An IoT enabled system for marine data acquisition and cartography

Rabab Al-Zaidi

Supervised by: Dr. John Woods

A thesis submitted for the degree of PhD

Department of Computer Science and Electronic Engineering

University of Essex

June 2018

Abstract

Traditional marine monitoring systems such as oceanographic and hydrographic research vessels use either wireless sensor networks with a limited coverage, or expensive satellite communication that is not suitable for small and mid-sized vessels. This thesis proposes an Internet of Marine Things data acquisition and cartography system in the marine environment using Very High Frequency (VHF) available on the majority of ships. The proposed system is equipped with sensors such as sea depth, temperature, wind speed and direction, and the collected data is sent through a Ship Ad-hoc Network (SANET) to 5G edge clouds connected to sink/base station nodes on shore. The sensory data is ultimately aggregated at a central cloud on the internet to produce up to date cartography systems. Several observations and challenges unique to the marine environment have been discussed and feed into the solutions presented. We have investigated the application of appropriate data quantization and compression techniques to the marine sensor data collected in order to reduce the size of transmitted data and achieve better transmission efficiency. The impact of marine sparsity on the network is examined and a marine Mobile Ad-hoc/Delay Tolerant hybrid routing protocol (MADNET) is proposed to switch automatically between Mobile Ad-hoc

Network (MANET) and Delay Tolerant Network (DTN) routing according to the network connectivity. The low rate data transmission offered by VHF radio has been investigated in terms of the network bottlenecks and the data collection rate achievable near the sinks. A sensory data management and transmission approach has also been proposed at the 5G network core using Information Centric Networks (ICN) aimed at providing efficient and duplicate less transmission of marine sensory readings from the base station/sink nodes towards the central cloud. Therefore, SANETs are realized as part of a 5G infrastructure for marine environment monitoring, paving the way to the Internet of Marine Things (IoMaT).

Acknowledgements

I would like to thank Almighty God and everyone that helped exceedingly in completing this thesis.

A special gratitude goes to my supervisor Dr. John Woods, for his patience, guidance, enthusiastic encouragement and useful critiques at all stages of this work, and also for encouraging me to published my work in high rank publication venues. Without him I believe this thesis would have been very different.

My special thanks is extended to the staff and colleagues at the School of Computer Science and Electronic Engineering (CSEE) for the enjoyable atmosphere and for their help, support, and perfect companionship.

I acknowledge the people who mean a lot to me, my parents, Mother and Father, for showing faith in me and giving me liberty to choose what I desire. I salute you both for the selfless love and care you provided to shape my life.

I owe special thanks to a very dear person, my loving husband Mohammed, who provided unending inspiration, support and understanding during my journey to get the Ph.D degree, which made the completion of this thesis possible. He was always around at times I thought that it was impossible to continue, he helped me to keep things in perspective. I greatly value his contribution and deeply appreciate his belief in me. Deepest gratitude to my children, Aws and Aya for the patience they showed during my PhD. Words would never say how grateful I am to all of you. I consider myself the luckiest woman in the world to have such a lovely family, standing beside me with their love and unconditional support.

Contents

List of Figures

List of Tables

Acronyms

AIMRP	Address light Integrated MAC and Routing Protocol
AIS	Automatic Identification System
AMDC	(Average Marine Data Compression
AODV	Ad-hoc On-Demand Distance Vector
AOMDV	Ad-hoc On-Demand Multipath Distance Vector
ASOTDMA	Ad-hoc Self-Organizing TDMA
BER	Bit Error Rate
BSW	Binary Spray and Wait protocol
CDNs	Content Distribution Networks
CDMA	Code Division Multible Access
CS-TDMA	Carrier Sensing TDMA
DSDV	Destination-Sequenced Distance Vector
DTNs	Delay Tolerant Networks
DYMO	Dynamic MANET On-demand
FID	Forwarding IDentifier
GPRS	General Packet Radio Service

GSM	Global System for Mobile Communications
GAF	Geographic Adaptive Fidelity
ICN	Information Centric Networks
IMO	International Maritime Organization
IoUT	Internet of Underwater Things
IoT	Internet of Things
IoMaT	Internet of Marine Things
ITU	International Telecommunication Union
MADNET	Marine Mobile Ad-hoc,Delay Tolerant Network
MANET	Mobile Ad-hoc Networks
M2M	Machine to Machine
MEG	Mobile Edge Computing
MWMNs	Maritime Wireless Mesh Networks
MTN	Maritime Tactical Networks
NAF	Network Abstraction Function
NMEA	National Marine Electronic Associate
Ofcom	Office of communications
OLSR	Optimized Link State Routing
PC	Personal Computer
PDR	Packet Delivery Ratio
PURSUIT	Publish Subscribe Internet Technology
RREQ	Route Request
RREP	Route Replay
RV	Rendezvous

RWM	Random Waypoint Mobility model
SANET	Ship Ad-hoc Network
SW	Spray and Wait protocol
SPAN	Smart Phone Ad-hoc Networks
TDMA	Time Division Multiple Access
TCL	Tool Command Language
TM	Topology Manager
UMNs	Unstructured Mobile Networks
VHF	Very High Frequency
VANET	Vehicular Ad-hoc Network
VBR	Variable Bit Rate
WMN	Wireless Mobile Network
WiMAX	Worldwide Interoperability for Microwave Access

Chapter 1

Introduction

Over the past decades, scientists have made every effort to study the weather and climate of the earth. However, it is not an easy task. For example, to measure the earths temperature and to characterize its warming speed in vast seas and oceans; temperature thermometers usually provide very accurate readings but they can not cover the aforementioned water surfaces. Satellites are widely used today to extract temperatures from radiative emission at microwave frequencies from oxygen in the atmosphere, but the contaminants in the atmosphere may corrupt the measurement, such as water droplets (either in clouds or precipitation) that can influence the temperature readings, causing satellite trends to be off (too cold) by up to 30 Percent [2]. Therefore, other solutions are required in order to provide more accurate readings and future estimates of the earth's climate. One favourable approach could be the exploitation of ship navigation paths and buoy deployments for collecting climate data.

Marine communication systems available today only provide the bare minimum essential services such as ship identification, positioning, location, course, heading, destination, tonnage, speed in the form of AIS (Automatic Identification System) using VHF radio frequencies. Inter-ship satellite communication is possible but is a costly option when compared to conventional wireless communications and not affordable by most small to medium seagoing vessels [3]. On the other hand, wireless Networks are not easily supported as there is no ready infrastructure in the marine environment to facilitate such an approach. The radio frequency spectrum allocated to maritime applications remains an essential limitation for network connectivity. The International Telecommunication Union (ITU) has defined marine band VHF radio to operate on internationally agreed frequencies in the band from 156 MHz to 163 MHz [4]. Achieving Ad-hoc Networks in the marine environment requires considering the characteristics of radio signal propagation at sea, and the need for repeaters and multi hop relays to achieve the Ad-hoc Networking objectives. The system design relies on the accurate modelling of the radio propagation.

The ITU has proposed Time Division Multiple Access (TDMA) as the channel access method for ship data communication over VHF channels. TDMA allows a number of users to use the same frequency channel by dividing the signal into several different time slots. However, TDMA cannot be directly applied to multi-hop Networks because its only suitable for single hop transmission between ships and base stations where ships can use it to send essential information (e.g., vessel number, location, and IMO number) [5]. Advances in the time slot allocation process to accommodate Ad-hoc Networks have been proposed including the the Ad-hoc Self-Organizing TDMA (ASOTDMA) protocol [5].

Mobile Ad-hoc Networks (MANETs) are a popular telecommunication technology that can easily be applied to almost any environment having fast configuration and no need for any underlying infrastructure support [6]. A MANET is a system of mobile nodes which are connected by wireless links. Each of these nodes acts as a router and is free to move in any direction independently. These characteristics are favourable to the marine environment, yet this environment possesses unique properties that differentiate its routing needs. When ships are moving, links can be obstructed by intervening objects. In addition, ship movement in sparse areas will likely lead to link disconnections due to communication range limitations. These events result in intermittent connectivity. For this reason, applications in the marine environment must tolerate delays beyond conventional IP forwarding delays, and these Networks are referred to as Delay Tolerant Networks (DTNs). MANET Routing protocols do not work properly in sparse areas because when packets arrive and no contemporaneous end-to-end paths for their destinations can be found, these packets are simply dropped [7]. On the other hand, DTN routing protocols do not work properly in dense areas, since they generate a high volume of packet copies and generate an unnecessary burden on the precious VHF link. Therefore, since the marine environment can be characterized as a mix between dense and sparse areas together, an alternative routing approach is needed that can fit the described scenarios.

1.1 Motivation

Marine cartography systems aim to provide a seamless, consistent and standardised database of marine climatic, environmental and navigational data using photogrammetric, sensing or laser scanning based approaches. However, the methods exploited so far are either remote monitoring based methods such as satellites which lack the required accuracy due to contaminants in the atmosphere that may corrupt the measurements [2]; or exploration based methods such as research vessels including hydrographic and oceanographic ships that can not cover the vast areas of oceans and seas due to the physical constrains, and limited number of deployed research vessels. Therefore, the marine cartography systems available usually either lack the required level of accuracy, geographic comprehensiveness or recency of the data samples. This necessitates exploring new alternative methods that can address these drawbacks, especially with the tremendous growth in data communications and the advancements in data analytics. For example, the emerging Internet of Things (IoT) aims to inter-connect physical devices, vehicles, home appliances and any other object with built-in sensors to gather data and take action on that data across a network. In preparation for the revolutionary 5G Networks, the Office of communications (Ofcom) has allocated more radio spectrum for the Internet of Things (IoT), specifically VHF spectrum aiming to encourage Machine to Machine (M2M) applications to use spectrum that will enable them to connect wirelessly over longer distances. This VHF spectrum has different properties to other frequencies, already in use for the IoT, and can reach distant locations which other frequencies may not [8]. This provides the motivation of this work to exploit marine band VHF for IoT applications and trigger the Internet of Marine Things (IoMaT).

1.2 Research methodology and relationship between chapters

The research in this thesis was conducted via simulation using the NS2 simulator, which is a discrete open source event simulator targeted at networking research. NS2 was chosen due to its broad range of routing protocols available for comparative evaluation. To set up the marine VHF data transmission environment, we input the actual transmission parameters that comply with the International Telecommunication Union (ITU) standards into the physical layer of our nodes (ships) to achieve realistic modelling of the marine environment. The experiments are based on real ship locations, speed and direction in the English channel and North sea extracted from the AIS data website in [9] for cargo vessels, passenger vessels, tankers and fishing boats.

Figure 1.1 provides an overview of the chapters contributions and the relationship between them. In Chapter 3, the application of MANETs in the marine environment is presented with the proposed physical, MAC, and Network layer protocols. The conducted evaluations and their results, form the basis for the proposed hybrid MANET/DTN routing protocol (MADNET) presented in Chapter 5. Chapter 4 discusses the application layer considerations and presents a quantization and compression method specific to the marine sensory data. Finally, Chapter 6 provides an overview of the IoMaT system and discusses the method proposed to collect the marine sensory data at the network core.

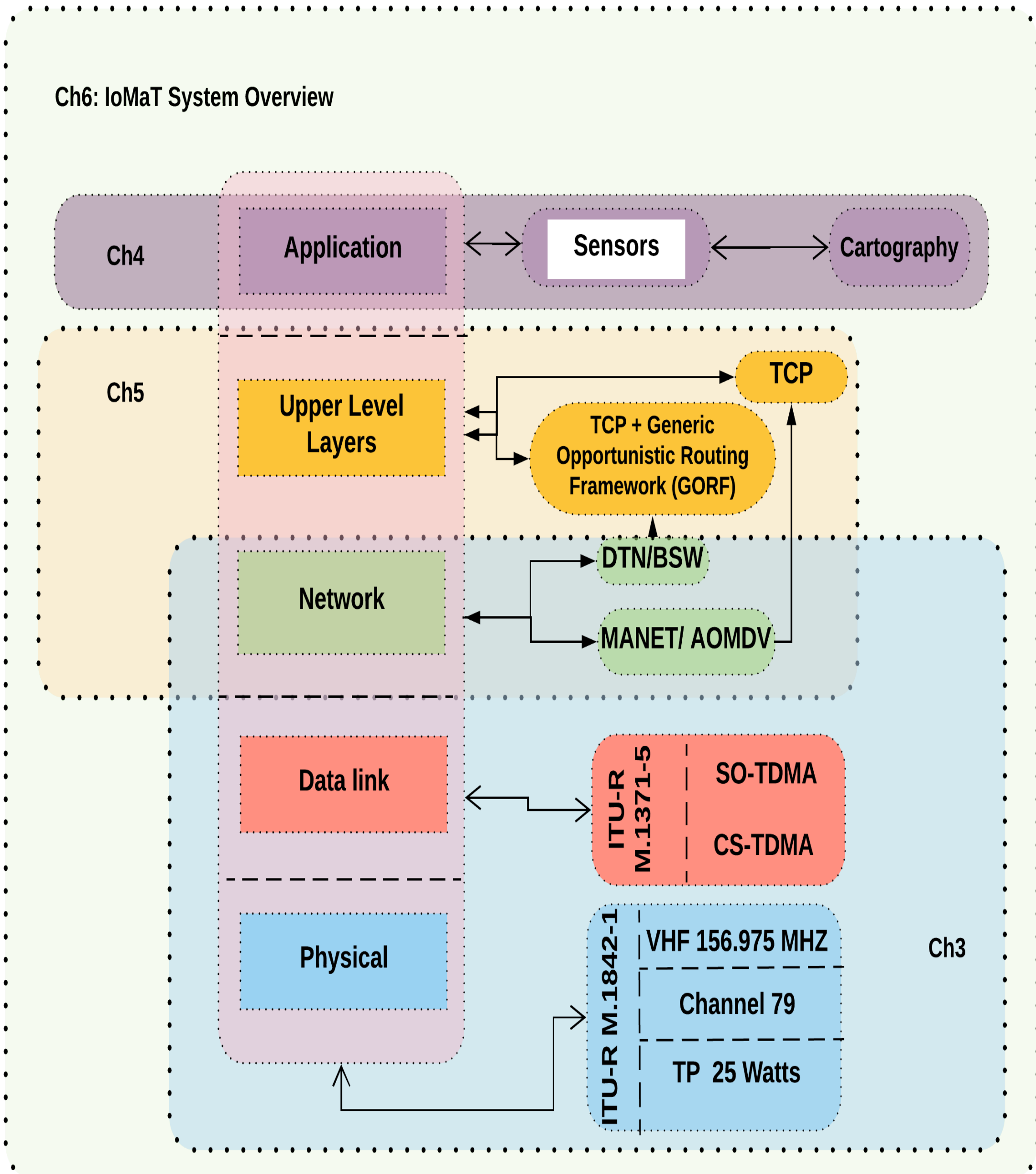


Figure 1.1: The relationship between chapters

1.3 Contributions of the Thesis

During the course of this thesis a number of significant contributions were made, as outlined below:

- This work has proposed the novel application of Ship Ad-hoc Networks (SANETs) in the marine environment as part of the Internet of Marine Things (IoMaT) and explores the use of MANET routing protocols to achieve a low cost AIS type system using existing on-board equipment. Implementing SANETs in the marine environment has been proven to be possible using the existing VHF communication infrastructure available on ships.

- The impact of ship density and mobility on SANETs has been investigated using real maritime traffic patterns. The performance of SANETs was evaluated using three different MANET routing protocols which are Ad-hoc On-Demand Distance Vector (AODV), Ad-hoc On-Demand Multipath Distance Vector (AOMDV) and Destination-Sequenced Distance Vector (DSDV) routing protocols.

