44, Issue:4, Apr 2006.

[16] Wei Ye, John Heidemann and Deborah Estrin, *"An Energy-Efficient MAC Protocol for Wireless Sensor Networks"*, IEEE INFOCOM, Volume:3, Page: 1567-1576, 2002.

[17] Ye Zhang, *"Modeling the S-MAC Protocol in Single-Hop Wireless Sensor Networks"*, International Conference on Communications, Page:321-327 , May 2008.

[18] P. Pannuto, and Y Lee *“M3: a mm-scale wireless energy harvesting sensor platform”*, ENSSys ’13 Proceedings of the 1st International Workshop on Energy Neutral Sensing Systems Article No. 17, 2013

[19] T. Chiras, M. Paterakis, and P. Koutsakis, *“Improved medium access control for wireless sensor networks - a study on the S-MAC protocol”*, IEEE workshop on Local and Metropolitan Area Networks, Page:5, Sept 2005.

[20] Ghenghea Mihai, Darie Angela Alina, and Bogdan Ion, *“Performance analysis on T-MAC protocol over a body area network”*, International Symposium on Electrical and Electronics Engineering, Page:224-227, 2010.

[21] Jelena Uljarevi, and Nataa Nekovi, *“A comprenhensive analysis of MAC protocols of wireless sensor networks”*, Telecommunications Forum, Page:481-484, 2011.

[22] Ji-Jun Zhao, and Xiang Sun, *“MAC protocol based on T-MAC multi-hop reservation for short-latency wireless sensor network”*, IEEE International Conference on Communication Technology, Page:114-117, 2008.

[23] Mahdi Lotnezhad and Peter Marbach, *“Throughput-delay trade-off of CSMA policies in wireless networks”*, Information Theory Workshop, Page:1-2, 2010.

[24] Y, L. Sun, H. Song, and X. Cao, *“Ubiquitous WSN for Healthcare: Recent Advances and Future Prospects”*, Internet of Things Journal IEEE, Volume 1, Issue 4, 2014

[25] A. Bachir, and M Dohle, *“MAC Essentials for Wireless Sensor Networks*, IEEE COMMUNICATIONS SURVEYS and TUTORIALS, Volume 12, No 2, 2010.

[26] Chen Xiaoming and Hong Geok-Soon, *“A simulation study of the predictive p-persistent CSMA protocol”*, Simulation Symposium, Page: 345-351, Apr 2002.

[27] Z Lu, and H. Yang *“Unlocking the Power of OPNET Modeler”*, Cambridge University Press, Page: 3-9, Jan 2012.

[28] http://en.wikipedia.org/wiki/OPNET.

[29] S. Adarshpal, Sethi, Y. Vasil, and Hnatyshin *"The Practical OPNET User Guide for Computer Network Simulation"*, CRC Press, Page: 12-15, Aug 2012.

[30] I. Katzela *"Modeling and Simulating Communication Networks: A Hands-on Approach Using OPNET"*, Prentice Hall, Page: 35-38, 1999.

[31] E Aboelela, *'Network Simulation Experiments Manual"*, Academic Press, Page: 185-188, 2003.

[32] C. Kong, *"OPNET Simulation of Signaling System No. 7 (SS7) Network Interfaces"*, Storming Media, Page: 256-258, Mar 2000.

[33] A. Adil, and N. Malik *"Modeling and Simulation of Computer Network Using Opnet"*, Lap Lambert Academic Publishing GmbH KG, Page: 132-135, 2012.

[34] F. Abdullah T. Al-Saedi, M. Mustafa, and A. Al-Adhamee, *"Performance Study of Mobile IPv6 Using OPNET Simulator"*, Lap Lambert Academic Publishing GmbH KG, Page: 102-104, 2013.

[35] Nikolaos M. Freris, *"Performance bounds for CSMA-based medium access control"*. IEEE Conference on Decision and Control and European Control, Page:5945-5950, Dec 2011.

[36] Mahdi Lotnezhad and Peter Marbach, *"Delay performance of CSMA policies in multihop wireless networks: A new perspective"*, Information Theory and Applications Workshop (ITA), 2010.

[37] Fei Tong, Wan Tang, Rong Xie, *"P-MAC: A Cross-Layer Duty Cycle MAC Protocol Towards Pipelining for Wireless Sensor Networks"*. IEEE International Conference on Communications, Page:1-8, 2011.

