INTELLIGENT MEDIUM ACCESS CONTROL PROTOCOL FOR UNDERWATER PIPELINE MONITORING NETWORKS

Ibrahim Bala Alhassan

PhD

Electronic Engineering
University of York

December 6, 2021

Abstract

This thesis studies the applications of Reinforcement Learning (RL) in designing an intelligent MAC protocols for linear chain Underwater Acoustic Sensor Networks (UASNs) suitable for marine pipeline monitoring. The key objective is to explore and devise simple strategies that re-imagine RL based algorithms with reduced inefficiencies due to overheads to improve channel utilisation and adaptability. Inspired by the successful implementation of RL on ALOHA in the recently proposed terrestrial ALOHA-Q, we explored the feasibility of applying similar approach in UASNs. The evaluation of ALOHA-Q in UASN, has shown the potential benefits to employing RL for adaptable underwater MAC design, however, new strategies on slot structure and method of feedback are needed for good utilisation.

Based on the relationship between packet duration and propagation delay, this thesis proposed two efficient slot structures. The viability of these slot structures are pictorially analysed and empirically evaluated for incorporation in MAC protocol implementation. The thesis presents novel RL based algorithms without any explicit feedback signal. Rather, it exploits packet flow in a two stage mechanism to simultaneously drive a slot selection Q-learning algorithm and a stochastic averaging function that heuristically measured the network wide optimal flow harmony, thereby, effectively creating a simple, powerfully adaptive intelligent scheduling with huge performance improvement.

Dedication

To *Fatima*. God Bless.

To *Hajia*. A great soul. God have mercy on you.

Declaration

All work presented in this thesis is original to the best knowledge of the author. References to other researchers have been acknowledged were appropriate. This work has not previously been presented for an award anywhere.

The material presented in Chapter 4 is presented as ”Monitoring free span sections of subsea pipeline with ALOHA-Q” at URSI Festival of Radio Science 2019.

The material presented in Chapter 5 is published as ”Packet Flow Based Reinforcement Learning MAC Protocol for Underwater Acoustic Sensor Networks” in Sensors 2021.

Acknowledgement

Gratitude to the Almighty God.

Special appreciation to my family for their immeasurable love and support during this journey.

My profound gratitude to Prof. Paul D. Mitchell for his boundless patience, consistent support and positive encouragement during this sapping yet fulfilling journey. And to my short lived second supervisor Tim Clark, your input and positive energy has surely been missed.

I would also like to acknowledge my sponsors PTDF and BUK for the financial backbone that make it possible for me undertake this research. To the staff and management of the University of York, your support during the raging pandemic has truly been remarkable.

Finally, to my friends, colleagues and experts from the communications and signal processing group, your presence has made positive impact.

Contents

List of Figures

List of Tables

Abbreviations

ACK Acknowledgement. 2

ADC Analog to Digital Converter. 18

AUV Autonomous Underwater Vehicles. 15

BER Bit Error Rate. 22

CPU Central Processing Unit. 19

CSMA Carrier sense Multiple Access Control. 25

CTS Clear To Send. 27

DAC Digital to Analog Converter. 18

Dec-POMDP Decentralised Partially Observable Environments. 35

DES Discrete Event Simulation. 37

EM Electromagnetic waves. 2, 10

FAMA Floor Acquisition Multiple Access. 28

FDMA Frequency Division Multiple Access. 25

HMM Hidden Markov Model. 35

LLC Logical Link Control. 21

MAC Medium Access Control. 2, 21

MACA Multiple Access Collision Avoidance. 27

MARL Multi Agent Reinforcement Learning. 4, 34

MDP Markov Decision Process. 31

POMDP Partially Observable Environments. 4, 35

RFID Radio Frequency Identification. 9

RIPT A Receiver-Initiated Reservation-Based Protocol for UANs. 29

RL Reinforcement Learning. 2, 7, 30

RTR Request To Receive. 29

RTS Request To Send. 27

SA Single Agent. 4

TARS A Traffic-Adaptive Receiver-Synchronized MAC Protocol for Underwater Sensor Networks. 29

TD Temporal Difference. 33

TDA Transmission Data Attributes. 42

TDMA Time Division Multiple Access. 24

UASN Underwater Acoustic Sensor Networks. 1

USN Underwater Sensor Networks. 10

WSNs Wireless Sensor Networks. 1, 8

List of Symbols

λ_t	Optimisation scale
α	Learning rate
γ	Discount factor
$\in$	Element
λ	Packet inter-arrival time
π	Policy function
γ_t	Tracking rate
$\mathbb{E}$	Expectation
τ_d	Data packet duration
τ_A	ACK packet duration
τ_g	Guard duration
τ_{pg}	Propagation delay
iFX	Interference range
rX	Reception range
A	Set of actions

a_i	i^{th} state action
B	Data rate
b_i	i^{th} belief state
Bs	Set of Belief states
d	Distance between nodes
e	Euler's number
F_D	Frame duration
G	Offered Load
$K\tau$	Ratio of τ_d -to- τ_{pg}
N	Number of nodes
N_p	Number of packets
N_s	Number of slots
N_{opt}	Optimum number of slots per frame
P_d	Packet length in bits
Q	Q function
r/Ψ	reward or punishment value
S	Set of states
s_i	i^{th} State
S_L	Slot duration/length
S_L^a	Slot size with ACK

S_L^n	Slot size without ACK
T	Time
U	Channel utilisation
W_n	Tracking window
Y	Set of Observations
y_i	i^{th} Observation
fl_τ	Packet flow average

Chapter 1

Introduction

1.1 Motivation

As an analogue of terrestrial Wireless Sensor Networks (WSNs), Underwater Acoustic Sensor Networks (UASN) are envisaged to enable a multitude of civilian and military applications, such as: environmental and infrastructural monitoring, to assisted navigation, surveillance, and exploration [1, 2, 3]. Monitoring of offshore infrastructure particularly, in the oil and gas sector is of immense economic valuable [4, 5], and in the event of sabotage or accidents, they pose huge risks of ecological damage. Moreover, the dwindling oil prices are making the industry resort to measures of extending the useful lifespan of offshore installations such as marine pipelines, arguably, that will increase the associated risks. An underwater acoustic multi-hop network of sensors deployed along the asset or distributed on site has the potential to monitor assets integrity and also proximity coverage for the whole site. To enable and advance these applications, sensor nodes are being developed to be small/compact for easy transport, given that the environment is characteristically challenging to access. There is interest in new sensor nodes being energy efficient for longer deployments; as currently, there is no viable energy harvesting technology. Nodes should also be inexpensive to lower the overall cost, since UASNs are envisaged to be deployed to cover substantial marine areas and require a large number of devices.

The use of acoustic waves in UASNs instead of Electromagnetic waves (EM) imposes some unique channel-centric constraints, such as limited capacity, long and variable propagation delay. Medium Access Control (MAC) is one of the key requirements for the development of UASNs. Moreover, in the highly dynamic underwater environment, MAC protocols need to be adaptive to changing conditions as well. As such, there is growing demand for efficient MAC solutions, especially adaptive MAC protocols for practical networks. Researchers have proposed several adaptive techniques that combined two or more existing/different schemes to create hybrid MAC protocols with some notable performance gains in utilisation, energy efficiency and delay performance. However, given the typical need for pre-planning of each potential situation, this approach is inherently complex.

Reinforcement Learning (RL) is an intelligent promising solution used in MAC protocols to provide adaptability and robustness in wireless sensor networks. The RL Algorithm employs an online learning that continually assess the network condition through feedback and appropriately responds with a view towards maintaining (as much as possible) a collision-free schedule. This intelligent solution inherently addressed the need for smart channel sharing policy, self-organisation due to dynamic changes in environment or network conditions and simplicity. however, owing to its reliance on Acknowledgement (ACK) to work the protocol is not suitable for UASNs. Notwithstanding the demonstrably degraded performance, there are adaptability benefits to the approach. The simplicity of random access strategy and distributive robustness of Q-learning, inspired this work in devising a set of simple strategies on efficient slot structuring, and exploitation of packets streams in a dual control stages to propose a novel MAC protocol capable of huge improvements in performance. Simulations results have empirically demonstrated that it is possible to for MAC protocols to achieve huge performance improvements whilst maintaining adaptability without waiting for an explicit feedback for nodes action.

1.2 Hypothesis

The cardinal function of MAC protocol is to efficiently share and regulate channel access between competing nodes by mitigating or eliminating collisions. All MAC schemes and design techniques aims to ensure a ”healthy” flow of data between source(s) and intended destination(s), with a quantifiable improvement in utilisation, delay and energy performance. Given no single MAC scheme is suitable for all scenarios nor efficient for all performance measures, for a specific scenario or target application, researchers have proposed several MAC strategies with the goal of optimising performance in terms of one or a combination of utilisation, delay or energy consumption. There are tradeoffs between these performance optimisations benefits and an increase in control signals/measures or overall protocol complexity, however.

This thesis is based upon the following hypothesis:
”It is possible to effectively achieve the optimal (achievable utilisation) network performance by devising new time slots based on the relationship between packet duration and hop propagation delay, coupled with intelligent MAC scheduling using packets flow in lieu of explicit reward signal to drive a reinforcement learning algorithm.”

The hypothesis has assisted us to design and propose MAC strategies that are inspired by reinforcement learning. Our MAC approach whilst distributively intelligent, autonomously robust is also simple and computationally cheap making it suitable for less powerful nodes.

1.3 Thesis Structure

This thesis is comprises of seven chapters and are structured as follows;

Chapter 2 presents the fundamentals and literature review of relevant fields. It opened with the general background on wireless sensor networks with emphasis on UASNs. The background highlights the similarities and differences in the enabling technologies and applications between WSN and UASNs. The chapter also introduces fundamentals on Medium Access Control protocols of sensor networks, covering aspects such as; classification, examples of several representative MAC protocols with a discussion on their merits demerits. The sections conclude with MAC design challenges. The chapter conclude with an introduction on theoretical Reinforcement Learning, by presenting the fundamental principles and techniques Q-learning, Single Agent (SA), Multi Agent Reinforcement Learning (MARL) and Partially Observable Environments (POMDP).

Chapter 3 presents the methodology and the simulation tools used in the design, implementation and validation of the network and elements. It particularly details the components of the simulation environment and modeling and of the network components. The performance metrics considered in this thesis are employed to validate the models by presenting the results of some well known protocols (namely ALOHA and ALOHA-Q).

Chapter 4 investigates the feasibility and performance evaluation of ALOHA-Q protocol in an underwater multi-hop chain network scenario. Variable frame sizes and multiple source operations are investigated to analyse the impact of long propagation delay on synchronization and performance. The novel contributions of this chapter are:

- Translation and investigation of the feasibility of an RL powered MAC scheme by evaluating ALOHA-Q protocol in linear chain underwater sensor network environment.
- Understands and highlights the effect of synchronisation and frame length on the performance of ALOHA-Q and ALOHA variants in underwater chain networks suitable for pipeline monitoring.
- Investigates the effect of multiple source operation on ALOHA-Q performance underwater and to propose the replacement of the existing FIFO queuing with RR queuing for improved fairness and overall performance.

The work performed in this chapter resulted in the following conference presentation:

Alhassan, Ibrahim, and Paul Daniel Mitchell. 2019. "Monitoring Free Span Sections of Subsea Pipeline with ALOHA-Q." URSI Festival of Radio Science ; Conference date: 16-12-2019 Through 16-12-2019.

Chapter 5 presents an analysis on a typical random access scheme to investigate the impact of frame structure on performance. The analysis involves simplifying the utilisation equation to inspire an aggressive frame duration. Whilst the size of the ACK packet might be small relative to the data packet size, however in UASNs the the impact of the incurred propagation delay on the frame duration is typically costly. Hence, a detailed pictorial analysis employed to drive the achievable utilisation of the representative protocol and network model is presented. The chapter proposes a novel MAC scheme (ALOHA-QUPAF) that exploits negative and positive RL on a dual Q functions to create the optimum schedule in a chain UASNs. The work presented in this chapter contains the following novel contributions:

- A set of proposed ***new frame sizes*** that can be preconfigured based primarily on the relationship between the propagation delay and the packet transmission duration for improved channel utilisation.
- ***Application of pictorial analysis for theoretical analysis*** to describe the achievable system utilisation levels using the proposed frame size in a random access scheme, thus providing the insight on the possible performance gains with intelligent MAC scheme.
- Employing packets ingress as a negative reinforcement signal for ***reception slot isolation using negative RL***.
- Employing packets flow (ingress and egress) at node level to implicitly derive the positive reinforcement signal for the transmission slot selection, to ultimately achieve a network-wide ***packet flow harmony***.

These contributions culminated in the the following journal article:

Alhassan, Ibrahim B., and Paul D. Mitchell. 2021. “Packet Flow Based Reinforcement Learning MAC Protocol for Underwater Acoustic Sensor Networks.” Sensors 21 (7). https://doi.org/10.3390/s21072284.

Chapter 6 presents areas of further research that expands on the scope of current work. These include specific recommendations on more complex scenarios, mathematical analysis on convergence properties and energy efficiency studies of ALOHA-QUPAF.

Chapter 7 draws the conclusions of the thesis. It summarises the contributions and presents the discussions on how the hypothesis was addressed.

Chapter 2

Literature Review

This chapter provides the review of the background research that inspired this thesis. Following a top-down approach, from network level down to protocol level; the chapter presents a brief history, applications and the principal components of underwater acoustic sensor networks. An overview is then given on the general classification of the medium access protocols with examples. The chapter then presents a focused review on underwater MAC protocols. Furthermore, a review of the concepts, applications and framework domain of RL are presented.

2.1 Wireless Sensor Networks

2.1.1 Introduction

By the turn of the century, the world saw an unprecedented leap in technological advancement, which in turn resulted in a paradigm shift on how we collate, transport and process information. Traditional information processing has been human-centric, then, embedded devices facilitate interacting with the environment to monitor and control physical processes. The massive demand for our need to monitor and control our environment sparked the development of sensors and actuators to meet the required tasks, but deployment of such devices demands a means of transporting (conveying) the accrued data to the desired destination. Wiring is the obvious interconnection option, but becomes impractical as internetwork (inter-device) distance, and number of participating devices scale higher. Thus, practicality dictates wireless interconnection between nodes/devices [6], and that was the birth of a new class of networks appropriately named; WSNs. As we push the boundary of what is possible, interactions at different levels manifest themselves; human-to-human,human-to-machine, machine-to-machine, and by extension human-to-physical world. WSNs are the networks deployed with components (nodes) configured to collaborate in tackling a common task; i.e. nodes could be of homogeneous or heterogeneous nature but their collective computation/processing resources are harmonized to optimally engage a common target scenario and fulfil task(s). Typically, these nodes are architecturally low power miniaturized devices endowed with a small processing unit, sensing unit and a transceiver. However, the huge computing power from the aggregated network entities ushers in a new frontier whereby potentially no task is too big or too complex to undertake, as such, the WSNs applications domain keeps expanding [2, 7, 8, 9].

2.1.2 Applications

Broadly speaking, the applications of WSNs fall under the domain of either monitoring or control, but a significant number of applications tend to have both monitoring and control combined. Some of the specific applications areas are:

- **Agriculture and conservation** Sensors can be attached directly to plants or soil and tag livestock, the resulting network can be employed to establish precision in; irrigation, fertilizing and pest control and monitor health status of livestock [10]. Similar to the WSNs applications in conservation [11], UASNs can also be adopted and used for marine biodiversity mapping, and in conservation whereby endangered sea animals can be tagged and tracked so they can be protected against unsustainable hunting practices.

- **Infrastructure monitoring** Another widely popular area of WSNs is application of sensors to wide range of facilities; pipelines, bridges,smart homes and building facilities. Depending on the target facility, WSN applications can range from monitoring structural integrity, localizing failure points to providing security, access control and movement of people and goods. Industrial facilities where safety is of major concern can be comprehensively covered with sensors to monitor and report status of processes from inaccessible locations to a central control on-site or off-site. This application area can be especially useful in monitoring subsea oil and gas assets that could prove to be costly ecologically and economically if left unmonitored.

- **Health care** Using smart wearable devices for monitoring personal health status for a doctor-patient tracking system [12]. With the number of injuries and fatalities directly attributable to a lack of up-to-date patient status and medical information, health care is an area where WSN is getting a lot of attention and acceptance. WSNs can eliminate the dangling wires from the sensors pads attached to patients either in recovery rooms or intensive care wards.

- **Disaster monitoring and early warning** WSNs can be deployed to disaster prone regions, for example networks on seismic and active volcanic sites could be deployed for monitoring and advance warning for preparation and evacuation to the authorities. Wild fires [13] and flood profiles (e.g temperature map) can be gathered for identification of early response and rescue mission entry points. This is another application that UASNs could be crucial in saving lives and property by providing an early warning system for tsunamis.

- **Mining** Mining sites represent a hazardous working environment, to which WSNs can play a major role in tracking miners (or equipment), monitoring of air pollution and prediction of seismic shifts induced by mining activity [14] or otherwise. With fatalities resulting from suffocation, tunnel collapse, lack of structural integrity information after collapse can often cause delay and uncertainty in proper rescue operation which can take months. With the proper network setup potentially disastrous situations could be avoided.

- **Logistics and Transportation** WSNs provide a nice platform for vehicle [15] and parcel tracking and smart inventorying at storage facilities and warehouses, as an upgrade to the popular passive Radio Frequency Identification (RFID) system which depends on the tagged object coming into the vicinity of the monitoring station, however this setup requires active system which can help track both stationary and mobile objects and additionally provide objects history (activity logs).

2.2 Underwater Acoustic Wireless Sensor Networks

As a complement of terrestrial WSNs, Underwater Sensor Networkss (USNs) are envisaged to be integrated into the conventional terrestrial radio networks, thus bridging both worlds, thus, finally enable access to the elusive underwater world [2, 3]. It should be noted that the choice of the choice of sound waves for USN over EM is out of practicality. Radio and experiences high attenuation [16, 17] and optical waves greatly affected by scattering. Nevertheless, successful transmission with both type of waves underwater has been reported [8, 18, 19, 9] albeit to limited distances (maximum 100m for radio and 10m for optical) thereby limiting their practical application. Sound may not offer the desired speed and datarate/bandwidth, but currently it is the least affected by attenuation and can therefore cover the distance of any relevance to practical application [17]. Broadly speaking, the applications of UAWSNs fall under the domain of either monitoring or control, but a significant number of applications tend to have both elements of monitoring and control. The applications areas based on deployment purpose/intent: Civilian or Military are given as follows:

2.2.1 Civilian Applications

A broad array of civilian applications are envisaged for UWASNs such as: Deploying undersea networks for conservation studies especially in mapping and tracking of marine biodiversity and pollution studies. Undersea infrastructural monitoring, fault localisation or to enable control of assets remotely. Lost of lives and property can also be alleviated or prevented by employing UAWSNs as a disaster and early warnings systems of undersea events, such as, an underwater earthquakes and tsunamis. Subsea mining and assisted navigation are other areas UAWSNs can be used.

2.2.2 Military Applications

UAWSNs can be deployed for both offensive and defensive purposes. Coastal defenses can be reinforced with subsea sensors against territorial breach or attack. Sensor nodes can potentially be deployed to provide assisted navigation for submarines, weapons guidance systems and also for tactical surveillance behind enemy lines.

2.2.3 Differences between WSNs and UAWSNs

Given the overlapping or similar application domain shared by terrestrial radio and underwater acoustic sensor networks it is important to highlight some of the key differences between them especially the enabling technologies, so as to appreciate the challenges of designing underwater networks.

- The power consumed by underwater nodes is typically higher than that consumed by terrestrial (RF) nodes for all operating modes i.e transmit, receive and idle, primarily due to the difference in physical layer technology.

- Terrestrial nodes are typically more densely deployed than underwater networks, whereas data correlation is common practice in WSN, large separation between nodes makes data correlation in UAWSNs unlikely.

- The cost of an underwater sensor node is more expensive than terrestrial node and the gap is expected to widen in the near future; this can be attributed to the extra level of protection measures that must be incorporated on acoustic nodes for the device to operate in the harsh environment, and significantly fewer number of vendors to effect competition in pricing.

- Terrestrial nodes are designed with limited on-board processing power and storage, whereas underwater nodes relatively require more on-board processing power, especially large memory so as to cache more data in anticipation of the highly probable link disruptions. However, radio nodes offer higher capacity.

- Spatio-Temporal Uncertainty is the uncertainty associated with packet arrival(s) at a receiver from the uncertainties in the relative propagation delays between separate transmitters to that receiver and the transmissions time. This uncertainty is especially prominent in UASNs, as the peculiar propagation delay underwater can potentially makes two different transmissions from unequally displaced transmitters to arrive at a receiver at the same time. Similarly, it also enables transmissions from equidistant nodes to arrive at different times due to the effect of dissimilar channel conditions on the two links (such as pressure/depth and temperature difference, salinity). Therefore, conventional MAC techniques that purely relies on transmitter-centric solutions (such as TDMA, CSMA, synchronised time slotting) to mitigate this uncertainty underwater need to be re-examined [20].

2.2.4 Offshore Pipeline Monitoring

Due to increasing global energy demand, the offshore oil and gas industry has a global market growth rate estimated at 6.8% CAGR [1] [21] and subsea pipelines are the lifeline of that industry[22, 23]. These pipelines are laid along the seabed with some sections buried or exposed. Uneven seabed geometry causes some of these exposed sections to have reduced support as they cross valleys along the seabed or pipeline sleepers[2]. These free span sections particularly become vulnerable to slugging, vibrations and ship anchors, all these can potentially damage the pipeline section, thereby disrupting pipeline operation and causing damage to the environment in the event of leak or rupture. To maintain pipeline integrity and to conform with regulations, various monitoring approaches are currently employed, including retrofitting sensors at appropriate intervals [24], guiding Remote Operated Vehicles (ROVs) or sending Autonomous Underwater Vehicles (AUVs) along the pipeline to inspect or measure: motion, vibrations, lateral displacement, and detect leaks [25]. However, sensors attached to pipelines to record data lack ”real time” reporting, provide outdated, incomplete and possibly corrupted data that is realised

[1] CAGR - Constant Annual Growth Rate
[2] Constructed to control displacement and buckling

at a much later time. Furthermore, the limited battery capacity impedes long term operation. While ROVs and AUVs are accurate in their task, however, they are notably expensive to procure and deploy with operational cost up to $5000 per day[26], they are slow during operation for in situ exercises and it is highly improbable for ROVs and AUVs to report incidents as they happen. Conventional methods of underwater monitoring activities are undertaken with sensing equipment that is placed in the location of interest to be retrieved later, or with equipment on ship(s) for data collation in situ. The first approach is susceptible to the problems of possible equipment failure, data corruption, or not been collected and late access to data, all of which might be realised much later. While the later approach is typically limited to short and sparse time intervals between data gathering and also depends on favorable weather conditions [27]. On the other hand, underwater exploration and undersea asset inspections employ highly specialised and expensive remote controlled machines with restrictive cabling or short range autonomous machines. Accounting for 30% [4, 5] of global oil production, these assets and their smooth operations are immensely valuable economically [28], and in the event of sabotage or accidents, they pose huge risks of ecological damage. Moreover, the dwindling oil prices are making the industry resort to measures of extending the useful lifespan of offshore installations such as marine pipelines, arguably, that will increase the associated risks. Typically, these assets covers a substantial marine area that is not readily accessible nor easily guarded. Arguably, an underwater acoustic multi-hop network of sensors deployed along the asset or distributed on site has the potential to monitor assets integrity and also proximity coverage for the whole site.

2.2.5 Network Deployment Scenarios and Topologies

Whilst UASN topologies are application specific, additional factors such as technological limitations, communication range, channel conditions, terrain layout and accessibility determine the manner in which a link is established between the source and the sink. However, a simplified categorization can be made based on the number of effective connections/links required for data from source node to reach its destination node in a given network [6].

- Single-Hop is a fully connected topology, where both source and sink can establish a direct link because they are within range despite the effect of prevailing channel conditions. This scenario can best suited for centralized networks whereby every node reports to (or controlled by) one central node, and distributed deployments in which case each node is its own master and connections between nodes of a given network is initiated individually.

- Multi-Hop is designed for a situation whereby establishing an effective source(s) to sink(s) link is not feasible various reasons. For example, due to transceiver power limitations, the sensor nodes are unable to cover the communication range necessary to setup a link, or poor channel conditions due weather, presence of shadow zones, and obstruction caused on the wave path by an underwater object(s) or general terrain topology. Application constraints can also demand a multi-hop implementation by imposing a certain network wide granularity in the collation of the sensed data and the direction of data flow. Typically, a relay network is setup with multiple nodes participating in the source to sink link creation. Notwithstanding the overall nodes density in this category, there are three approximate realisations of multi-hop topologies; random, linear-chain, or cross-chain. It is possible in some applications to reconfigure and aggregate multiple existing single hop networks to realize a multi-hop network, thereby upscaling the network to cover larger area or adapt it to new task.

2.2.6 Network Architecture

Underwater networks are designed and deployed in accordance with the application and location. Notwithstanding the diversity of applications, two architecture options are available to suit all implementations. Presumably, no matter the engaged underwater phenomena, the sensed or collated data has to be transported ashore, a common denominator found in these networks is a surface gateway. The node designated as the surface gateway is afloat and situated within the range of onshore station(s) and submerged node(s), it is therefore equipped with both and acoustic transceivers. The RF transceiver partake in relaying signal from onshore or satellite and the node (surface link) whilst the acoustic transceiver maintains the underwater link. This

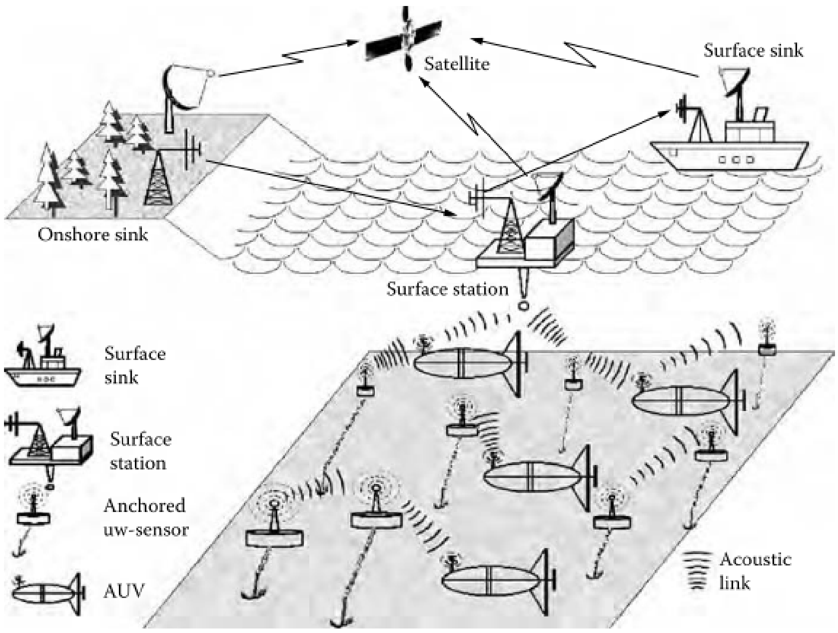

Figure 2.1: 3D Underwater Network Architecture [Directly reproduced from [17]]

gateway serves as the relay between the surface network and the undersea network by accepting acoustic signal from submerged nodes and converting to RF then transmitting onshore/satellite and vice versa. The nodes and relevant underwater network components are secured directly to the seafloor seabed and anchored at varying depths [17], the depths can be regulated electronically when the node is equipped with a motorized system. In this case however, due to the presence of undersea activity (related to:currents,creatures, ships) node motion is unavoidable which can be detrimental to network connectivity. Additionally, Autonomous Underwater Vehicless (AUVs) can be incorporated into this architecture for added robustness, sensing and communication range. Typically, when a large area is to be covered which is beyond the range of the nodes, a cluster-based or multihop approach is used. In a cluster based scenario, a group of nodes within range of each other in a particular area are considered a cluster and one of the nodes is chosen as the head (cluster-head) than can communicate with neighboring cluster-heads and this node may (or may not) be more powerful than the other members of the same cluster. In other cases, polling or the role of the cluster-head can be rotated between the nodes based on some elaborate election algorithm (e.g based on power reserve or fault detection).