- A novel marine MANET/DTN hybrid routing protocol (MADNET) has been proposed and evaluated. An important aspect of it, is that it combines synchronous (Ad-hoc) and asynchronous (opportunistic) delivery mechanisms. Furthermore, it is able to switch between the two different forwarding techniques according to the network conditions encountered and thus outperforms the other approaches in the marine environment.

- Novel data quantization and compression techniques specific to the marine sensor data collected have been proposed and evaluated in order to reduce the burden on the VHF channel links and achieve better transmission efficiency.

- The low rate data transmission offered by VHF radio has been investigated in terms of the SANET bottleneck links and the data collection rate per ship achievable near the sinks. Evaluation results show that despite the low data rate of VHF channels, high volumes of sensory data per ship have been observed at the bottleneck links.

- A management and transmission approach specific for marine sensory data has been proposed at the IoMaT network core using Information Centric Networks (ICN). This allows for efficient data synchronization between edge and central clouds and elevates any possibility of duplicated data transmission.

- A system prototype has been designed that integrates the IoMaT Software module proposed to be installed on a Personal Computer (PC) at each ship. The modules internal components have been described in addition to their roles and interactions with each other and with the main AIS system. The proposed data format of the new IoMaT packet has also been described and how the packet can be embraced by the AIVDM/AIVDO protocol that defines the AIS packet format.

1.3.1 List of Publications

- Rabab Al-Zaidi, John Woods, "Performance evaluation of MANET routing protocols in a maritime environment" , Computer Science and Electronic Engineering Conference (CEEC) IEEE, 2014.

- Rabab Al-Zaidi, John Woods and Mohammed Q. Shawkat, "Density and Mobility Impact on MANET Routing Protocols in a Maritime Environment" , Science and Information Conference (SAI) IEEE, 2015.
- Rabab Al-Zaidi, John Woods and Mohammed Q. Shawkat, "(AMDC) Algorithm for wireless sensor networks in the marine environment" , International Journal of Advanced Computer Science and Applications(IJACSA), (6), 2015.
- Rabab Al-Zaidi, John Woods, Mohammed Al-Khalidi, Huosheng Hu, "An IOT-enabled System for Marine Data Aquisition and Cartography", Transactions on Networks and Communications, (5), 2017.
- Rabab Al-Zaidi, John Woods, Mohammed Al-Khalidi, Khattab M. Ali and Klaus McDonald-Maier, "Next Generation Marine Data Networks in an IoT Environment", IEEE Conference on Fog and Mobile Edge Computing (FMEC), IEEE, 2017.
- Rabab Al-Zaidi, John C. Woods, Mohammed Al-Khalidi and Huosheng Hu, "Building Novel VHF-based Wireless Sensor Networks for the Internet of Marine Things", IEEE Sensors Journal, 2018.

1.4 Thesis Outline

The remainder of this chapter describes the layout of the thesis, which is organised into seven further chapters.

- Chapter 2 presents a background on the marine communication technology, outlining the components of the system such as the marine band VHF radio, the AIS system and the marine channel access method. It also provides an overview of Unstructured Mobile Networks and the routing protocols suitable for the marine environment including MANET and DTN routing protocols. Furthermore, it presents a comprehensive review of the research efforts related to the contributions of this thesis and aimed at realizing data networks at sea.

- Chapter 3 presents the novel application of MANETs in the marine environment and explores the use of MANET routing protocols to achieve a low cost AIS type Ship Ad-hoc Network (SANET) using existing on-board equipment. It also investigates the impact of marine traffic patterns on the performance of MANET routing protocols in the marine environment. The solution is aimed at mid to small size vessels and uses VHF technology available on the majority of ships.

- Chapter 4 presents a quantization and compression method specific to the marine sensory data in order to minimize transmission data redundancy overhead for efficient use of the transmission channel. The predictability of gathered sensor data makes it beneficial to quantize and compress the data to reduce the amount of bits needed to represent each reading in the binary representation.

- Chapter 5 discusses the behaviour of SANETs observed from the evaluations conducted in Chapter 3, high lighting the problem of sparsity which necessitates the inclusion of selective opportunistic and MANET routing. It presents the proposed hybrid MANET/DTN routing protocol (MADNET) with evaluations in both dense and sparse scenarios. It also investigates the drawbacks of low rate data transmission offered by VHF radio and the achievable data rate collected at the sink for each ship within the network.

- Chapter 6 presents the proposed marine cartography system as the realization of 5G IoMaT connectivity. It proposes a data management and transmission approach at the 5G Network core using ICN. This is aimed at providing efficient and duplicate-less transmission of marine sensory readings from the base station/sink nodes towards the central cloud in order to preserve Network resources and guarantee efficient delivery. The chapter also presents an overview of the cartography system architecture in the form of four distinctive functional layers. An IoMaT software module design is also presented with a discussion of the AIS system level considerations and how the IoMaT software components are proposed to interact with the existing system. The existing AIS message format is also discussed with the proposed extensions for IoMaT support.

- Finally, in Chapter 7, the conclusions drawn are highlighted, and areas for possible future study are suggested.

Chapter 2

Background Study and Literature Review

2.1 Introduction

Data networks at sea are an increasing but at the same time challenging application. There are a number of obstacles facing current marine wireless networking, such as lack of network coverage and sparsity. Available and reliable solutions such as satellite communication still remain very expensive and impractical for most small to mid size vessels. In the literature, only a few related works can be found in the area of marine IoT, especially away from the shore. In this chapter, we provide an insight into the marine radio transmission technology, the routing protocols suitable for the marine environment and the most important efforts in the literature that attempt to solve the problems of data networks at sea. We analyse and summarize each of these efforts and explain their contributions and shortcomings leading to the conclusions drawn and the solutions proposed in the following chapters.

2.2 The Marine Communication System

This section explains the essential components of the marine communication system that is mainly designed to support ship-to-land and ship-to-ship communication and provide an aid to ship navigation, location awareness and collision avoidance.

2.2.1 Marine-Band VHF Radio

Marine-Band VHF [10] (Very High Frequency) describes a particular part of the radio spectrum allocated to boats and ships. Marine band VHF radio operates on internationally agreed frequencies between 156 MHz and 163 MHz. It provides a means for vessels to communicate with each other and with shore stations (e.g. ports, locks, bridges and marinas) about operational, navigation and safety matters. A range of VHF radio channels exists, with specific permitted use in each case; for example, the channels allocated to port operations and ship movement may only be used on matters concerning navigation, vessel operations and safety. Other channels are allocated for ship to ship use, safety transmissions and contacting marinas [10]. The system is effective over distances, up to 100 km at sea in ideal circumstances. The transmission range achieved depends on several criterions, such as the antenna height, transmitting power, receiver sensitivity, and distance to the horizon. The propagation of VHF signals is near line-of-sight in ordinary conditions although atmospheric ducting is possible under certain conditions; VHF radio waves at these frequencies are bent back slightly toward the Earth by changes in atmospheric density. Accordingly, the distance to the radio horizon is extended slightly over the geographic line-of-sight to the horizon. A formula to calculate the earth line-of-sight horizon distance is given below [11]:

$$d = \sqrt{12.746 * hb} \tag{2.1}$$

where hb is the height in meters. In communications systems, to assess the probable coverage area of a proposed transmitter station, a more complex calculation is required [11].

2.2.2 The Automatic Identification System (AIS)

AIS [12] is a VHF radio broadcasting system that transfers packets of data over a VHF Data Link (VDL) and enables vessels and shore-based stations to send and receive identification information that can be displayed on an electronic chart, or compatible radar. This information can help in location awareness, collision avoidance and provide an aid to navigation, by displaying emergency buoys and beacons. AIS is designed with the capability to provide information about the ship to other ships and coastal authorities automatically. It is a Self Organized Time Division Multiple Access system representing the core technology of the communication and navigation equipment. AIS can automatically and continuously issue static information (ship name, length, breadth, ship types, etc), dynamic information (real-time position, speed, course, heading, navigation state, etc.) and voyage information (draft, port of destination, estimated time of arrival, etc.). It also automatically receive this information from neighbouring ships.

The International Maritime Organization (IMO) requires AIS to be installed aboard all ships of 300 gross tonnes and upwards engaged in international voyages, cargo ships of 500 gross tonnes and upwards not engaged in international voyages and all passenger ships irrespective of size. The requirement became effective for all ships by 31 December 2004 [13]. In efforts to standardize VHF data network communication at sea, the ITU (The International Telecommunication Union) has defined Recommendation ITU-R M.1842-1 'The Characteristics of VHF Radio Systems and Equipment for the Exchange of Data and Electronic Mail in the Maritime Mobile Service Radio Regularization (RR) Appendix 18 Channels' [14]. They also provide a guideline on the use of digital technologies by VHF systems of different bandwidths.

2.2.3 Marine Channel Access Method

Time Division Multiple Access (TDMA) is the channel access method proposed by the ITU for ship data communication over VHF channels. TDMA allows a number of users to use the same frequency channel by dividing the signal into several different time slots. The users transmit in rapid sequence, one by one, each using its unique time slot. This allows several ships to share the same transmission medium while consuming only a part of the channel capacity. There are two variants of TDMA for the marine communication system as proposed by ITU Recommendation ITU-R M.1371-4 [15], which are Self Organizing-TDMA (SO-TDMA) and Carrier-Sensing TDMA (CS-TDMA). However, they cannot simply be applied to Ship Ad-hoc Networks directly because they are only suitable for single hop transmission between ships and base stations where ships can use it to send essential information (e.g., vessel number, location, and IMO number) [5].

The Ad-hoc Self-Organizing TDMA (ASO-TDMA) protocol has been proposed for Ad-hoc Networks in the marine environment. In ASO-TDMA each frame is divided into a number of sub frames, and a network is divided into a number of hops. A hop is defined as a zone area specified by its distance from a base station. A ship positioned in any hop can only seize time slots in available sub-frames according to the defined rules for the time slot allocation, resulting in less receiver collisions. Thus, ASO-TDMA performs better than SO-TDMA and CS-TDMA with respect to receiver collisions and total delays [5].

2.3 Unstructured Mobile Networks (UMN)

Unstructured Mobile Networks (UMNs) are mobile networks in which there is little or no pre-installed infrastructure (access points, antennas), and as such message forwarding is performed among the mobile nodes or within the wireless infrastructure. UMNs are suitable to a wide range of applications, from environmental monitoring to vehicular networks, mesh networks, among others [16]. Routing in UMNs can be classified into two categories: MANET [17] and DTN [18] routing. In MANETs, once the network topology is decided, end-to-end transmissions can be carried out by using various routing protocols. Therefore, a MANET has high reliability in high density and low mobility areas. On the other hand, in low density or high mobility areas, the reliability of this network decreases because of the unstable paths due to frequent disruption. In such an environment, Delay Tolerant Networks (DTNs) perform more effectively. DTN does not use route information like MANET. DTN provides interoperable communication between a wide range of networks which may have exceptionally poor transmission and be disparate. It has the ability to store the message whose destination is another node. If a message cannot be delivered immediately due to network partition, the best carriers for a message are those that have the highest delivery probabilities [19].

2.3.1 MANET Routing Protocols

Routing protocols in MANETS can be classified into three main types according to the way routes are maintained in the network [20].

- Proactive Routing Protocols

  Proactive or table driven protocols maintain routing information even before this information is required. Each node maintains routing information to every other node in the network. Route information is generally stored in routing tables and is periodically updated with any change in the network topology. The protocols that fall under this category maintain different numbers of tables. Also, they are not suitable for large scale networks, because they need to maintain entries for each node in the routing table [21].

- Reactive Routing Protocols (On-Demand)

  In reactive or On-Demand protocols, nodes initiate route discovery throughout the network, only when they want to send packets to the destination. For this purpose, a node initiates a route discovery process through the network. The route discovery process is completed once a route is established or all possible variations have been examined. Once a route has been established, it is maintained by a route maintenance process until either the destination becomes inaccessible along every path from the source or until the route is no longer desired through the use of timers [20].

- Hybrid Routing Protocols

  Hybrid routing protocols are a new protocol generation; they are together proactive and reactive in nature. These protocols are designed this way to increase their scalability by allowing the nodes that have close proximity to work together forming some sort of a backbone to reduce the route discovery overheads. This is accomplished by proactively maintaining routes to all nearby nodes and establishing routes to far away nodes using a route discovery strategy. The Majority of hybrid protocols proposed are zone-based, meaning that the network is partitioned or seen as a number of zones while others group nodes into trees or clusters [22].

2.3.2 MANET Routing Protocols in the Marine Environment

One of the main difficulties in MANETs, and also inherent in SANETs due to the mobile nature of these networks, is the problem of routing. Topologies can change due to node movement, radio interference and network partitions [23]. Many routing protocols have been proposed to deal with the routing challenges and we present the most popular of these protocols:

- Ad-hoc On-Demand Distance Vector Protocol (AODV)

  AODV is an On-Demand (reactive) routing protocol that uses a backward learning procedure in order to record the previous hop in the routing table. During the backward learning procedure, and upon receipt of a broadcast query RREQ (Route Request) that contains the source and destination addresses, the sequence numbers of those addresses, the request ID, message lifespan, and the address of the query sending node will be recorded in the routing table. Recording the information of the previous sender node into the table makes the destination able to send the reply packet (RREP) to the source through the path learnt from backward learning. A full duplex path is setup by flooding the query and sending reply packets. The path will be maintained as long as the source uses it. Upon receiving a link failure report (RERR) message which is forwarded recursively to the source, the source may trigger the establishment of another query-response procedure in order to find a new path. The on-demand establishment of a new route from source to destination enables the AODV protocol to be effective in both unicast and multicast routing. Multiple RREP messages can be delivered to the source through different routes but updating the routing entries will only occur if the RREP has the greater sequence number. The message with the higher sequence number represents the more accurate and recent information. A number of enhanced approaches have been proposed to reduce the large overhead and high latency which result when encountering high packet loss in the AODV routing protocol [24]. Route discovery procedures are used by nodes to obtain routes on an as needed basis. RREQ duplicate copies are immediately discarded upon reception at every node. On receiving the first copy of a RREQ packet, the destination forms a reverse path in the same way as the intermediate nodes, it unicasts the RREP back to the source along the reverse path. As the RREP moves towards the source, it sets up a forward path to the destination at each hop [25].

- Ad-hoc On-Demand Multipath Distance Vector Protocol (AOMDV)

  AOMDV [26] is an AODV protocol extension for providing multiple loop-free and link disjoint paths. The routing entries for each destination in the routing table contain a list of the next-hops along with the total hop counts to reach the destination. All the next hops hold the same sequence number. This facilitates keeping track of a route. For every destination, a node maintains the advertised hop count. Each duplicate in the route advertisement received by a node forms an alternate path to the destination. Loop freedom is maintained for a node by accepting alternate paths to the destination if it has a hop count less than the advertised hop count for that particular destination. As a result of using the maximum hop count, the advertised hop count does not change for the same sequence number [27]. When receiving a route advertisement for a destination with a larger sequence number, the list of next-hops and the advertised hop count are reinitialized. AOMDV can be utilized to find node-disjoint or link-disjoint routes. To find node-disjoint routes, every node does not reject duplicate RREQs immediately. Each RREQs that arrives through a different neighbour defines a node-disjoint path. This is because nodes cannot broadcast duplicate RREQs, therefore any two RREQs arriving at an intermediate node via a different neighbour of the same source could not have passed through the same node. After the first hop, the RREPs traverse the reverse paths, which are node disjoint and thus link-disjoint. The trajectories of every RREP may intersect at an intermediate node, but still each takes a different reverse path to the source to ensure link disjointness [27]. One of the main advantages of using AOMDV is that it allows intermediate nodes to reply to RREQs, while still selecting disjoint paths. But also, AOMDV has more message overheads during route discovery because of the increased flooding and due to the multipath nature of the routing protocol. The destination replies to multiple RREQs resulting in a greater overhead in keeping track of a route.