[38] Shangwei Duan and Xiaobu Yuan, *"Exploring hierarchy architecture for wireless sensor networks management"*, IFIP International Conference on Wireless and Optical Communications Networks, Page:6, 2006.

[39] Xueqing Wang , *"Comparison of several sensor deployments in wireless sensor networks"*, International Conference on E-Health Networking, Digital Ecosystems and Technologies (EDT), Volume:1, Page: 236-239, 2010.

[40] Xiaocheng Du, *"Researching of adaptive Traffic-Aware in wireless sensor networks MAC protocol"*. International Conference on Information Science and Engineering, Page: 4629-4633, Dec 2010.

[41] Haifang Feng, Lixiang Ma, Supeng Leng, *"A low overhead wireless sensor networks MAC protocol"*. International Conference on Computer Engineering and Technology, Volume:4, Page: 128-131, Apr 2010.

[42] Ming Zhang, *"An Novel Energy-Efficient MAC Protocol Based on Collision Avoidance for Wireless Sensor Networks"*, International Conference on Wireless Communications, Networking and Mobile Computing, Page:1-4, Sept 2009.

[43] Yongrui Chen, Weidong Yi and Yang Yang, *"A Global Time Synchronised MAC for Wireless Sensor Network*, WiCOM 08, pp. 1-6, Dalian, 2008.

[44] N. AL-KARAKI and A. E. KAMAL, *"Routing techniques in sensor networks: A survey"*, IEEE Wireless Communications December, pp. 6-28, 2004.

[45] Zhiwei Zhao, Xinming Zhang, Peng Sun, Pengxi Liu, *"A Transmission Power Control MAC Protocol for Wireless Sensor Networks"*, International Conference on Networking, Page:5, 2007.

[46] Yunhuai LIU and Lionel M. NI, *"A new MAC protocol design for long-term applications in wireless sensor networks"*, International Conference on Parallel and Distributed Systems, Volume:2, Page:1-8, Dec 2007.

[47] Yi Zhou, *"Enhanced Slotted Aloha Protocols for Underwater Sensor Networks with Large Propagation Delay"*, Vehicular Technology Conference, Page:1-5, May 2011.

[48] C.Steliosand, and A.Thomopoulos *"A simple and versatile decentralized control for slotted ALOHA, reservation ALOHA, and local area networks"*, IEEE Transactions on Communications, Volume:36, Page:662-647, 1988.

[49] Osama Kubbar, *"An Aloha-based channel access scheme investigation for broadband wireless networks"*, IEEE International Symposium on Computers and Communications, Page:203-208, 1999.

[50] Nitthita Chirdchoo, Wee-Seng Soh and Kee Chaing Chua, *"Aloha-Based MAC Protocols with Collision Avoidance for Underwater Acoustic Networks"*, IEEE International Conference on Computer Communications, Page: 2271-2275, May 2007.

[51] Akassh A Mishra , *"Enhanced slotted aloha"*, International Conference on Electronics Computer Technology, Volume:5, Page 278-282, 2011.

[52] F.L.Lo, *"Delay-Throughput Comparison of Single and Multi-Channel Slotted ALOHA Networks"*, International Conference on Communication Technology, Volume:2, Page:932-937, 1996.

[53] Parak Kulkarn, *"Introduction to Reinforcement and Systemic Machine Learning"*, Wiley-Blackwell, Page:1-21, Sept 2012.

[54] Chunxuan Ye, Yan Li and Alex Reznik, *"Performance Analysis of Exponential Increase Exponential Decrease Back-Off Algorithm"*, Global Telecommunications Conference, Page:1-6, Dec 2010.

[55] Mohamed Arabi and Bogdan Ghita, *"Transmission rate control through MAC layer acknowledgements feedback"*, Wireless and Mobile Networking Conference, Page:1-7, Oct 2011.

[56] Rani, Shalli, Ahmed, Syed Hassan, *"Multi-hop Routing in Wireless Sensor Networks"*, SpringerBriefs in Electrical and Computer Engineering, 2006

[57] Wuyungerile Li, *"Trade-offs among Delay, Energy and Accuracy of Partial Data Aggregation in Wireless Sensor Networks"*, IEEE International Conference on Advanced Information Networking and Applications, Page:917-924, Apr 2010.

[58] I. Rhee, A. Warrier, M. Aia, J. Min, and M. L. Sichitiu, *"Z-MAC: A hybrid mac for wireless sensor networks"*, IEEE/ACM Transactions on Networking, vol.16, pp. 511-524, 2008.