2.2.7 Data Collection Modes

Sensor networks are all about interacting with the environment that is under observation. A cardinal aspect of such interaction involves the manner in which data is collated at the source. In order to appropriately and adequately meet/represent the varying applications, one of the following devised strategies are employed [29, 30]:

- Polling entails taking turns in data sensing/requests from sensor nodes by a central application or a master/central node. This can also be expanded to include on-demand data collection based on proximity to node in the case the requesting node (master node) is mobile.
- Event-driven data is only gathered based on a triggering event sensed by the node. Here the nodes are typically on high alert to any activity regarding a change in the observed environmental conditions. Examples include seismic activity monitoring, smoke/fire systems and intrusion systems.
- Periodic sampling is performed at regular intervals on the environment and the accrued data can also be requested from the node in similar fashion. This strategy is suitable for environmental monitoring applications, such as pollution monitoring and climate change monitoring.

2.2.8 Acoustic Sensor Node Technology

Embedded in a typical sensor node is an acoustic modem. This is the device that establishes and maintains the link(s). The fresh attention garnered by UAWSN in recent years, spurred researchers to focus on proposals and development of various standards, simulators, platforms and test-benches for acoustic modems [31]. Figure 2.2 illustrates the general block constituents parts of a typical acoustic sensor node used for UAWSN setup. The components are:

- **Acoustic modem** is responsible for converting the generated electrical data traffic from the node into modulated acoustic analog signal and finally transmitting the signal, also

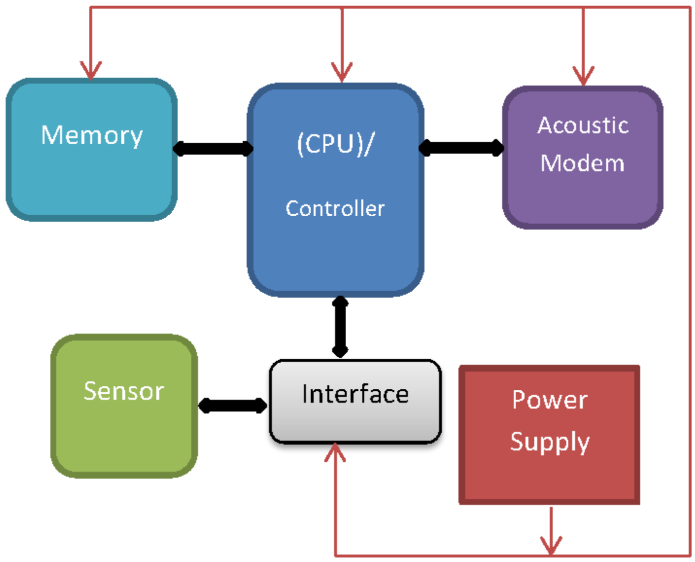

Figure 2.2: Block Diagram of Underwater Sensor Node

receiving the acoustic signal from other transmitters demodulating and converting the signal to electrical signal suitable for the node to process.

- **Power supply** all the components draw their power from this unit, and is typically battery based which is one of the challenging aspects of UWASN design, since the size/capacity of this unit plays a key role on the lifespan of the node.

- **Sensor and sensor interface circuitry** depending on the phenomena under observation the sensor(s) output (response) is appropriately filtered, formatted and conveyed to the processing unit (or controller) by the sensor interface circuitry. For example a temperature transducer's output is analog signal with lots of noise to which the interface circuitry will filter out the noise and correctly feed the controller the appropriate digital levels.

- **Processor and Memory units** whereas the processor and associated controlling unit co-ordinate and monitor the activities of the other components, the memory unit stores the processor's instructions and also provides a permanent or temporary location for storage of sensed data in-transit prior to transmission.

All things considered, among the above sensor node components, the acoustic modem is of

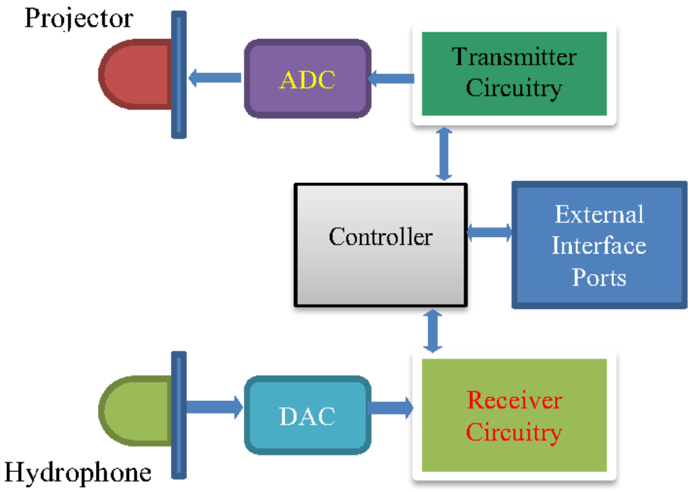

Figure 2.3: Block Diagram of Underwater Acoustic Modem

special interest to researchers and underwater acoustic network designers. Functionally it is the most distinctive component that define this class of networks. Architecturally, acoustic modems can comprise of everything found in underwater sensor node but the sensor unit. Figure 2.3 depicts the architecture of a typical acoustic modem. A brief description of their functions is given below. Some of the units are grouped together owing to the similarity in their functions (differ only in direction of data flow).

- **Projector and Hydrophone** these are the analogues of a loudspeaker and microphone; the set makes up the transceiver interface of an acoustic modem. In an underwater environment, the projector is an electro-acoustic transducer is used to map the generated electrical signal into sound signal suitable for the water medium. Conversely, the hydrophone picks up and converts the incident sound signals into electrical signals for further processing in the internal circuitry.

- **ADC/DAC** units convert between the two basic means of signal representation. Both the projector and hydrophone are analog devices whilst the internal circuitry of the modem itself works with digital signals. An Analog to Digital Converter (ADC) accepts the hydrophone's analog output and converts it to digital form for the receiver circuitry. Digital to Analog Converter (DAC) converts the transmitter digital output into analog for the projector.

- **Transmitter/Receiver circuitry** these units are functionally the mirror identical to each other. Depending on the direction of data movement. Signals from the Central Processing Unit (CPU) are passed to the transmitter for coding and modulation and then send out, likewise receiver accepts the incoming signal for demodulation and decoding and subsequently passed on to the CPU.

- **External interface** provides the means to upload programs (or reprogramming and diagnostics) the modem or routine download of stored data from on-board flash storage as is common in most of the commercially available devices. These ports can be USB or RS-232.

- **Controller(CPU)** this unit controls and coordinates all the other components especially with respect to data flow and also housed the storage space for the ROM programs. This unit may be optional especially if the modem is part of a sensor node.

2.2.9 Practical Implementations, and Deployment

A fairly exhaustive study of acoustic modems [32] both commercially available and from researchers reveals a wide range of design goals and philosophies. Whilst the experimental modem implementations have varied design goals with some focusing on cost [33, 34], adaptability [35], offered flexibility through reprogramming [36, 37, 38], or dual operation modes with variable bit rates [39]. On the other hand, the commercial offerings are generally designed for specific application(s) in mind, thereby achieving a compromise of optimising the critical parameters and tuning other quoted parameters to an acceptable levels. Due to the challenges of supplying power to the devices, majority of current modems are designed with power efficiency considerations. Even though the relatively more expensive commercial modems are generally superior to the experimental modems, however, comparison between the two options based on maximum transmission range is more definitive: Commercial modems cover longer ranges

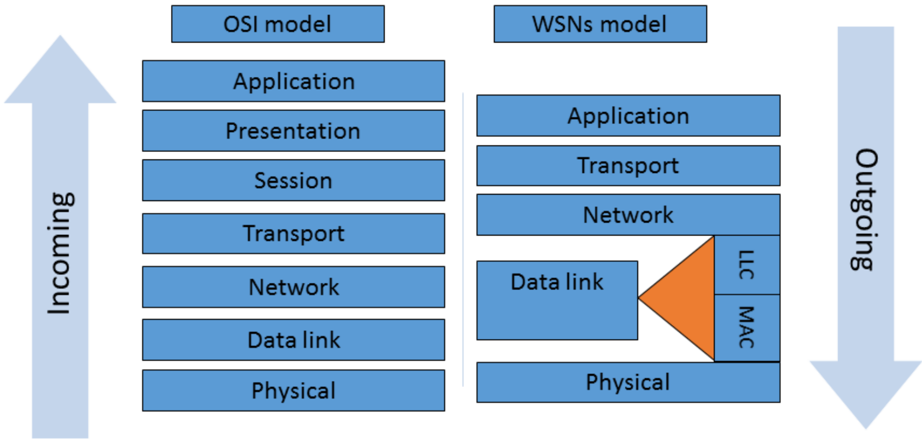

Figure 2.4: OSI vs WSNs models

0.5km to 25km versus experimental 100m to 2km respectively, however, at the expense of more energy consumption [32, 40, 41].

2.2.10 Communication Protocol Stack

The OSI model developed by the ISO in the 80's consists of seven groups of logically defined functions in a stratified (layered) manner. This model decomposes the complex communication tasks into functionally dependent sub-tasks [42, 43, 44]. The OSI model inspired the existing practical internetwork model employed in WSNs. The model used by sensor networks is a simplified version made up of five layers, two factors influenced the modified model namely; the overall reduced functions and size requirements of nodes, and the feasibility of seamless merger of closely related layers while maintaining the abstracted functionality. Figure 2.4 is an illustration of both the OSI and the sensor network model.

The Application layer is where the application interface resides, here the information (data) is manipulated and given access to the network. The Transport layer delivers data from process to process that by mainly establishing and maintaining a logical route between the source(s) and the destination(s) nodes. The Data link layer is sandwiched between the network and the physical layer, overall this layer is in-charge of link establishment, link termination and channel access control between devices. This layer is further sub divided into two layers, the Logical Link Control (LLC) and MAC respectively. The MAC sub-layer is responsible for orchestrating channel access, and collision resolution. This sub layer is the focus of this research, and more details on it will follow in the subsequent sections. The physical layer is the bottom layer of the stack tasked with basic services needed by the node to properly send and detect signal over the physical link. These functions range from coding, modulation, frequency selection and generation, and selection of transmission and reception power.

2.2.11 Challenges

Noise, cost, interference, and power are some of the major constraints to a varying degrees affecting all communications systems [2, 45, 7, 46, 1]. Hence, communication system design practices involve a compromise on amongst the constraints. However, the harshness of underwater environment and the peculiarity of the acoustic channel amplifies and introduces fresh challenges that demand new strategies and solutions in achieving a functional and practical UASNs. Some of the impacting factors are highlighted below:

- Noise/Interference: Sounds originating from sea ambiance, aquatic life, and shipping and other man made activities are some of the various types of noises in the underwater environment that poses a serious challenge to acoustic network link stability and reliability. These sounds can be in the form of short impulses (e.g shrimp snapping), or longer duration (e.g ship propellers, drilling or mining activities) that could interfere and disrupt connections.

- Power : The Power supply is generally an issue in sensor networks, this is because nodes are mainly powered by batteries. once the battery power has been depleted. the current energy scavenging techniques have not adequately matured for recharging to be practical. Indeed, developers are resorting to exploring new sustainable energy saving solutions in both hardware and software to optimise and reduce energy consumption in devices. For example, employing shorter communication ranges (shorter hop distance) for reduced transmission power requirements per node and incorporating sleep cycles to reduce idle listening.

- Cost : Despite the huge application potential, UASN nodes unit cost is expensive to procure because of the technology is young (relative to terrestrial sensor nodes). Also nodes have to be ruggedly encased for protection against fouling from both environmental factors (such as the water pressure with increasing depth) and animal activities. With few manufacturers means there is limited competition to drive prices down as the industry is currently a niche market. UASNs are deployed to cover substantial area, hence, increasing the network size (number of nodes required) increases the overall nodes unit cost as well. Furthermore, network deployment is an expensive undertaking requiring specialised equipment for transportation and personnel protection.

- Environmental Dynamics/Conditions : The underwater environment is both harsh and predominantly unpredictable, therefore location access and targeted deployment of sensor nodes are equally challenging.

- Channel Characteristics : The greatest challenge to designing UASNs is arguably the acoustic channel, as the long propagation delay, multipath limited bandwidth, and high Bit Error Rate (BER) are all linked to a channel property.

2.3 Medium Access Control

This section introduces the Medium Access Control (MAC), including some classic MAC schemes, and some specific radio MAC schemes because they inspire some of the UASN schemes. It also presents the current state-of-the-art solutions in UASN MAC protocol implementations.

The MAC protocols define the set of rules that control channel access to networked devices. This is necessary because the medium is a constrained resource and unsystematic use of the channel by multiple devices will result in partial or total communication failure due to interference. Hence, the medium needs to be efficiently shared amongst the competing devices. That is why the cardinal function of MAC layer is to prevent/alleviate collision and/or resolve contention during communication [47]. Supplementary network tasks of MAC protocols include: improving throughput, energy efficiency, robustness, latency, and scalability [48]. In reality, no single MAC scheme is sufficient to satisfy all applications requirements. Therefore, when designing MAC protocols, optimum weights are identified for each metric through appropriate trade-offs based on the requirements of the target application. Although the application design goals dictate the choice of scheme and the overall MAC protocol design, the following key properties are deemed important for any good MAC protocol;

- **Effective channel sharing** through proper protocol initialisation and effective capacity allocation.
- **Robustness** to failure of equipment or channel conditions.
- **Efficiency** with respect to throughput, latency and energy consumption.
- **Flexibility** to different types of traffic.
- **Stability** to changing load conditions.
- **Fairness** in that protocols should have a justified means of channel access among nodes.

In addition, limitations of power supply to the sensors place constraints on lifetime of sensor networks. Many MAC protocol designs advocate a special focus on energy efficiency. The following were identified as some of the primary causes of energy waste in sensor networks [45, 49].

- Deafness due to the half-duplex nature of UASNs a deafness situations occurs when a node misses an incoming packet(s) while the node itself is busy transmitting.
- Overhearing whereby node keeps receiving packets not destined to it. The energy consumed in processing the packet before discarding becomes a serious source of energy

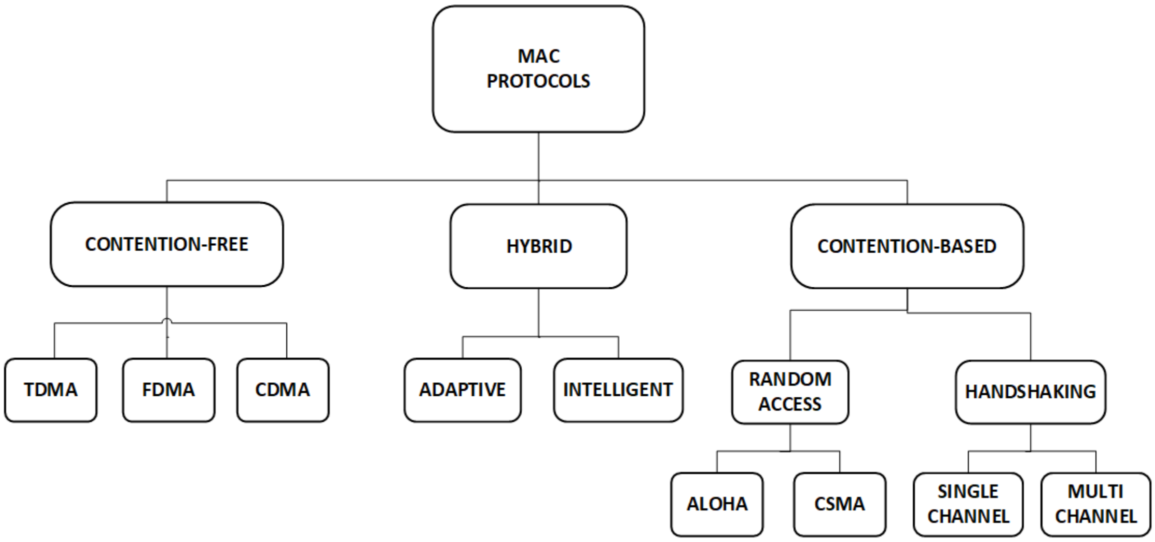

Figure 2.5: Classification of MAC protocols

wastage as the network becomes dense. Proposed solutions include sleep cycles, wake-up tones, and multichannelling.

- Collisions of two or more packets transmitted simultaneously are susceptible to overlap at the receiver and when recovery attempts failed additional energy has to be consumed in retransmission and reception.

- Idle listening is also another significant source of energy waste in sensor networks, where the transceiver unit is kept active for both receive and transmit phases, as the unit consumes relatively equal energy (in UASNs the transmit power is significantly greater than the receive power).

In the literature MAC protocols are conventionally categorised based on the whether channel access is contention-free or contention-based. Another popular category in the research community is the hybrid MAC protocols as included in Figure 2.5.

2.3.1 Contention-free

The techniques under this category ensure collision free access on a shared channel based on assigning distinctive: time slots, frequency bands or codes to each participating device in a network. Accordingly, three(3) classic schemes found here: Time Division Multiple Access

(TDMA), Frequency Division Multiple Access (FDMA) and Carrier sense Multiple Access Control (CSMA). The nature of pre-allocating a particular channel resource in this category, makes the schemes inherently centralized. TDMA is one of the major scheme used in UASNs. Due to the long and variable propagation delay, node mobility, and the complexity of implementing distributed TDMA, this scheme in its purest form is not entirely good/appropriate for UASNs [50]. However, its potential for true collision-free transmission, simplicity, flexibility and sleep-cycle incorporation has inspired development of many enhanced and hybrid MAC strategies for UASNs.

2.3.2 Contention-based

In this class, protocols are designed such that nodes compete for channel access and control on-demand. Consequently, pre-allocation of channel access is eliminated. Because the protocols are distributive architecturally, contention-based strategy is one of the major approaches in developing UASN MAC protocols.

The core idea is to simply grant unrestricted channel access to users with little or no coordination. The disadvantage of this freedom is reduced performance at high loads due to rampant collisions. Random access protocols are predominantly ALOHA-based [51]. CSMA is one of the most popular protocols used in terrestrial networks. Carrier sensing is offered to address the rampant collisions suffered by blind transmissions in ALOHA variant protocols. By allowing nodes to listen for channel activity before transmitting, users will make an informed decision prior to engaging the channel and avoid potential interference. However, because of the long propagation delay in underwater environment this carrier sensing is ineffective and fewer CSMA inspired protocols have been developed.

Handshaking protocols/Reservation based schemes are based upon the principle of exchanging short messages between devices to acquire the channel prior to data transfer. This dialogue between the communicating devices to secure the channel is entirely contention-based. Some strategies utilise a TDMA like approach of dividing time into periodic frames, whereby in each frame there are separate fixed slots for data and reservation messages, as such nodes compete for channel access using the reservation slots and if successful, data is transferred in the reserved slot otherwise the node will try again in the next frame. Other schemes employ ALOHA approach to transmit request packets and if unsuccessful, transmission is deferred with a back-off timer before trying to request again. The reservation schemes are typically implemented on a single channel due to the limited underwater acoustic bandwidth, notwithstanding multi-channel techniques are being proposed. Essentially, in the multi-channel approach the data and control packets occupy different channels [52], whilst in the single channel case both handshaking and data exchange will happen on the same channel.

2.3.3 Hybrid

Hybrid protocols are gaining wide attention in the research community especially in the UASNs because they provide a needed versatility for changes in networks status effected by a dynamic environment, traffic or power reserves. A protocol can combine elements of different MAC schemes to achieve improved performance. The approach tends to be more complex and computationally intensive, requiring more capable nodes. There are a variety of realisations which include: switching and activating the most suitable component protocol, or optimum settings aided by an intelligent learning algorithm to adapt to changing conditions. The following are some examples of the relevant MAC protocols in the literature:

The ALOHA Protocol

In an attempt to connect the remote terminals located on different islands with a central terminal of the university of Hawaii over a packet radio network, Abramson in the 70's created the ALOHA protocol [51]. In its purest form ALOHA used two distinct frequency bands $f1$ and $f2$ (similar to uplink and downlink channels in satellite system), one band (say $f1$) shared by the remote terminals to connect with the central terminal, and the other band($f2$) for the central terminal to broadcast messages to the remote nodes. A terminal with packet to send simply do so immediately, the terminals monitor the central terminal broadcast channel for acknowledgement (ACK) indicating the transmission was successful for a period equal to the maximum round trip time of packets between the nodes, failure to receive an ACK packet after the round trip timed-out the node assumes collision has occurred and retransmission is initiated. The ALOHA scheme is a fully distributed but the problem with this scheme is the rampant destruction of packets due to collisions as multiple nodes attempt simultaneous transmissions.

For a packet A of fixed duration (τ) any other transmission originating within time T-τ and T+τ will collide with A. This period of 2τ is called the vulnerable period. Analytically the throughput defined as the average number of successful transmissions (S) is a function of the offered load (G) and is given by; $S = Ge^{-2G}$, based on this, the maximum theoretical throughput achievable by ALOHA is 18.4% at 50% offered load [51]. This translates to under utilization at lower offered loads and above the 50% load too frequent collisions degrade the throughput. This poor performance called for an improvement and come in the slotted ALOHA scheme which introduced two modifications; creating time slots equal to the packet duration and forcing nodes to only transmit at the beginning of a slot. These two modifications effectively reduced the vulnerable period by half and more than doubled the throughput (37%) of pure ALOHA at full load. However, such techniques employed in improving ALOHA in radio were found to be ineffective underwater, for example, slotted ALOHA which has twice the channel utilisation of pure ALOHA, was found to under-perform and lose its advantage underwater due to the lack of synchronisation and the presence of spatial-temporal uncertainty (See Section 2.2.3).

-

Multiple Access Collision Avoidance (MACA)

MACA is one of the original approaches that utilises handshaking between devices before data is transmitted [53, 54]. Prior to data transfer, the transmitting node initiates the session by sending a Request To Send (RTS) control packet to the receiver, which in turn replies with a Clear To Send (CTS) packet. Nodes in the neighborhood that hear the RTS signal will wait for the CTS response, upon hearing the CTS nodes will defer their transmissions for a preset time (adequate for the transmission duration). However, not hearing the CTS signal signifies to the neighboring nodes the receiver is outside their range and are free to compete for channel access themselves. This handshaking exchange will reserve the channel for the pair of nodes until data transmission is completed and the mechanism addresses the shortcomings suffered by CSMA in ad-hoc sensor networks, especially the hidden/exposed terminal problems, and the degrading inefficiency with increasing propagation delay. MACA-U [55] is a proposed MACA variant for multi-hop UASNs. In this adaptation, unlike MACA whereby when nodes that have sent RTS packets and overhear another RTS/CTS packet automatically defer their own transmissions (reservation deemed unsuccessful), MACA-U introduces a waiting states for both the CTS at the transmitter and the data packets at the receiver to address collisions of control packets due to the effect of propagation delay. In UASNs the handshaking is still ineffective in eliminating collisions and the schemes suffer from reduced channel utilisation as a result of the added handshaking overheads, latency, and energy inefficiency from excessive overhearing.

Slotted Floor Acquisition Multiple Access

Slotted-Floor Acquisition Multiple Access (FAMA) [46] is based on the FAMA [56]. FAMA scheme itself is an improvement over MACA. It incorporates carrier sensing (CS) and handshaking as defined in MACA. However, to achieve collision-free transmission using CS, Slotted FAMA address the two conditions that must be satisfied: the RTS packet lengths must be greater than the maximum propagation delay, and the CTS packet must be greater than twice the propagation delay plus the RTS packet combined. In order to meet the collision-free conditions, the long propagation delay in UASN will result in an unpractical RTS and CTS packet lengths, thereby rendering FAMA highly ineffective. Therefore, the slotting in Slotted FAMA removes the asynchronous aspect of FAMA and limits the length of the control packets that may become excessively long underwater. The slots are structured such that the length is duration of CTS plus the maximum propagation delay, and transmissions are restricted to the beginning of time slots. Additionally packet trains, ARQ and backoff strategies in a high BER environment are also included to improve the overall performance. These modifications in Slotted FAMA added to its complexity, reduced system utilisation from the control packets and long guard duration overheads. The propagation delay makes carrier sensing ineffective as a channel maybe sensed idle whilst another transmission is active. Furthermore, at high load and connectivity, the channel reservation becomes highly contentious and difficult achieve leading to waste of power due to retransmission attempts.

A Receiver-Initiated Reservation-Based Protocol for UANs (RIPT)

RIPT is another handshaking protocol proposed for multihop UASNs [57]. However, unlike in MACA and its variants, the handshaking is initiated by the receiver. It polls multiple nodes by broadcasting Request To Receive (RTR) packets to its neighboring nodes, and the transmitters then respond with a packet with their individual number of intended packets to send. The receiver then creates and broadcasts the schedule. In this way, the receiver accepts a synchronised packet stream and since collisions occur at the receiver, this approach eliminates the transmit-receive type of collision suffered by transmitter initiated sessions. By requesting and coordinating several packet trains the network performance is improved. The main disadvantages of this protocol is the dynamics of acoustic channel and variable traffic conditions limits its efficacy in UASNs.

A Traffic-Adaptive Receiver-Synchronized MAC Protocol for Underwater Sensor Networks (TARS)

TARS [58] is an adaptive protocol that leverages the long propagation delay between nodes to create a receiver synchronised schedules. Each node gathers and stores the propagation delay information between its neighbors, this delay is then used to compute the transmission phase (within a slot) between any transmitter-receiver pair. Generated traffic is assigned to separate queues for each outgoing receiver. This queues are then used to generate Q-tables for both incoming and outgoing transmissions. The protocol employs a traffic-adaptive algorithm to create the transmission schedules. This algorithm is a probabilistic routine that also relies on a both local Q-tables and shared Q-tables from all nodes in the neighborhood. Although, this approach has demonstrated how to avoid cross-slot reception, the reliance on shared information between nodes for the algorithm to work is complex, unreliable and will reduce the efficiency.

Furthermore, description on how to compute the slot size which is an integral part of the protocol was not given.

ALOHA-Q

ALOHA-Q is an intelligent hybrid MAC protocol initially developed for terrestrial sensor networks [59]. It employs Q-learning to intelligently create collision-free schedules in framed-ALOHA. The basic idea is to create a set of un-assigned time slots for nodes to independently find and occupy. The frame is formed by grouping a fixed number of time slots. Each slot is then assigned a unique index in a Q-table. During protocol execution, at the beginning of each frame, every node with data to transmit will looked in the the Q-table and schedule transmission in the slot with the highest Q-value. The result of each transmission attempt (ACK signal) is used as the reward/punish signal (± 1) that updates the Q-table. Therefore, successful slots are positively reinforced and unsuccessful slots negatively reinforced, this process which initially starts in trial and error eventually enables each node to occupy a unique transmission slot. The overall result is dramatic improvement in channel utilisation as the final collision-free schedules offers TDMA like performance without the investment in planning resources. However, the protocol suffers from reduced performance due to the effect of propagation delay on the slot size and is unfair under multiple sources. Nevertheless, due to its less complexity and adaptability has been considered for further studies in Section 4.2.

2.4 Reinforcement Learning

RL is a class of problems defined by learning through experience that is found in the natural world. Living organisms learn by interacting with the environment, and the feedback received shapes their behaviour. RL traced its roots in psychology and has since gained popularity in Computer science and Engineering applications. The principal idea is for an agent interacting with an unknown or dynamic environment to learn how to behave by trial-and-error [60, 61].