- Destination Sequenced Distance Vector Protocol (DSDV)

DSDV is a hop-by-hop distance vector proactive routing protocol. Each node in the network maintains a routing table containing the next-hop, and the number of hops to all reachable destinations. Periodic repeated broadcasts of routing updates keep the routing table updated at all times. To maintain the coherence of routing tables in a dynamic topology, each node periodically transmits updates, and also transmits updates instantly when important new information is available [28].

Routing information is exchanged between nodes by sending chunks infrequently and smaller updates more frequently. Routing information is distributed by broadcasting or multicasting the packets as topological changes are detected to guarantee freedom of loops, DSDV uses sequence numbers to indicate the freshness of a route. A route is considered more favourable than another if it has a greater sequence number or, if the routes have the same sequence number, it has a lower hop-count. The sequence number for a route is established by the destination node and incremented for every new originating route advertisement. When a node along a path detects a broken route to a destination, it advertises its route to that destination with an infinite hop-count and a sequence number increased by one [29].

Route loops can occur if incorrect routing information exists in the Network after a change in the network topology, e.g., a broken link. In this context using sequence numbers makes DSDV adaptive to a dynamic network topology such as in Ad-hoc Networks. DSDV requires regular updating of its routing tables, which consumes battery power and a small amount of bandwidth even when the network is in idle state. Whenever the network topology changes, a new sequence number is necessary before the network changes; therefore, DSDV is not suitable for highly dynamic networks [30].

2.3.3 Delay Tolerant Networking (DTN)

DTN is a type of Wireless Mobile Network (WMN) that does not guarantee the existence of a path between a source and a destination at any time. When two nodes move within each others transmission range during a period of time, they can contact or meet each other. When they move out of their transmission ranges, the connection is lost. If there is no contact with the other node, the message to be delivered needs to be stored in the local buffer of the current node until a connection comes again [31]. To support the heterogeneity of different Networks, the DTN architecture is designed to run as an overlay Network over the Network layer. To do so, two new layers are added: The bundle layer and the convergence layer [32]. The bundle layer encapsulates application data units into bundles, which are then forwarded by DTN nodes following the bundle protocol. The convergence layer abstracts the characteristics of lower layers to the bundle layer. The convergence layer does not need to run over the Internet protocol stack, thus allowing for the implementation of DTNs over any type of network.

- Bundle Protocol

  It stores and forwards bundles between DTN nodes, which is performed by hop-by-hop forwarding instead of end-to-end forwarding used by MANETs. To deal with network disruption, the Bundle Protocol can store bundles in storage buffers until a new transmission opportunity arises. A DTN node will not remove a bundle from its buffer until another node has taken custody of it. This is ensured through a reliable custody transfer mechanism. The Bundle Protocol operation depends on contacts (connections) between DTN nodes. The contact type depends on the type of network. It may be deterministic, as in Interplanetary networks, opportunistic, as in Vehicular Networks, or persistent, as in the Internet [32]. If the size of a bundle exceeds the maximum data rate of a contact, the bundle protocol must perform fragmentation. Fragmentation is supported in two different ways: (i) proactive, where a DTN node may fragment an application message into different bundles and forwards every bundle independently, and (ii) reactive, where bundles are fragmented during transmissions between nodes.

- Convergence Layer

  The convergence layer abstracts the characteristics of lower layers to the bundle protocol and is in charge of sending and receiving bundles on behalf of the bundle protocol. The convergence layer allows for any set of lower protocols to be used to reliably transfer a bundle between two DTN nodes. For example the TCP/IP convergence layer [33] uses a TCP connection between two DTN nodes to transfer bundles. The TCP connection can be established via the Internet. To implement a DTN over other technologies, new convergence layers are needed. Convergence layers must provide the bundle protocol with a reliable delivery and reception mechanism.

Two of the most popular DTN routing protocols are the *Epidemic* and *Spray and Wait* protocols [34] [35].

- The Epidemic protocol is a stochastic routing algorithm for DTNs where the message is replicated to all nodes, it is expensive and does not appear to scale well with increasing load. It can however, operate without any prior network configuration. The alternatives, by requiring a priori connectivity knowledge, appear infeasible for a self-configuring network.

- Spray and Wait protocol (SW) is n-copy routing protocol. This routing algorithm consists of two phases: *spray* followed by *wait*. Here, the number of copies to be created is decided beforehand. Suppose n copies are sprayed to relay in the network, then they enter the wait phase until they meet the destination and the message is finally delivered. Two Spray and Wait models are suggested by authors: (i) Normal mode: In this case, the sender node replicates a message to all nodes that are encountered. Only n nodes get copies because there are only n message copies available. (ii) Binary mode: In this case, out of n copies, n/2 copies are stored by a sender node and the remaining copies to all first encountered nodes. These n/2 stored copies are then relayed until a single copy is left and the last copy is forwarded to the final destination [36].

2.4 Realizing Data Networks at Sea

With the increasing demand of data networks at sea, there has been a large number of attempts at realising this objective especially for Wireless Sensor Networks. The main drawback of the majority of these research efforts is the low coverage area of the proposed wireless transmission technologies compared to the vast areas of oceans and seas in scales of thousands of kilometres. Therefore, these efforts have only been considered for network connectivity near to shore. Summarizing the available proposed technologies for marine networks, various IEEE 802 standards have been proposed in the literature to build wireless networks at sea. The highest communications range proposed within the IEEE 802 family is the Worldwide Interoperability for Microwave Access (WiMAX) standard which barely exceeds 10 km. Other proposals involve Global System for Mobile Communications (GSM) or General Packet Radio Service (GPRS) that provide a higher communication coverage of about 35 km, but are impractical for marine environments that constitute seas and oceans not covered with any GSM base stations except for shore areas that are close to urban land. On the other hand, satellite communication proposals can be considered the most reliable solutions for such vast unpopulated areas, but their main drawback is the high cost of satellite communication that makes it a choice only for large vessels that carry valuable cargoes.

Some of the most mature research efforts in the field of marine communication networks and applications are summarized below:

- An environment monitoring framework was proposed in [37], based on Wireless Sensor Network technology. The proposed Ad-hoc system was based on clusters relying on a star topology, and encompassed a sensing activity, a one-step local transmission from sensor nodes to the gateway (using ZigBee technology) with transmission range of 30 m and a remote data transmission from the gateway to the control centre (using 2.4-GHz Xstream Radio technology) with transmission range of no more than 16 Km.

- A wireless sensor network was proposed for monitoring a coastal shallow water marine environment [38]. It was composed of several sensor nodes or buoys which collect oceanographic data and sent it (using ZigBee technology) to coordinator nodes that transmit the messages to a data server situated on a remote station. It was assumed that GPRS communication was used between the coordinator nodes and the data server and therefore only applicable at coastal areas where Cellular base stations are available.

- The coverage of existing terrestrial wireless broadband networks was extended to the sea so that cost-effective wireless access was available to the ships near the shore [39]. The coverage extension was achieved through a multi-hop WiMAX wireless mesh networks where ships, maritime beacons and buoys were the nodes. Ship to ship communication transmission range was specified as 20 km at maximum.

- The authors in [4] provide a detailed research study on maritime communication technologies to achieve reliable and resilient maritime wireless mesh networks. Comprehensive guidelines have been outlined to critically assess the different deployed maritime communications networks, and identify the milestones in the process of developing Maritime Wireless Mesh Networks (MWMNs).

- In [40] a method of producing a realistic model for constructing a Vehicular Ad-hoc Network (VANET) for shipping in a sea environment was presented. They have gathered data from the international standard Automatic Identification System (AIS) used by commercial shipping and processed it to produce a realistic movement model. They use the model in an NS2 simulation to evaluate the application of MANET technologies to shipping. Two means of transmission rate control have been proposed maximizing the system throughput subject to the reservation-based random access inherited from the ad-hoc self organizing TDMA (ASO-TDMA) protocol in a SANET. It has been shown that the proposed control schemes can maintain the maximum system throughput adaptively as the number of ships varies for the different hop ranges and their target attempt rates.

- In [41] the author has introduced a novel Genetic based approach for optimized routing Maritime Tactical Networks (MTN). It finds the most efficient path having enough resources. Optimized path selection criteria should maximize resources (bandwidth, battery power etc.) should be available, delay should be low and a path with a low hop count should be selected.

- In [42], the author used a wireless Ad-hoc Network that provides an efficient positioning service and restores the weak sea-to-land communication link from small fishing boats to central base-stations. The proposed network combines the Global Positioning System (GPS) service with a wireless Ad-hoc Network.

2.4.1 Density and Mobility Impact on Ship Ad-hoc Networks Performance

Much research has been focussed on MANET routing protocol performance under different operational constraints such as energy consumption, node density, mobility patterns, traffic type, network size, quality of service etc. Focusing on density and mobility constraints, in [43] the authors have calculated the hop progress by taking into account the role of node density and obtained the required hop count for a multi-hop path. Based on the result, they discuss the scaling relationship between node density, throughput and delay in multi-hop wireless networks. In [44] the authors provide a framework for specifying the connectivity that reflects the underlying network architecture and protocols. Based on this framework, they define and analyse connectivity requirements for two network architectures which are GAF (Geographic Adaptive Fidelity) with Manhattan routing, mainly proposed for Ad-hoc Networks and AIMRP (Address-light Integrated MAC and Routing Protocol) which uses tier based routing in sensor networks. Also in [45] the authors disclose an analytical framework for the hop count distribution in a multi hop wireless network with an arbitrary node density. They analyse the average progress per hop and obtain the path connectivity probability in a network. On the other hand in [46] the authors present a detailed analytical study of the clustering overhead associated with node mobility in wireless Ad-hoc networks, and propose a way to understand how changing the node mobility influences the clustering overhead. While in [47] the authors discusses the effect of mobility on parameters such as Bit Error Rate (BER) of a multi-hop route joining a source-destination pair, and minimum required node density of an Ad-hoc wireless network for complete connectivity.

2.4.2 Hybrid MANET/DTN Routing

Several Ad-hoc and DTN Networking environments have been recognized and addressed in the literature by various routing protocols, each protocol addressing the specific requirements of the investigated environment. However, some environments demand a hybrid MANET/DTN routing behaviour due to the hybrid nature of the environment itself. Here we shed some light on the most interesting efforts for such environments.

The authors in [48] suggest a hybrid scheme that combines AODV and DTN-based routing and allows keeping the AODV advantage of maintaining end-to-end semantics whenever possible while, at the same time, also offering DTN-based communication options whenever availableleaving the choice to the application. However, this scheme proposes that a subset of the Ad-hoc nodes are DTN routers, scattered among the rest of the nodes that have AODV routing capabilities. This intuitively requires prior planning of DTN router nodes distribution and is not possible in all MANET/DTN environments.

In [49], the authors propose an adaptive scheme that uses local information to transmit the messages from source to destination using either AODV or DTN routing, depending on current node density, message size, and path length to destination. The main drawback of this proposal is the use of a global density threshold reference in terms of number of neighbouring nodes. This reference number is a global parameter not directly available to the node making the routing decision in a local manner.

Also in [50], the authors present DT-DYMO, a combination of Ad-hoc routing based on Dynamic MANET On-demand Routing protocol and mechanisms for in-network storage and delivery likelihood prediction. DT-DYMO uses a route discovery mechanism in order to find the destination or nodes that are likely to be able to deliver the message in the future based on a minimum delivery probability for the potential message carrier nodes. Efficiency of the protocol performance depends greatly on the prediction accuracy of delivery probability. Unfortunately, the delivery prediction comes with the cost of an active beaconing mechanism, which creates large overheads in dense traffic patterns.

DTN-OLSR proposed in [51] builds a DTN overlay Network that supports the exchange of bundles on top of OLSR. Nodes use modified OLSR control messages to build and maintain an overlay network which uses the bundle protocol for data delivery. Nodes not supporting the full DTN stack use the nearest DTN-capable node for transmitting and receiving bundles. Evaluations show that the overhead for building and maintaining the overlay decreases the performance in well-connected networks.

2.4.3 Link Congestion and Bottlenecks in Ship Ad-hoc Networks

Networking in the ship environment through Ship Ad-hoc Networks (SANET) is still in its infancy compared to other environments such as Vehicular Ad-hoc Networks (VANETs), Smart Phone Ad-hoc Networks (SPAN) and Military/Tactical Ad-hoc Networks that have received a lot of attention in the research community. Therefore the most relevant related work is discussed in terms of identifying and mitigating link bottlenecks in Mobile Ad-hoc Networks in general. The authors in [52] introduce the concept of Node Dependence to characterize how much a node relies on each of its parent nodes. They model the routing process as a Hidden Markov Model, and use a machine learning approach to learn the state transition probabilities in this model based on the observed traces. They also utilize Node Dependence to explore the hidden bottleneck nodes in the network. The authors in [53] present a new solution for available bandwidth estimation on MANET. In this solution, the destination node will send the probe packets to the source node so that the source node could estimate the available bandwidth between them. A new formula for available bandwidth estimation of 802.11 Ad-hoc Network is also presented. In [54] the authors propose a MANET routing protocol based on AODV considering link stability and bandwidth efficiency.

The protocol uses distributed Q-Learning to infer network status information and takes in to consideration link stability and bandwidth efficiency while selecting a route.

2.4.4 The Internet of Marine Things (IoMaT)

As IoT applications and use cases are emerging, marine sensory systems are increasingly being foreseen as an integral part of the IoT picture. Therefore, efforts are arising to define the protocols, standards, architectures, and data acquisition and analysis technologies that comprise the marine IoT use case. Although most of these efforts are still in their infancy, we have managed to summarize the most interesting of them below.

- A new low cost technology for sensing in oceans is described in [55]. The technology is based on readily available commercial electronic devices and uses an Internet of Things approach whereby data is transmitted using internet protocols between sensor devices that are deployed in large numbers. However, it is optimized for short duration ocean measurements of medium spatial frequency close to shore only.

- A marine environmental monitoring system based on the Internet of Things technology is demonstrated in [56]. At first, the system requirements and the overall framework are introduced. Then, the paper discusses how surface and underwater wireless sensor networks contribute to network building. The paper proposes a ZigBee communication model with 30 to 300 meters coverage range relaying sensory data to a CDMA mobile network onshore.

- A comprehensive study of Internet of Underwater Things (IoUT) is provided in [57]. The paper introduces and classifies the practical underwater applications that can highlight the importance of IoUT; and points out the differences between underwater Wireless Sensor Networks and traditional territorial Wireless Sensor Networks. The paper also investigates and evaluates the channel models in IoUT, but does not discuss any higher layer communication protocols.

2.5 Summary

This chapter has presented a review of state-of-the-art work related to the contribution of this thesis. It started with a general overview of the marine communication technology, outlining the components of the system such as the marine band VHF radio, the AIS system and the marine channel access method. An overview of unstructured mobile networks and the routing protocols suitable for the marine environment has been introduced including MANET and DTN routing protocols. Furthermore, a comprehensive review of the research efforts related to the contributions of this thesis and aimed at realizing data networks at sea has been presented. The review has also elaborated on many of the efforts attempting to solve the various marine networking challenges including the impact of density and mobility on SANETs, hybrid MANET/DTN routing, link congestion and bottlenecks in SANETs, and finally, the Internet of Marine Things (IoMaT).