[59] Pei Huang, Li Xiao, S. Soltani, and M.W Mutka, *"The Evolution of MAC Protocols in Wireless Sensor Networks: A Survey"*, Communications Surveys and Tutorials, IEEE, page:101-120, Feb 2013

[60] Yang Xiao, Miao Peng, J. Gibson, and G.G. Xie, *"Tight Performance Bounds of Multihop Fair Access for MAC Protocols in Wireless Sensor Networks and Underwater Sensor Networks"*, Mobile Computing, IEEE Transactions on, Page:1536-1233, Oct. 2012.

[61] Pochiang Lin and Tsungnan Lin, *"Machine-Learning-Based Adaptive Approach for Frame-Size Optimization in Wireless LAN Environments*, Vehicular Technology, IEEE Transactions on, Page:5060 - 5073, June 2009.

[62] Thomas Beluch, Daniela Dragomirescu, Florian Perget, Robert Plana, *"Cross-Layered Synchronization Protocol for Wireless Sensor Networks"*. International Conference on Networks, Page:167-172, 2010.

[63] Qinghua Wang, *"Traffic Analysis and Modeling in Wireless Sensor Networks and Their Applications on Network Optimization and Anomaly Detection"*, Network Protocols and Algorithms, Vol.2, 2010.

[64] Pu Wang and Ian F. Akyildiz *"Spatial Correlation and Mobility Aware Traffic Modeling for Wireless Sensor Networks"*, Global Telecommunications Conference, 2009

[65] Frits C.Schout, *"Dynamic Frame Length ALOHA"*, IEEE Transactions On Communications, vol.31, no.4, Apr 1983.

[66] Yi Chu, Mitchell P.D and Grace D, *“ALOHA and Q-Learning based medium access control for Wireless Sensor Networks”*, Wireless Communication System (ISWCS), 2012 International Symposium on, pp. 511-515, 28-31, Aug 2012.

[67] Kosunalp Selahattin, Mitchell P.D, Clarke T and Grace D *“Practical Implementation Issues of Reinforcement Learning Based ALOHA for Wireless Sensor Networks”*, Wireless Communication Systems (ISWCS 2013),the Tenth International Symposium on, Aug. 2013

[68] Yan, Yan, Mitchell P.D, Clarke T and Grace D, *”Distributed Frame Size Selection for a Q learning based Slotted ALOHA Protocol”,* Wireless Communication Systems (ISWCS 2013), Proceedings of the Tenth International Symposium on, Aug. 2013.

[69] Yan, Yan; Mitchell, Paul; Clarke, Tim; Grace, David, *”Adaptation of ALOHA-Q protocol to Multi-hop Wireless Sensor Networks”*, European Wireless 2014; 20th European Wireless Conference; Proceedings of, 2014

[70] K Jain, J Padhye, V. N. Padmanabhan, L Qiu, *”Impact of Interference on Multi-Hop Wireless Network Performance”,* Journal of Wireless Networks, vol. 11, pp. 471-487, 2005.

[71] R Mine, G Hasegawa, Y Taniguchi, H Nakano, *”Effect of Radio Interference Models on the Performance of Time Slot Assignment Algorithms in IEEE 802.16j Relay Networks”,* Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on, pp. 107 - 113, Mar 2013.

[72] Dongxu. S, Li V.O.K, *”Performance analysis for a stabilized multi-channel slotted ALOHA algorithm”,* Personal, Indoor and Mobile Radio Communications 2003, PIMRC 2003, vol.1, pp 249-253, Sept. 2003.

[73] Galluccio. L, Morabito. G, Palazzo. S, *”TC-Aloha: A Novel Access Scheme for Wireless Networks with Transmit-Only Nodes”,* Wireless Communications, IEEE Transactions on, vol.12, issue.8, Aug. 2013.

[74] Sylvia Tai, Robert Benkoczi, Hossam Hassanein, and Selim Akl, *"An Energy Consumption Study of Wireless Sensor Networks with Delay-Constrained Traffic"*, Global Telecommunications Conference, Page:1-6, 2006,

[75] Yunbo Wang, *"Cross-Layer Analysis of the End-to-End Delay Distribution in Wireless Sensor Networks"*, IEEE/ACM Transactions on Networking, Volume:20, Issue:1, Page:105-118, 2012.

[76] Hadi S. Aghdasi and Maghsoud Abbaspour, *"ET-MAC: An Energy-Efficient and High Throughput MAC Protocol for Wireless Sensor Networks"*, Communication Networks and Services Research Conference, Page:526-532, May 2008.