RL differs significantly from other machine learning paradigm namely, supervised and un-

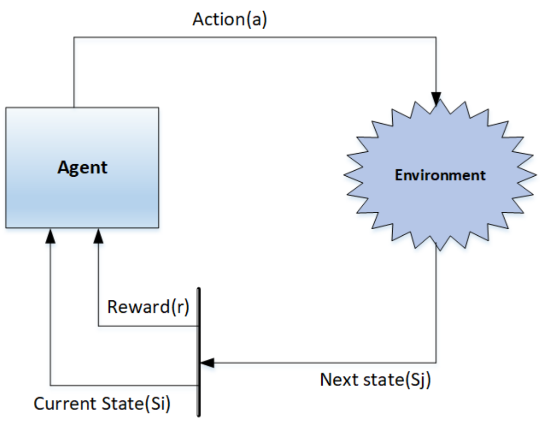

Figure 2.6: RL elements in agent and environment interaction

supervised learning. Whereas both receive training data, in the case supervised learning, full feedback is provided for its actions at all times, while unsupervised learning receives no feedback for its actions. No training data is given to the learning agent in RL, thus, it must learns from experience through its action and the associated feedback signal. The fact that RL does not require a priori knowledge of the environment makes it an attractive approach in designing adaptable and resilient sensor networks such as the UASNs, where the environment is statistically un-classified and currently un-modeled [62, 17, 63].

RL is comprises of three core and one optional sub-components as used in its standard formulation: A policy function, a value function, reward signal and the model of the environment respectively [60]. Figure 2.6 is a representation of the agent-environment interaction. The formal RL model comprises of a set of environment states $s_i \in S$, a set of available actions per state $a_i \in A$ and a set of scalar reward signals per action per state $r_i \in r(s_i, a_i)$. At each time step, an agent perceives the current state of the environment (s_t) and performs an action (a_t) to interact with and change the environment state (s_i to s_{i+1}) and receives a consequence of that action (r_i). Thus the process is described as a Markov Decision Process (MDP). The agent is principally tasked with finding the optimal policy π^* (*policy function*); mapping of states with actions that maximizes the overall received rewards. The *reward signal* is typically a scalar number responded to the agent by the environment that defines the quality or polarity (in terms of good or bad) of an action taken by the agent. This reward signal is immediate and depends on both the current state and the action taken. On the other hand, the *value function* defines the long term attractiveness of states with respect to the following states and their actions. The *value function* (also Q-function) is formulated as the expected cumulative future rewards in a state-action pair mapping from each state onward. The downside of accumulating the maximum rewards (r) to achieve optimality is that in an unbounded RL problem this becomes problematic as $r \rightarrow \infty$. Hence the rewards are typically bounded by a discount factor $\gamma < 1$ [60, 61]. In this way the choice of whether immediate reward or future reward should dictates the agent's action can be asserted. The *value function* is given by [64]:

$$Q(s_i, a_i) = \mathbb{E}[r + \gamma Q(s_{i+1}, a_{i+1})] \tag{2.1}$$

where $Q(s_i, a_i)$, $Q(s_{i+1}, a_{i+1})$, r, γ denote the Q-value of the current state, Q-value of the next state, the reward and the discount factor respectively.

Following the Bellman's optimality condition (Equation 2.2) [64], the optimal policy π^* can be directly derived by greedily choosing the action with the maximum reward.

$$a \in \arg\max_{a \in A} Q(s, a) \tag{2.2}$$

Hence,

$$\pi^* = \arg\max_{a \in A} Q^*(s, a) \tag{2.3}$$

However, the problem of exploration and exploitation poses a major dilemma to the RL paradigm. This is because, if an agent greedily decides to always chooses the best rewarding action based on its experience it exploits the system and risks (by not exploring) missing a potential new and better rewarding action. Some of the exploration-exploitation strategies developed in balancing this dilemma includes, $\varepsilon-$greedy which chooses the best action with probability $1-\varepsilon$ and a random action with probability ε ($\varepsilon \in [0, 1]$) the disadvantage of this approach is there is a risk of choosing the worst action, in random walk strategy the agent discards the any relevant experience and chooses new action always, and softmax (Boltzmann exploration) is one of the advanced strategy that balance exploration based on individual actions utility, and therefore minimises the risk of choosing a bad action [65, 61].

2.4.1 Learning an Optimal Policy

There are two main approaches for reinforcement learning algorithms to learn the optimal policy. In model-based the algorithm learns/builds a model of the environment and use it to computes the optimal policy. Therefore, the state transition function (S) and the reward function (r) are known beforehand. However, obtaining the optimal policy from a developed model in advance deviate from the core premise of RL. The scope of this thesis is concerned with the interaction between multiple agents (sensor nodes) interacting in a dynamic environment, and thus we consider the model-free implementation.

Q-Learning

Q-Learning is one of the most influential and popular off-policy Temporal Difference (TD) algorithms of RL. The algorithm learns and directly approximates the optimum value function regardless of the policy followed by the agent. A necessary condition for an optimal policy to converge in an MDP is that each state-action pair must be continually visited and updated, Q-learning has been shown to simplify analysis and is proven to converge with certainty [60, 64]. The standard formal Q-learning is recursively updated using Equation 2.4 [60].

$$Q(s_i,a_i) \rightarrow Q(s_i,a_i) + \alpha[r_j + \gamma \max_{a_j} Q(s_j,a_j) - Q(s_i,a_i)] \tag{2.4}$$

where $Q(s_i,a_i)$, (s_i,a_i), $\max_{a_j} Q(s_j,a_j)$, (s_j,a_j),r_j,γ, and α denote the Q-value of current state, current state-action pair, maximum Q-value of next state actions, next state-action pair, the reward signal, discount factor, and the learning rate.

Single State Q-Learning

A environment is classified as stationary, when the complete history of the environment can be sufficiently described by the information in the current state. As such, the environment states reduce to one and the learning agents become stateless. Accordingly, the Q-learning is significantly simplified with only single state recursive update with the reward one action. Since, the policy estimation in the standard Q-learning requires significant recursive updates of multiple states Q-values, the single-state update will potentially reduce the computational cost and the associated number of trial iterations needed to for the policy approximation. In order to study the multi-agent cooperative learning problem the standard Q-learning algorithm is redeveloped and simplified to remove the state dependency for solving stateless/single-state game problems(Equation 2.5) [66], however, the technique has since then being applied to similar single state multi-agent problems [67, 68] and in development of intelligent MAC protocols [59, 69].

$$Q(a) = (1-\alpha)Q(a) + \alpha r \tag{2.5}$$

where $Q(a)$ is the Q value of current action a, α is the learning rate and r is the reward of the chosen action.

Multi Agent Reinforcement Learning

MARL describes a learning involving multiple agents in a single environment. This presents additional challenges and complexity to the learning problem of varying degrees depending on the learning objective(s). An important consideration is that, the environment is no longer static, as actions of other agents affects others. In a situation whereby the learning is cooperative, such as the sensor networks, formulating an optimum policy becomes complicated. Most of the MARL algorithms are inspired by the early studies of MARL problems as reported in cooperative game theory [66, 70, 67]. To extend RL into MARL problem the works in [66, 67] both assumes independent learners to justify reducing the MARL to a single state MDP. This is because an independent learner can perform actions and update the Q-learning algorithm directly without any regard to the actions of others. Although, the assumption of stationary state has been shown to generate remarkably good results in developing intelligent MAC protocols, to prove convergence and computational tractability some aspects of the environment have to be explicitly modeled [69, 59]. In reality, the environment in MARL is definitely unpredictable and information regarding other agents and the environment are shrouded by noise.

2.4.2 Partially Observable Environments

The agent-environment interaction in RL problem is modeled as an MDP because it certifies the Markov property [60]. In most practical systems perceiving the state of the environment is not always possible. Interacting with these environment states only emits probabilistic observations. Similar to the MDP, it consists of finite set of discrete states(S), probabilistic state and action transitions, however the current state is uncertain and all actions result in a noisy observations that are probabilistic function of the states [71]. POMDP are then treated as a continuous MDP with belief state (b) that is the probability distribution over the entire states [72]. Figure 2.7 depicts the elements of POMDP. At each discrete time step, the objective of the controller is to compute an optimal policy π^* that controls the transitions of the states and observations ($s_i \in S$ and $y_i \in Y$) in the Hidden Markov Model (HMM) by choosing actions($a_i \in A$) that that maximise the expected reward based on the perceived belief state(b_i).

Decentralised Partially Observable Environments (Dec-POMDP) is a formal framework that extends the POMDP framework to cooperative/social multi-agent systems [73]. As stated in MARL, the environment is not stationary, and the noisy observations include contributions of other agents actions in the system, hence, the interaction between agents influence the decision of others. De-POMDP is specially designed to address at least three class of uncertainties: due to action outcome, environment state and multi-agent. In UASNs where explicit communication between agents may not possible or ineffective. Agents have to locally make decisions by estimating the state conditions from the noisy observations without any explicit knowledge of other agents actions. There is a lot of potential source of uncertainty. Given the techniques applied to solving MDPs such as the Bellman's equation are computationally intractable in POMDP models, stochastic approximation strategies and heuristcs algorithms are principally used instead [71]. Algorithms developed in Dec-POMDP framework have the potential to provide the missing component in developing effective and practical system protocols and controllers for sensor networks [73]. The work in [74] demonstrated the feasibility of using stigmergy (implicit) communication in MAC protocol.

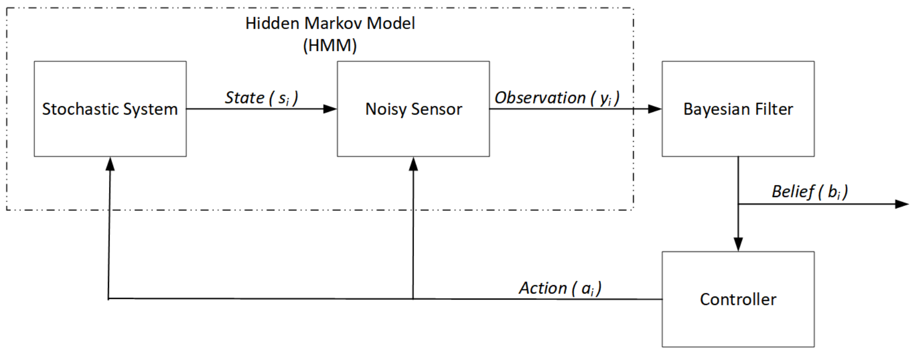

Figure 2.7: Block representation of POMDP components

2.5 Summary

This chapter presented a review into wireless sensor networks from terrestrial to Underwater Acoustic Sensor Networks (UASNs), Medium Access Control (MAC) protocols and Reinforcement Learning and Partially Observable Markov Decision Processes(POMDP). The chapter starts by introducing the fundamentals on wireless sensor Networks and UASNs, from background concepts, applications to deployment challenges. It follows with a review on MAC protocols, whereby elements of MAC designs considerations, and classifications, some specific example protocols from the literature and challenges were presented. While contention-based protocols currently dominates the UASNs MAC space, hybrid MAC are gaining traction in recognition of their versatility and better performance capability. In particular, because of its simplicity and intelligent collision avoidance and adaptive properties, ALOHA-Q has been single out as a candidate for adapting in UASNs and further studies. Finally, the chapter presents a sections on Reinforcement Learning, with the aim of introducing Q-learning as formulated in a cooperative domain similar to sensor networks. The section on POMDP framework gives an insight into a better approach of looking at multi-agent learning. The background knowledge on POMDP is contextualised in our proposed MAC algorithm in Section 5.4.

Chapter 3

System Modeling and Methodology

3.1 Introduction

This thesis present intelligent algorithms for UASN MAC protocols. Simulation is principally employed to empirically evaluate and demonstrate operation of the developed protocols. This chapter presents Riverbed modeler (formerly, OPNET) as the simulation environment used in this exercise. It follows with the methodology and the performance measures considered for the performance evaluation. Finally, the performance metrics used to demonstrate how the developed models have been validated are described based on known baseline results.

3.2 Simulation Environment And Protocol Development Using OPNET

Discrete Event Simulation (DES) is one of the popular paradigms developed to simulate the complex operations of discrete-event systems such as WSNs.

3.2.1 Riverbed Modeler

Riverbed Modeler is an efficient, complete and powerful industry-leading object-oriented network simulator and analyser [75]. Because of its versatility and accuracy in modeling, simulation and analysis the tool has been embraced by network operators, equipment manufacturers, civilian and military research institutions. As a compiled environment with parallel simulation capability, the Modeler offers relatively faster execution speed compared to other interpreted simulators. It is feature-packed with a substantial built-in library of ready to use, and fully editable models of established networking devices and protocols for fast modeling and design. Riverbed Modeler offers three interactive tools; graphical interface, run and debug, and the dynamic observer [76, 77]. Furthermore, the Modeler is currently capable of interfacing with other popular modeling software and with other programming languages such as Python through Riverbed's Open APIs for added portability.

Flexible and scalable system modelling is achieved through three distinctive domains: the process, the node and network. The object-oriented programming facilitates precise parameter definitions and settings of models. Although, the top-down approach (network-node-process) is typically followed during modelling, only the final harmonisation is essential. The components are described in a bottom-up approach as follows:

3.2.2 The Process Domain

This is where the underlying behavior of the modules present in the upper level node domain is designed and implemented. The process models are translated into PROTO-C, which is a specially developed programming language that incorporates and accepts C/C++ library and syntax [77]. The set of tasks a model is to perform defines the complexity of the model. Typically, individual tasks are represented by state objects. This approach mirrors a finite state machine structure and achieves event based modelling through the scheduling of interrupts and the defined transitions between states. Switching between states or procedures is event-driven when the system deliver events to the process models in the form of interrupts. During system run, the predefined sequences create events that are queued in an event list (ordered based on execution time and/or priority) and removed from the list for execution. In this regard, only one state/procedure is activated at a time. Simulation is successfully terminated when either the allocated simulation time has elapsed, or until the event list has been exhausted, or when a specified number of events have occurred. In both multi state and single states aggregation, task switching is invoked by distinctly associating the right interrupts with the task functions or code branches. These states (or functions and code branches) may be conditionally interlinked or otherwise. Figure 3.1 is an example of a multi state process along side a functionally equivalent condensed single state version. The states are colored with red state as unforced state representing and green states representing forced states.Whereas an event in needed to transitioned out of an unforced state, forced state transitioned out immediately after executing its routine.

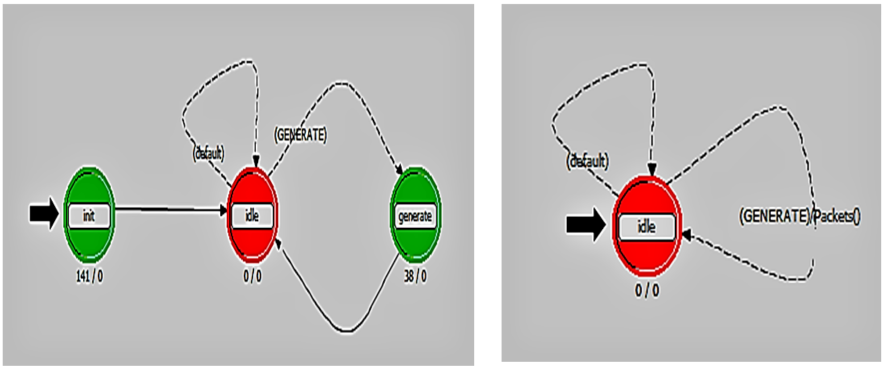

Figure 3.1: Process model of a packet generator: right: Multi-State, left: Single State.

3.2.3 The Node Domain

Defines and models the internal modules and connections of a node. These internal modules represent the different functionalities supported by the node, such as transmission/reception, storage, processing and interfacing. Therefore, for proper operation each module holds the corresponding process model implementation from the lower level. Processor, Queue and the External System(esys) modules are the three containers provided in the node domain to implement various functional entities. These modules can be linked by either packet streams for transporting packets or statistics wires for values/signals between modules. Esys modules are models that represent the behaviour of objects external to the modeler, in this way the modeler can interface with an external object to exchange data. Figure 3.2 depicts an example node model with four(4) modules. The Processor module (**GENERATOR**) models the packets generator which outputs a stream of packets to the Queue module (**MAC**) for storage, sorting in queues/sub-queues and processing. The transceiver modules (**TX** and **RX**) is represented by two separate wireless radio modules one each for the transmitter and receiver. Because our models are for acoustic media, some modifications have to be made to the default radio pipeline element of the Communication link domain, as discussed in Section 3.2.5.

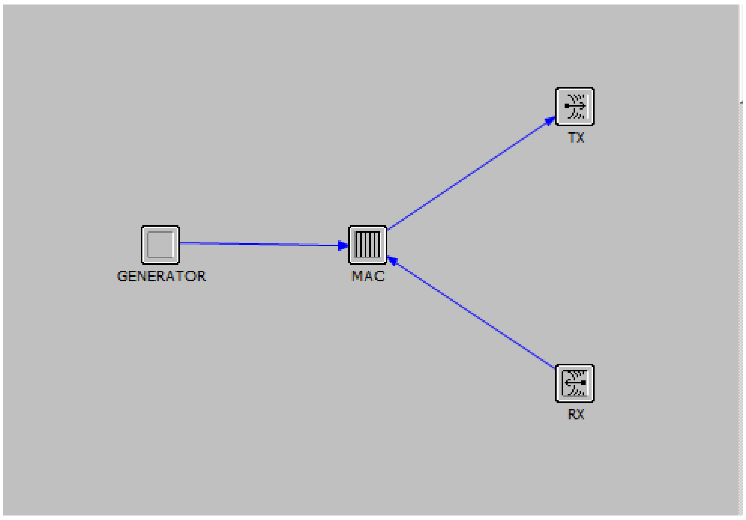

Figure 3.2: Node model of an acoustic transceiver

3.2.4 The Network Domain

The network domain is the top level description of the entities, their locations, links and configurations in the complete simulated system. Hence, the network model implement the familiar aspect of network objects and systems, such as nodes, subnetworks, topology, scenario and type of physical interconnections for conveying data/information between devices. Figure 3.3 shows an example of a linear topology of five(5) nodes in a wireless system. At this level nodes can be defined as static or mobile. Furthermore, non-network objects present in the physical environment (such as, buildings and machinery) and their interactions with the devices can be modelled to accurately represent the scenario. The modeler enables various simulations results to be defined and collected either network wide or on individual per node selections.

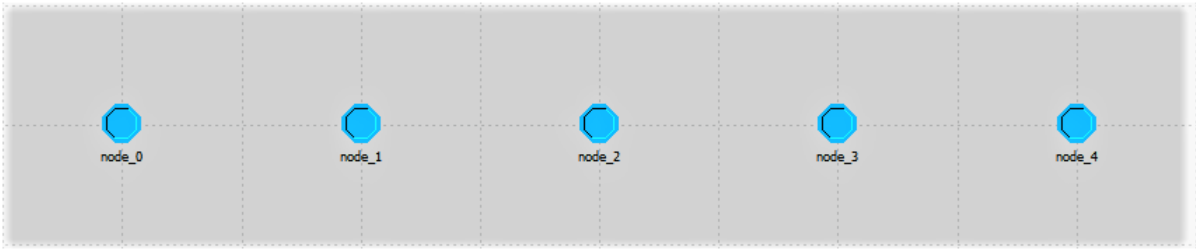

Figure 3.3: Network model of a linear chain network

3.2.5 Pipeline Stages

The communication link models support various types of standardized wired (point-to-point and bus) and wireless (radio broadcast) links [75]. Similar to the models in the above domains, the link models have an open architecture for modification by the developer to suit the application. The radio transceiver pipeline define a sequence of computational stages modelled to reflect specific behaviour involved in establishing and transferring data/signals in a radio link. Figure 3.4 represent the standard radio link computational stages in the Modeler.

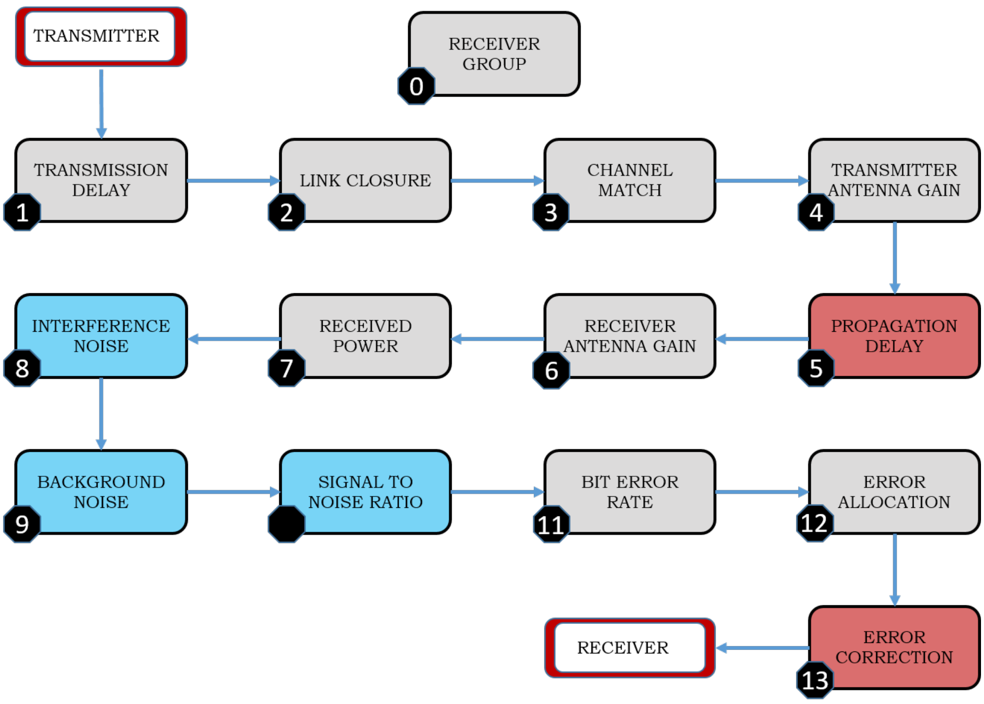

Figure 3.4: Wireless Transceiver Pipeline Stages

There are thirteen stages that are executed each time data is to be sent from a transmitter (six stages including the "receiver group" shown in Figure 3.5) to receiver (eight stages shown in Figure 3.6). The stages have access to the Transmission Data Attributes (TDA) to read/write in data for the relevant processing. To simulate the acoustic networks in this thesis, the propagation delay (stage 5) and error correction (stage 13) were identified as two relevant stages for modification. This setting primarily simulate the network with the average acoustic speed in water of 1500m/s [62] and the assumption of a collision model which ensures that any packet(s) that overlap at the receiver are dropped.

Further modifications were made to interference noise, background noise and signal to noise ratio (stages 8, 9 and 10 respectively) with data obtained from BELLHOP of simulated sea conditions to reflect practical conditions (See Appendix A).Specifically, the gain and noise outputs from BELLHOP are employed effect the modifications. The following sections provide description, functionality and where applicable the detailed modifications to the default pipeline stage as used in this work.

Transmitter Module Attributes

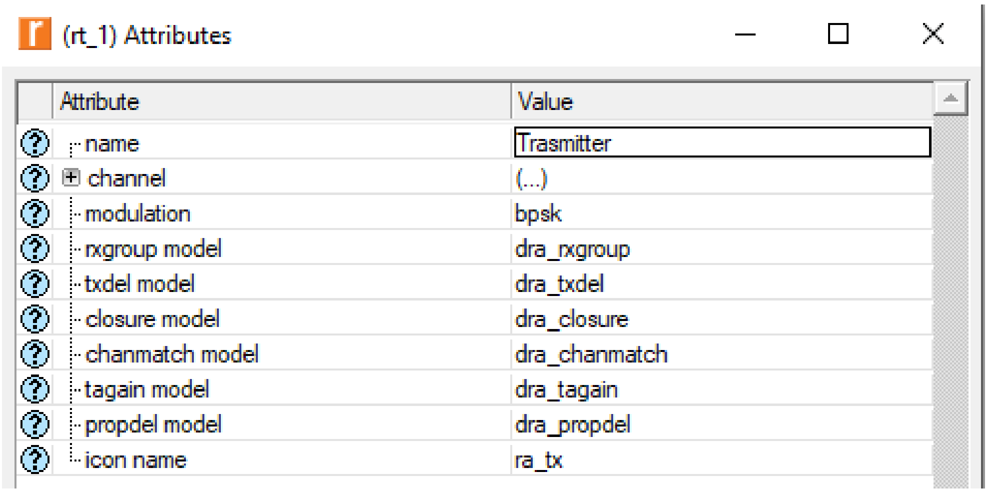

Figure 3.5: Default Radio Transmitter Module Attributes

1. **Receiver Group (Stage 0)** is the initial stage that and a non dynamic component of the radio transceiver pipeline. In order for the Simulation Kernel to models the broadcast behaviour of wireless (radio) each individual transmitter channel is linked with a set of

receiver channels. Therefore, the primary function of the receiver group stage is the computation and assemblage of a group of potential receiver channels that is then maintained by the transmitter channel for future transmissions. The "rxgroup model" attribute of the transmitter module implements the stage. The following are the source code sections of the default and the modified receiver group stage implementation used in this thesis.

- **Default:**

```
//////////////////////////////////////////

int rx_group_template (Objid tx_channel_objid, Objid rx_channel_objid)
{
int result;
FIN (rx_group_template (tx_channel_objid, rx_channel_objid))

FRET (result)
}
```

- **Modified:**

```
//////////////////////////////////////////

/*setting the receiver group for one receiver scenario*/

/* Extract this node (X, Y) position*/
op_ima_obj_attr_get(node_id,"x position",&x);
op_ima_obj_attr_get(node_id,"y position",&y);

/*Reseting the default receiver group*/
```

```
op_radio_txch_rxgroup_set(txch_id,0,OPC_NIL);

/*LOOP THROUGH ALL NODES IN THE NETWORK*/

for(i=0;i¡nodes;i++)
{
other_node_id=op_topo_child(subnet_id,OPC_OBJTYPE_NDFIX,i);
op_ima_obj_attr_get(other_node_id,”user id”,&other_user_id);
/* Extract neighbor node (X, Y) position*/
op_ima_obj_attr_get(other_node_id,”x position”,&sx);
op_ima_obj_attr_get(other_node_id,”y position”,&sy);
/* Compute distance between this node and other node*/

dist = sqrt(pow(x-sx,2)+pow(y-sy,2));

}

/* Add receiver channel of any node to this node’s receiver group if within interfer-
ence rage*/

if ((dist ≤ infx_range))
{
other_rx_id=op_topo_child(other_node_id,OPC_OBJTYPE_RARX,0);
other_rxcomp_id=op_topo_child(other_rx_id,OPC_OBJTYPE_COMP,0);
other_rxch_id = op_topo_child(other_rxcomp_id,OPC_OBJTYPE_RARXCH,0);
op_radio_txch_rxch_add(txch_id,other_rxch_id);

}
```

2. **Transmission Delay** is defined by the “txdel model” attribute of the transmitter module. The stage is executed at the beginning of every packet transmission. Given the invocation of this stage is handles the computation of the complete packet transmission duration from the packet’s first bit to the last bit, the resulting output from this stage is shared with all the remaining pipeline stages to support their operation. The default model has not been altered in this thesis.

3. **Link Closure** is defined by the “closure model” attribute of the transmitter module and its function is to compute the effect of any transmissions on a particular receiver. In essence, the closure model determines whether a transmission will affect a receiver channel, regardless of the transmission’s validity or viability. The default model has not been altered in this thesis.

4. **Channel Match** is specified by the “ channmatch model” attribute of the transmitter module. Its function is the classification of packet transmissions as one of three possible outcome for each receiver channel, namely; valid, noise or ignored. This model has not been altered in this thesis.

5. **Transmitter Antenna Gain** is defined by the “tagain model” attribute and its function is to computing the transmitter gain from the antenna using the resultant vector between the transmitter and the receiver. This output is employed in the calculations of the received power. This model has not been altered in this thesis.