Chapter 3

Ship Ad-hoc Networks (SANETs) over VHF radio frequency

3.1 Introduction

DSDV (Destination Sequenced Distance Vector), AODV (Ad-hoc On-Demand Distance Vector) and AOMDV (Ad-hoc On-Demand Multipath Distance Vector) are among the most widely used routing protocols in Mobile Ad-hoc Networks (MANET) due to their compatibility with multi hop routing environments and scalability towards increased traffic and mobility. On the other hand due to the differences in the routing schemes where DSDV is a table driven routing scheme while AODV and AOMDV are on-demand routing schemes, each routing protocol has its advantages and disadvantages. In this chapter, we investigate the performance of these routing protocols in Ship Ad-hoc Networks (SANETs) in a maritime environment where ships communicate using Very High Frequency (VHF) as the physical layer. Figure 3.1 shows a SANET over VHF physical links using the marine communication system described in (Section 2.2).

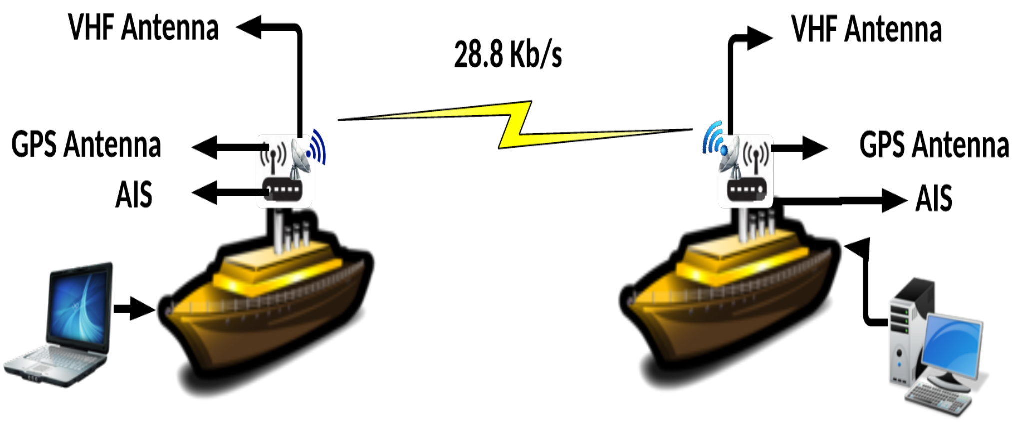


Figure 3.1: Establishing a SANET over VHF Physical Links.

A mobile Ad-hoc Network is achieved so that ships can use it to send particular sensor data such as sea depth, temperature, wind speed and direction, etc. to a central server to produce a public information map. We also investigate the effect of different maritime traffic patterns on the performance of the three different MANET routing protocols. The traffic patterns are represented by different node densities and mobility behaviours which are likely to be found in the marine environment. The novel application of MANET routing protocols in the marine environment using available technology is the main contribution of this chapter. The high cost of other available technologies make our approach an attractive proposition for small and mid sized craft.

3.2 Marine Communication Limitations and SANET Challenges

Networks in the marine environment share a number of characteristics and constraints with MANETs, VANETs (Vehicular Ad-hoc Networks) and other packet Networks which effect the efficiency of network establishment and packet delivery ratio. Also marine networks pose a series of unique characteristics of communication problems at sea that add a number of complexities and design constraints that are specific to ship Ad-hoc Networking:

- First of all, marine networks are low bandwidth due to VHF channel limitations with packet data rates below 30 Kbps (Kilobits per second).
- Power consumption is not an issue in marine communication as power usage is very low and available either from small batteries or using the ships power.
- Very large network areas covering vast oceans and seas, but with relatively limited transmission ranges provided by VHF seldom reaching beyond 40 Km for each ship.
- Marine networks can be described as sparse networks comparing the number of ships at sea to the area the sea covers.
- Marine networks are weather governed networks as weather conditions will affect Network density and mobility.
- Marine networks do not depend on any established infrastructure or centralized administration. Every node operates in a distributed peer-to-peer mode.

Network management is distributed among different nodes, which brings added difficulty in fault detection and management.

- In Ship Ad-hoc Networks, due to the nodes arbitrary movement, the network topology, which is usually multi-hop, can change unpredictably, resulting in route changes, frequent network partitions, and possibly packet loss.

- Scalability is essential for the successful deployment of these networks. The steps toward a large network consisting of nodes with limited resources are not straightforward, and hold many challenges that are still to be solved in areas such as: addressing, routing, location management, interoperability, security, etc..

3.3 SANET Network Architecture

Figure. 3.2 shows the proposed SANET Network architecture for monitoring Marine environments, which consists of sensor nodes (ships), sink nodes (base stations) and edge clouds. Communication between ships is usually Ad-hoc (point-to-point). A sink node collects data from the ships, and transmits the collected data to an edge cloud that stores and processes the data received from the sink. Details of proposed sensory data management and synchronization at the Network core will be discussed later in (Chapter 6). Due to the sink based nature of the proposed SANET and the packet routing contention towards the sink, the network topology can be described as a tree topology which has a hybrid (star−mesh) architecture. It takes advantage of the low power consumption and simple architecture of a star topology, as well as the extended range and fault tolerance of a mesh one [58]. The density distribution of nodes and their pattern of mobility also have distinctive influences on the network behaviour. In a sparse and scattered Network, route formation could be very difficult because the nodes may not be in each other's wireless range. Nevertheless a dense or congested network may suffer high interference between the network nodes. Node mobility can affect the overall network performance as high mobility rates can lead to repeated link breakages which increase packet drop rates and delay route formation [59]. The marine environment is usually described as dense in some areas (shores, gulfs and major ship routes) and sparse in other areas (vast oceans and seas).

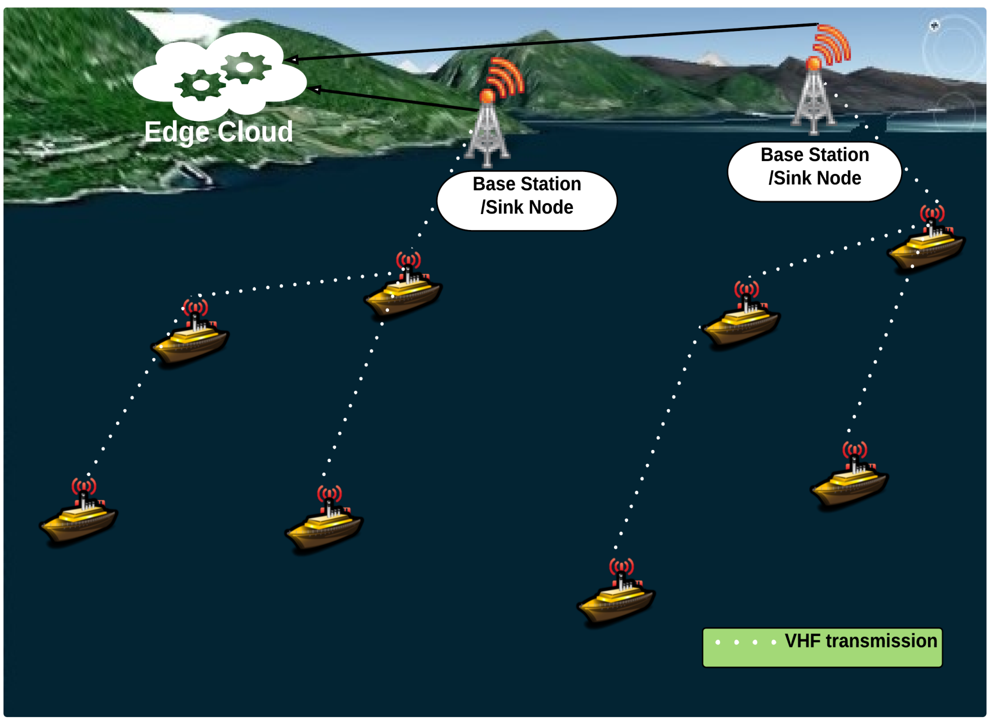


Figure 3.2: SANET Network Architecture

3.4 Marine Buoys for Connectivity and Data Acquisition

Buoys are floating devices that can be anchored (stationary) or allowed to drift with the ocean current. They are low power consumption devices that are usually solar cell powered and use energy harvesting schemes where energy is derived from solar power, captured, and stored for consumption. There are differences in shapes and types as showed in Figure 3.3. Buoys can be used as data collecting systems for Meteorological/Oceanographic parameters and can accommodate a selection of sensors depending on the application needs. Data from the buoy can be transmitted real time and/or stored internally depending on network connectivity.

Figure 3.3: Types of Bouys

Buoys in general can have many communication solutions including VHF, UHF, GSM and others, while in the proposed SANET solution, VHF data transmission is assumed for all participating ships and buoys as a unified, low cost and highly available means of communication at sea. The fully equipped buoys are used to collect data collectors. Therefore they will not only collect and send sensory data, but will also forward data collected by other ships/buoys if the forwarding buoys lie on the data's shortest path towards the sink. Figure 3.4 shows a typical design of the proposed buoys' data collection and transmission system.

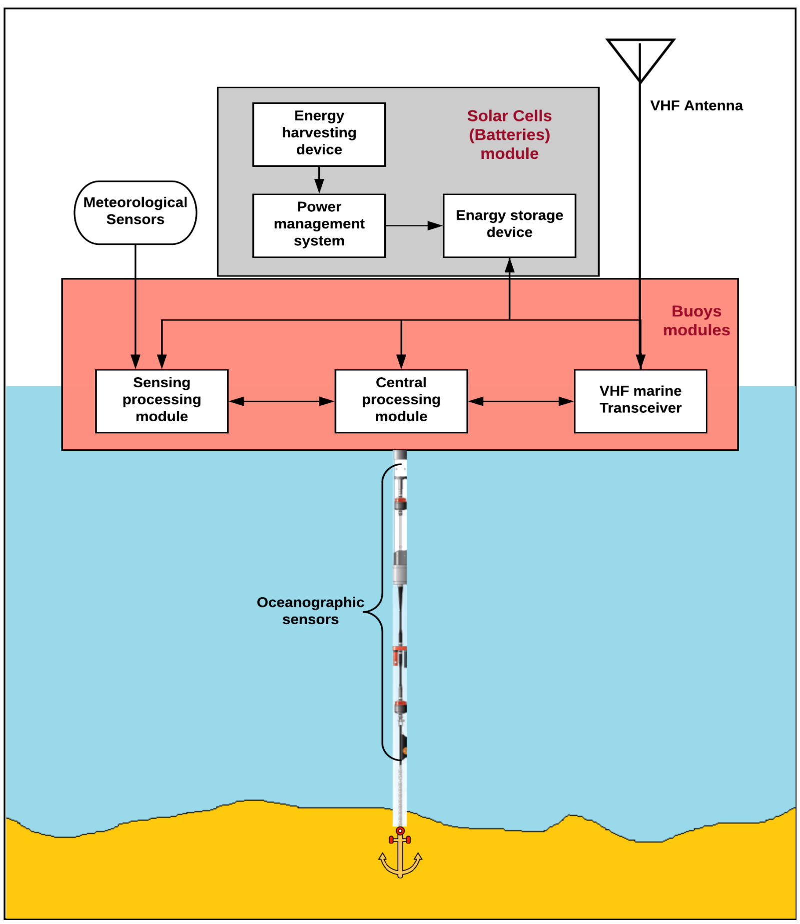


Figure 3.4: Buoys Component and Functional Block Design for Data Acquisition and Forwarding

3.5 Node Density and Mobility Characteristics in SANETs

The number of nodes in the network (node density) and their movement patterns (node mobility) significantly affect the performance of the network. In a sparse scattered network, route establishment is usually difficult due to divisive communication ranges. While a dense network may cause increased interference among the network nodes. It is also essential to study the effect of node mobility on network efficiency as high mobility rates may induce frequent link breakages leading to packet drops and delays in route establishments [59] .

3.5.1 Node Density

In reality large mobile networks do not have uniform sets of nodes populating the network area as assumed in a significant number of research efforts. Instead, node groups varying in numbers are to be found scattered around the network area. This situation can be found in the normal mobile network environment or other environments such as disaster areas, vehicular networks and others [60]. Therefore evaluating the performance of MANETs in an environment without considering the challenges mentioned previously in (section 3.2) would not provide a clear picture of scalability of the MANET routing protocols. Node density is an important factor in the process of route selection and route repair. Increasing node density may bring more opportunities for selecting a route or repairing a route along with other benefits. If nodes are too crowded in the network, the hops between them increase the network overhead, and potentially cause the problem of load imbalance. Therefore using a suitable node density within a realistic situation, network performance can be improved [61] .

3.5.2 Network Mobility Rate

Mobility models are designed to picture the movement patterns of mobile nodes, and how their location, velocity and accelerations change over time. Since mobility patterns may play a fundamental role in determining the protocol's performance, it is required for the mobility models to describe the movement pattern of targeted real life applications in a reasonable way. Otherwise, the observations and conclusions drawn from the simulation studies may be misleading. Therefore, when evaluating MANET protocols, it is essential to choose the proper underlying mobility model [62]. Mobility modelling of nodes in Ad-hoc Networks is an immature science, because there are very few real-world Ad-hoc Networks with which models of node mobility can be compared [63]. Various mobility models have been categorised into several classes based on their specific mobility characteristics. For certain mobility models, the movement of a mobile node is likely to be affected by its movement history. These types of mobility models are referred to as having temporal dependency. In some mobility scenarios, the mobile nodes tend to move in a correlated manner. Such models are referred to as having spatial dependency. Another class is the mobility model with geographic restriction, where the movement of nodes is restricted by streets, freeways or obstacles [62]. When simulating MANETs to measure the performance of various protocols, the majority of studies use the Random Waypoint Mobility Model [64]. While this mobility changes the network topology over time, it is often criticized as being unrealistic because actual node movement is not random in real life networks [63]. Also due to the non-uniform node speed distribution, random waypoints show speed decay. Another issue is that the node movement has bias towards the centre of the simulation area which is as a result of the next destination selection by the moving nodes [65]. The nodes in the Random Waypoint Model behave quite differently compared to nodes moving in groups [66]. It is not reasonable to evaluate the applications where nodes tend to move together using the Random Waypoint model. Therefore, there is a real need to develop a deeper understanding of mobility models and their impact on protocol performance [62]. The Random Waypoint Mobility model includes pause times that separate between changes in direction and/or speed. A mobile node stays in one location for a certain period of time (i.e., a pause time). Once this time expires, the mobile node chooses a random destination in the simulation area and a speed that is uniformly distributed between min-speed and max-speed. The mobile node then travels towards the newly chosen destination at the selected speed. When arriving, the mobile node pauses for a specified time period before starting the process again as seen in Figure 3.5 [67].

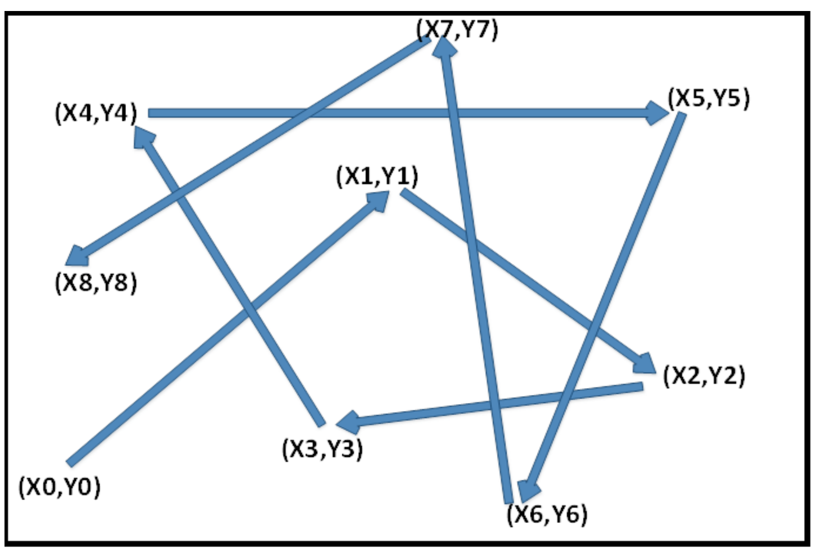


Figure 3.5: Node Movement example in the Random Waypoint Model

In the Random Waypoint model, Velocity (V) and Time (T) are the two main parameters that determine the mobility behaviour of nodes. If the V is small and the pause time T is long, the topology of the Ad-hoc Network becomes relatively stable. On the other hand, if the node moves fast (V is large) and the pause time T is small, the topology is expected to be highly dynamic. Varying these two parameters, especially the V parameter, the Random Way point model can generate various mobility scenarios with different levels of nodal speed.