[77] C. Jandaeng and W. Suntiamontut, *"Throughput Improvement of Collision Avoidance in Wireless Sensor Networks"*. International Conference on Wireless Communications Networking and Mobile Computing, Page:1-5, Sept 2010.

[78] Flavio Fabbri, *"Area Throughput and Energy Consumption for Clustered Wireless Sensor Networks"*, Wireless Communications and Networking Conference, Page:1-6, Apr 2009.

[79] Roberto Verdone, *"Maximizing area throughput in clustered wireless sensor networks"*, IEEE Journal on Selected Areas in Communications, Volume:28, Issue:7, Page: 1200-1210, Sept 2010.

[80] Mostafa Fazeli, *"Assessment of Throughput Performance Under OPNET Modeler Simulation Tools in Mobile Ad Hoc Networks (MANETs)"*. International Conference on Computational Intelligence, Communication Systems and Networks, Page 328-331, July 2011.

[81] I. S. Hammoodi, *"A Comprehensive Performance Study of OPNET Modeler for ZigBee Wireless Sensor Networks"*, International Conference on Next Generation Mobile Applications, Services and Technologies, , Page: 357-362, Sept 2009.

[82] C. Marghescu, M. Pantazica, A.Brodeala and P. Svasta, *"Simulation of a wireless sensor network using OPNET"*, IEEE International Symposium for Design and Technology in Electronic Packaging, Page:249-252, Oct 2011.

[83] Norman Abramso, *"THE ALOHA SYSTEM: another alternative for computer communications"*, Proceeding AFIPS, Page:218-258, Nov 1970.

[84] T. Richard, B. Ma, Vishal Misra and Dan Rubenstein, *"An Analysis of Generalized Slotted-Aloha Protocols"*, IEEE/ACM Transactions on networking, Volume:17, Issue:3, Page:936-949, June 2009.

[85] R. Rom and M. Sidi, *"Multiple Access Protocols: Performance and Analysis"*, New York: Springer-Verlag, 1990.

[86] Yi Chu, *"Application of Reinforcement Learning to Medium Access Control for Wireless Sensor Networks"*, Ph.D. dissertation, Dept. Elect. Eng., York Univ., York, United Kingdom, 2013

[87] L. P. Kaelbling, M. L. Littman, A. W. Moore, *"Reinforcement Learning: A Survey"*. Journal of Artificial Intelligence Research, Volume 4, Page:237-285, 1996.

[88] Tom M. Mitchell, *"Machine Learning"*. 1st Edition, McGraw-Hill, 1997.

[89] George F. Luger, *"Artificial intelligence Structures and Strategies for complex problem Solving"*, 6th Edition, New Mexican University, 2006.

[90] Yi Shi, Y. Thomas, and Hou Jia Liu, *"How to Correctly Use the Protocol Interference Model for Multi-hop Wireless Networks"*, Proceedings of the tenth ACM international symposium on Mobile ad-hoc networking and computing, 2009.

[91] N. Al-Karaki, and A. E. Kamal, *"Routing techniques in sensor networks: A survey"*, IEEE Wireless Communications December, 2004.

[92] Mehmet C. Vuran and Ian F. Akyildiz, *"Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks"*, IEEE Transactions on Networking, vol.14, no.2, 2006.

[93] S. P. Meyn and R.L. Tweedie, *“Markov Chains and Stochastic Stability”*, Second edition, Cambridge University Press, 2008.

[94] P. Prasad, *“Recent trend in wireless sensor network and its applications: a survey”*, Sensor Review, Vol.35, Issue 2, pp.229 - 236, 2015.

[95] C. Alcaraz, P. Najera, J. Lopez, and R. Roman, *“Wireless Sensor Networks and the Internet of Things: Do We Need a Complete Integration?”*, 1st International Workshop on the Security of the Internet of Things (SecIoT10), Tokyo, 2010.

[96] https://en.wikipedia.org/wiki/Communications_protocol.

[97] J.M. Gilbert, and F Balouchi, *“Comparison of energy harvesting systems for wireless sensor networks”*, International Journal of Automation and Computing, Volume 5, Issue 4, pp 334-347, 2008.

[98] O.S. Kosovych, *“Fixed Assignment Access Technique”*, Communications, IEEE Transactions on, Volume 26, Issue 9, 1978.