6. **Propagation Delay** is specified by the “propdel model” attribute of the transmitter module. The primary purpose of this stage is the computation of the elapsed time for packet to transfer between the transmitter and the receiver. Therefore, the result of this stage is typically dictated by the transmitter-receiver separation. Primarily, the propagation speed of radio is replaced by the speed of sound underwater from the default radio “propdel model” attribute.

Receiver Module Attributes

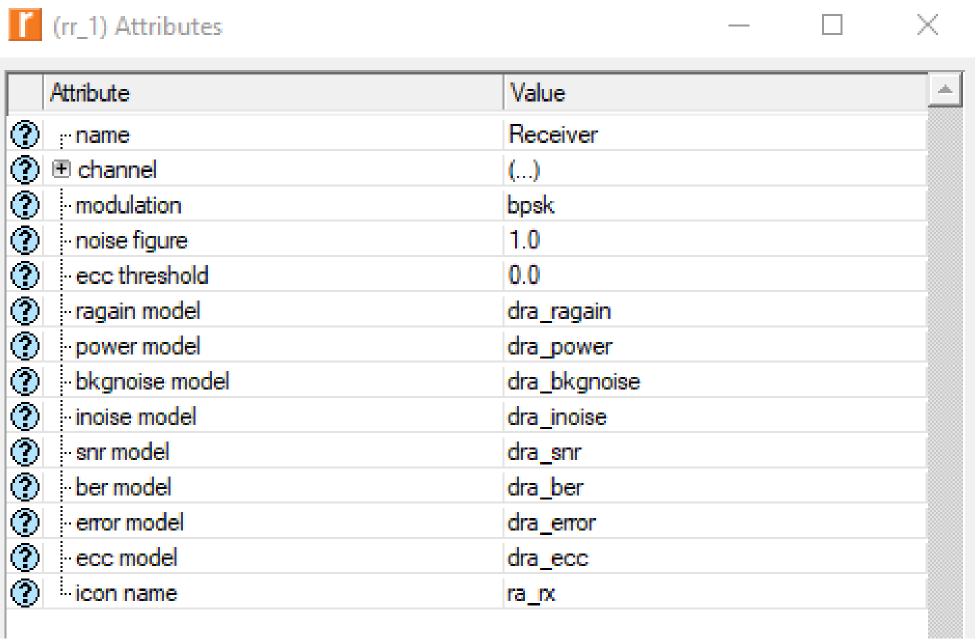

Figure 3.6: Default Radio Receiver Module Attributes

1. **Receiver Antenna Gain** is the first pipeline stage of the receiver module. The function of this stage is similar to the transmitter antenna gain, and is specified by the "ragain model" attribute of the receiver module. This model has not been altered in this thesis.

2. **Receiver Power** is defined by the "power model" of the receiver module, and its function is the computation of the received power on an incoming packet. The output of this stage is critical in further stages (snr and error allocation) for reading and differentiating valid packets from noise. The default source code and the modified sections of the receiver power used in this thesis are given below.

   - **Default:**

     ```
     /* Compute the amount of in-band transmitter power. */
     in_band_tx_power = tx_power * (band_max - band_min) / tx_bandwidth;

     /* Get antenna gains (raw form, not in dB). */
     tx_ant_gain = pow (10.0, op_td_get_dbl (pkptr, OPC_TDA_RA_TX_GAIN) / 10.0);
     ```

```
rx_ant_gain = pow (10.0, op_td_get_dbl (pkptr, OPC_TDA_RA_RX_GAIN) / 10.0);

/* Calculate received power level. */
rcvd_power = in_band_tx_power * tx_ant_gain * path_loss * rx_ant_gain;

/* Assign the received power level (in Watts) */
/* to the packet transmission data attribute. */
op_td_set_dbl (pkptr, OPC_TDA_RA_RCVD_POWER, rcvd_power);
```

- **Modified:**

```
//////////////////////////////////////////

//Get tx and rx node IDs

tx_node_id=op_topo_parent(op_td_get_int (pkptr,OPC_TDA_RA_TX_OBJID));
rx_user_id = op_topo_parent(op_td_get_int (pkptr, OPC_TDA_RA_RX_OBJID));

/*/Read the file bellhop file
gain_list_ptr = op_prg_gdf_read ("bellhop_tab");
lst_size = op_prg_list_size (gain_list_ptr);

/* Test for error in reading. ////
if (gain_list_ptr == OPC_NIL)
{
sprintf (err_str, "File Name: %s", "bellhop_tab");
}
else
{
```

```
//Initialize to below the noise level(81.5755db from bellhop in niger delta) -100db
tx_gain_db=-100;

//Skip the first tittle row

for(iter=1;
iter¡lst_size;
iter++)
{

field_list_ptr = op_prg_str_decomp (op_prg_list_access (gain_list_ptr, iter), " ,/");

//Extracting columns elements from the bellhop table(source, destination, gain)
tx_s = atoi (op_prg_list_access (field_list_ptr, 0));
rx_d = atoi (op_prg_list_access (field_list_ptr, 1));
gain_col = atof (op_prg_list_access (field_list_ptr, 2));

/*Extracting gain for the exact transmitter-receiver pair*/

if(tx_s != tx_node_id)

continue;

else if (rx_d == rx_user_id)
{

tx_gain_db = gain_col;
```

```
break;
}

}

}

/*Convert gain in dB to Watt*/

tx_gain_w = pow(10, tx_gain_db/10.0);

/*Computing the recieved power using the using the */

rcvd_power = tx_gain_w *tx_power*tx_ant_gain*rx_ant_gain;

////////////////////////////////////////////////////////////

/* Assign the received power level (in Watts) */
/* to the packet transmission data attribute. */

op_td_set_dbl (pkptr, OPC_TDA_RA_RCVD_POWER, rcvd_power);
```

3. **Interference Noise** is defined by the “*inoise model*” attribute of the receiver module, and its main function is to monitor the concurrent receptions activity at a given receiver channel. This model has not been altered in this thesis.

4. **Background Noise** functioned as the stage that accounts for all noise sources other than the incoming packets which are already monitored by the interference model, asnd it is

defined by the "bkgnoise model". The output of this stage is used in the calculations of the signal-to-noise ratio at the next stage. The noise from the Bellhop simulation is the sum total of the ambiance and background noises, hence it is used accordingly in the model as follows:

- **Default:**

  /* Put the sum of both noise sources in the packet transmission data attr.*/

  //op_td_set_dbl (pkptr, OPC_TDA_RA_BKGNOISE, (amb_noise + bkg_noise));

- **Modified:**

  ///Noise power in watts -81.2db from "Bight of Benin"

  bellnoise = pow(10,BELLHOPNOISE/10.0);

  op_td_set_dbl (pkptr, OPC_TDA_RA_BKGNOISE, bellnoise);

5. **Signal to Noise Ratio** is defined by the "snr model" attribute of the receiver module, and its main function is to used the outputs of the earlier pipeline stages such as received power and the noise sources to calculate the current average SNR power of an incoming packet. The output of this stage is crucial in the correct reception of packet's contents by the receiver. Because the ouptut of this stage is entirely dependent on the results from the preceding stages, this model has not been altered in this thesis.

6. **Bit Error Rate** is defined by the "ber model" attribute of the receiver module. The function of this stage is the computation of the expectation of bit error rate for the previous constant SNR readings. The output of this stage is also dependent on the type of modulation being used. This model has not been altered in this thesis.

7. **Error Allocation** is defined by the "error model" attribute of the receiver module. The

function of this stage is to used a constant bit error rate (from the output of the *ber model*) to estimate the number of errors in a given packet segment or its entirety. this model has not been altered in this thesis.

8. **Error Correction** is defined by the “ecc model” attribute of the receiver module. The function of this stage is to ascertain the acceptance or otherwise of a received valid packet. This acceptance criteria is computed from number of collisions, error rate experienced by the packet and the ability of the receiver to correct the bit error detected in the packet. However, the in this thesis, the primary error correction is implemented based on the number of collisions to reflect the collision model used. Below is the default code section of the “ecc model” with the only corresponding modifications made to the default “ecc model” source file.

   - **Default:**

```
////////////////////////////////////////////

/* Obtain number of errors in packet. */

num_errs = op_td_get_int (pkptr, OPC_TDA_RA_NUM_ERRORS);

/* Test if bit errors exceed threshold. */
if (pklen == 0)
accept = OPC_TRUE;
else
accept = ((((double) num_errs) / pklen) ≤ ecc_thresh) ? OPC_TRUE : OPC_FALSE;
```

   - **Modified:**

```
////////////////////////////////////////////
```

```
/* Obtain number of collisions experienced by packet. */

num_colls = op_td_get_int (pkptr, OPC_TDA_RA_NUM_COLLS);

/* Test if number of collisions is at least 1. */
if (pklen == 0)
accept = OPC_TRUE;
else
accept = (int) num_colls > 0 ? OPC_FALSE : OPC_TRUE;
```

3.2.6 Traffic Modelling

The traffic generator models the distribution of packet generation that load the network. A saturated model considers that there is always at least one packet ready to send in the queue, whilst Poisson traffic offers variable load with an exponential distribution. The Erlang is the unit of offered traffic in a fixed capacity network. Its value ranges from zero (0) to unity (1), respectively equating to unloaded and fully loaded channel in any given time.

Saturated Traffic

This traffic is modelled to ensure constant generation of packets by source nodes. It is employed to demonstrate the network ability to handle cope when subjected to constant stream of packets. Therefore, by maintaining a non-empty queue (at least one packet in the queue) at all times source nodes schedule and transmit packet during each transmission cycle. A saturated model maximises monitoring rates as a new measurement is sent whenever the opportunity arises. However, when an event occurred and the sensors are triggered, delivery of data may then be restricted by the MAC layer, ultimately, the network cannot offer better performance than achievable with a saturated model. Furthermore, it is useful to ascertain the resilience of the network when a critical event loads the system. Hence, we use the saturated traffic principally in evaluating our proposed protocol.

Poisson Traffic

Poisson traffic model is originally developed to model traffic in the telephone networks [78]. However, the model is now well established and extensively used traffic generator that models events with varying magnitudes and duration in sensor networks [79, 59]. As a Poisson process, it is described by the Poisson distribution with mean arrival rate of events translated to the packet inter-arrival times. For a given node density in a neighborhood N, bit rate B bps, offered load G Erlang, and Packet size P_d bits, the mean inter-arrival time λ of the Poisson process is described by:

$$\lambda = \frac{P_d N}{GB} \tag{3.1}$$

The packets generation event can be scheduled in the Modeler from the output of the exponential distribution function: *op_dist_exponential(λ)*. Poisson traffic is demonstrated in Section 4.2.

3.3 Performance Measures

The following are some of the typical system parameters and performance metrics considered in the evaluation and design of the protocols in this work.

- **Utilisation** measured in Erlangs, it describes the percentage of time useful data is successfully received at the designated sink node. For example, a measured utilisation of 0.5 Erlangs means, the network has delivered data 50% of the time to the sink node. A number of factors, such as communication overheads, network topology and the interference model determine the achievable utilisation. For a given a number of packets (N_p) recieved at the sink node in the simulation time (T), The Utilisation (U) is computed using:

$$U(Erlang) = \frac{N_p P_d}{TB} \tag{3.2}$$

Where: P_d,and B denote the data packet length and the datarate respectively.

- **System Channel Capacity** is the quoted datarate of the system in bits per second (bps). Therefore, it defines the maximum amount of data the acoustic channel medium can accommodate, typically in the kbps range.
- **Offered Load** is the average amount of active data traffic placed on the channel. The offered traffic comprehensively covers both the freshly generated data plus any re-transmitted data as a result previously unsuccessful transmissions. differs with the generated traffic, because, some packets could be re-transmitted due to failed transmission attempts at high contention. Similar to the utilisation, a 1 Erlang traffic corresponds to a fully engaged channel.
- **End-to-end Delay** End-to-End (E2E) delay is one of the most important metrics in UASNs performance, especially with the long propagation delay involved underwater. The E2E delay gives packets' latency as they traverse the network. Generally, low E2E delay is desirable, however, some non critical applications may prioritise other metrics such as energy efficiency and utilisation.
- **System Convergence Time** Learning is a gradual process that requires an exploratory phase prior to finding an optimal solution by the algorithm. This average time whereby the initial learning completes is an important metric that affects the effectiveness of the algorithm, when convergence is achievable. Whilst some algorithms stop further searches after finding an initial solution, others continually look and update the solution dynamically.In the later case, the system convergence time is useful in determining the average time the algorithm makes no discernible improvements to the initial solution. Traditionally, in Q-learning MAC protocols, the solution refers to when all nodes find and occupy

a unique collision-free transmission slots. However, the solution herein, refers to the optimum channel utilisation, since, unique transmission slots is not a prerequisite for a collision-free transmission in UASNs due to the pronounced spatio-temporal uncertainty (see Section 2.2.3) and the nature of our proposed slot structure (see Fig. 4.3).

To adequately represent the system performance, the time to start collecting the network data for the performance metrics is usually stated. In this thesis, the developed algorithms never settle, as they are continuously adapting to the dynamic underwater acoustic environment. Hence, the initialisation stage is included in the data collection. In the beginning this may negatively affect the overall performance, however, this effect is cancelled out when the network is allowed to run for a sufficient time. The system models, simulation environments and the results obtained are predicated upon the following assumptions:

- All nodes are homogeneous in a network.
- The network layer handles packet routing based on Dijkstra's shortest path algorithm.
- In synchronous operations, all nodes are time synchronised across the network.
- All nodes operate on a half duplex mode.
- Average underwater acoustic propagation speed is 1500m/s.

3.4 System Models and Result Validations

In order to validate the simulated networks, firstly, the correctness of the implemented system models need to verified. This is achieved by matching the simulation results of the network running an established MAC protocol, namely the popular Aloha protocol[51] in both radio and acoustic medium. Secondly, the simulated results of the developed protocols are then validated through comparison with the results of some analytical models. Simulation parameters are given in Table 4.3 The raw data is exported into MATLAB for plotting and visual representation.

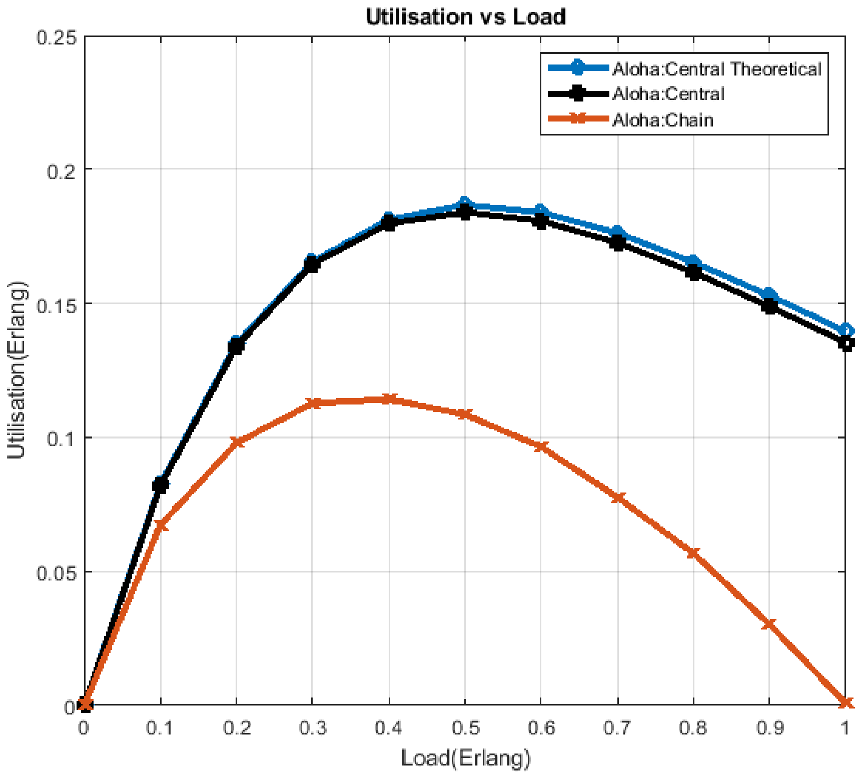

Figure 3.7: Model Validations: Aloha Protocol in Radio and Sound Channels

3.5 Comparison Schemes

Framed ALOHA, and its reinforcement learning powered variant : Aloha-Q modified with underwater parameters were chosen for comparison with our proposed protocol in the later chapters(see Section 5.4). Whilst ALOHA suffers from low utilisation and notoriously unstable at high loads, its simplicity and benefits of been extensively studied is still highly regarded and it continues to provide a solid foundation for developing complex/sophisticated schemes. ALOHA-Q is one of such intelligent variant developed for WSNs, as the concept it employs in part inspire this work, we feel it provides a reasonable baseline for comparison.

3.6 Summary

In this chapter, we present the need for simulation in UASNs development and the powerful software tool we employ in this endeavor. An expanded overview on the Modeler's systematic modelling approach with the three main hierarchical modelling domains, namely the process, node and the network domains was also presented. It follows with the metrics used in performance evaluation of the developed MAC protocols based on the simulated system wide assumptions. Finally, the validation process and techniques are also discussed.

Chapter 4

Reinforcement Learning in Underwater Acoustic Sensor Networks

4.1 Introduction

The previous chapter presented pre-requisite tools, methodology and models of the network, sub-systems and basic process entities employed in execution and evaluation of the developed networks and protocols in this thesis. Generally, the target application has the specifications that define the scope and hence guide the choice of an appropriate scenario for network implementation and evaluation. The scenarios are typically chosen between single-hop and multi-hop (refer 2.2.5). The use of a single-hops for networks is appropriate for short-range communication networks and this scope limits its usage in UASNs. Underwater applications are largely envisaged to be implemented using multi-hop scenarios, since UASNs typically cover a substantial marine area. There are several advantages of using multi-hop networks, such as large scale coverage, and increased connectivity for efficient and robust communication routing as nodes can discover alternative and better or new routes. Underwater pollution monitoring, marine bio-diversity, rescue missions, tsunami/disaster early warning, and offshore infrastructure monitoring are some of the examples applications that can benefit from multi-hop However, multi-hops network topologies have characteristic transmissions and interference patterns that increase the challenges and complexity of developing network protocols, especially underwater MAC protocols, whereby low capacity, long and variable propagation delay, high dynamic nature of the acoustic channel demand MAC schemes that are highly efficient, robust and adaptive for optimum performance and best Quality of Service delivery. Topology consideration is an important natural next step prior to network deployment. This is because the topology represents the applications network deployment by translating the scenario to an approximate spatio-temporal positional layout and orientation of components and network integration.

In this chapter, the original Aloha-Q is presented. The purpose is to adapt the Q-learning MAC scheduling onto a multi-hop linear chain underwater network. Aloha-Q was initially developed for terrestrial radio WSNs, therefore to extend its utility underwater, its performance is evaluated in both synchronous and asynchronous operation modes. The evaluation study gives us an insight on how to introduce some modifications to achieve better performance in terms of both utilisation and fairness.

4.2 ALOHA WITH Q-LEARNING

This section presents the background and working principles of Aloha-Q protocol. The protocol performance in terms of the channel is evaluated and compared in both the terrestrial and underwater acoustic network environments.

4.2.1 Operational Principles

Aloha-Q (introduced in Section 2.3.3) employs Q-learning as a reinforcement learning technique with the goal of optimizing and learning distinct transmission schedule/slots by nodes that are initialized with framed slotted Aloha. To ensure reliable communication between any two nodes, provisions for an acknowledgement (ACK) response is typically made following a successful reception. The ACK signal is translated in ALOHA-Q to provide the essential reward parameter (Equation (4.1)) for the scheme to work. Therefore, a slot must be structured to be wide enough to accommodate at least: the transmission time of a data packet, an acknowledgement packet, and propagation delays. A guard interval is also incorporated into the slot size such that a timeout period is defined within the slot boundary whereby acknowledgement packets not received before the timeout will incur a packet retransmission in the next transmission cycle. Scheduling is achieved by grouping a contiguous pre-specified number of slots into a repeating block (Fig. 4.6); the frame.

In addition to having an appropriate slot size, the frame has to be composed with the optimum number of slots as well, in order for the for the nodes to find and occupy unique transmission slots. A vector of values is maintained representing the Q values of the slots per frame, whereby each slot is assigned a Q-value which is updated according to the stateless Q-function rule (4.1). At the beginning of each frame, nodes will choose the slot that corresponds to the highest Q value and transmit in that slot provided there is a packet in the queue to send. However if there is more than one highest Q value, one will be chosen at random amongst them and the node will transmit in that slot. The reward is given as the outcome of the transmission: successful transmission earns $(+1)$ and failed transmission gets punished (-1). Therefore, rule 4.1 is continuously updated each time a node receives an acknowledgement packet or a timeout occurs. Initially, nodes will be competing for transmission slots until eventually every node manages to occupy a distinct slot when the protocol converges. Convergence means, as nodes find unique slots the Q-values of each selected slots will continue to rise and approach one (1) and Q-values of the unselected slots either remain unaffected from the initialized values or decrease and approach zero (0). However, convergence may not always be possible if the environment is constantly changing (as is the case underwater), nevertheless, as online learning has the ability to track and adapt to the changing environment, using the Q learning algorithm offers a lot of benefits as demonstrated in Aloha-Q [69, 59] and subsequently in this work. Therefore, employing Aloha-Q on sensor nodes for an underwater pipeline monitoring network can provide the benefits of TDMA like scheduling, improved channel performance without the constraints of central controller, precise knowledge of the environment, and also adaptability to changing conditions and robustness against changes in network topology due to nodes removal and/or addition. The complete Aloha-Q algorithm is given in Algorithm 1.

$$Q_i = (1-\alpha)Q_i + \alpha r \tag{4.1}$$

Where: Q_i is the Q value of i^{th} slot, α is the learning rate and r is the reward/punishment.

Algorithm 1: ALOHA-Q algorithm.

```
Initialization;
Q_values, learningrate, and reward;
while node is online do
    // Previous transmission outcome;
    if ACK then
        // Rewarded;
        reward ← +1;
    else
        // Punish;
        reward ← −1;
    end
    // Update Q value of transmission slot;
    Q_i ← Q_i + learningrate(reward − Q_i);
    // Next transmission slot selection;
    next slot ← [i | i ∋ argmax_{i∈𝒮} Q_i];
    //Resetting the reward;
    reward ← 0;
end
```

Q-Learning Update Example

This section demonstrates the underlying Q-learning update procedure as employed in the original Aloha-Q protocol. To develop a MAC protocol, this is translated to a node taking the action of transmitting the data packet, and the successful/unsuccessful reception of an ACK packet represents the reward/punish signal. Each node is given a vector of Q-values (Q-table), and each Q-value is in turn assigned to one slot in the frame. At the beginning of each frame, a node will scan the Q-table and select the slot with the highest Q-value to schedule transmission in that slot. Successful transmissions are rewarded and unsuccessful transmissions punished based on the reception or otherwise of an ACK packet and updating the Q-value of the transmission slot using Equation (4.1).

Table 5.3 illustrates an example of the Q-learning as implemented in ALOHA-Q. Consider an initial situation (Frame 0) whereby a node $\boldsymbol{i}$ with data to send randomly chooses Slot 2 (because all slots have equal Q-values) at the beginning of a frame to schedule transmission and the transmission was unsuccessful.

- The new Q-value of Slot 2 becomes;
  $Q_2 \leftarrow 0 + 0.1(-1-0)$; updated to -0.1
- In the next frame, Slot 2 has the lowest Q-value and is not considered, and the node again chooses Slot 1 randomly (among Slots 0, 1 and 3). Following a successful ACK reception, the new Q-value of Slot 1 is updated.
  $Q_1 \leftarrow 0 + 0.1(+1-0)$; updated to 0.1
- For Frame 2, the node chooses Slot 1 as it has the highest Q-value (0.1) and sends data; with successful ACK reception, the Q-value is updated accordingly.
  $Q_1 \leftarrow 0.1 + 0.1(+1-0.1)$; updated to 0.19

The table gives the Q-values up to twenty frames assuming Slot 1 continues to be successful. This simple, yet effective recursive Q-learning update bootstraps the trial-and-error mechanism to a robust collision-free schedule as each node will eventually and independently occupy a unique transmission slot. For the purpose of implementation in the simulation environment, a 1-D array is generated to store the Q-values, where the Q index represent the slot number, the following notation is employed accordingly: $Q[i]$ represents/returns the Q value of i^{th} slot. However, while the ACK signal is crucial to the Q-value update operation, owing the long

Table 4.1: Example of Q-value update in Aloha-Q

Frame/Q-values	$Q[0]$	Q[1]	Q[2]	Q[3]
FRAME 0	0	0	0	0
FRAME 1	0	0	-0.1	0
FRAME 2	0	0.1	-0.1	0
FRAME 3	0	0.1900	-0.1	0
FRAME 4	0	0.2710	-0.1	0
...	...	...	...	...
FRAME 20	...	0.8499	-0.1	0

$slot_0$	$slot_1$	...	$slot_n$	$slot_0$	$slot_1$	...	$slot_n$	$slot_1$	$slot_2$	...
FRAME 0				FRAME 1				FRAME 2		

Figure 4.1: Frame/Slot Structure

propagation delay, it extends the frame duration to a level that puts an additional burden on the scarce network resources underwater with the downside of: reduced utilisation due to overheads and increased delay due to the ACK signal wait times. one of our goals is to address the reduced performance by devising a scheme that exploit/utilise the wide underutilised frame gap to improve the channel utilisation and end-to-end delay.

4.3 Modeling Aloha-Q in a Linear Chain Underwater Network

This section presents the scenario and parameters employed in the evaluation of the Aloha-Q protocol by adaptation and modeling in the underwater network. The adaptation and evaluation of Aloha-Q is performed on a linear chain multi-hop topology.

4.3.1 Linear Chain

A multi-hop chain network can be applied to several underwater applications by placing sensor nodes along a target general trajectory (in this case linearly spread nodes), or an established multi-hop route through a distributed network (randomly spread nodes). Fig. 4.2 depicts an example 10-hop linear chain network, whereby the source and the destination nodes are at the opposite ends and data is routed hop-by-hop along the chain via the relay nodes. While the straight line and the equidistant regularity of node positioning may seems an idealistic simplification given that nodes are typically deployed in a random manner to cover the site, such as in ad-hoc networks for in disaster monitoring. Multi-hop linear chain can be found in oil/gas pipeline monitoring networks. Notwithstanding, the more complex interference signature of randomly distributed networks, a linear chain route abstraction can typically be established from the source to the destination, as such linear chain networks are popularly used for analysis and evaluations of scenarios and protocols [74].

In Fig. 4.2, **rX** and **iFX** illustrate the reception and the interference ranges of a transmitting node (node 5 in this illustration) respectively. The protocol model is used to characterise the interference relationship of the network. As defined in the protocol model, a transmission is successful if the receiving nodes fall within **rX** of the intended transmitter and outside **iFX** of any non-intended transmitters [80]. This simplicity of the protocol model relative to the considered complex reference SINR model made the protocol model widely used by researchers in characterising the behaviour of wireless interference for developing network protocols [81]. Since accurate definition of the communication and interference ranges (i.e **rX** and **iFX**) in real applications especially in underwater may not be feasible, therefore we follow the convention of defining the ranges in terms of hops. This boundary assumption in terms of hops is reasonable in idealised simulation conditions for a network with homogeneous wireless nodes having identical hardware features and specifications.