3.6 Network Setup and Performance Evaluation

For the implementation of this study, the network simulator chosen is NS2.35 running under LINUX UBUNTU 12.04 LTS operating system. Network Simulator Version 2 (NS2) is a discrete open source event simulator targeted at networking research. For simulating the routing protocols, the Tool Command Language (TCL) is used to initiate the event scheduler, set up the network topology, and tells traffic sources when to start and stop sending packets through the event scheduler. The final result is drawn from the execution of a TCL code written for this purpose. In order to extract the results from the trace files (.tr) we used an AWK script which is a powerful control command in LINUX that can process the rows and columns of the file to calculate the network performance metrics. To set up the marine VHF data transmission environment, we input the actual transmission parameters that comply with the International Telecommunication Union (ITU) standards recommendation ITU-R M.1842-1 and maritime mobile service Radio Regularization (RR) appendix 18 into the physical layer of our nodes to achieve realistic modelling of the marine environment.

The transmission frequency used is 156.975 MHz corresponding to channel 79 in the marine VHF band. This channel uses the same transmitting frequency from ship and coast stations and also supports inter ship communication on the same channel as described in the ITU standard. The transmission power used is 25 Watts which is used by most ships. The transmit and receive gain was set to 3. The antenna height was set to 12 meters. We also use two transmission ranges in our simulation: one with transmission range of 30 Km for some ships and another with 40 Km for other ships as published in [68]. The transmission ranges were calculated using the Free Space Propagation model as in equation (3.1).

$$pr = \frac{pt * Gt * Gr * (\lambda)^2}{(4\pi)^2 * d^2 * L} \tag{3.1}$$

where pt is the transmitted power, pr(d) is the received power, Gt is the transmitter antenna gain, Gr is the receiver antenna gain, d is the Tx-Rx separation and L is the system loss factor [69]. AODV, AOMDV and DSDV routing protocols have been used in this simulation. Packet Delivery Ratio (PDR) is the metric used to evaluate the performance of SANETs in the marine environment. PDR is the ratio of successfully arrived data packets at the destination and can be calculated as in equation (3.2).

$$PDR = \frac{\sum Number\ of\ packets\ received}{\sum Number\ of\ packets\ sent} \tag{3.2}$$

Table 3.3 shows a summary of the simulation parameters used in our simulation. Also the hardware and operating system configuration used are listed in Table 3.1. We have performed three different simulation experiments, each with specific density and mobility characteristics as follows:

Table 3.1: Hardware and Operating System configuration.

Indicator	**Description**
Processor	Core(TM)i73537U , CPU 2.50 GHz
RAM	8.00 GB
OS	Ubuntu 12.04
Kernel	Linux
System type	x64-based processor

3.6.1 Using Random Way Point Mobility

A random way point mobility pattern was used with different node density values of 50, 100, 150 and 200 nodes respectively. The simulated nodes move within an area of 400 Km x 300 Km during a period of 2000 Seconds simulation time, during which the nodes pause every 10 seconds to change direction. Table 3.2 shows a summary of the simulation parameters used in this simulation.

Table 3.2: Simulation Parameters for Random Way Point Mobility Scenario

Parameters	Value
Simulator	NS2.35
Routing Protocols	AODV, DSDV, AOMDV
Simulation Time	2000 sec
Traffic Type	CBR
Pause Time	10 Sec
No of Nodes	50, 100, 150, 200
Simulation Area	400 km X 300 Km
Propogation Model	Free Space
Transmission Range	30km, 40 km
Node Movement Model	Random Way Point

3.6.2 Using Real Density and Mobility Patterns in a High Density Area

This experiment is based on real ship locations, speed and direction in the English channel extracted from the AIS data website in [9] for cargo vessels, passenger vessels, tankers and fishing boats as shown in Figure 3.6. Parameters of this experiment have also been summarized in Table 3.3.

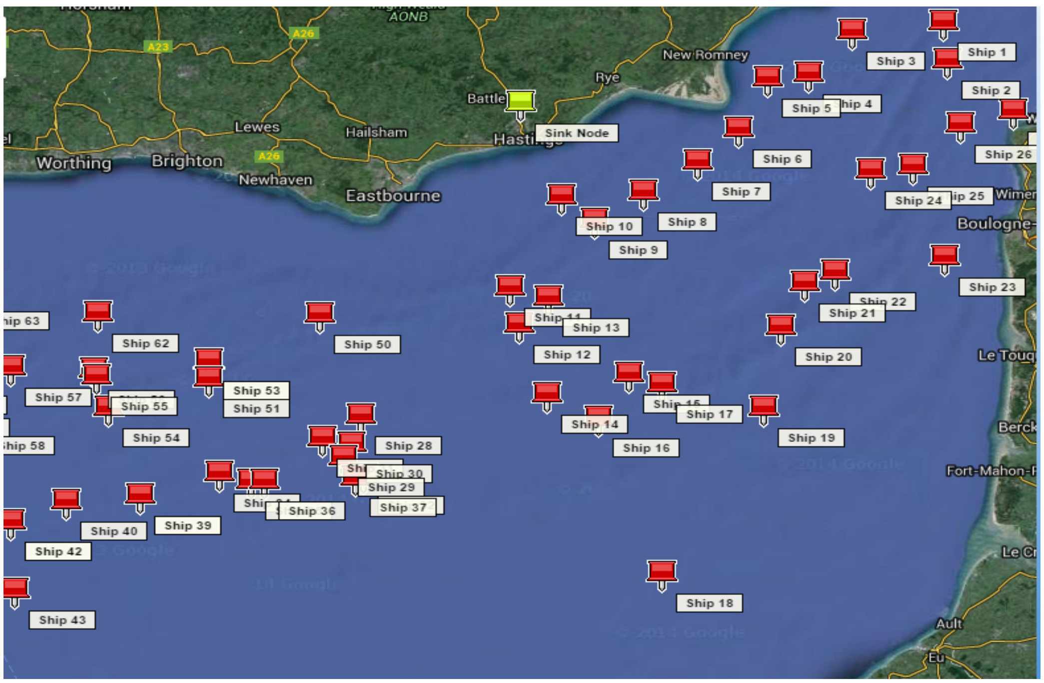


Figure 3.6: Real Ship Locations in English Channel

3.6.3 Using Real Density and Mobility Patterns in a Low Density Area

This experiment utilizes the vessel traffic in the North Sea as shown in Figure 3.7. 19 real ships have been simulated over area of 500 Km x 400 Km extracted from the real

AIS data website in [9] as in experiment 1. This experiment reflects a sparser network environment than the previous scenario. The parameters used for this simulation are also shown in table 3.3.

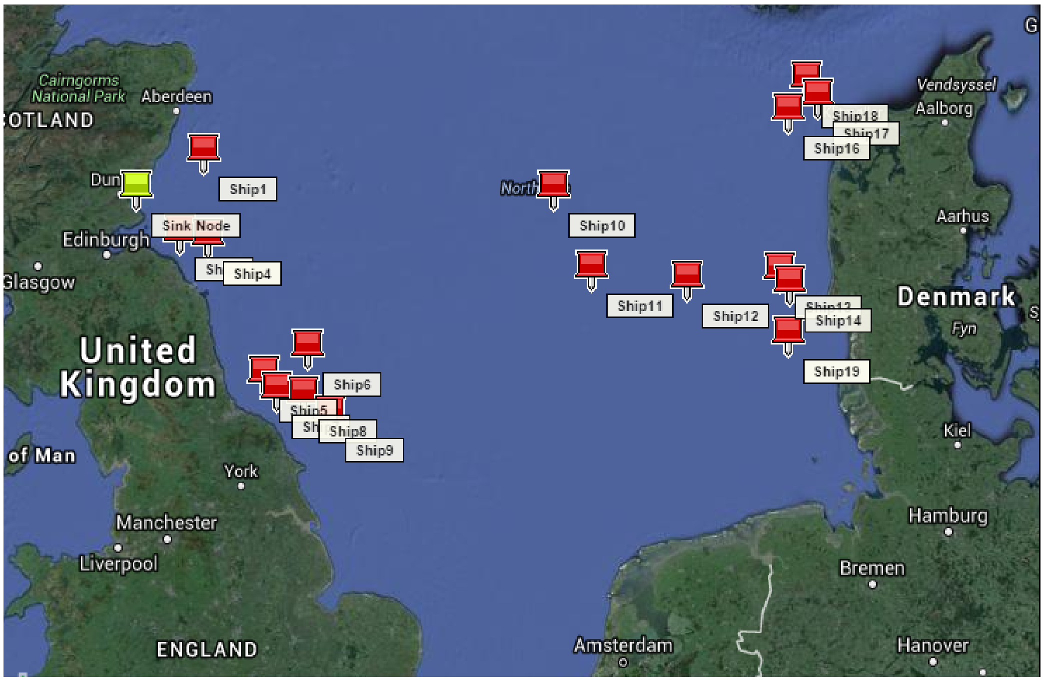


Figure 3.7: Real Ship Locations in the North Sea

3.6.4 Using Marine Buoys to Reduce Isolated Sparse Areas

This experiment proposes the use of buoys to increase the density of forwarding nodes in sparse areas in order to increase the connectivity between ships. We use the buoys in this experiment with real location of ships for Iceland area as shown in Figure 3.8. The simulation was run with 5, 10, 15 and 20 buoys respectively located at highly disconnected locations. Table 3.3 shows the parameters used in this experiment.

Table 3.3: Simulation Parameters

Parameters	**Real Mobility in a Dense Area**	**Real Mobility in a Sparse Area**	**Real Mobility in a Sparse Area with Buoys**
Simulator	NS2.35	NS2.35	NS2.35
Routing Protocols	AODV, DSDV, AOMDV	AODV, DSDV, AOMDV	AODV, DSDV, AOMDV
Simulation Time	2000 sec	2000 sec	2000 sec
No of Nodes	63 nodes with real locations	19 nodes with real locations	37 nodes with real locations with 5, 10, 15, 20 buoys
Simulation Area	380 Km X 200 Km	500 Km X 400 Km	600 Km X 400 Km
Propagation Model	Free Space	Free Space	Free Space
Transmission Range	30 Km, 40 Km	30 Km, 40 Km	30 Km, 40 Km
Movement Model	Real mobility from live AIS website	Real mobility from live AIS website	Real mobility from live AIS website
Bandwidth	28.8 Kb/s	28.8 Kb/s	28.8 Kb/s

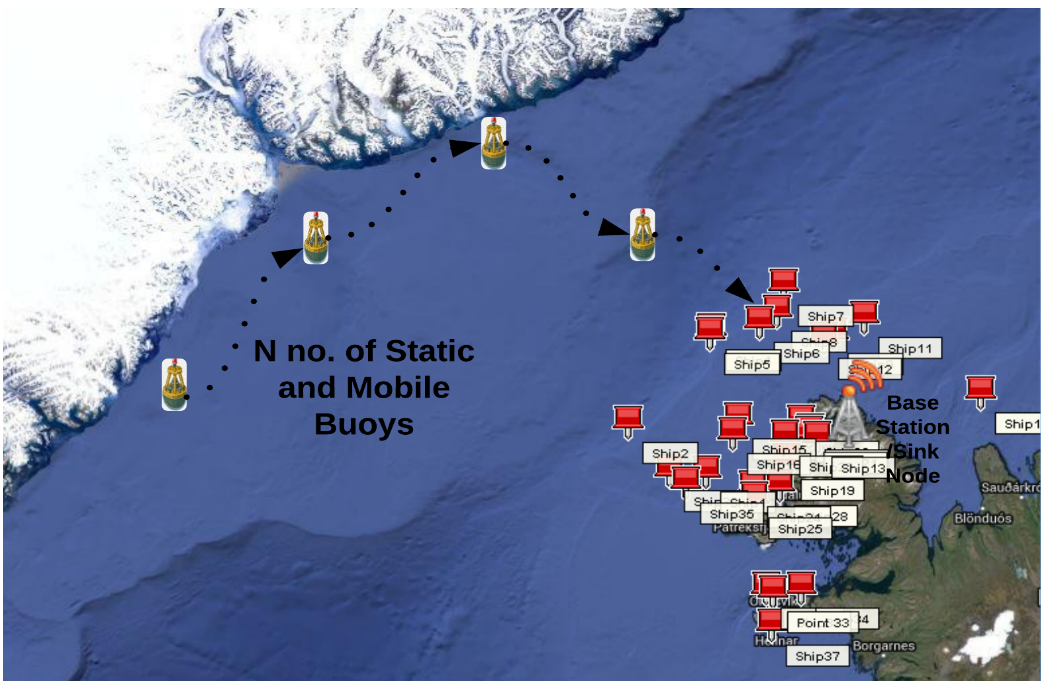


Figure 3.8: Real Ship Locations in Iceland

Disconnections in sparse areas can have a brutal affect on network performance where forwarders are forced to drop packets after the packet lifetime expires before finding a suitable candidate to forward the packet onto. This is an essential limitation in SANET's and is discussed in detail with our proposed solution in Chapter 5.

3.7 Results and Comparison

Figure 3.9 shows the comparison of Packet Delivery Rate (PDR) for AOMDV, AODV, and DSDV protocol in the first simulation scenario. The maximum PDR reached almost 77% in AOMDV protocol while DSDV produced the lowest packet delivery ratio of about 61% when 200 nodes were simulated. It can be clearly seen from the figure that reducing the density by 50% (to 100 nodes) for the same simulation area, lowers the PDR rates to about 49%, 42% and 37% for AOMDV, AODV and DSDV respectively. Although Random Way Point mobility was used in this simulation scenario, it can be observed that node density has a very high impact on the PDR rates achieved, and that reducing the density by 50% has reduced the PDR rates by approximately 30% for the same simulation area.