[99] https://en.wikipedia.org/wiki/Frequency-division_multiple_access

[100] https://en.wikipedia.org/wiki/Time_division_multiple_access

[101] https://en.wikipedia.org/wiki/Code_division_multiple_access

[102] A. Bottcher, *“Performance of TDMA and demand assignment multiple access protocols in mobile radio channels”*, Telecommunications Symposium, SBT/IEEE International, 1990.

[103] https://en.wikipedia.org/wiki/Polling_(computer_science).

[104] https://en.wikipedia.org/wiki/Token_passing.

[105] `https://en.wikipedia.org/wiki/Resource_Reservation_Protocol`.

[106] J Yang, L. Yang, G. Xiong, F. Zhu, *"Performance evaluation of IEEE 802.15.4 MAC layer"*, Vehicular Electronics and Safety (ICVES), 2013 IEEE International Conference on, 2013.

[107] F. Shu, T. Sakurai, H.L. Vu, M. Zukerman, *"Optimizing the IEEE 802.15.4 MAC"*, IEEE Region 10 Conference, 2006.

[108] J. Gao, H. Jia, G. Min, *"Performance Modelling of IEEE 802.15.4 MAC in LR-WPAN with Bursty ON-OFF Traffic"*, TComputer and Information Technology, Ninth IEEE International Conference on, 2009.

[109] L. Riliskis, and E. Osipov, *"Symphony: Simulation, emulation, and virtualization framework for accurate WSN experimentation"*, Software Engineering for Sensor Network Applications (SESENA), 4th International Workshop on, 2013.

[110] A. Intana, M.R. Poppleton, G.V. Merrett, *"Adding value to WSN simulation through formal modelling and analysis"*, Software Engineering for Sensor Network Applications (SESENA), 4th International Workshop on, 2013.

[111] Q.I Ali, A. Abdulmaowjod, H.M. Mohammed, *"Simulation and performance study of wireless sensor network using MATLAB"*, Energy, Power and Control (EPC-IQ), 1st International Conference on, 2010.

[112] T. Antoine-Santoni, J.F. Santucci, E.de. Gentili, B. Costa, *"DEV-SWSN : A discrete event approach for Wireless Sensor Network simulation"*, Computer Systems and Applications, IEEE/ACS International Conference on, 2008.

[113] D.V. Manatakis, *"Simulation-driven emulation of collaborative algorithms to assess their requirements for a large-scale WSN implementation"*, Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on, 2014.

[114] S.Maragatha Sundari, and S.Srinivasan, *“M/M/C Queueing Model For Waiting Time of Customers In Bank Sectors”*, International Journal of Mathematical Sciences and Applications, 2011.

[115] https://en.wikipedia.org/wiki/Pigeonhole_principle.

[116] https://en.wikipedia.org/wiki/Liebig27s_law_of_the_minimum.

[117] http://www.libelium.com/smart_cities_wsn_overwiew/.

[118] http://www.graystone-ind.com/wireless.html.

[119] http://lrd.buffalohair-jage.com/2013/04/21/tracking-us-with-powder-smart-dust/.

[120] http://www.infotech.oulu.fi/Annual/2007/opme.html.

[121] M.G.C Torres, *“Energy Consumption In Wireless Sensor Networks Using GSP”*, Master Thesis, University of Pittsburgh, 2006.

[122] M. N. Halgamuge, M. Zukerman, and K. Ramamohanarao, *“An Estimation Of Sensor Energy ConsumptionAN”*, Progress In Electromagnetics Research B, vol. 12, pp. 259-295, 2009.

[123] http://www.eol.ucar.edu/isf/facilities/isa/internal/CrossBow/DataSheets/mica2.pdf.

[124] http://www.memsic.com/userfiles/files/Datasheets/WSN/IRIS_Datasheet.pdf.

[125] Q. Wang, *“Traffic Analysis and Modelling in Wireless Sensor Networks and Their Applications on Network Optimisation and Anomaly Detection”*, technical report, Research at MidSweden University, 2010.

[126] A. Bagula, *"Modelling and Implementation of QoS in Wireless Sensor Networks: A Multiconstrained Traffic Engineering Model"*, EURASIP Journal on Wireless Communications and Networking, 2010.

[127] `https://en.wikipedia.org/wiki/Constant_bitrate`.

[128] `https://en.wikipedia.org/wiki/Poisson_distribution`.

[129] `https://en.wikipedia.org/wiki/Pareto_distribution`.