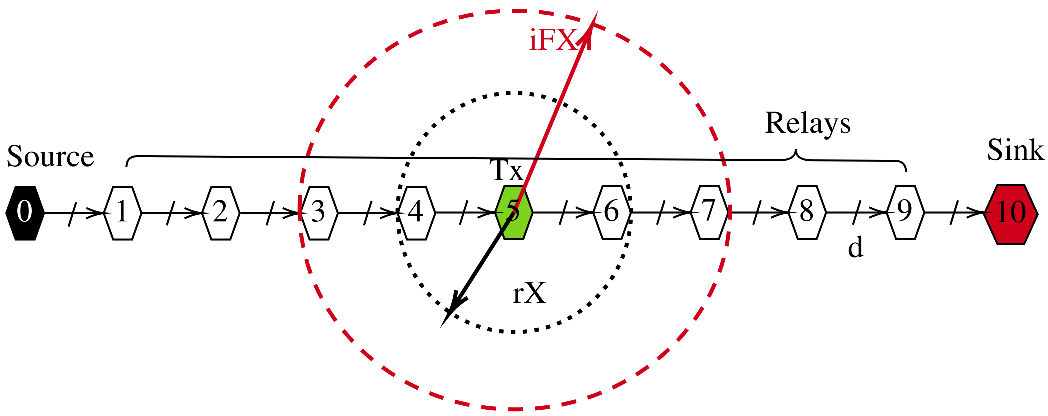

Figure 4.2: A simple 10-Hops Linear Chain Network, showing both the reception and interference ranges.

The interference characterisation is of importance in creating time slots for TDMA like scheduling in multi-hop networks whereby the structure of the time slot is defined by taking the effect of interference into consideration with the goal of establishing collision free schedule. Moreover, spatial reuse of the same time slot(s) by multiple nodes for concurrent transmissions

Table 4.2: Theoretical Throughput of Interference Ranges

Interference Range	Unit Population	Theoretical Throughput (Erlangs)
1-hop	3	0.333
2-hops	4	0.250
3-hops	5	0.200

can be achieved when the physical separation between the nodes is sufficiently large to remove interference. Referring to Fig. 4.2, nodes can communicate with each other within **rX** 1 hop away and potentially interfere with nodes **iFx** 2-hop away, as such we can derive the theoretical maximum/achievable throughput of the system. Consider the situation whereby node 5 is transmitting packet to the downstream node 6, for a successful reception the none of the three nodes (4, 7 and 8) should transmit during that period, because as per the model their transmissions will interfere and cause collision(s) at node 6. Furthermore, for a successful reception node 6 itself must not be transmitting since, the nodes cannot transmit and receive at the same time. Thus, in a given interference neighborhood only 1 in 4 nodes is guaranteed to successfully transmits data. This translates to a theoretical throughput of 0.25 Erlangs. Varying the interference ranges to cover different number of hops also varies the theoretical throughput. For example when the interference range covers 1-hop, then 1 in 3 nodes can transmit and hence with achievable throughput of 0.33 Erlangs. Table 4.2 summarises three examples for completeness.

Therefore, for a sufficiently long chain so the interference model is applicable, we can create a periodic schedule (frame) that assign distinctive time slots to nodes in a given unit of interference locality and systematically be reused in the subsequent units in such a fashion that the overall network schedule is collision free and achieves the optimum throughput.

4.3.2 Optimum Frame Length

The determination and pre-allocation of the appropriate number of slots per frame is critical to the optimum channel utilization [59, 69]. For every interference model there is an optimum number of slots necessary to guarantee the best performance. Both over-allocating and under-allocating the number of slots per frame against the number of nodes in a network will have a degrading effect on performance. Based on the assumptions of treating each individual packets originating from different sources along the chain as a separate flow, [59] postulated the generalised frame duration F_D for each node with Eq. (4.2).

$$F_D \geq F_{Dmin} = \begin{cases} ((S_j + S_{j-1} + \sum_{i=0}^{iFx-1} S_{j-2-i})S_L, & N \geq 2. \\ S_L, & N = 1. \end{cases} \tag{4.2}$$

where: S_j, S_L, and N denotes the jth source node, slot duration, and the number of nodes along the chain respectively.

However, by marshalling the understanding from Section 4.2.1 and Section 4.3.1, the F_{Dmin} is principally dictated by the interference range (iFx). That is, the interference range (hops) will provide the minimum necessary scaling factor to the slot duration that guarantee a particular node will have a successful transmissions. Furthermore, since the Aloha-Q algorithm is inherently designed to choose and transmit only one packet per frame, regardless of the number of sources along the chain, it puts an upper limit on the amount of data packets that gets delivered to the sink per frame. The system throughput (utilisation) is defined as the ratio of delivered data against time, which can be then expressed in terms of total received packet duration per frame duration.

$$Utilization = G\frac{P_d}{F_D} \tag{4.3}$$

where: P_d, and G denote the packet duration (sec) and the normalized offered load in Erlangs (i.e percentage of the total time the network is actively loaded by each source). From 4.3 the lower the value of F_D the better for the utilisation, however, allocating lower value below what is allowable per interference range (fewer slots for competing nodes) will potentially cause collisions especially with increase in traffic loads, in the same token, large value F_D in excess of the optimum will allocate extra/unused slots that ultimately degrade the throughput. For example, referring to Table 4.2 when neighboring nodes can only interfere with their closest neighbour (1-hop) the ideal optimum frame duration (F_D) must be three slots wide as only 1-in-3 nodes can successfully transmit, similarly for 2-hop interference range F_D should be four slots wide.

4.3.3 Modes of Operation

This section presents an investigation on the performance Aloha-Q in a simulated asynchronous operation underwater, for the purpose of gaining an insight with respect to performance by evaluating the protocol whereby network wide synchronisation may not be feasible.

Synchronicity

The original Aloha-Q assumes all nodes in the network to be synchronised, however, at the current level of technological advancement this is not always practical in underwater networks due to the lack of: GPS like positional capability, high precision clocks on nodes, and the high dynamics of the environment. Therefore, to fairly reflect a potential out of sync situation occurring underwater, additional simulation was conducted with the nodes initialized asynchronously with a uniformly distributed random frame start times as depicted in Fig 4.3. The network is in synchronous operation when each node initialises with the global reference τ (i.e 0 or any other starting value), and asynchronous when nodes individually choose τ randomly. The consequence of initializing the frames asynchronously means packets will arrive at the receiver out of sync, thereby increasing the risk of collisions in the network. However, in the case of underwater acoustic communication, due to the substantial portion of the slot occupied by the propagation delay, even packets transmitted in the same numbered slot(s) could potentially be received at the receiving node without collision.

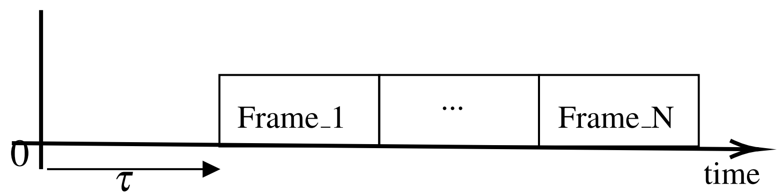

Figure 4.3: Frame initialisation

Multi source

The protocol is further simulated with all nodes acting as sources nodes, this scenario can be found on a flowline connecting a subsea wellhead to the base of a vertical riser to an offshore platform is retrofitted with acoustic sensor nodes for monitoring operations, such as of free span sections along the flowline in a multimodal fashion. Since, the setup is for monitoring purposes every node but the sink is considered a source node and can relay packets from upstream sources along the chain. To achieve this multi source scenario, we propose a modification to the original Aloha-Q by incorporating a systematic round robin update in the queue of each node. This is necessary given that Aloha-Q is only extensively evaluated for a single source in chain network and thus, unfair with the First In First Out (FIFO) queuing in a multisource chain network, especially in high loads. Therefore, as an improvement to make the protocol suitable for a pipeline monitoring in a chain network where multiple nodes could potentially be both sources and relays we replace the FIFO queuing with Round Robin (RR) queuing [82]. The problem with the FIFO queuing is that, the upstream sources will be starved by the downstream nodes acting as both sources and relays, therefore nodes closer to the sink will capture the channel and degrades further as the network load increases. Figure 4.4 illustrates the insertion of the flow de-multiplexer and an RR scheduler on the existing FIFO mechanism to implement the fair queuing block. Therefore, sources/relay nodes will separate and insert incoming packets from different sources into the corresponding flows and so that packets are forwarded by the round robin scheduler. Treating packets originating from different sources as separate flows and taking turns in forwarding packets in a round robin fashion all packets from all sources are treated fairly. Furthermore, when the flow queue for the next packet is empty, the RR scheduler maintains efficiency by skipping flow queue(s) that are empty and schedules transmission from the next available queue.

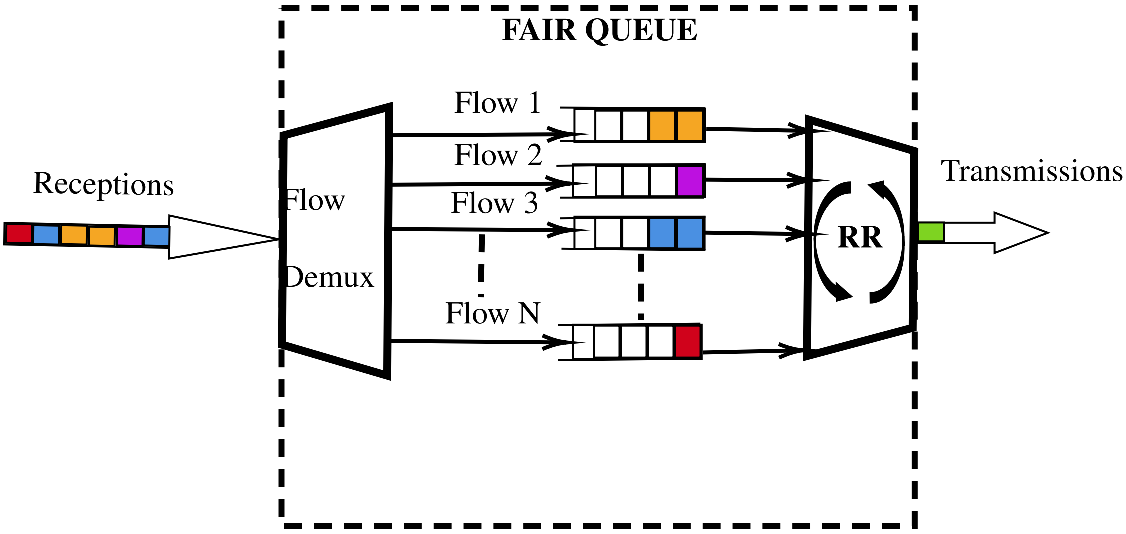

Figure 4.4: Fair Queue by FIFO to RR Queue Modification

Jain's fairness index [83] is a performance metric used to show fairness in network with multiple flows in a shared communication network.

$$J = \frac{(\sum_{i=1}^{n} x_i)^2}{n * (\sum_{i}^{n} x_i^2)} \tag{4.4}$$

where: x_i and n are the recorded utilization of i^{th} source at the sink and the number of source nodes in the network respectively.

4.3.4 Network Scenario

Basic Scenario and Assumptions

This section presents the topology setup and basic parametric configurations for the performance evaluation of the original Aloha-Q as simulated in the underwater environment. Fig. 4.5 depicts the simulated scenario for a chain network consisting of nine (9) nodes along a chain network of 1600m length with inter-hop distance of 200m. This type of scenario is applicable to a pipeline monitoring network. Data flows from the source node(s) and is relayed hop-by-hop along the path to the sink node *node_8*. The size of Data packets and Acknowledgement packets are fixed and the optimum data route is assumed to be established in the network layer using the shortest path algorithm. In terms of network connectivity, in this work, nodes transmit and receive on the same channel therefore with half duplex operation within 1-hop reception range and can interfere with nodes within 2-hop distance. Network wide synchronisation is established and maintained unless otherwise stated (such as for asynchronous mode), transmitting nodes can only transmit one packet per frame and finally all internal processing delays are ignored. Finally, to avoid accepting overlapped packets during reception a protocol model is used.

Figure 4.5: Simulated Scenario

4.3.5 Traffic Models

In this chapter both Poisson arrival and saturated traffic were employed in the simulations. While our eventual objective is focus on Saturated traffic (refer Section 3.2.6), we include the Poisson traffic model because a major contribution of this chapter is to explore the feasibility of extending and adapting the original Aloha-Q underwater (this entails validation and comparison with the original Aloha-Q protocol model) to which the Poisson traffic is an important validation component [59]. For the rest of this thesis however, we primarily use the saturated traffic, whereby new packet is generated whenever a source node is ready to transmit, this is particularly suited to real-time monitoring applications such as in oil and gas pipeline monitoring whereby in the event of leaks or other compromises in structural/operational integrity, timely and robust relaying of data is critical for accurate analysis and reaction to the problem.

Table 4.3: Simulations Parameters

Parameter	**Radio Channel**	**Acoustic Channel**
Transmission/Reception Data Rate	25000bps	640bps
Data Packet Size	1060bits	512bps
Acknowledgement Packet Size	16bits	16bits
Slot Size	1100bits	710bits
Slots per frame	4	4
Reception Range	200m	200m
Interference Range	400m	400m
Propagation Speed	3E8m/s	1500m/s

4.3.6 Simulation Parameters

The overall simulation parameters are given in Table 4.3. The acoustic channel utilises the quoted parameters of the Newcastle University's nano modem [84] and the radio channel parameters are from the original Aloha-Q. These parameters values are employed in all simulation in this chapter. Whilst the propagation speed in the radio channel is negligible, the impact of the slow propagation speed on the slot size of the acoustic channel can be seen on the corresponding calculated value of the acoustic slot size.

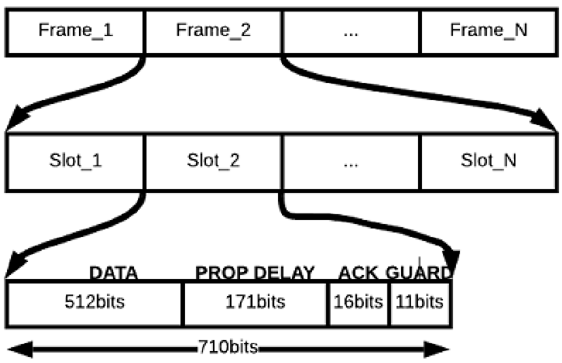

Figure 4.6: Acoustic Frame/Slot Structure

4.3.7 Performance Evaluation and Results

The simulation was run for 15000 frames each for both radio and acoustic channels. Packets are generated according to Poisson arrivals and saturated load, nodes are allowed to send only one packet per frame. The statistical data is collated from the beginning of the simulation, primarily because we are interested in the overall performance including the learning phase as in reality the network will be deployed long enough for the effect of the learning process to be neutralised. To evaluate the effect of frame size, the overall network throughput is plotted against the frame duration for each operation mode run.

Fig. 4.7 compares the performance of pure Aloha and Aloha-Q in both radio and acoustic channels as measured at the sink node. As can be seen from Fig. 4.7, Aloha-Q manages to achieve 0.24 Erlangs in a terrestrial chain network and it only manages 0.18 Erlangs in an underwater network. However, Aloha-Q outperforms pure Aloha because of the blind transmission strategy in pure Aloha, thereby only achieving maximum utilization of 0.11 Erlangs. The computed results in the simulation based on (4.5) is consistent with the expected results of Equation (4.3) and illustrate the impact of the overhead due to the long propagation delay on the channel utilisation in the underwater environment. Overall, Q-learning helps elevate and stabilise Aloha to achieve near perfect channel utilization in a radio channel and offers similar performance in an acoustic channel.

$$Utilization = \frac{Total\ packets\ received \times packet\ size}{data\ rate \times simulation\ time} \tag{4.5}$$

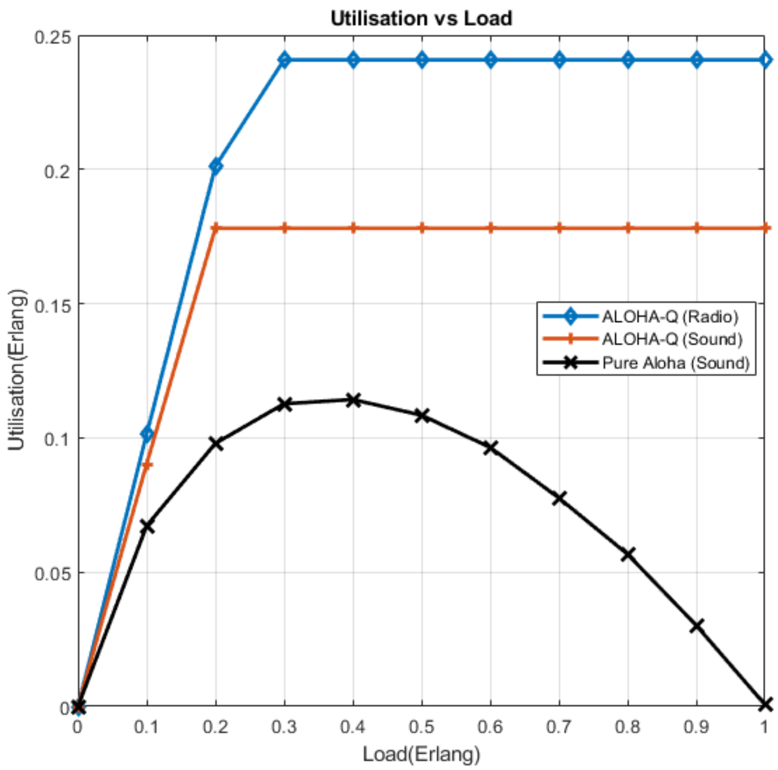

Figure 4.7: Underwater Performance

Fig. 4.8 compares the maximum channel utilization against number of slots per frame from simulating Aloha-Q at full load (G=1). The Aloha-Q ideal plot is based on Equation (4.3) and is the result without any interference taken into account with the expected inverse relation between maximum utilization and the number of slots per frame. The Sync:Aloha-Q curve shows how the interference model affects the protocol to under-perform versus the ideal case. For example 2-slots (0.36 Erlang versus 0.3 Erlang). Simulating Aloha-Q asynchronously gives the Async:Aloha-Q plot and shows how the protocol performance degrades and swings significantly to almost zero for 2 slots and 3 slots (0.16 Erlang). Overall, however, the protocol pivots at the optimum 4-slots (0.18 Erlang) and remains unaffected by the interference model and the lack of synchronisation as the number of slots increases due the sufficient number of slots per frame and slot size. This is a significant result that highlights feasibility of exploiting the long propagation delay to achieve good performance whilst violating the recommended F_D. Due to the intelligent scheduling strategy, the asynchronous Aloha-Q protocol still outperforms both synchronous and asynchronous framed Aloha on the pipeline network except for the 2-slots case.

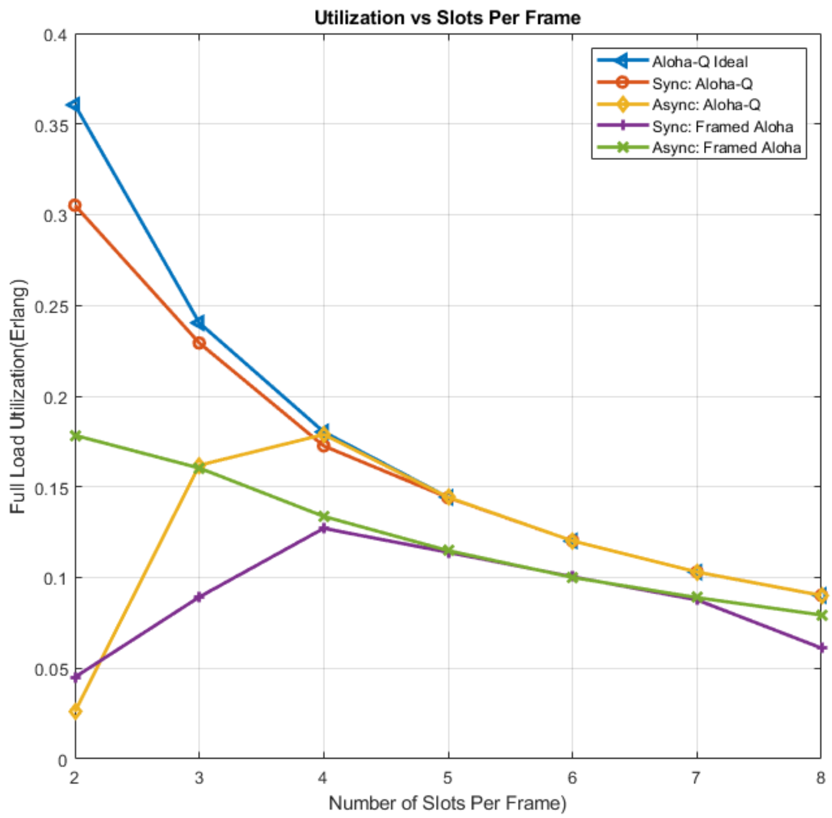

Figure 4.8: Variable Frame Size Operation Modes

Fig. 4.9 is the result that compares the fairness achieved by the protocol, FIFO queuing and Round Robin(RR) queuing, which clearly shows a significant improvement as the modification achieved 100% fairness while the original Aloha-Q achieves less than 40% fairness.

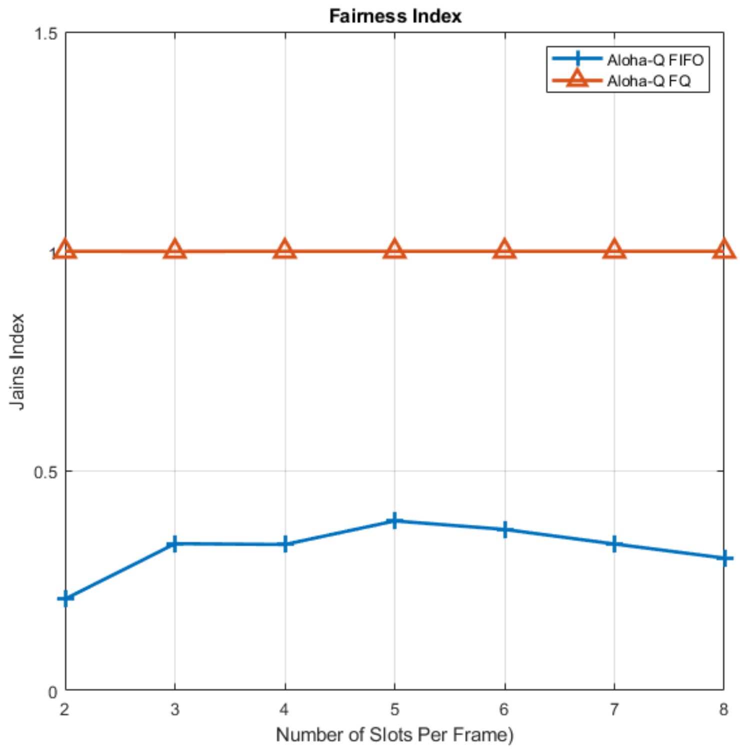

Figure 4.9: Jains Fairness Index

4.4 Summary

In this chapter the Q-learning based Aloha protocol modelled and evaluated in underwater chain network. The protocol has also been shown to be suitable for applications in underwater chain network, whereby the protocol achieve good performance. Whilst it is possible to leverage the long propagation delay to run the protocol asynchronously, nevertheless the fundamental frame duration employed by the Aloha-Q protocol needs to be reduced as this can become costly to the overall performance. Furthermore, for the protocol to be fair and useful in multi source applications the first-in-first-out (FIFO) queue needs to be replaced with a Round Robin (RR) Queue to achieve 100% fairness.

Chapter 5

Dual Control Q-learning for Medium Access Control in Multi hop chain UASNs

5.1 Introduction

In the previous chapter, a translation and evaluation of the ALOHA-Q protocol in an underwater scenario was presented. In principle, the intelligent Q learning approach has a clear advantage and promise to transform MAC protocols in UASNs. However, the preliminary evaluations have demonstrated that the original ALOHA-Q implementation has some inherently limiting attributes that called for a new approach in employing the Q-learning paradigm in the underwater acoustic sensor network application domain. Whilst scoring high on the simplicity, decentralisation and adaptability metrics, the original ALOHA-Q is demonstrably inefficient underwater. Given the limited capability of acoustic nodes in terms of capacity and energy efficiency, it is imperative the devices are deployed with the most efficient protocols in order to have a practical network.

This chapter presents a novel MAC protocol for underwater acoustic monitoring chain network. This new MAC scheme is based upon two core approaches: a new and improved time slot structure and a reformulated Q-learning coupled with implicit feedback mechanism. With the aid of a simple network model, we analyse and identify the limitations of frame based random access scheduling (in terms of achievable channel utilisation) to re-imagine a new slot structure that minimises the underutilised gap in the traditional framed based slot structures. The intelligent scheduling is achieved by coupling two separate reinforcement learning (RL) functions: Q-learning for the slot selection and another to serve as an averaging function for overall packet flow detection. Results show a substantial utilisation improvement and resilience with increasing network range.

The rest of the chapter is structured as follows. Section 5.2 presents a baseline frame based random access MAC scheme with our proposed slot structure. The scheme is pictorially analysed to ascertain the viability of the proposed time slots and characterise the achievable utilisation and followed by a discussion on the theoretical and the simulation results. In Section 5.4, we proposed and detailed a dual-control intelligent MAC scheme, and the results obtained when applied to varying lengths of chain networks.

5.2 Frame Based Random Access MAC Protocol

Framed ALOHA is not just one of the baseline protocols we compare against our proposed intelligent scheme. Framed ALOHA provides the underlying scheme to which the Q-learning was applied. In contrast to slotted ALOHA, whereby time is divided into slots and nodes can only transmit at the beginning of each slot, a frame is used in framed ALOHA, which comprises a fixed number of contiguous slots N_s. In the framed ALOHA random access strategy, each node independently and randomly chooses one of the transmission slots at the beginning of each frame.

Typically, a slot is structured such that it accommodates: a data packet of duration (τ_d), an acknowledgement packet of duration (τ_A if required), the associated propagation delays of each packet (τ_{pg}) and a small guard band (τ_g): the band is essential to correct and guard against drifts in clock precision and synchronisation. The slot structure is shown in Figure 5.1, for cases with and without acknowledgements. Whereas, in radio networks, the overheads due to the wait period between successive data transmissions in a slot/frame can be of negligible length with respect to the packet duration, in an underwater acoustic channel however, the physics impose a long propagation delay, plus low capacity (bandwidth and therefore data rate), making the overheads significant, thus negatively impacting the channel utilization and end-to-end delay.

Defining the channel utilization (U) as the rate of delivering data at the designated sink node (Equation (5.1)), then, in frame/slot based protocols, the utilization is also a function of the number of slots (N_s) in the frame. For example, if a node is allowed to transmit N packets per frame, then the maximum effective utilization at the sink is going to be upper bounded at N/N_s. The value of N_s is determined from the topology and interference population of the network. Setting N_s inappropriately will negatively affect not just the utilisation, but potentially the stability of the MAC protocol as well. For example, in a star topology, N_s is equal to the number of transmitting nodes (N_n); as each node should have a unique transmitting slot, setting $N_s > N_n$ adds extra un-utilised slot(s), and $N_s < N_n$ will cause contention as some nodes will not exclusively own a slot. Therefore, for a particular topology and interference model, there is an optimum N_s (N_{opt}) [69]. Erlang [78] is a dimensionless unit that represents continuous channel usage (for example 0E = zero channel activity, 0.5E = half channel activity and 1E = full channel usage).

$$U_{normalised}(Erlang) = \frac{N\tau_d}{N_s S_L} \tag{5.1}$$

therefore, the optimum utilization is:

$$U_{normalised}(Erlang) = \frac{N\tau_d}{N_{opt} S_L} \tag{5.2}$$

where S_L, τ_d denote the slot duration and packet duration in seconds respectively.