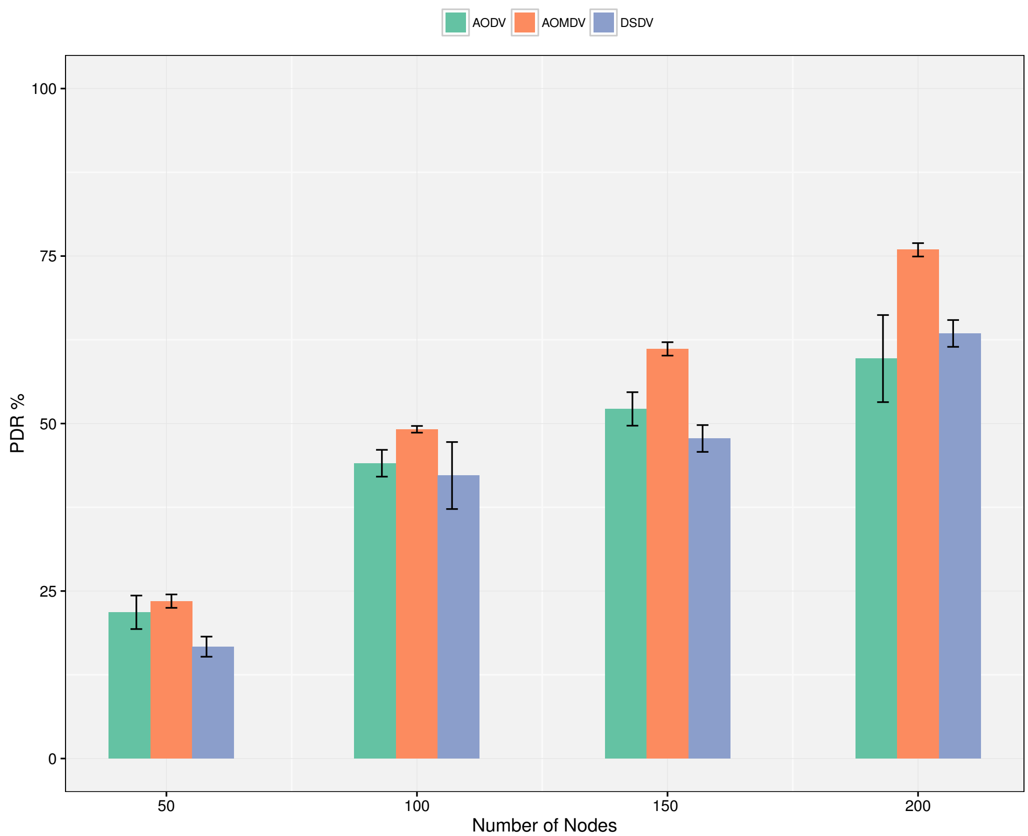


Figure 3.9: PDR Using Random Waypoint Mobility

In the second simulation scenario when real mobility patterns in dense areas are applied to the nodes, high PDR rates are observed reaching a maximum of about 80% with AOMDV protocol and a minimum of 68% with DSDV protocol as shown in Figure 3.10.

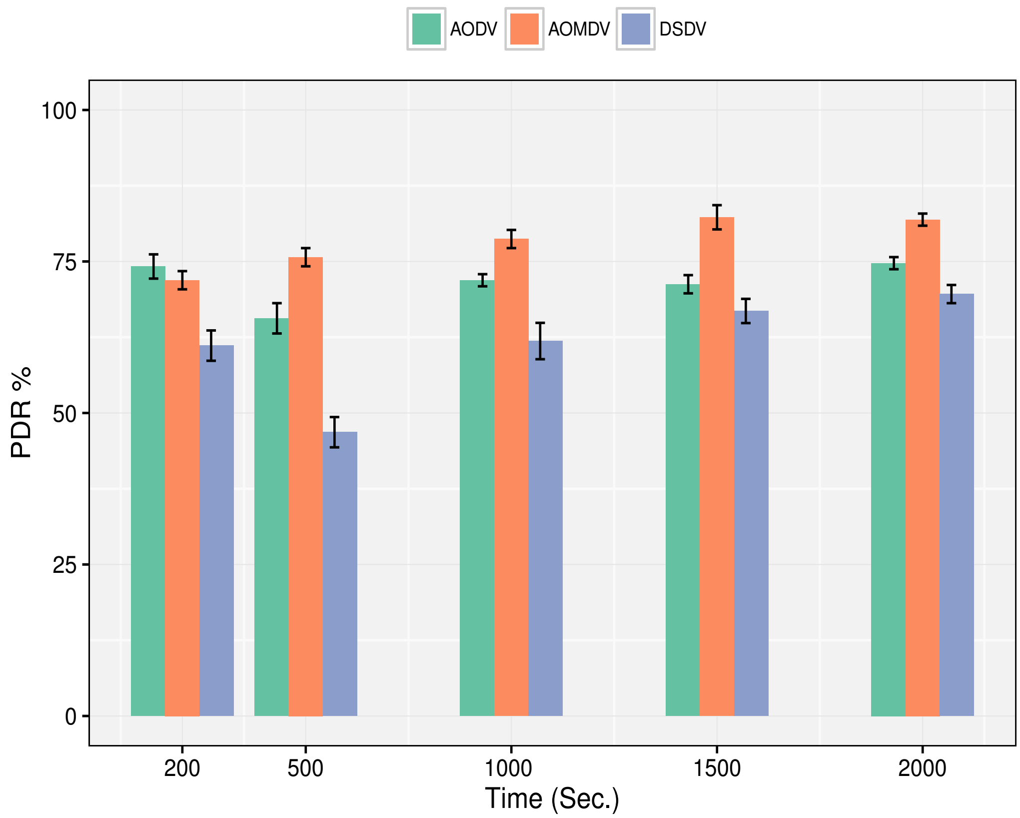


Figure 3.10: PDR of Real Density and Mobility in a Dense Area

Figure(3.11) shows the effect of sparseness on the packet delivery ratio for the three protocols vs. time in the third simulation scenario. PDR for AOMDV protocol records the highest rate as in previous scenarios reaching about 35% while DSDV again records the lowest rate reaching approximately 16%. Considering the VHF marine communication environment proposed and the performance of the routing protocols applied in this chapter, it can be seen that the performance of MANETs in the marine environment decreases in low density situations where huge portions of sea have few vessels that can be used as multi hops to forward data packets. Therefore to increase node density for a higher potential of PDR rates, we have used buoys in our last simulation experiment which are distributed in the most disconnected areas of the water surfaces studied according to real maps of vessel tracks that where observed over long periods of time.

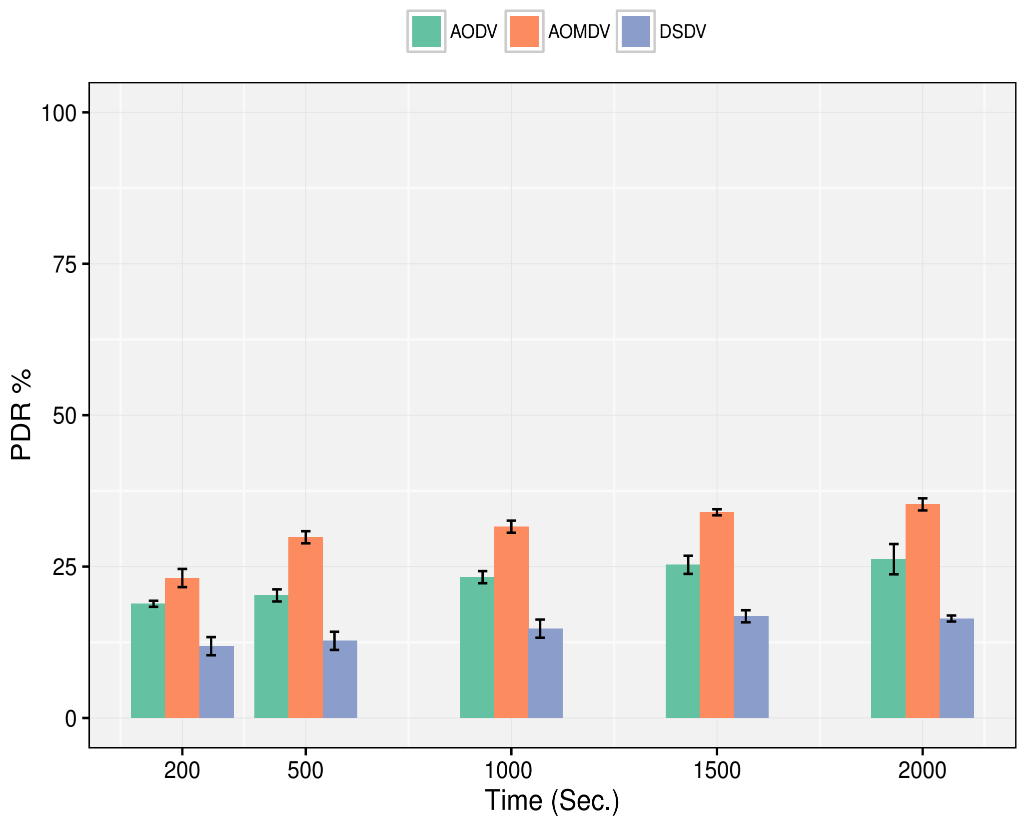


Figure 3.11: PDR of Real Density and Mobility in a Sparse Area

Figure (3.12) shows the PDR rates vs. time according to distributing 5, 10, 15 and 20 buoys respectively. It's can be seen from the figure that applying 20 buoys to support the network connectivity increases PDR rates for the sparse area in Figure (3.11) reaching about 54% for AOMDV.

From the evaluated scenarios and obtained results, it can be clearly seen that real-life marine density and mobility patterns show better PDR rates than the Random

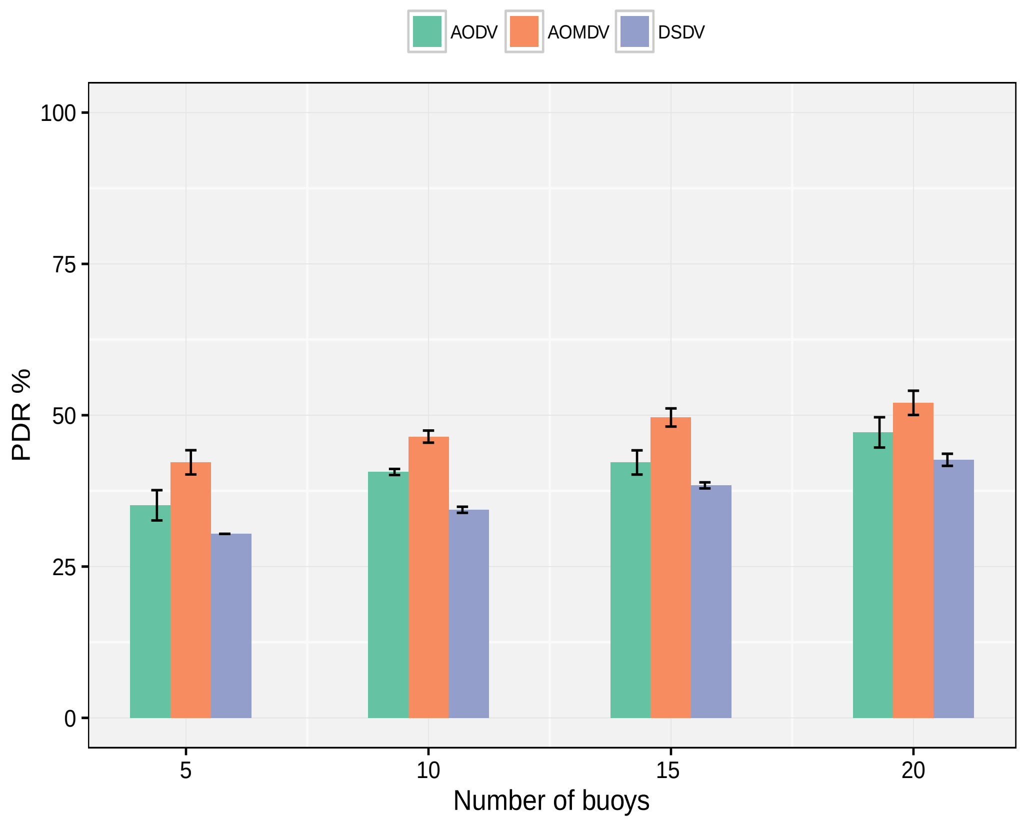


Figure 3.12: PDR of Real Density and Mobility in a Sparse Area with Buoys

Waypoint Mobility model (RWM). This is because the navigational routes drawn by most vessels at sea tend to follow pathways which aid multi hop connectivity for data packets hopping towards the destination. This is not always the case as there are ships that can be scattered sparsely so such pathways do not exist. It can also be seen that in high density and mobility environments, AOMDV achieves higher packet delivery rates than AODV and DSDV and is a better choice for marine applications due to its multipath route discovery process which takes into account frequent link breakages in networks with high mobility rates. On the other hand, AODV achieves higher packet delivery rates than DSDV, since the latter is a proactive routing protocol that is slower in recovering forwarding state and converging routing table changes in high mobility environments. MANET protocol performance has shown a positive relationship with density and an inverse relationship with mobility and sparseness. PDR rates increase when density increases and on the contrary decrease when mobility or sparseness increases.

3.8 Summary

In this chapter, We have proposed and evaluated the novel application of MANETs in the marine environment and explored the use of MANET routing protocols to achieve a low cost AIS type Ship Ad-hoc Network (SANET) using existing on-board equipment. Our solution is aimed at mid to small size vessels and uses VHF technology available on the majority of ships. The impact of marine traffic patterns on the performance of MANET routing protocols has been investigated. It has been concluded that AOMDV is the most efficient MANET routing protocol for Ship Ad-hoc Networks (SANETs). This is due to its multipath route discovery process which takes into account frequent link breakages in networks with high mobility rates, maintaining alternative routes whenever needed.

Chapter 4

Average Marine Data Compression (AMDC) Algorithm for SANETs in the marine environment

4.1 Introduction

Oceans have abandoned resources, wide spaces and play important roles in the activities of the Earth's environment and climate [68]. Oceanography is very rich, involving marine physics, marine chemistry, marine biology, marine geology and many other research fields [70]. How to collect data effectively to understand the marine environment, so as to exploit marine resources, has become one of the most important technologies in the oceanic areas [68].

Marine sensor data come from sensor networks deployed in a marine environment. Types of marine environments include rivers, seas and oceans. In most cases, the raw data stored in databases are first retrieved and processed using mathematical and statistical tools and are then visualized dependent on the user requirements.

In this chapter, we categorize the most important sensory data to be gathered by ships, we analyse the data's characteristics in terms of sensor reading range and acceptable decimal place accuracy. We then employ this study to obtain a quantization and compression algorithm by using our model (Average Marine Data Compression (AMDC)) to reduce the traffic size on our low data rate VHF channel for SANETs in the marine environment. We evaluate and compare the proposed data compression techniques with other known techniques and evaluate its suitability for deployment in the marine environment.

4.2 Digital Signal Processing

The basic communication problem may be posed as conveying source data with the highest fidelity possible without exceeding an available bit rate, or it may be posed as conveying the source data using the lowest bit rate possible while maintaining a specified reproduction fidelity. In either case, a fundamental trade-off is made between bit rate and signal fidelity. The ability of a source coding system to suitably choose this trade-off is referred to as its coding efficiency or rate distortion performance. To represent a signal in the digital domain, it has to go through a number of steps as shown in Figure (4.1) which are described in turn [71].

4.2.1 Sampling

A digital signal is formed from an analogue signal by the operation of sampling, quantizing, and encoding. The analogue signal, denoted x(t), is continuous in both time and amplitude. The result of the sampling operation is a signal that is still continuous in amplitude but discrete in time. Such signals are often referred to as sampled-data signals. A digital signal is formed from a sampled data-signal by encoding the time-sampled values onto a finite set of values [72].

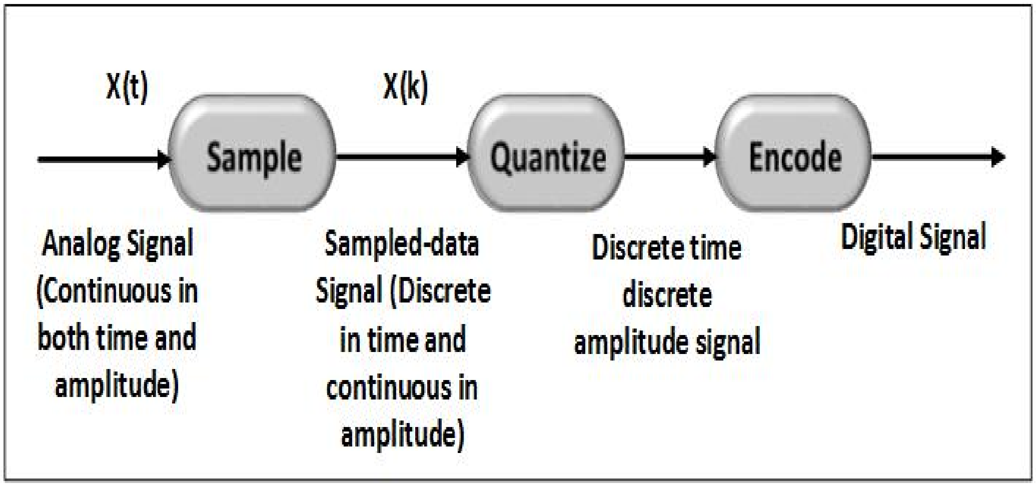


Figure 4.1: Sampling, Quantizing and Encoding.