One of the consequences of having low capacity is the long transmission duration, which presents two situations for a given transmitter and receiver pair: the transmission duration could be greater, equal or less than the propagation delay between the nodes. For practical purposes, since each slot will ultimately be padded with a small guard band, henceforth, we are treating the the rare ”equal” situation as a special case of the ”greater than” situation. Following [85], if we introduce the parameter K_τ (Equation (5.3)), then the resulting slot structure can have either of two sets of transmission-reception patterns: overlapping and non-overlapping based on the value of K_τ, as shown in Figure 5.1.

$$K_\tau = \frac{\tau_d}{\tau_{pg}} \tag{5.3}$$

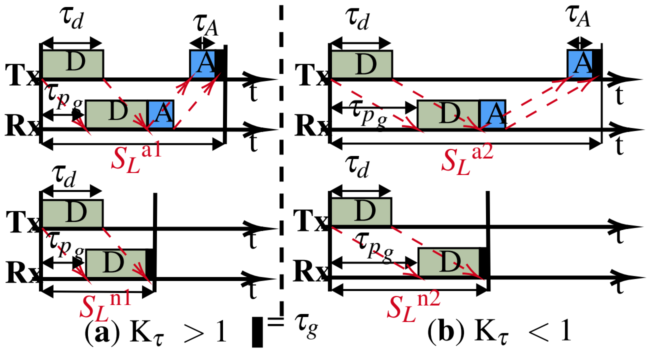

Figure 5.1: Typical slot structures: (**a**) Overlapping, transmission-reception occurs concurrently for the data packet. (**b**) Non-overlapping, data transmission completed before reception occurs.

S_L^{a1} and S_L^{a2} represent the slots' length with ACK and are typically used by slotted protocols employing an ACK signal such as ALOHA-Q. Similarly, S_L^{n1} and S_L^{n2} are the slots without ACK as used in framed ALOHA and TDMA. Equations (5.4) and (5.5) are used to calculate the slot sizes.

$$S_L^a = \tau_d + \tau_A + 2\ \tau_{pg} + \tau_g \tag{5.4}$$

$$S_L^n = \tau_d + \tau_{pg} + \tau_g \tag{5.5}$$

In this slotted concept, nodes are allowed to transmit only one packet per frame (i.e., N = 1), and the expression of maximum utilisation (U) can be simplified to the ratio of packet duration-to-frame size (Equation (5.6)). We can combine Equations (5.2) and (5.6) to calculate the expression of the utilisation below:

$$U = \begin{cases} \frac{\tau_d}{N_{op_t}\ (\ \tau_d + 2\tau_{pg} + \tau_A + \tau_g\ )}, & S_L^a \\ \frac{\tau_d}{N_{op_t}\ (\tau_d + \tau_{pg} + \tau_g\ )}, & S_L^n \end{cases} \tag{5.6}$$

As $\tau_d, \tau_{pg} >> \tau_A, \tau_g$, Equation (5.6) approximates to:

$$U \approx \begin{cases} \frac{\tau_d}{N_{op_t}\ (\ \tau_d + 2\tau_{pg}\ )}, & S_L^a \\ \frac{\tau_d}{N_{op_t}\ (\tau_d + \tau_{pg}\ )}, & S_L^n \end{cases} \tag{5.7}$$

From Equation (5.7), it can be seen that, since τ_d and τ_{pg} dominate, the value of K_τ will guide us on how to improve channel utilisation by restructuring the slot size. For $K_\tau > 1$, we are constrained with respect to any change to the slot size. Any reduction will create overlapping slot reception that will effectively render the slotting meaningless, as demonstrated with the downgrade of slotted ALOHA to pure ALOHA underwater [85].

In most UASN applications, the propagation delay is longer than the transmission duration because of sparse connectivity. Therefore, $K_\tau < 1$ best describes such scenarios. We propose the slot structure in Figure 5.2. The slot size is now reduced to approximate the propagation delay ($S_L \approx \tau_{pg}$), which is possible since with $K_\tau < 1$, the data packet can be safely accommodated in τ_{pg}. This simple slot structure aims to reduce and fill the otherwise wide gap in the conventional slots with useful data (compared to Figure 5.1). Therefore, for a given chain UASN, designed with nodes separated by a dm transmission range, we demonstrate that there are advantages to the performance improvements of using our slot structure; for example, the peculiar characteristic of the underwater communication channel in terms of its distance dependent capacity, that is, the acoustic transmission bandwidth and data rates decrease with increasing transmission distance [86]. Thus, given the flexibility, instead of employing relatively fewer hops to transmit over longer ranges (requiring high power) with low capacity, we can potentially achieve higher capacity transmissions with additional hops added to route data over shorter ranges (suitable for low power and cheaper nodes). To investigate the achievable utilisation, the slot structure shown in Figure 5.2 is based on $K_\tau \approx 1$: a special case of $K_\tau < 1$. This is purely to limit the overhead in the slot, as increasing the slot size beyond τ_{pg} negatively affects the utilisation according to Equation (5.6).

5.2.1 Scenario and Network Model

Consider a scenario comprising quasi-stationary equally spaced nodes in an N hop underwater network chain topology, with data delivered along the chain from one end to the other. Figure 5.3 depicts an example of such a network with N = 4 and hop distance d. This topology is representative of pipeline monitoring. As such, during the reporting cycle, the network can be considered loaded to capacity; accordingly, this work is primarily concerned with the achievable utilisation. To aid the analysis, the following assumptions are made:

1. All nodes are homogeneous and communicate over a single channel, half-duplex mode.
2. The collision model (non-capture) is used, i.e., if two or more packets overlap at the receiver, they are discarded.
3. Nodes are globally synchronised, an assumption commonly employed to simplify analysis and applicable to quasi-stationary nodes synchronised before deployment.
4. The interference range (Ifx) is twice the reception range (Rx); this model is typically employed for chain networks as an illustrative model to incorporate the effect of interference from nodes that are two hops away.
5. A source node has saturated traffic, i.e., always has a packet to send, to provide the maximum monitoring rate based on the transmission opportunities offered by the MAC layer. Similar research papers are concerned with achievable utilization [87, 88, 74].
6. All source/relay nodes can only transmit one packet per frame, a consequence of Assumption (4) yielding a frame consisting of four slots [69], as only one of four connected nodes can transmit successfully at a given time.

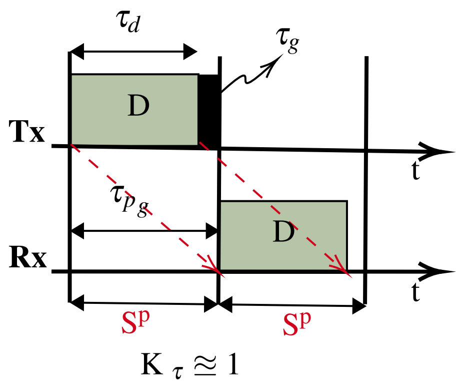

Figure 5.2: Proposed slot structure.

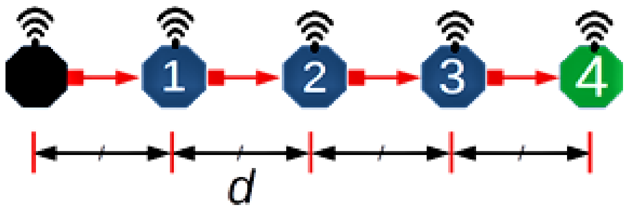

Figure 5.3: An example scenario.

We re-write Equation (5.7) of S_L^n to get the new utilisation for the proposed slot structure:

$$U_{normalised}(Erlang) = \frac{\tau_d}{N_{op_t}\tau_{pg}} \tag{5.8}$$

and in terms of K_τ, it becomes:

$$U_{normalised}(Erlang) = \frac{K_\tau}{N_{op_t}} \tag{5.9}$$

In summary, while the traditional slot structure that incorporates the propagation delay and/or ACK packet within the constraints of the available channel resources, we show that with $K_\tau < 1$, the propagation delay is sufficient to accommodate the data packet, and it is then possible for the slot size to be effectively reduced and restructured (by at least 50% of the cases in the $K_\tau < 1$ regime), and as long as a protocol does not require an ACK packet, there is a potential for a dramatic improvement in performance (Equation (5.9) vs. Equation (5.7)).

5.3 Model Analysis

To analyse the network with the proposed slot structure (Figure 5.2), we consider a baseline scheme whereby each node initialises by randomly choosing a transmission slot. The purpose of considering this scheme is first to demonstrate the inefficiency of a random access scheme by analysing the distribution of the achievable channel utilization, second to investigate the feasibility of applying intelligent techniques to the model that could lead to a significant performance improvement and, finally, to evaluate the efficacy of the proposed slot structure coupled with the intelligent techniques relative to similar intelligent approaches and random access baseline schemes.

To build the frame, we start with the optimal number of slots per frame N_{opt}. In a linear chain network (such as Figure 5.3 and longer,) N_{opt} is four as computed according to the two hop interference model [69]. This is because in a linear topology with the two hop interference model, technically only one in four nodes can successfully transmit at a given time. Similarly, for one hop and three hop interference conditions, one in three and one in five nodes can transmit successfully [69, 74]. Therefore, for a distributed MAC protocol, such as framed ALOHA employed in this setup, each node is free to chose any of the available four slots in the frame, resulting in $4^4 = 256$ ways for nodes to independently select and occupy transmission slots. Table 5.1 lists the range of the 256 possible slot combinations in a four column array of 64 unique patterns, with each column vector signifying the transmission slot pattern from Node 0 to Node 3. That is, the vector [0000] denotes all nodes selecting and occupying Slot 0; likewise, slot sequence [2210] signifies both Nodes 0 and 1 choosing Slot 2, while Nodes 2 and 3 choose Slot 1 and Slot 0, respectively.

5.3.1 Pictorial Analysis

Pictorial timing depictions are employed to observe and obtain the theoretical bounds of the scheme in terms of channel utilisation. The diagrammatic method provides a visual intuition of our core idea. For each pattern, N_0 is the source node; it generates and transmits data in every

Table 5.1: Possible Slot Permutations

S/N	SLOT SEQUENCE			
	SEQ_0XXX	SEQ_1XXX	SEQ_2XXX	SEQ_3XXX
0	[0 0 0 0]	[1 0 0 0]	[2 0 0 0]	[3 0 0 0]
1	[0 0 0 1]	[1 0 0 1]	[2 0 0 1]	[3 0 0 1]
...	[...]	[...]	[...]	[...]
...	[...]	[...]	[...]	[...]
62	[0 3 3 2]	[1 3 3 2]	[2 3 3 2]	[3 3 3 2]
63	[0 3 3 3]	[1 3 3 3]	[2 3 3 3]	[3 3 3 3]

frame to N_1, which forwards the packet (if successfully received) to N_2 in the next frame, and so on. Overall, individual packets are traced frame-by-frame as they traverse the network from source to sink (N_0 *to* N_4). The final utilisation is measured when an overall periodic pattern emerges at the sink node (vertical red lines in each example figure; refer to Section 5.3.1).Figures 5.5-5.10) are provided to illustrate the process as follows:

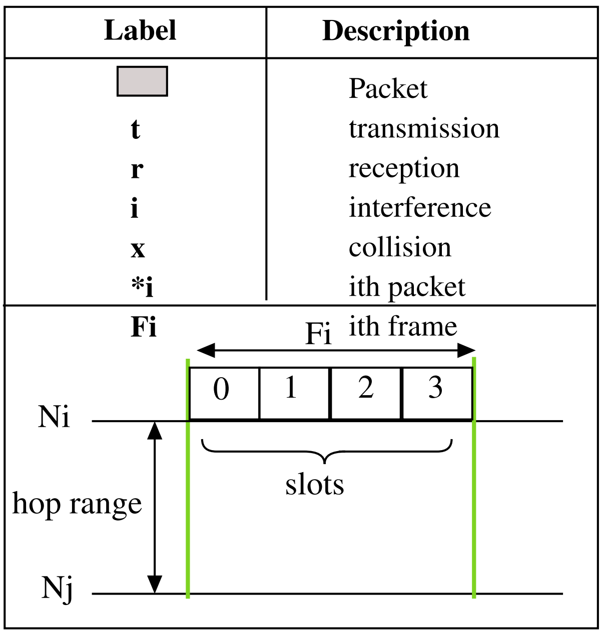

Figure 5.4: Legend for packet labels and illustrations.

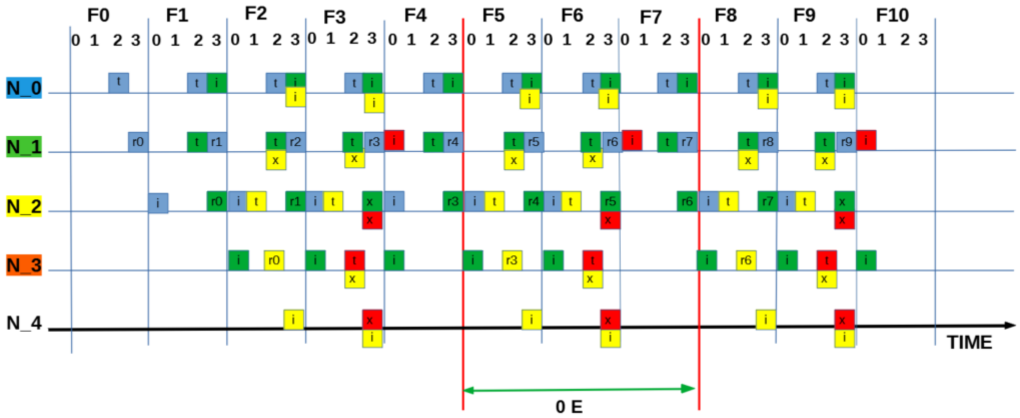

Figure 5.5: SEQUENCE:[2 2 1 2]: "Worst" measured utilisation based on zero packets being delivered = 0.0 E.

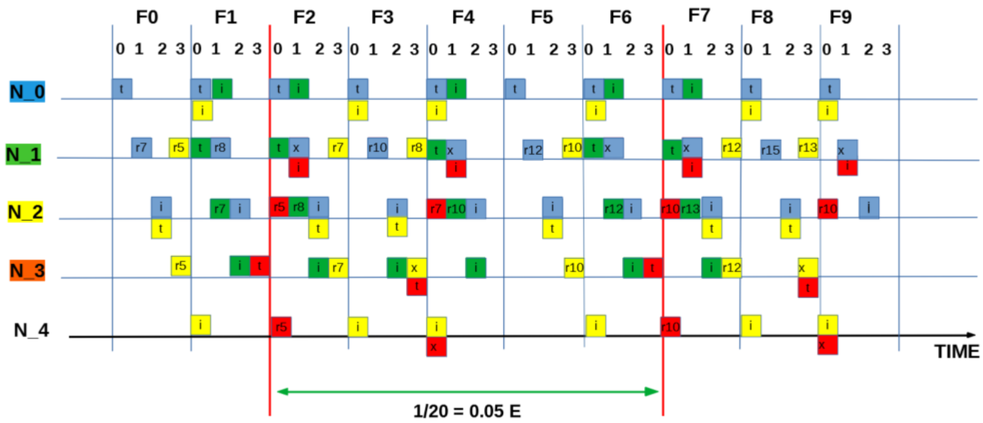

Figure 5.6: SEQUENCE:[0 0 2 3]: "Intermediate" measured utilisation based on one packet in five frames (20 slots) = 0.05 E.

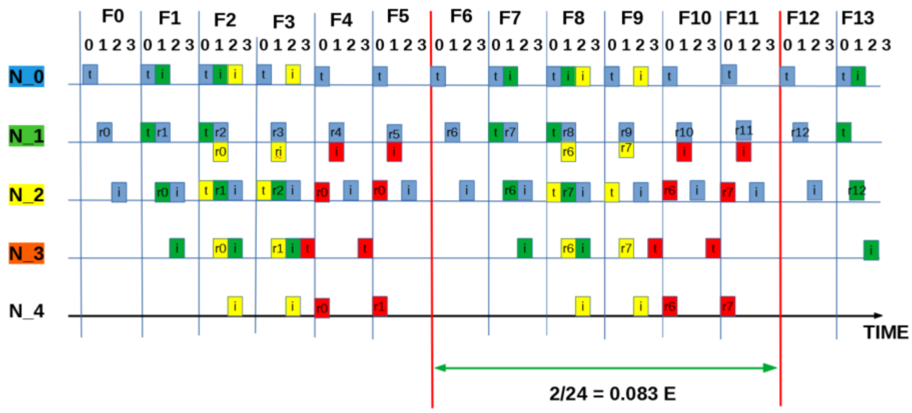

Figure 5.7: SEQUENCE:[0 0 0 3]: "Intermediate" measured utilisation based on two packets in six frames (24 slots) = 0.083 E.

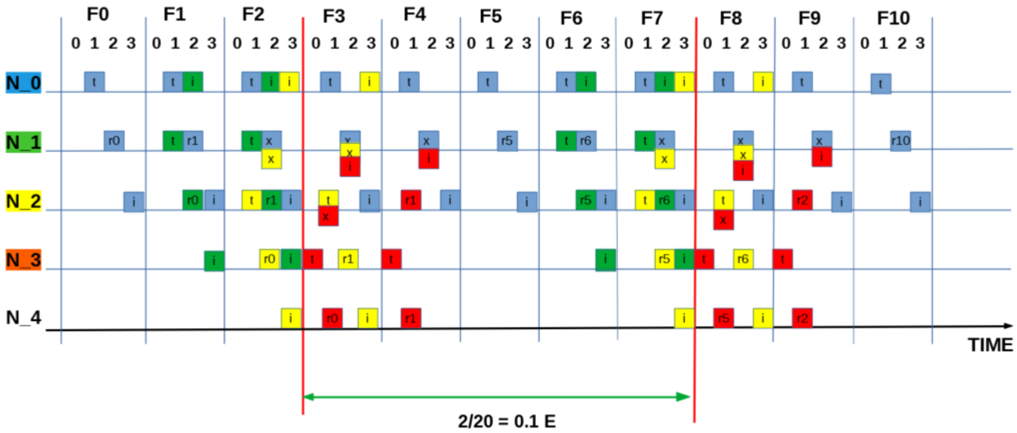

Figure 5.8: SEQUENCE:[1 1 1 0]: “Intermediate” measured utilisation based on two packets in five frames (20 slots) = 0.1 E.

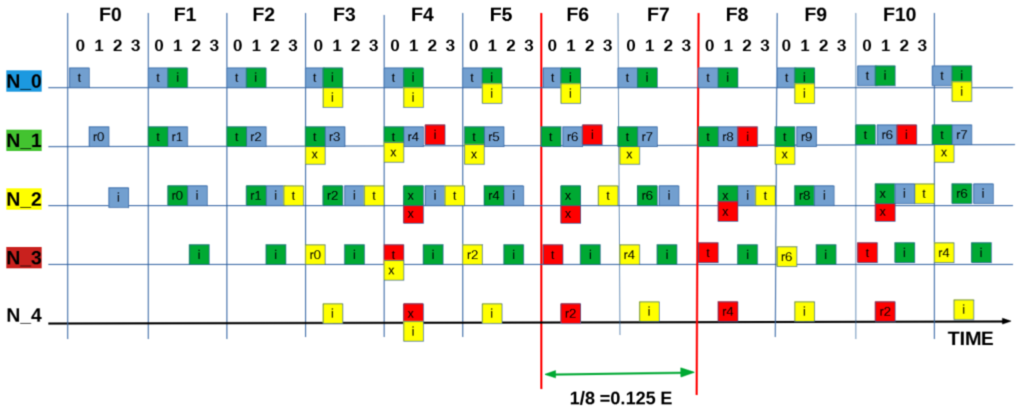

Figure 5.9: SEQUENCE:[0 0 3 0]: “Half” measured utilisation based on one packet every two frames (8 slots) = 0.125 E.

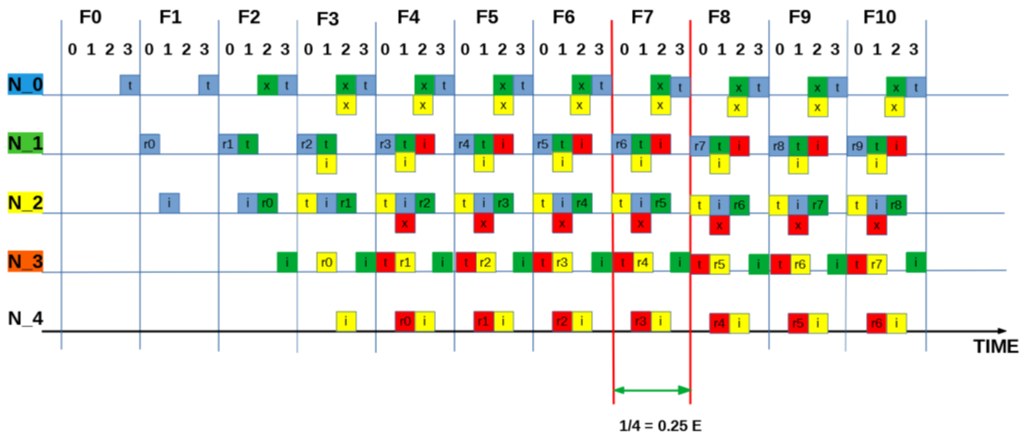

Figure 5.10: SEQUENCE:[0 2 1 1]: “Best” measured utilisation is one packet in every frames (4 slots) = 0.25 E.

5.3.2 Results

In order to empirically evaluate the performance of the above random access scheme, we ran a simulation on a network of five nodes (Figure 5.3) configured with the proposed slot structure analysed in Section 5.3. Each node is pre-configured to run a MAC protocol that randomly selects and maintains a transmission slot at the beginning of each simulation run. It should be noted that in this simulation, since $K_\tau \approx 1$, the transmission delay and propagation delay are abstracted to 1:1 for the best results. Moreover, the choice of four transmitting nodes in this analysis is based upon the minimum number of nodes for the chosen two hop interference range model effect to manifests, thus, the analysis is not limited to four transmitting nodes, but can be scaled to network(s) with different number of hop and/or different interference ranges. Theoretically, replicating slot patterns with perfect collision free schedules along longer chain should have proportional effect on average end-to-end delay without affecting utilisation.

Figure 5.11 shows and compares the utilisation results from both the analytical distributions of the slot patterns and the simulations. Overall, there are three prominent utilisation levels and some spurious intermediate levels, as summarised in Table 5.2. The summary provides individual proportions of levels in each plot, and the overall column is the contribution of each sequence in the combined set of 256 slots.

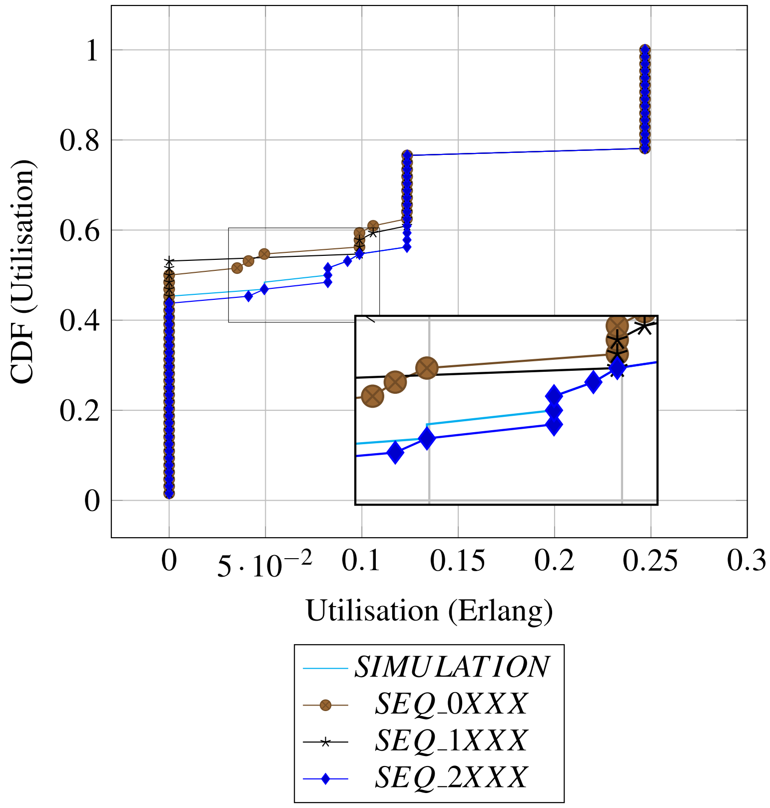

Figure 5.11: Distributions' comparison.

Table 5.2: Summary of Utilisation Levels

Level	**Proportions(%)**				
	SEQ_0XXX	SEQ_1XXX	SEQ_2XXX	SEQ_3XXX	Overall
Worst case (0E)	45.3	50.0	53.1	43.8	48.1
Intermediate (0.03E - 0.1E)	9.4	10.4	6.3	10.9	9.4
Half (0.125E)	21.9	15.6	17.2	21.8	19.1
Maximum (0.25E)	23.4	23.4	23.4	23.4	23.4

Depending on the chosen slot by the source node, transmissions could be initiated from either the frame edge (Slots 0 or 3) or mid-frame (Slots 1 or 2), and to some degree, the results show how the position of a chosen slot affects the utilisation. As shown in the result summary (Table 5.2), there is a subtle, but clear advantage in performance when the source node initiates transmissions with emerging slot patterns at frame edges (i.e. SEQ_0XXX, SEQ_3XXX) relative to the mid frames (i.e. SEQ_1XXX, SEQ_2XXX) or there is at least an 8% better chance of getting a packet received at the sink node when the source node transmits at the edges of a frame compared to when source node uses mid frame (in terms of the worst case utilisation levels).

Intuitively, the distribution of the utilisation of the patterns can be assumed to be similar, since it can be demonstrated that each column sequence can be translated to another corresponding sequence in the remainder of the columns (Table 5.1). However, due to the transmission strategy of the protocol of scheduling packet transmission at the beginning of each frame, the simple slot structure guarantees that packets transmitted at $slot_i$ be received at $slot_{i+1}$. This means sequence translations will result in packet reception/interference across frames, consequently causing the distribution of the utilization outcomes to vary. For example, consider the corresponding slot selection sequences: [0 0 3 0], [1 1 0 1], [2 2 1 2] and [3 3 2 3]. [0 0 3 0] and [3 3 2 3] both have cross-frame receptions and have a similar utilization of 0.125 Erlangs (Figure 5.10). In contrast, [1 1 0 1] and [2 2 1 2] have no cross-frame reception and yield 0 Erlangs (Section 5.3.1: Figure 5.5). Only 60 out of the total 256 slot sequences yield the maximum utilization level as a whole and remain immune to the slot sequence translations because they are perfectly collision-free. In Figures 5.5–5.10 (Section 5.3.1), we show how we computed six of the ten prominent utilisation levels for brevity.

The simulation results are in agreement with our analytical results, as they show that no data is delivered 48% of the time. This corresponds to the average of the possible 43–53% worst cases in the given original slot patterns, as expected. Most importantly, the simulation result confirms that the full channel utilization is achievable with the exact proportion of 23%. Finally, the simulation result shows the average performance of the random slot selection protocol and will serve as a baseline with which to demonstrate the merit of slot based learning in the new protocol ALOHA-QUPAF. However, as previously stated, while the ACK signal is crucial to the Q-value update operation, it puts an additional burden on the scarce network resources underwater: reducing utilisation due to overheads and increased delay due to the ACK signal wait times. Our goal is to implement a novel Q-learning approach that maintains the level of intelligence without this explicit ACK signal, thereby maximising the channel utilisation and improving end-to-end delay.

5.4 Underwater Packet Flow ALOHA-Q: ALOHA-QUPAF

The proposed slot structures in Figure 5.2 pose a critical question: how do we apply a simple reinforcement learning algorithm to ultimately achieve collision-free scheduling without an ACK packet? In this section, we present a two stage solution using a reformulated Q-learning coupled with a simple stochastic averaging expression inspired by [89], where the harmonised stages are succinctly described in **Algorithm 2**. We demonstrate the efficacy of our dual-mode learning approach in improving performance in a chain network as introduced in Section 5.2.1.

5.4.1 Protocol Design

In order to achieve the goal of realising a collision-free schedule without an explicit ACK signal, we modified the Q-value update process (Section 4.2.1) while maintaining the remaining protocol settings and assumptions (Section 5.2.1). Specifically, at the beginning of each frame, a relay node chooses the slot with the highest Q-value (if more than one slot has the highest Q-values, one is chosen at random) to forward a received packet on to the next hop. In the case of the source node, it initialises by randomly selecting and maintaining a constant slot for transmission. This is because we employ a Q-learning process that utilises packet receptions to update and reinforce transmission slot selection. Our solution involves a two stage approach based on the following intuitions:

1. In a network with half-duplex nodes, they cannot transmit and receive at the same time (in the same slot); therefore, we employ Q-learning to isolate all reception slots by punishing those slots to lower their Q-values. As such, when a node scans the Q-table, reception slots will have low Q-values and are unlikely to be selected for transmission.

2. A continuous flow of packets over the chain is expected in saturated traffic with a healthy channel. Thus, a relay/sink expects a new packet(s) in every frame after receiving the first packet, and a packet collision is inferred whenever that stream of packets gets disrupted. To exploit this realisation, every time a relay node transmits a packet, it rewards the

chosen transmission slot (positively updates the slot's Q-value) if and only if a new packet is received afterwards.

We denote the two stages in the dual mode control as slot selection and flow harmony, and a detailed description of the process is given below:

- **Slot selection:** This is implemented by Q-learning to eliminate the reception slot(s). When a source node generates a packet and transmits, upon receiving the packet, the receiver (relay node) will record the reception slot (rx_s) and update the Q-value of the slot according to (Equation (4.2.1)). Specifically, each slot in a frame is mapped to a value in the vector of Q-values ($Q[ns]$), and the Q-values are initialised with a uniform random number less than one, whereby for each reception, the node computes rx_s and updates $Q[rx\_s]$ accordingly with $\psi = -1$. Consequently, this continual negative reinforcement of reception slots isolates those slots, and the slot(s) with the highest Q-value(s) signifies a probable collision-free slot at the local level, therefore a good candidate(s) slot(s) for transmission. For a relay node, at the beginning of each frame, if a node has a packet(s) in its queue, it will schedule a packet transmission in a slot with the maximum Q-value; however, if more than one slot shares the maximum Q-value, one will be chosen at random from amongst them. Whilst the Q-value of the reception slot is always punished following any reception, the Q-value of the transmission slot is only updated after every transmission. If there is a subsequent packet reception, the transmission slot is rewarded ($\psi = 1$), otherwise it is punished ($\psi = -1$). However, since this scheme lacks a definitive feedback signal based on this node action(s) of transmissions, the success of any transmission in the chosen slot is uncertain. This is because, unless the packet flow is network wide, a continuous transmission and reception by a relay node does not mean that a given node's transmissions are not interfering with some other transmissions especially for the downstream links. Therefore, to avoid nodes from getting stuck in local minima, a control mechanism has to be devised to regulate the Q-values especially of the transmission slot.

- **Flow harmony:** Although we devise a means to obtain feedback from the environment (reward/punishment), the node cannot directly link these signals to its own action(s); hence, at any given time during the network run, we only have a partial observation of the channel condition; this type of process is best modelled as a partially observable Markov decision process (POMDP) (see in Section 2.4.2). This is because, instead of certainty in the network wide flow, the packet flow experienced by each node gives us a partial observation on the channel at the local level. The POMDP framework enables us to model the local observations by agents to generate a probability distribution of a belief state (in our case, settled or unsettled flow). The network can be in either stable or unstable packet flow states, and we therefore designate two belief states accordingly. We employ a simple heuristic strategy based on stochastic averaging [90], whereby each node independently tracks its overall local packet flow in a given window, which we then translate as the distribution of the belief state. The distribution of the belief states is computed with Equation (5.10). For each reception in a frame, fl_τ is updated by λ_t steps at the tracking rate γ. While the expression monotonically approaches one, it is continually windowed every (W_n) frames and compared to a fixed threshold ($thresh$). Based on our simulation experiment, ideally, fl_τ will reach 98% by the 20th frame; hence, we heuristically set ($W_n = 20$) to check for fl_τ with a tolerance of $thresh = 95\%$, which should be achieved at ($W_n = 14$).

If we designate the belief states $S1$ and $S2$ respectively as the initial state (both Q-values and fl_τ reset; the network is assumed to have no stable flow during learning) and the flow harmony state, $S1$ is decided when the averaging function exceeds the threshold, which indicates that flow harmony has been achieved at least in the node's local interference group, otherwise the node resets to $S2$. In essence, every node has a window of 20 frames to isolate incoming reception slots and settle on a transmission slot. Whenever a particular node(s) fails to settle and join the flow, the reset will make the node switch to another slot and potentially notify other nodes in the neighbourhood as well.

$$fl_\tau \leftarrow (1-\gamma_t) fl_\tau + \lambda_t \tag{5.10}$$

where fl_τ, γ_t and λ_t denote the flow averaging, the learning/tracking rate and the increment scale, respectively.

In order to infer and glean the signal that drives fl_τ from the flow of packets, the function is heuristically chosen to have an identical trend to the underlying slot selection Q-learning. In Fig. 5.12, Q_s and Q_f are an idealised/stable trajectory of continuously successful and failing slot respectively, whilst Fl is the trend of fl_τ for different learning rates. However, in practice the trajectory of the slot selection is not expected to be smoothly reinforced initially, since at that stage the node is still learning and exploring the slots. Running fl_τ at the same or lower learning rate as Q means either delaying the flow harmony checks or lowering the threshold. On the other hand, setting fl_τ to track at a faster learning rate means earlier checks. Therefore, fl_τ has to balance fast checks whilst avoiding unnecessary resets to Q.

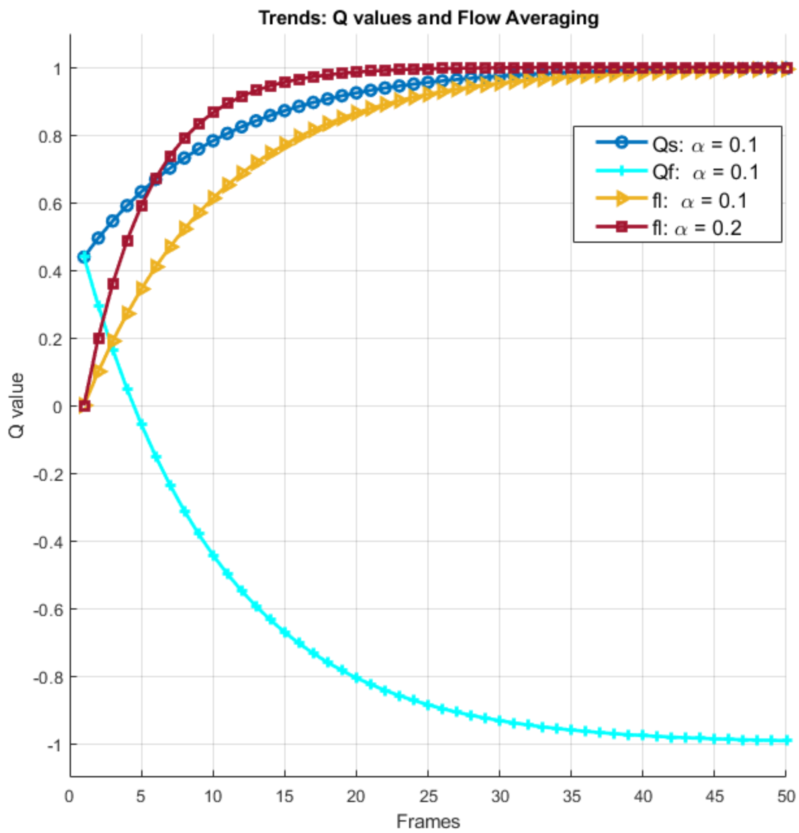

Figure 5.12: Trends of both Q value and fl_τ

By using this two stage solution, ALOHA-QUPAF unlike ALOHA-Q effectively isolates

Algorithm 2: ALOHA-QUPAF algorithm.

```
Result: S1,or S2
S1;
Initialization;
α, γ_t, λ_t, ψ // From Table 5.4;
// For all n ;
Q[n] ← rand([0,1));
W_n ← 20, thresh ← 0.95, fl_τ ← 0;
S2;
while node is online do
    if Reception then
        get rx_s ;
        //Activating the packet reception flag;
        Rx_τ ← True;
        Q[rx_s] ← Q[rx_s] + α(ψ − Q[rx_s]);
    end
    // Frame Block;
    W_n− = 1;
    if Rx_τ then
        fl_τ ← (1 − γ_t) fl_τ + λ_t;
        Q[tx_s] ← Q[tx_s] + α(ψ − Q[tx_s]);
    end
    // Belief State Block: compares flow rate with threshold;
    if W_n == 0 then
        if fl_τ < thresh then
            // Node resets parameters;
            node ← S1;
        else
            // Maintain parameters;
            node ← S2;
        end
        W_n = 20;
    end
    // Transmission slot selection;
    tx_s ← [x | x ∋ argmax_{x∈X} Q[x]];
    //De-activating the packet reception flag;
    Rx_τ ← False;
end
```

both reception slots from the transmission slots and finds an implicit way of getting the feedback signal of the node's action based on the individual nodes experiencing successful reception of a continuous stream of packets. Furthermore, it differs from framed ALOHA, since it can intelligently create and maintain a robust collision-free schedule. The complete ALOHA-QUPAF algorithm is given below.

Aloha-QUPAF Update

Table 5.3 illustrates an example of the learning implementation in ALOHA-QUPAF. Consider for example a situation whereby a node ***i*** that has in the previous frame received two packets in Slots 0 and 3. To forward a data packet the node chooses Slot 1 (0.9649, the highest assigned with uniform random Q-value, in accordance with Section 4.2.1) at the beginning of the current frame to schedule transmission. The Q-values of Slots 0 and 3 are updated below.

- The new Q-value of Slots 0, 1 and 3 becomes;
  $Q[0] \leftarrow 0.1576 + 0.1(-1 - 0.1576)$; [0.0418]
  $Q[3] \leftarrow 0.9572 + 0.1(-1 - 0.9572)$; [0.7615]

- In the next frame, Slot 0 has the lowest Q-value and is not considered, and the node chooses Slot 1. Since there was packet(s) receptions in Slots 0, and 3, previous transmission is assumed successful given no apparent disruption to the packet stream. Hence, both reception and transmission slots Q-values will be updated.
  $Q[0] \leftarrow 0.0418 + 0.1(-1 - 0.0418)$; [-0.0623]
  $Q[1] \leftarrow 0.9649 + 0.1(1 - 0.9649)$; [0.9684]
  $Q[3] \leftarrow 0.7615 + 0.1(-1 - 0.7615)$; [0.5853]

- Similarly, for Frame 2, with successful packets receptions at Slot 0 and 3, the node chooses Slot 1 again as it has the highest Q-value (0.9684) and sends data; , the Q-values are updated accordingly.
  $Q[0] \leftarrow -0.0623 + 0.1(-1 - 0.9652)$; [-0.4073]
  $Q[1] \leftarrow 0.9684 + 0.1(1 - 0.9684)$; [0.9716]
  $Q[3] \leftarrow 0.5853 + 0.1(-1 - 0.5853)$; [0.4268]

- Finally, with each reception/transmission cycle fl_τ is also updated to track and average the flow of packets (Eq. (5.10)), following our example, fl_τ becomes 0.9820 which has exceeded the threshold of 0.95 needed to maintain the current state of transmission slot selection.

The table gives the Q-values up to twenty frames assuming both receptions and transmission patterns is maintained; packets continually received at Slot 0 and Slot 3. This simple, yet effective recursive Q-learning update bootstraps the trial-and-error mechanism to a robust collision-free schedule as each node will eventually and independently occupy a transmission slot that will join and maintain the flow.

Table 5.3: Example of Q-value update in Aloha-QUPAF

Frame/Q-values	Q[0]	Q[1]	Q[2]	Q[3]
FRAME 0	0.1576	0.9649	0.7572	0.9572
FRAME 1	0.0418	0.9684	0.7572	0.7615
FRAME 2	-0.0623	0.9716	0.7572	0.5853
FRAME 3	-0.1561	0.9744	0.7572	0.4268
FRAME 4	-0.2405	0.9770	0.7572	0.2841
...	...	...	...	...
FRAME 20	-0.8436	0.9953	0.7572	-0.7356

5.4.2 Simulation Parameters

Table 5.4 provides the parameters used in this thesis to simulate and evaluate ALOHA-QUPAF and the comparison protocols.

Table 5.4: Simulation Parameters

Parameter	**Value**
Transmission/Reception Data Rate	640bps
Data Packet Size	632bits
ACK Packet Size	16bits
Slot Size	640bits
Slots per frame	4
Reception Range	200m
ψ	± 1
α	0.1
λ_t	0.2
γ_t	0.2
1 hop Propagation Delay (Relative to packet size)	1s

5.4.3 Results

Since the focus of this work is principally to improve performance in terms of channel utilization measured at the sink, ALOHA-QUPAF is compared to a state-of-the-art ALOHA-Q, which employs a similar Q-learning technique, and a baseline framed ALOHA scheme in terms of the normalised utilization. We simulated networks of varying hop lengths with the protocols configured with respect to the structures in Figure 5.2. For a fair comparison, as our proposed slot structure is constrained to $K_\tau > 1$, we only compare ALOHA-QUPAF with the other protocols in the $K_\tau > 1$ regime. The network was simulated in the Riverbed Modeler (formerly OPNET) environment, and the setup used the parameters given in Table 5.4, which were based on a modem developed by Newcastle University [84]. In all cases, the network was simulated for 15,000 frames, with a single saturated source at one end of the network and a sink at the other end. In terms of result collection, due to the continuous nature of the learning of the ALOHA-QUPAF algorithm, the results were collected from the beginning of the simulation. Figures 5.13 and 5.14 are the results obtained when the network was simulated on four and eight hop networks, respectively. The figures compare the performance of ALOHA-QUPAF with ALOHA-Q and framed ALOHA. This comparison is particularly important as the protocols share similar reception conditions in the $K_\tau > 1$ scenario; transmission and reception occur in the same slot (Figure 5.1). Evidently, in this setup, both ALOHA-QUPAF and ALOHA-Q are dramatically affected as the network size increases (four hops to eight hops). The maximum utilisations of ALOHA-QUPAF (0.217 Erlang) and ALOHA-Q (0.191 Erlang) are both sharply halved for about 40% and 58% of the simulated cases, respectively. This performance drop can be explained by the presence of the hidden node phenomenon [43, 44]. This is simply the situation whereby a particular communication between any two nodes is interfered by another transmission within range of the receiver.

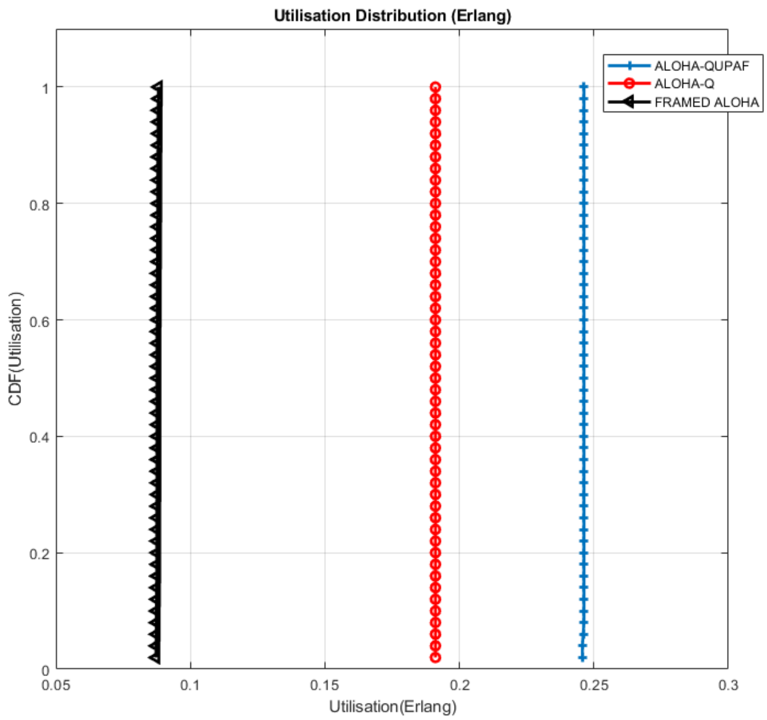

Figure 5.13: $K_\tau > 1$: 4 hops utilisation performance comparison.

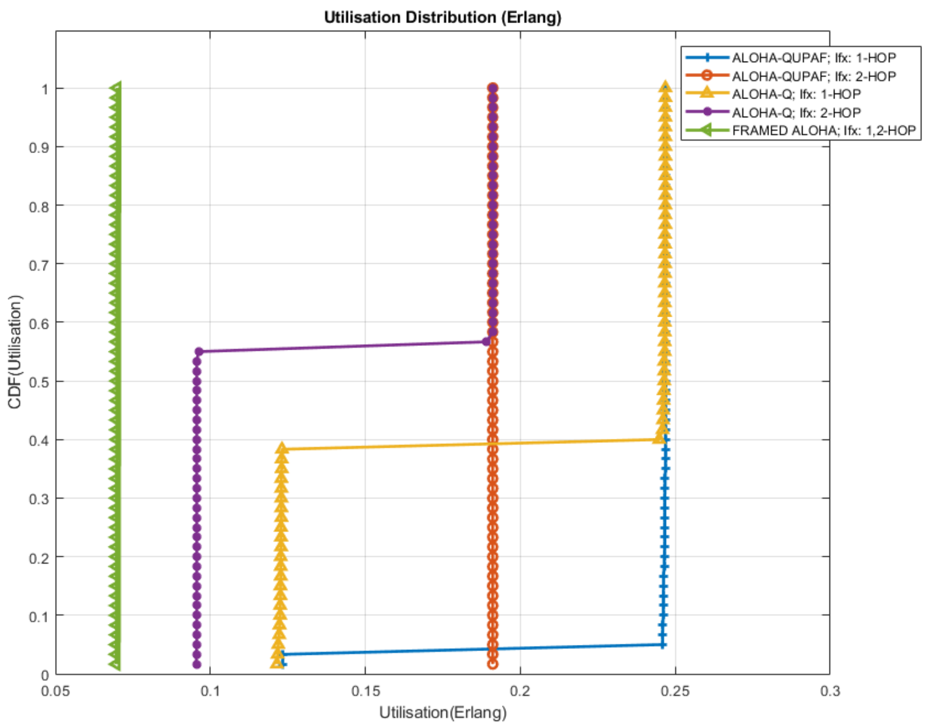

Figure 5.14: $K_\tau > 1$: 8 hop utilisation performance comparison.

Figure 5.15 depicts the hidden node problem in an eight hop chain network, in a situation whereby both $N2$ and $N5$ share the same transmission slots; thus, transmission from $N2$ to $N3$ will be periodically interfered by $N5$ transmitting to $N6$, as packets are relayed along the chain. The effect of the hidden node problem as the reason for the performance degradation is confirmed by the agreement shown in the simulation results obtained when the interference range (*Ifx*) is reduced from two hops to one hop in the eight hop chain (Figure 5.14) with the results in the four hops network (Figure 5.13). This is because, in a two hop interference range model, a four hop range chain network is of insufficient length for the issue to manifest. Mitigating the hidden node issue is a subject of further work. Another important metric worth mentioning is the end-to-end delay; however, it is not presented here, since ALOHA-QUPAF does not implement packet retransmissions. Therefore, neglecting any processing and queuing delays in the nodes, the E2E delay is fixed as a function of the number of hops in the network. The simulations show that ALOHA-QUPAF achieves 0.124 Erlangs at its worst and 0.248 Erlangs at its best, outperforming both ALOHA-Q (0.19 Erlangs best) and framed ALOHA (0.069 Erlangs) respectively by at least 13% and 148% in all simulated scenarios.

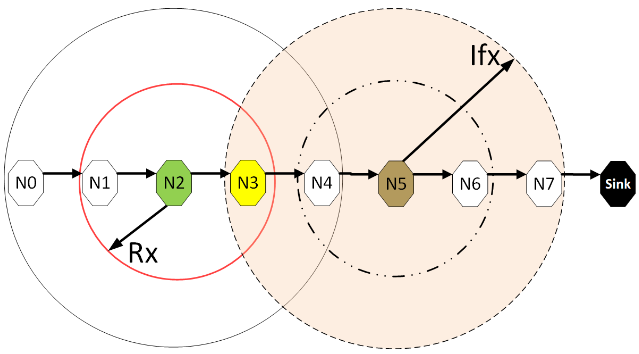

Figure 5.15: The hidden node problem.

Figure 5.16 presents the performance of ALOHA-QUPAF with our proposed slot structure (Figure 5.2) in the $K_\tau < 1$ scenario. To demonstrate how the ALOHA-QUPAF protocol is affected by the network length, we extend the range to 16 hops and evaluate its performance. The results show a subtle drop in the overall performance from four to 16 hops. The decrease in performance is attributable to the increase in the hidden node spots (bottlenecks points) and the time needed for the protocol to find a collision-free schedule as the network size increases.

Each time a node switches to a different transmission slot, this will have a ripple effect across the neighbouring nodes, causing others to potentially switch slots as well, essentially resetting the process. Despite the lack of an explicit acknowledgement signal, the protocol demonstrates significant performance improvement with more than 90% of cases achieving 0.24 Erlangs for networks in the 4–12 hop range and 80% for the 16 hop range.

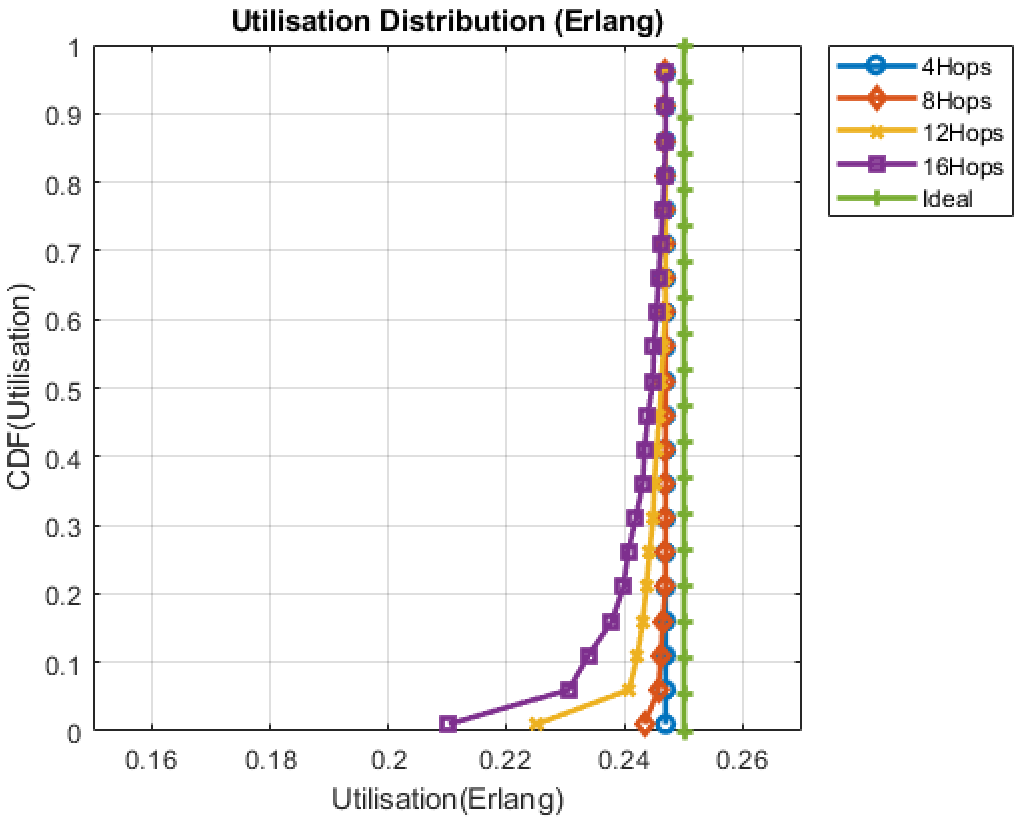

Figure 5.16: ALOHA-QUPAF utilisation for 4, 8, 12 and 16 hops networks using the proposed slot structure.

5.5 Summary

This chapter presents a simple slot structure based on the relationship between packet transmission duration and propagation delays in conjunction with two stage reinforcement learning techniques to develop a novel MAC protocol (ALOHA-QUPAF) that can achieve near channel capacity utilisation in a UASN chain topology. Our solution addresses the excessive overhead required in slot structures used by typical slotted/framed protocols. Incorporating Q-learning in the protocol makes it robust against network and channel changes due to the high dynamic underwater environment. Furthermore, one of the primary goals is for the protocol to be distributed, adaptive, simple and computationally inexpensive so that it is suitable for use in inexpensive and low capacity modems.

To implement our solution, firstly, we analyse the slot structure using an intuitive diagrammatic representation to map the achievable channel utilisation levels. We then reformulate a Q-learning routine that exploits an implicit feedback signal to negatively reinforce and isolate reception slots in the slot selection phase. Secondly, by averaging the packet flow rate, we are able to generate a distribution for belief states that control and consolidate the choice of transmission slot to achieve overall network wide packet flow. We finally evaluate and demonstrate that ALOHA-QUPAF significantly outperforms the comparable protocols with similar Q-learning and slotting concepts.

Chapter 6

Future Work

This chapter presents a some additional related areas of future research that will potentially enhance and extend the work in this thesis.

6.1 Mathematical Translation of the Pictorial Analysis

In chapter 5 a pictorial analysis is employed to analytically describe utilisation levels and visualise the operation of a baseline random access scheme. Through a combination manual hop-by-hop tracing of packets along the network and simulation runs, the complete achievable utilisation performance can be established. However, whilst the approach is arguably effective, accessible and intuitive, the technique's efficiency and universality is sub par to a mathematical formulation. For example, only a few results of the analysis was given in Section 5.3.1, as an exhaustive picture is not feasible. Whereas mathematical equation(s) could concisely describe the entire idea. By translating the pictorial analysis in to the corresponding mathematical framing will comprehensively improve and convey the core idea and potentially facilitate better integration with additional performance metrics.

6.2 Power saving measures

Although the work presented in this thesis is principally concerned with applying intelligent scheduling to maximise utilisation efficiency of linear chain multi-hop UASNs. The significance of MAC protocol on energy efficiency is of major importance as well. Owing to the constraints of power supply and the need to deploy sensors for extended monitoring period, power efficient MAC protocols are critical to UASNs. The use of negative reinforcement in isolating slots can be exploited to incorporate power management techniques such as sleep cycles. For example, since ALOHA-QUPAF uses incoming packets as an aversive stimulus for the reception slots, by identifying the incoming and transmission slots all other slots can potentially be used for sleep modes.

6.3 Aloha-QUPAF Implementation on Mobile Nodes

Monitoring of subsea assets is the focused application scenario in this thesis. For example, the network is envisaged to be deployed by retro-fitting sensor nodes on/along a pipeline, consequently, ALOHA-QUPAF is designed and simulated on a quasi-stationary network. This enabled us make some reasonable assumptions: synchronisation and fixed transmission ranges. In other applications nodes could be intentionally mobile, such as an Autonomous Underwater Vehichle (AUV) or unitentionally displaced nodes by sea currents or marine life. Hence, global synchronisation and transmission ranges (thereby changing the interference population of neighborhood (s)) initially assumed are no longer the case. Given ALOHA-QUPAF is already light and simple with an aggressive no overheads design, it is possible to greatly enhance the capability of ALOHA-QUPAF to cover mobile nodes in a network by incorporating robust synchronisation and adaptive frame sizing techniques base on the prevalent interference population.