4.2.2 Quantization

Quantization is the division of a quantity into a discrete number of small parts, often assumed to be integral multiples of a common quantity. The oldest example of quantization is rounding off, which was first analysed by Sheppard [73] for the application of estimating densities by histograms [74]. Quantization makes the range of a signal discrete, so that the quantized signal takes on only a discrete, usually finite, set of values. Unlike sampling (where we saw that under suitable conditions exact reconstruction is possible), quantization is generally irreversible and results in loss of information. It therefore introduces distortion into the quantized signal that cannot be eliminated. One of the basic choices in quantization is the number of discrete quantization levels to use. The fundamental trade-off in this choice is the resulting signal quality versus the amount of data needed to represent each sample [75].

4.2.3 Encoding

Encoding is a digital symbol processing operation in which the digital form of the information is changed for improved communication. In general, encoding contains many different processes, such as ciphering, compression, and error control coding. One of the main purposes of encoding is compressing information. By using data compression we can reduce the disk space needed to store data in a computer. In the same way we can decrease the required data rate on the line to a small fraction of the original information data rate. We could, for example, use very short codes for the most common characters instead of the full seven-bit ASCII code. Rarely needed characters would use long codes and the total data rate would be reduced [76].

4.3 Data Compression

Data compression is the art of reducing the number of bits needed to store or transmit data. There are two types of compression, lossy and lossless. Lossy compression reduces file size by eliminating some unneeded data that will not be recognized by the human after decoding, this is often used in video and audio compression. Losslessly compressed data on the other hand, can be decompressed to exactly its original value. This is important because if a file is lost even a single bit after decoding, will mean the file is corrupted [77].

These steps can be used to reduce the transmission overhead attributed to data transmission.

4.4 Arithmetic code

Arithmetic coding is a technique for coding that allows the information from the messages in a message sequence to be combined to form a single bit stream. A code word is not used to represent a symbol of the text. Instead it uses a fraction to represent the entire source message [77]. The technique allows the total number of bits sent to asymptotically approach the sum of the self information of the individual messages (recall that the self information of a message is defined as $log_2(1/P_i)$).

In the following discussion we assume the decoder knows when a message sequence is complete either by knowing the length of the message sequence or by including a special end-of-file message. We will denote the probability distributions of a message set as p(1), . . , p(m), and we define the accumulated probability for the probability distribution as in equation 4.1

$$f(j) = \sum_{i=1}^{j-1} p(i)(j = 1, ..., m). \tag{4.1}$$

The main idea of arithmetic coding is to represent each possible sequence of n messages by a separate interval on the number line between 0 and 1.The occurrence probabilities and the cumulative probabilities of a set of symbols in the source message are taken into account [78]. The cumulative probability range is used in both compression and decompression processes. In the encoding process, the cumulative probabilities are calculated and the range is created in the beginning. While reading the source character by character, the corresponding range of the character within the cumulative probability range is selected. Then the selected range is divided into sub parts according to the probabilities of the alphabet. Then the next character is read and the corresponding sub range is selected. In this way, characters are read repeatedly until the end of the message is encountered. Finally a number should be taken from the final sub range as the output of the encoding process. This will be a fraction in that sub range. Therefore, the entire source message can be represented using a single fraction. To decode the encoded message, the number of characters of the source message and the probability/frequency distribution are needed [79] .

4.5 Benefits of the Proposed System

Our proposed work is a real world environmental sensor application surveillance system in a marine environment. The purpose of our network is to collect environmental information from different ships. Each ship has a box for AIS and a VHF transceiver. A number of sensors will be placed on the ship to get useful information of (Position, Velocity, Humidity, Temperature, Wind speed, Wind direction, Barometric Pressure, Salinity, Depth and PH). This data is then sent through a Ship Ad-hoc Network (SANET) by multi hop transmission over VHF radio to a destination cloud server, where the accumulated collected data can be processed for end user applications. Because of the low bandwidth available, low complexity and efficient data compression algorithms are essential for SANETs. The proposed algorithm; Average Marine Data Compression (AMDC) addresses this problem by reducing the amount of data transmitted over the Network channel.

4.6 Summarization and Analysis of Marine Sensory Data

The most important sensors applied in our proposed sensor data acquisition network are shown in table 4.1. For each of the sensors which are explained below, we have set the extreme lower and upper limits of the sensors readings likely to be found in the marine environment.

4.6.1 Position

Any location on Earth is described by two numbers, latitude and longitude. If a ship wants to specify position on a map, these are the coordinates they would use. Actually, these are two angles, measured in degrees, ”minutes of arc”and ”seconds of arc.”These are denoted by the symbols (, ’, ”) e.g. 35 43’9”means an angle of 35 degrees, 43 minutes and 9 seconds. A degree contains 60 minutes of arc and a minute contains 60 seconds of arcand you may omit the words of arc where the context makes it absolutely clear that these are not units of time. Calculations often represent angles by small letters of the Greek alphabet, and that way latitude will be represented by l (lambda, Greek L), and longitude by f (phi, Greek F) [80].

4.6.2 Velocity

Velocity is the rate of change of the position of a ship, equivalent to a specification of its speed and direction of motion e.g. (60 km/h to the north). The applicable range in the marine environment would be between 0 and 76 Km/h.

4.6.3 Humidity

Humidity of air is a function of both water content and temperature. The relative humidity of an air-water mixture is defined as the ratio of the partial pressure of water vapour H_2O in the mixture to the saturated vapour pressure of water at a given temperature. The applicable range in the marine environment would be between 0 and 100 %.

4.6.4 Temperature

Temperature is a comparative objective measure of hot and cold. It is measured, typically by a thermometer, through the bulk behaviour of a thermometric material, detection of heat radiation, or by particle velocity or kinetics. It may be calibrated in any of various temperature scales, Celsius, Fahrenheit, Kelvin, etc. The applicable range of air temperature in the marine environment would be between -50° and 50° C. While sea water temperature is inclusive between -2° and 36° C.

4.6.5 Wind Speed

Wind speed is the measure of motion of the air with respect to the surface of the earth covering a unit distance over time. The applicable range would be between 0 and 110 mph.

4.6.6 Wind direction

Wind direction is an indicator of the direction that the wind is heading and is usually measured in a degree between 0 and 360.

4.6.7 Barometric pressure

Barometric pressure (also known as atmospheric pressure) is the force exerted by the atmosphere at a given point. It is known as the ”weight of the air”. A barometer measures barometric pressure. Measurement of barometric pressure can be expressed in millibars (mb) or in inches or millimetres of mercury (Hg). The applicable range would be between 800 and 1100 mb.

4.6.8 Salinity

Salinity precisely measures the total dissolved salt content of ocean or brackish water. The applicable range would be between 0 and 44%.

4.6.9 Depth

A depth sensor measures sea level close to the shore and in the deep ocean. The highest applicable reading in the marine environment is about 10,925 m.

4.6.10 PH

A PH sensor measures sea and ocean water acidity in the range between 0 and 14. A neutral reading would be around 7.

Table 4.1: Marine sensors resolution and quantization.

Parameters	**Lower value**	**Upper value**	**Quantized bits no.**	**Resolution**
Position Longitude	0	90	12	1 s
Position Latitude	0	180	13	1 s
Speed Velocity	0	76	7	1 km/h
Direction Velocity	0	8	3	1
Weather temperature	−50	50	8	0.5 °C
Weather humidity	0	100	7	1.0 %
Wind direction	0	360	9	1.0 deg
Wind speed	0	110	7	1.0 m/s
Water temperature	−2	36	6	1 °C
Pressure barometric	800	1100	9	1 mb
Salinity	0	44	7	0.5%
Depth	0	10925	14	1.0 m
PH sensor	6.9	7.2	2	0.1

4.7 Quantization of Marine Sensor Data

For each of the sensors mentioned previously we have set the extreme lower and upper limits of the sensors readings likely to be found in the marine environment as well as the level of accuracy required to represent each reading. This would enable us during the quantization process to reduce the number of bits required for representing the readings of each sensor limited to the predefined ranges, and accuracy steps within those ranges. Table 4.1 shows each sensor measurement upper and lower limits, the corresponding accuracy step level and the number of quantized bits required to represent the sensory reading.

All the bit calculations were done according to the quantization rules in a straightforward manner where we use the range of readings for each sensor and the required steps within that range to calculate the exact number of possible readings that should be represented as binary bits. The only exceptions are the positioning readings (longitude and latitude) which were represented so as to reduce even more the bit representation required. In all cases linear quantisation is used. Ships latitude [81] is represented in degrees and tenths of a degree, measured in terms of degrees north or south of the equator. Latitudes are determined using standard shipboard methods i.e. a GPS receiver. Tenths are obtained by dividing the number of minutes by 6, and disregarding the remainder (Ignoring seconds). Coding is done with three digits; the first two digits are actual degrees, the last digit for tenths of a degree. Code 46° 41'as 466 (46° is coded as is, 41'divided by 6 is 6 5/6, 5/6 is disregarded); 33° 04'as 330 (33° is coded as is, 04'divided by 6 is 4/6 which is disregarded and coded as 0 in this case); 23° 00'as 230. Latitude can vary from 0° (coded 000) to 90° (coded 900). Quadrant of the globe (Qc) is used to specify whether the latitude is north or south. Ships Longitude [81] is also represented in degrees and tenths of a degree, measured in degrees east or west of the Greenwich Meridian. Values reverse at the international dateline. Tenths are obtained by dividing the number of minutes by 6, and disregarding the remainder (Ignoring seconds). Coding is done with four digits, with the leading (hundreds) figure coded as 0 or 1. The first three digits are actual degrees, the last digit for tenths of a degree. Code 142° 55'as 1429 (142° is coded as is, 55'divided by 6 is 9, the remainder is ignored); code 60° 31'as 0605 (60° is coded as 060, 31 divided by 6 is 5, the remainder is ignored); code 9° 40"as 0096 (9° is coded as 009, 40"is coded as 6); code 0° 16'as 0002 (0° is coded as 000, 16"is coded as 2). Longitude can vary from 0° (coded 0000 on the Greenwich Meridian) to 180° (coded 1800 on the dateline). Quadrant of the globe (Qc) is used to specify whether the longitude is east or west. Quadrant of the globe [81] varies according to your position with respect to the equator (0° latitude) and the Greenwich Meridian (0° longitude). If you are north of the equator (north latitude), Qc is coded as 1 when east of the Greenwich Meridian (east longitude), or as 7 when west of the Greenwich meridian; If you are south of the equator (south latitude), Qc is coded as 3 when east of the Greenwich meridian, or as 5 when west of the Greenwich meridian as shown in Figure(4.2). For positions on the equator, and on the Greenwich or 180th meridian, either of the two appropriate figures may be used.

4.8 The proposed compression algorithm

In order to further reduce the amount of data required to represent the sensor readings after quantization, the Average Marine Data Compression (AMDC) algorithm is proposed. In this lossless data compression algorithm described in Figure 4.3, the

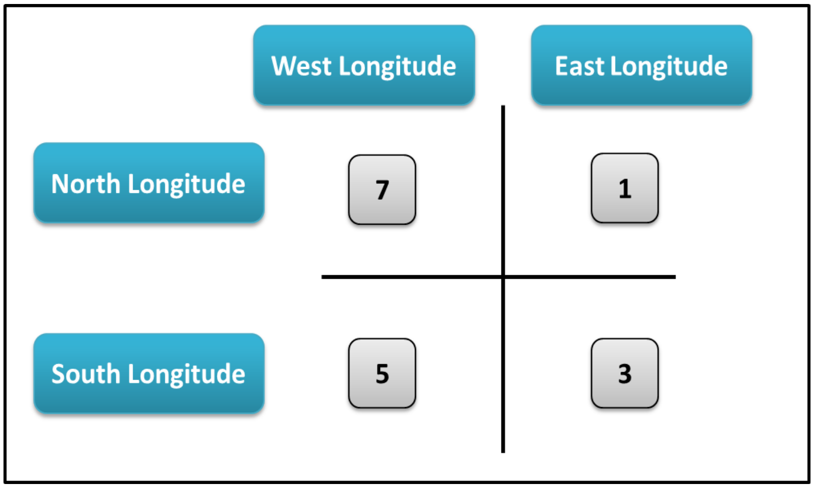


Figure 4.2: Positioning according to quadrant of the globe

first sensor reading in every five readings block is considered as a reference for the block. The AMDC algorithm consists of two phases:

- Average Reading (AR) and Deviation (DV) values: The AR value for every block is calculated by summing the four readings after the reference reading (R_i) and dividing by four as in equation 4.2 below:

$$AR = \sum_{j=1}^{4} R_{i+j}/4 \tag{4.2}$$

  Then the deviation from the 1st reference reading D_v is calculated as in equation 4.3 below:

$$D_v = R_i - AR \tag{4.3}$$

- Arithmetic Coder: The coder applies arithmetic code compression for both R_i and D_v values. After compression, the data is transmitted to the channel.

Using arithmetic coding, a message is usually represented by an interval of real numbers in the range between 0 and 1. And the interval needed to represent the message becomes smaller as the message becomes longer, and the number of bits required to specify that interval grows. Consecutive symbols of the message reduce the size of the interval according to the symbol probabilities generated by the model. The more frequent symbols reduce the range by less than the less frequent symbols and hence add fewer bits to the message. Before transmission, the range for the message is the whole interval [0, l), denoting the half-open interval $0 \leq x < 1$. As each symbol is addressed, the range is narrowed to that part of it allocated to the symbol [82] [83].

To encode a message using arithmetic coding, the following steps are followed [84] [85]:

1. The "current interval" [L, H] is initialized to [0, 1].

2. For each occurrence in the message, two steps are performed:

   (a) The current interval is subdivided into subintervals, one for each possible occurrence. The size of a occurrence's subinterval is proportional to the probability that the occurrence will be the next event in the message.

   (b) The subinterval corresponding to the occurrence that actually happens next is selected, and is considered the new current interval.

3. Enough bits are needed in the output to distinguish the final current interval from all other possible final intervals. The length of the final subinterval is clearly equal to the product of the probabilities of the individual occurrences, which is the probability p of the particular sequence of occurrences in the message.

   In step 2, only the subinterval corresponding to the occurrence that actually happens needs to be computed. For this purpose, two probabilities are used: the cumulative probability

For a more comprehensive explanation of arithmetic code compression, the reader is referred to [82]. According to the AMDC phases described above, the AMDC algorithm has been formulated as follows:

Algorithm 1 Proposed AMDC algorithm

Input : Sensory Data Readings $R_i \rightarrow R_{i+4}$
Output : Compressed Sensory Data
Variables: R_i, AR, Dv, i, H, L, R,$C(x)$,$P(x)$, t
Step 1: Compute Average Reading AR
$AR \leftarrow Sum(R_{i+1} : R_{i+4})/4$
Step 2 Compute Deviation from R_i
$Dv \leftarrow R_i - AR$
Step 3: Compute Arithmetic Code for Dv and AR.
Initialize L:=0 and H:=1;
for $i \leftarrow 1$ to n **do**
  $R \leftarrow H - L$
  $L \leftarrow L * C(x_i)$
  $H \leftarrow L + R * P(x_i)$
end for
$t \leftarrow (H + L)/2$

The AMDC algorithm is particularly suitable for the limited transmission resources of the proposed SANET according to the results showen later in 4.10 that show a gain of approximately 90 % in the number of bits required to represent the readings of the deployed sensors.