6.4 Practical Evaluation of ALOHA-QUPAF

Arguably, the evaluation of ALOHA-QUPAF in a simulation environment with preset or computed set of parameters and assumption is limited. Implementing and running the protocol in real underwater scenario will provide an essential validation. The empirical evaluation of ALOHA-QUPAF has shown remarkable improvement relative to the comparison schemes (ALOHA-Q and framed ALOHA). Practical tests will be helpful in evaluating how ALOHA-QUPAF performs when subjected to both hardware and other real world random variables. The test will additionally provide an insight on potential areas/variable modifications for improvements.

Chapter 7

Conclusion

The harsh and extremely dynamic underwater environment makes the application of adaptive and intelligent techniques an essential approach for efficient implementation of viable underwater acoustic sensor networks (UASNs). Despite sharing similar underlying concepts and application paradigm with terrestrial wireless sensor networks (WSNs), adoption of conventional schemes to the underwater networks have largely been found to be ineffective. Consequently, new strategies that takes into account the peculiar characteristics of underwater environment are required to enable the UASNs technologies. Reinforcement Learning (RL) is a promising approach that has been demonstrated to be powerfully robust and able to achieve good performance in WSN. RL exhibits distributed properties that eliminates the need for typically challenging and advance planning of resource allocations whilst intelligently self organise to efficiently manage the network resources. Given the majority of challenges faced by UASNs can be linked to the physics of the acoustic channel and limitations of current technologies, RL is well suited to equip MAC protocols with an effective mechanism of collisions resolution and flexible transmission capabilities specially needed for practical and successful operation of UASNs.

ALOHA-Q was a recently proposed MAC protocol that employs the RL paradigm in WSNs, despite the original incompatible assumption about propagation delay made in implementing ALOHA-Q, owing to its inherent simplicity and adaptability it has the potential to be a contender for mainstream adoption and in UASNs. However, ALOHA-Q rely on a crucial explicit feedback signal to drive the RL algorithm which can be unreliable underwater and further constraints the capacity and delay performance of the network thereby rendering the protocol inefficient.

The work undertaken in this thesis focused on devising a restructured time slot that eliminates overheads and an implicit feedback signal thereby improving the overall utilisation and delay in a linear chain UASNs. Firstly, two tight hop time slot structures based on the relationship between data packet duration and propagation delay was proposed to improve data to slot utilisation and delay. This is followed by a pictorial analysis of random access scheme to empirically evaluate the achievable utilisation and thus, the feasibility of utilising the new structure with practical gains in performance. Secondly, we proposed a two stage RL mechanism that uses a Q-learning algorithms for incoming packets as a negative feedback signal to reinforce arrival slots and a second one dimensional averaging function to continuously track packet flow in a given window. The proposed two stage strategy is driven by consistent stream of packets across the n effectively achieve collision free scheduling by isolating reception slots from

7.1 Original Contributions

The main contributions of this thesis are summarised as follows:

7.1.1 Adaptation of ALOHA-Q to Linear Chain UASNs

We implemented and evaluated the performance of ALOHA-Q in linear chain UASNs. The results obtained primarily demonstrated the impact of underwater parameters particularly the long propagation delay on the performance of the protocol under ideal conditions and asynchronously. The simulation additionally validates the relationship between utilisation and the number of slots per frame (frame size). While accurate predetermination of such number in an uncertain environment is challenging, however, since the target application scenario of this thesis is concerned with quasi-stationary nodes in monitoring network, devising dynamic frame adaptation mechanism will introduce an unnecessary complexity to the ALOHA-Q protocol. In a monitoring applications each node is a potential source/relay, therefore MAC protocols has to guarantee fair access, this is to avoid having few nodes from capturing the channel and starving others. The fairness of ALOHA-Q has not been studied in the previous works, and our evaluation has shown that in multiple source setting the protocol is not fair overall. By simply modifying the default FIFO queuing system with a RR the protocol is able to achieve 100% fairness in all settings.

7.1.2 New frame size for UASN

MAC protocols use an acknowledgement to signify successful communication session(s), the absence of which triggers a re-transmission. This is essential for applications that require guaranteed delivery, for most UASNs monitoring applications, however, are more concerned with prompt delivery of the most up to date data. This is because, freshly generated data is usually the most relevant in describing the current status of the system, in which case MAC protocol with best effort delivery is sufficient. Therefore, we proposed two aggressive no overhead time slot structures based on the ratio of packet duration to one hop propagation delay for maximal slot utilisation. This approach greatly improves on the existing state-of-the-art techniques that account for longest possible propagation delay in creating the frame size, such as [91, 92].

7.1.3 Application of Pictorial Analysis for Theoretical analysis

By using diagrammatical representation we visualise the operation of a random access scheme based on our proposed slot structure. This enables us to describe the idealised frame-by frame system transitions and traced the packets as they traversed the network. The utilisation describing the system can easily be computed at the designated sink node by observing the periodic pattern arrival that eventually manifest. Similarly, the average delay ca also be discerned for each slot selected pattern across the network. There is an excellent agreement between the simulation results and the analytical results when compared. The pictorial method provide an alternative approach to the conventional mathematical especially in conducting network analysis that may require advance techniques to describe.

7.1.4 Reception slot isolation using negative RL

The reformulation and application of Q-learning algorithm in MAC protocol has introduced a novel approach that solve the problems of excessive collision and low utilisation of framed ALOHA based protocols. Instead of nodes learning unique transmission slots through exploratory trial-and-error by receiving feedback of their actions, in this reversed strategy, streams of packets provide an aversive stimulus that continuously reinforces the reception slot(s) and therefore those slots will be effectively isolated and unlikely to be used as transmission slot during selection by the Q-algorithm. Since nodes can only receive or transmit at a time, the goal of this branch of the algorithm is to avoid collisions at the node level. Due to the new slot structure perfect scheduling can still be accomplished whilst two adjacent nodes in a particular interference population share same transmission slot. By making the RL signal reception-centric most of the data transfer uncertainty is improved and robust compared to the explicit ACK needed in [59, 74, 69].

7.1.5 Packet Flow Harmony

In a multi hop networks data flow along the route from source(s) to destination(s), thus, nodes are typically expected to forward data from other upstream nodes to downstream nodes. Essentially when a route is established in an active network, a consistent stream of data along the route can provide a fair indication of certain type of network stability, and perfect/collision free scheduling is attained when the measured flow across the network approximates the optimum channel capacity. Taking into account the initial learning stage of the Q-learning, we have proposed a second averaging function that will track the flow of data. A perfect scheduling will results in node independently recording flow rate that exceeds the threshold, otherwise some nodes will lag and hence indicates a flow disharmony somewhere in the network.

7.1.6 Revisiting the Hypothesis

The following is the hypothesis stated for work in this thesis:

"It is possible to effectively achieve the optimal (achievable utilisation) network performance by devising new time slots based on the relationship between packet duration and hop propagation delay, coupled with intelligent MAC scheduling using packets flow in lieu of explicit reward signal to drive a reinforcement learning algorithm."

We hereby summarise the main contributions presented in Section 7.1 of this thesis in contextualised to the above hypothesis in the following:

- The Chapter 4 evaluation of ALOHA-Q in UASNs provides proof of concept study for the feasibility and suitability of adopting and extending the RL paradigm in the design of MAC protocols underwater. RL has the capacity to enable adaptability in the dynamic underwater environment and potentially improve performance with the right time slots and feedback approach.

- Following the assumption of an interference range and the associated computation of the optimum frame duration. The proposed time slot structures in Chapter 5 aggressively reduce overheads based on a relationship between two components of communications: packet duration and hop propagation delay. When the packet duration is longer than the propagation delay the slot size incorporates the propagation delay, otherwise the slot size is approximated to the propagation delay. The design of these slots aims for the most maximum achievable in-slot utilisation in a flexible manner relative to the conventional slot structures employed in framed-ALOHA and ALOHA-Q.

- ALOHA-QUPAF is an intelligent protocol that employed dual control to achieve perfect scheduling without explicit feedback of interaction with the network. Packet arrivals provide the negative reward for the Q-learning algorithm to punish receptions slots, in this way the reception slots will be negatively reinforced and become unlikely to be selected by the Q-policy for transmissions, hence local level collision is eliminated. The slot selection component provide local level intelligent scheduling.

- Flow harmony is needed to fairly judge the overall condition of the network. Since no explicit feedback is received following each transmission, the success of the chosen transmission slot is difficult to ascertain. Therefore, subsequent flow of packets after each transmission is used as an implicit feedback signal to reinforce the transmission slots, however, a separate averaging function that tracks the overall flow with respect to the optimum channel utilisation is employed to validate and control the choice of current transmission slot over a given window. Flow harmony achieves global level scheduling.

The contributions outlined above have been empirically shown to dramatically achieve the optimum channel capacity, whilst intelligently maintaining adaptability at varying length, hence, proving the hypothesis of this thesis.

Appendix A

BELLHOP TABLE

Table A.1: BELLHOP GAIN DATA OF NIGER DELTA

SRC_INDEX	RX_ INDEX	CH_GAIN	CH_DELAY	SPREAD
1	2	-54.9	0.1323	0
1	3	-64.52	0.2744	0.0001
1	4	-82.44	0.4022	0
1	5	-93.6	0.5357	0
1	6	-104.07	0.669	0.8119
1	7	-114.3	0.805	0.7424
1	8	-123.89	0.9368	0.6827
1	9	-133.57	1.0732	0.6283
2	1	-54.9	0.1323	0
2	3	-53.2	0.1421	0.0005
2	4	-68.56	0.2699	0.0003
2	5	-82.55	0.4034	0
2	6	-93.67	0.5367	0.8893
2	7	-103.66	0.6727	0.0002
2	8	-114.58	0.8045	0

2	9	-124.19	0.9409	0.6815
3	4	-54.3	0.1279	0
3	2	-56.18	0.1421	0
3	5	-69.43	0.2614	0
3	1	-70.71	0.2744	0
3	6	-80.39	0.3946	0.0002
3	7	-91.57	0.5306	0.8935
3	8	-101.64	0.6625	0.0001
3	9	-111.31	0.7988	0.0001
4	3	-54.3	0.1279	0
4	5	-55.07	0.1336	0
4	6	-69.96	0.2668	0
4	2	-70.28	0.2699	0
4	1	-82.44	0.4022	0.9795
4	7	-82.24	0.4028	0.0002
4	8	-93.51	0.5346	0.8919
4	9	-104.22	0.671	0.8124
5	6	-54.11	0.1333	0.0007
5	4	-54.13	0.1336	0.0007
5	3	-69.43	0.2614	0
5	7	-70.21	0.2693	0
5	8	-82.35	0.4011	0
5	2	-82.55	0.4034	0
5	1	-93.6	0.5357	0
5	9	-93.74	0.5375	0
6	5	-55.03	0.1333	0
6	7	-55.39	0.136	0

6	4	-69.96	0.2668	0
6	8	-70.07	0.2679	0
6	3	-81.78	0.3946	0
6	9	-82.62	0.4042	0
6	2	-93.67	0.5367	0.8906
6	1	-104.07	0.669	0.8129
7	8	-54.84	0.1319	0
7	6	-55.39	0.136	0
7	9	-70.11	0.2682	0
7	5	-70.21	0.2693	0
7	4	-79.2	0.4028	0.0001
7	3	-91.2	0.5306	0.0001
7	2	-104.35	0.6727	0.8104
7	1	-114.3	0.805	0.7426
8	7	-54.84	0.1319	0
8	9	-55.44	0.1364	0
8	6	-66.92	0.2679	0.0003
8	5	-82.35	0.4011	0
8	4	-91.96	0.5346	0.0002
8	3	-103.57	0.6625	0
8	2	-108.64	0.8045	0.7432
8	1	-123.89	0.9368	0.6838
9	8	-55.44	0.1364	0
9	7	-70.11	0.2682	0
9	6	-82.62	0.4042	0
9	5	-93.74	0.5375	0.8896
9	4	-104.22	0.671	0.8114

9	3	-113.84	0.7988	0.7456
9	2	-124.19	0.9409	0.6817
9	1	-133.57	1.0732	0.629

References

[1] A Gkikopouli, G Nikolakopoulos, and S Manesis. “A survey on Underwater Wireless Sensor Networks and applications”. In: 2012.

[2] R Otnes et al. *Underwater Acoustic Networking Techniques*. Springer, 2012.

[3] J Guan et al. “The underlying design in underwater acoustic wireless sensor network”. In: *2013 IEEE International Conference of IEEE Region 10 (TENCON 2013)*. Oct. 2013, pp. 1–5.

[4] *Offshore production nearly 30% of global crude oil output in 2015 - Today in Energy - U.S. Energy Information Administration (EIA)*. `https://www.eia.gov/todayinenergy/detail.php?id=28492`. Accessed: 2018-9-28.

[5] Jude Clemente. “The Quiet Rise In U.S. Offshore Oil Production”. In: *Forbes Magazine* (Apr. 2018).

[6] H Karl and A Willig. *Protocols and Architectures for Wireless Sensor Networks*. Wiley, 2005.

[7] Mandar Chitre, Shiraz Shahabudeen, and Milica Stojanovic. “Underwater acoustic communications and networking: Recent advances and future challenges”. In: *Mar. Technol. Soc. J.* 42.1 (2008), pp. 103–116.

[8] J Heidemann et al. “Research challenges and applications for underwater sensor networking”. In: vol. 1. 2006.

[9] S Climent et al. “Underwater Acoustic Wireless Sensor Networks: Advances and Future Trends in Physical, MAC and Routing Layers”. In: *Sensors* 14.1 (2014), pp. 795–833.

[10] Antonio-Javier Garcia-Sanchez, Felipe Garcia-Sanchez, and Joan Garcia-Haro. “Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops”. In: *Comput. Electron. Agric.* 75.2 (Feb. 2011), pp. 288–303.

[11] W Dargie and C Poellabauer. *Fundamentals of Wireless Sensor Networks: Theory and Practice*. Wiley, 2010.

[12] João Martinho, Luís Prates, and João Costa. “Design and Implementation of a Wireless Multiparameter Patient Monitoring System”. In: *Procedia Technology* 17 (Jan. 2014), pp. 542–549.

[13] Kechar Bouabdellah, Houache Noureddine, and Sekhri Larbi. “Using Wireless Sensor Networks for Reliable Forest Fires Detection”. In: *Procedia Comput. Sci.* 19 (Jan. 2013), pp. 794–801.

[14] Sudipta Bhattacharjee et al. “Wireless sensor network-based fire detection, alarming, monitoring and prevention system for Bord-and-Pillar coal mines”. In: *J. Syst. Softw.* 85.3 (Mar. 2012), pp. 571–581.

[15] A Pascale et al. *Motorway speed pattern identification from floating vehicle data for freight applications*. 2015.

[16] Leonid Maksimovich Brekhovskikh. *Fundamentals of ocean acoustics*. Springer Science & Business Media, 2003.

[17] Yang Xiao. *Underwater acoustic sensor networks*. CRC Press, 2010.

[18] Almir Davis and Hwa Chang. “Underwater wireless sensor networks”. In: *2012 Oceans*. IEEE. 2012, pp. 1–5.

[19] M Chitre and W S Soh. “Reliable Point-to-Point Underwater Acoustic Data Transfer: To Juggle or Not to Juggle?” In: *IEEE J. Oceanic Eng.* 40.1 (2015), pp. 93–103.

[20] Affan A Syed et al. "Understanding spatio-temporal uncertainty in medium access with ALOHA protocols". In: *Proceedings of the second workshop on Underwater networks*. ACM, 2007, pp. 41–48.

[21] "Offshore Pipeline Market – 2019 Global Industry Analysis By Size, Growth, Merger, Share, Trends, Revenue, With Regional Forecast To 2024 - Reuters". In: *Reuters* ().

[22] James G Speight. *Handbook of Offshore Oil and Gas Operations*. en. Elsevier, Oct. 2014.

[23] Huacan Fang and Menglan Duan. *Offshore Operation Facilities: Equipment and Procedures*. en. Gulf Professional Publishing, Sept. 2014.

[24] Ahmed A Elshafey, MR Haddara, and H Marzouk. "Free spans monitoring of subsea pipelines". In: *Ocean Systems Engineering* 1.1 (2011), pp. 59–72.

[25] Subsea Asia 2015-Jakarta. "Structural Monitoring of Subsea Pipelines and Role in Reducing Mitigation Costs". In: ().

[26] C Mai et al. "Subsea infrastructure inspection: A review study". In: *2016 IEEE International Conference on Underwater System Technology: Theory and Applications (USYS)*. Dec. 2016, pp. 71–76.

[27] Nord Stream Ag. *Maintenance - Nord Stream AG*. `https://www.nord-stream.com/operations/maintenance/`. Accessed: 2019-11-4.

[28] Cheng Hong et al. "Subsea production layout: design and cost". In: *International Conference on Offshore Mechanics and Arctic Engineering*. Vol. 57694. American Society of Mechanical Engineers. 2017, V05AT04A053.

[29] Elena Gaura et al. *Wireless Sensor Networks: Deployments and Design Frameworks*. en. Springer Science & Business Media, 2010.

[30] Yueh-Min Ray Huang. *Sensors: Advancements in Modeling, Design Issues, Fabrication and Practical Applications*. en. Springer Science & Business Media, 2008.

[31] Javier Poncela-Gonzalez et al. "Investigation of Underwater Acoustic Modems: Architecture, Test Environment & Performance". In: (2016).

[32] S Sendra et al. “Underwater Acoustic Modems”. In: *IEEE Sens. J.* 16.11 (2016), pp. 4063–4071.

[33] Antonio Sánchez et al. “An ultra-low power and flexible acoustic modem design to develop energy-efficient underwater sensor networks”. en. In: *Sensors* 12.6 (2012), pp. 6837–6856.

[34] G Cario et al. “SeaModem: A low-cost underwater acoustic modem for shallow water communication”. In: 2015.

[35] L Wu et al. “Designing an Adaptive Acoustic Modem for Underwater Sensor Networks”. In: *IEEE Embedded Sys. Lett.* 4.1 (Mar. 2012), pp. 1–4.

[36] Ethem Mutlu Sözer and Milica Stojanovic. “Reconfigurable acoustic modem for underwater sensor networks”. In: *Proceedings of the 1st ACM international workshop on Underwater networks - WUWNet ’06*. 2006.

[37] N Nowsheen, C Benson, and M Frater. “A high data-rate, software-defined underwater acoustic modem”. In: *OCEANS 2010 MTS/IEEE SEATTLE*. 2010, pp. 1–5.

[38] M S Martins et al. “High data rate acoustic modem for underwater aplications”. In: 2014.

[39] L Freitag et al. “The WHOI micro-modem: an acoustic communications and navigation system for multiple platforms”. In: *Proceedings of OCEANS 2005 MTS/IEEE*. 2005, 1086–1092 Vol. 2.

[40] *Underwater Acoustic Modem 6 - Sonardyne*. `https://www.sonardyne.com/product/underwater-acoustic-modems/`. Accessed: 2020-7-2.

[41] *Acoustic Modems — EvoLogics*. `https://evologics.de/acoustic-modems`. Accessed: 2020-7-2.

[42] James Kurose and Keith Ross. *Computer Networking: A Top-Down Approach, Global Edition*. en. Pearson Higher Ed, 2017.

[43] Tanenbaum. *Computer Networks*. en. Pearson Education(singapore) Pte. Limited, 2011.

[44] Larry L Peterson. *Computer Networks - A Systems Approach 3rd Edition*. en.

[45] K Chen et al. “A Survey on MAC Protocols for Underwater Wireless Sensor Networks”. In: *IEEE Communications Surveys & Tutorials* 16.3 (2014), pp. 1433–1447.

[46] M Molins and M Stojanovic. “Slotted FAMA: a MAC protocol for underwater acoustic networks”. In: 2006.

[47] S Jiang. “State-of-the-Art Medium Access Control (MAC) Protocols for Underwater Acoustic Networks: A Survey Based on a MAC Reference Model”. In: *IEEE Communications Surveys Tutorials* 20.1 (2018), pp. 96–131.

[48] Shengming Jiang. *Wireless Networking Principles: From Terrestrial to Underwater Acoustic*. en. Springer, Apr. 2018.

[49] S Shahabudeen, M Chitre, and M Motani. “MAC protocols that exploit propagation delay in underwater networks”. In: 2011.

[50] Y Noh and S Shin. “Survey on MAC protocols in Underwater Acoustic Sensor Networks”. In: 2014.

[51] N. Abramson. “Development of the ALOHANET”. In: *IEEE Transactions on Information Theory* 31.2 (1985), pp. 119–123. ISSN: 0018-9448. DOI: 10.1109/TIT.1985.1057021.

[52] M S Gao et al. *A Multi-Channel MAC Protocol for Underwater Acoustic Networks*. 2015 Ieee 20th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks. New York: Ieee, 2015.

[53] Phil Karn and Others. “MACA-a new channel access method for packet radio”. In: *ARRL/CRRL Amateur radio 9th computer networking conference*. Vol. 140. 1990, pp. 134–140.

[54] Vaduvur Bharghavan et al. “MACAW: a media access protocol for wireless LAN’s”. In: *SIGCOMM Comput. Commun. Rev.* 24.4 (Oct. 1994), pp. 212–225.

[55] H H Ng, W S Soh, and M Motani. “MACA-U: A Media Access Protocol for Underwater Acoustic Networks”. In: 2008.

[56] Chane L Fullmer and J J Garcia-Luna-Aceves. “Floor Acquisition Multiple Access (FAMA) for Packet-radio Networks”. In: *SIGCOMM Comput. Commun. Rev.* 25.4 (Oct. 1995), pp. 262–273.

[57] N Chirdchoo, W s. Soh, and K C Chua. “RIPT: A Receiver-Initiated Reservation-Based Protocol for Underwater Acoustic Networks”. In: *IEEE J. Sel. Areas Commun.* 26.9 (Dec. 2008), pp. 1744–1753.

[58] Y Han and Y Fei. “TARS: A Traffic-Adaptive Receiver-Synchronized MAC Protocol for Underwater Sensor Networks”. In: 2015.

[59] Yi Chu et al. “Application of reinforcement learning to medium access control for wireless sensor networks”. In: *Engineering Applications of Artificial Intelligence* 46.A (Nov. 2015), pp. 23–32. ISSN: 0952-1976. DOI: 10.1016/j.engappai.2015.08.004.

[60] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[61] M L Littman and A W Moore. “Reinforcement learning: A survey”. In: *Journal of artificial intelligence* (1996).

[62] M Stojanovic and J Preisig. “Underwater acoustic communication channels: Propagation models and statistical characterization”. In: *IEEE Commun. Mag.* 47.1 (2009), pp. 84–89.

[63] *Oceans: The Great Unknown*. https://www.nasa.gov/audience/forstudents/5-8/features/oceans-the-great-unknown-58.html. Nasa.gov, 2009.

[64] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning, second edition*. en. MIT Press, Dec. 2018.

[65] A D Tijsma, M M Drugan, and M A Wiering. “Comparing exploration strategies for Q-learning in random stochastic mazes”. In: *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. Dec. 2016, pp. 1–8.

[66] Caroline Claus and Craig Boutilier. "The dynamics of reinforcement learning in cooperative multiagent systems". In: *AAAI/IAAI* 1998.746-752 (1998), p. 2.

[67] Spiros Kapetanakis and Daniel Kudenko. "Reinforcement learning of coordination in cooperative multi-agent systems". In: *AAAI/IAAI* 2002 (2002), pp. 326–331.

[68] Tao Jiang et al. "Single-state Q-learning for self-organised radio resource management in dual-hop 5G high capacity density networks". In: *Trans. Emerging Tel. Tech.* 27.12 (Dec. 2016), pp. 1628–1640.

[69] Y Yan et al. "Adaptation of the ALOHA-Q protocol to Multi-Hop Wireless Sensor Networks". In: *European Wireless 2014; 20th European Wireless Conference*. May 2014, pp. 1–6.

[70] Craig Boutilier. "Sequential optimality and coordination in multiagent systems". In: *IJCAI*. Vol. 99. 1999, pp. 478–485.

[71] Vikram Krishnamurthy. *Partially Observed Markov Decision Processes*. en. Cambridge University Press, Mar. 2016.

[72] Mohssen Mohammed, Muhammad Badruddin Khan, and Eihab Bashier Mohammed Bashier. *Machine Learning: Algorithms and Applications*. en. CRC Press, Aug. 2016.

[73] Frans A Oliehoek and Christopher Amato. *A Concise Introduction to Decentralized POMDPs*. Springer, Cham, 2016.

[74] Tautvydas Mickus, Paul Daniel Mitchell, and Timothy Clarke. "The emergence MAC (E-MAC) protocol for wireless sensor networks". In: *Engineering Applications of Artificial Intelligence* 62 (June 2017). © 2017 Elsevier Ltd. This is an author-produced version of the published paper. Uploaded in accordance with the publisher's self-archiving policy., pp. 17–25. ISSN: 0952-1976. DOI: `10.1016/j.engappai.2017.03.003`.

[75] *Riverbed Modeler*. `https://www.riverbed.com/gb/products/npm/riverbed-modeler.html`. Accessed: 2020-12-22.

[76] Adarshpal S Sethi and Vasil Y Hnatyshin. *The practical OPNET user guide for computer network simulation*. CRC Press, 2012.

[77] Zheng Lu and Hongji Yang. *Unlocking the power of OPNET modeler*. Cambridge University Press, 2012.

[78] T Rappaport. "Wireless communications principles and practice edition". In: (2001).

[79] S Shahabudeen, M Chitre, and M Motani. "Adaptive Multimode Medium Access Control for Underwater Acoustic Networks". In: *IEEE J. Oceanic Eng.* 39.3 (July 2014), pp. 500–514.

[80] A Iyer, C Rosenberg, and A Karnik. "What is the right model for wireless channel interference?" In: *IEEE Trans. Wireless Commun.* 8.5 (May 2009), pp. 2662–2671.

[81] Yi Shi et al. *How to correctly use the protocol interference model for multi-hop wireless networks*. 2009.

[82] Guowang Miao et al. *Fundamentals of Mobile Data Networks*. en. Cambridge University Press, Mar. 2016.

[83] R Jain, D Chiu, and W Hawe. "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems, DEC Research Report TR-301". In: *Digital Equipment Corporation* (1984).

[84] *USMART - Newcastle University*. `https://research.ncl.ac.uk/usmart/newsevents/october2018v3modemsreadyforfieldtesting.html`. Accessed: 2019-10-17.

[85] P Mandal, S De, and S S Chakraborty. "Characterization of Aloha in underwater wireless networks". In: *2010 National Conference On Communications (NCC)*. Jan. 2010, pp. 1–5.

[86] D E Lucani, M Stojanovic, and M Medard. "On the Relationship between Transmission Power and Capacity of an Underwater Acoustic Communication Channel". In: *OCEANS 2008 - MTS/IEEE Kobe Techno-Ocean*. Apr. 2008, pp. 1–6.

[87] S Sen et al. “Analysis of Slotted Aloha with multipacket messages in clustered surveillance networks”. In: *MILCOM 2012 - 2012 IEEE Military Communications Conference*. Oct. 2012, pp. 1–6.

[88] R T B Ma, V Misra, and D Rubenstein. “An Analysis of Generalized Slotted-Aloha Protocols”. In: *IEEE/ACM Trans. Netw.* 17.3 (June 2009), pp. 936–949.

[89] Herbert Robbins and Sutton Monro. “A stochastic approximation method”. In: *Herbert Robbins Selected Papers*. Springer, 1985, pp. 102–109.

[90] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. en. OUP USA, 1999.

[91] Nils Morozs, Paul D Mitchell, and Yuriy Zakharov. *Linear TDA-MAC: Unsynchronized Scheduling in Linear Underwater Acoustic Sensor Networks*. 2019.

[92] Lawrence G. Roberts. “ALOHA Packet System with and Without Slots and Capture”. In: *SIGCOMM Comput. Commun. Rev.* 5.2 (Apr. 1975), pp. 28–42. ISSN: 0146-4833. DOI: 10.1145/1024916.1024920. URL: http://doi.acm.org/10.1145/1024916.1024920.