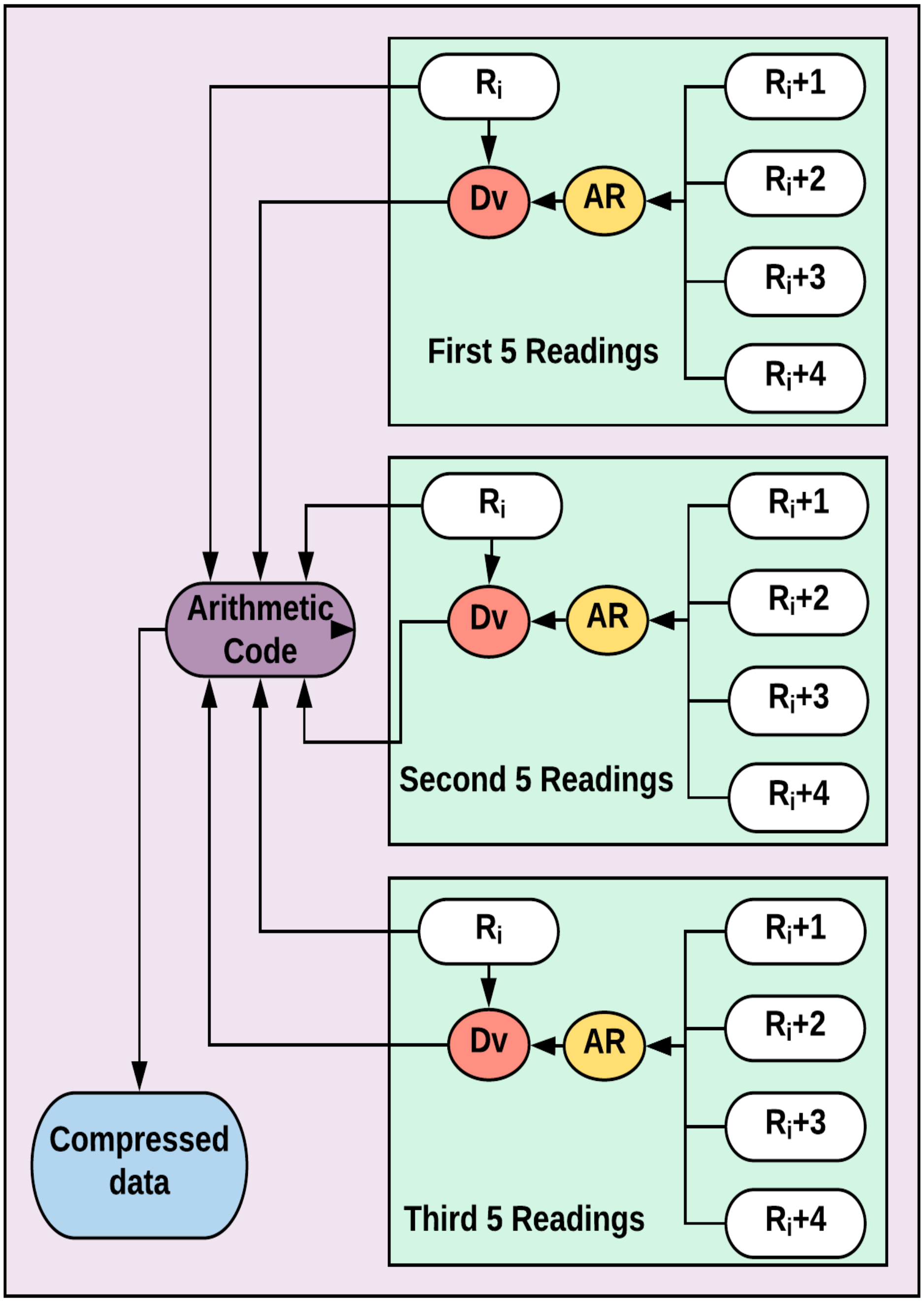


Figure 4.3: AMDC Proposed Model

4.9 Experiment

In our proposed SANET over VHF radio frequencies, the transmission bandwidth used is 28.8 kbps. By reducing data size, less bandwidth is required for sending and receiving data. The data compression is one effective method to utilize the limited bandwidth of the VHF channel, therefore it is crucial to compress the data before sending it over the transmission media. We have simulated a lossless data compression algorithm and used it on our marine data that was obtained from an AIS live system [9], [86]. We compare between our proposed AMDC algorithm and the Arithmetic coding algorithm and evaluate the performance of the algorithm in terms of compression ratio for the compressed data at the originator node.

4.10 Results and comparison

The scheme presented can be implemented on sensors in a marine SANET. In our application, 11 sensors have been used, each producing sensory values once per minute. The performance of the scheme was analysed according to the number of bits required to transmit the acquired data after quantization and compression, in addition to the compression ratio achieved. Ten sets of data were considered, representing the position, velocity, humidity, weather temperature, water temperature, wind speed, wind direction, barometric pressure, salinity, depth and PH values. The proposed sets represent sensory readings collected during 15 minutes in the marine environment. Considering the acquisition time of 15 minutes, each sensor should acquire 15 values (one value for each minute). For 11 sensors we have 165 readings in total. The equivalent number of bits to represent the total original readings is 2085, while

1560 bits are required to represent the quantized data and an average of 223 bits is required to represent the AMDC compressed data. For the ten samples of readings, the results can be seen in figure 4.4.

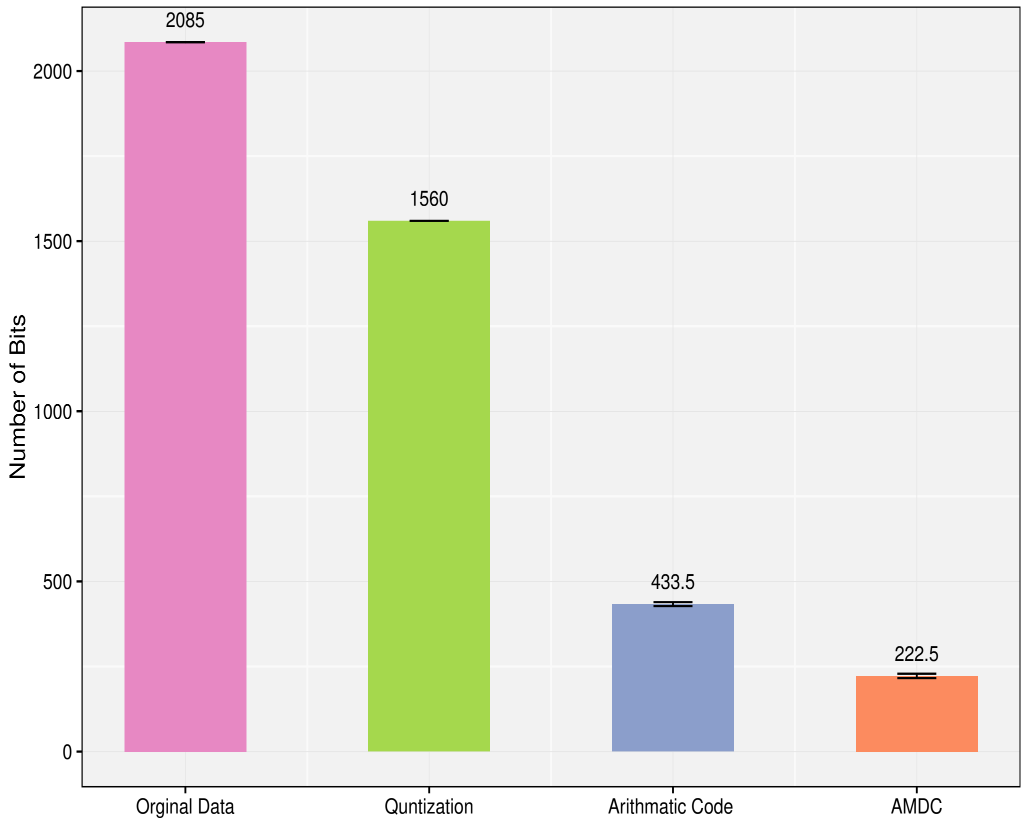


Figure 4.4: comparison between Arithmetic code and AMDC

The metric used to compute the performance of the data compression algorithm is the Compression Ratio (CR) and is defined as the ratio between the size of the compressed file and the size of the source file as shown in equation 4.4 :

$$CR = 100 * (1 - C_{Size}/O_{Size}) \tag{4.4}$$

where C_{Size} represents the Compression Size and O_{Size} represents the Original size.

Table 4.2 summarizes the average compression ratio obtained by applying the proposed compression algorithm in contrast with applying arithmetic code compression for ten data sets that represent ten different input streams of marine sensor data. The table shows clearly that the proposed (AMDC) algorithm outperforms arithmetic code in compression rate for the data sets applied.

Table 4.2: Comparison Ratio for average ten Data sample

	Arithmetic code compression	**Proposed AMDC compression**
Average Data	79.1%	89.3%

4.11 Summary

MANET networks in the marine environment using VHF technology available on the majority of ships and vessels in order to gather important sensor data is a promising research field to overcome the high cost burden of satellite communications currently in place. But on the other hand due to bandwidth limitations of the VHF channel, minimizing transmission data redundancy overhead is essential for efficient use of the transmission channel. For our marine application, the predictability of gathered sensor data makes it beneficial to quantize the data to reduce the amount of bits needed to represent each reading in the binary representation. Applying this quantization in conjunction with the proposed compression algorithm (AMDC) has proved effective data compression rates in comparison with the major known compression method (Arithmetic Code). We have shown in this chapter that the proposed quantization and compression method reduces the number of bits required to represent the marine sensory readings by 89.3% compared to 79.1% when replacing AMDC with Arithmetic Code.

Chapter 5

Hybrid MANET/DTN (MADNET) Routing in SANETs

5.1 Introduction

Mobile wireless Ad-hoc Networks were first studied under the assumptions of moderate node mobility and sufficient density to ensure end-to-end connectivity. Both conditions are necessary for traditional MANET approaches, either proactive or reactive described in Section (2.3.1). When the density of nodes diminishes, end-to-end connectivity can disappear. In such sparse networks, nodes have very few, if any, neighbours within their transmission ranges. The topology eventually splits into several non-communicating connected components [87]. This is typically the domain of Delay Tolerant Networking [1].

One can characterize the relevant routing paradigms in mobile wireless networks along the two main parameters of node density and node mobility. The relation ship between the node density and the MANET routing is positive relationship, and with the node mobility is negative, in contras with DTN which deals with the node mobility as a positive relationship while with node density as negative relationship.

5.2 Hybrid MANET/DTN Routing to overcome SANET Dis-connectivity

As an effective solution to the well-known sparsity problem that works against MANETs in real marine scenarios as described in Chapter 3, we propose the inclusion of store and forward on network nodes (ships) to make each ship retain packets for as long as possible until another suitable candidate is available to pass the packet onto, and to prevent the node from dropping the packet which would be the norm in a MANET. We have found that the concept of Delay Tolerant Networking (DTN) is the most suitable candidate for our application. In DTN routing, a next hop may not be always available for the current node to forward the message to. The node will then have to buffer the data until it gets an opportunity to forward the message, and it should be able to buffer the message for a considerable duration. The complexity in designing a protocol for efficient and successful message delivery is to decide for each message, the best nodes and time to forward. If a message cannot be delivered instantly because of network detachment, then the best carriers for a message are those that hold the highest chance of successful delivery, in other words the highest delivery probabilities [19]. Several DTN approaches have been studied and lead us to a number of conclusions.

As SANETs can be very dense in certain busy locations across main shipping channels and very sparse in others such as deep ocean and shallows, a new hybrid

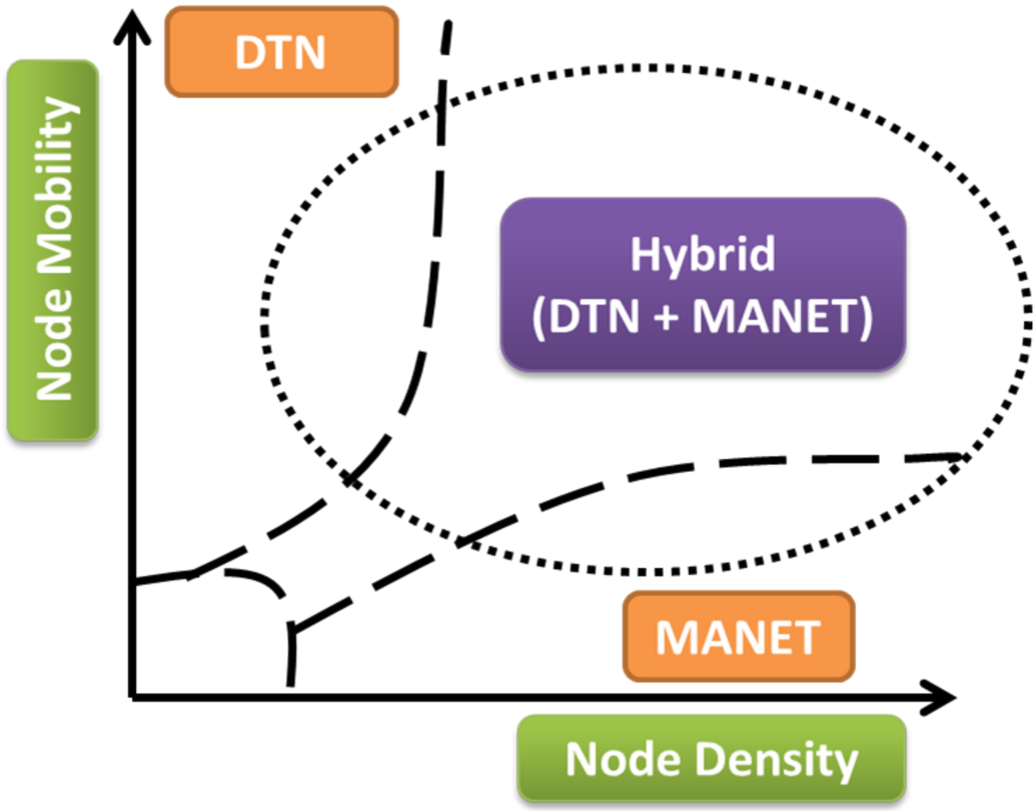


Figure 5.1: The relationship between node density and mobility and its effect on the hybrid routing protocol functionality [1]

routing protocol is proposed that switches automatically between AOMDV (the best performing routing protocol in dense locations as found in (section 3.8) and published in [68] [88]) and DTN (the best performing routing protocol in sparse situations [89] [90]). Figure 5.1 shows the inverse relationship between node density and mobility and how that affects the routing decision in our proposed hybrid routing protocol. While our proposal is not the first to combine DTN functionality with other routing protocols [87], it is the first novel application of Binary (SW)/AOMDV hybrid routing over VHF Ad-hoc Networks in a marine environment. Figure 5.2 shows the proposed protocol stack for SANETs and gives an overview of the solution paths for each of the corresponding layers. As DTN routing is a desired feature in areas with scattered network forwarding nodes, it can overload the network with undesired duplicate traffic when the number of forwarding nodes exceeds a certain limit. Thus placing an unrequired burden on the network bottlenecks as the traffic approaches the network sinks collecting the sensory data. Therefore in dense areas, AOMDV routing substitutes Binary (SW) to provide better network efficiency.

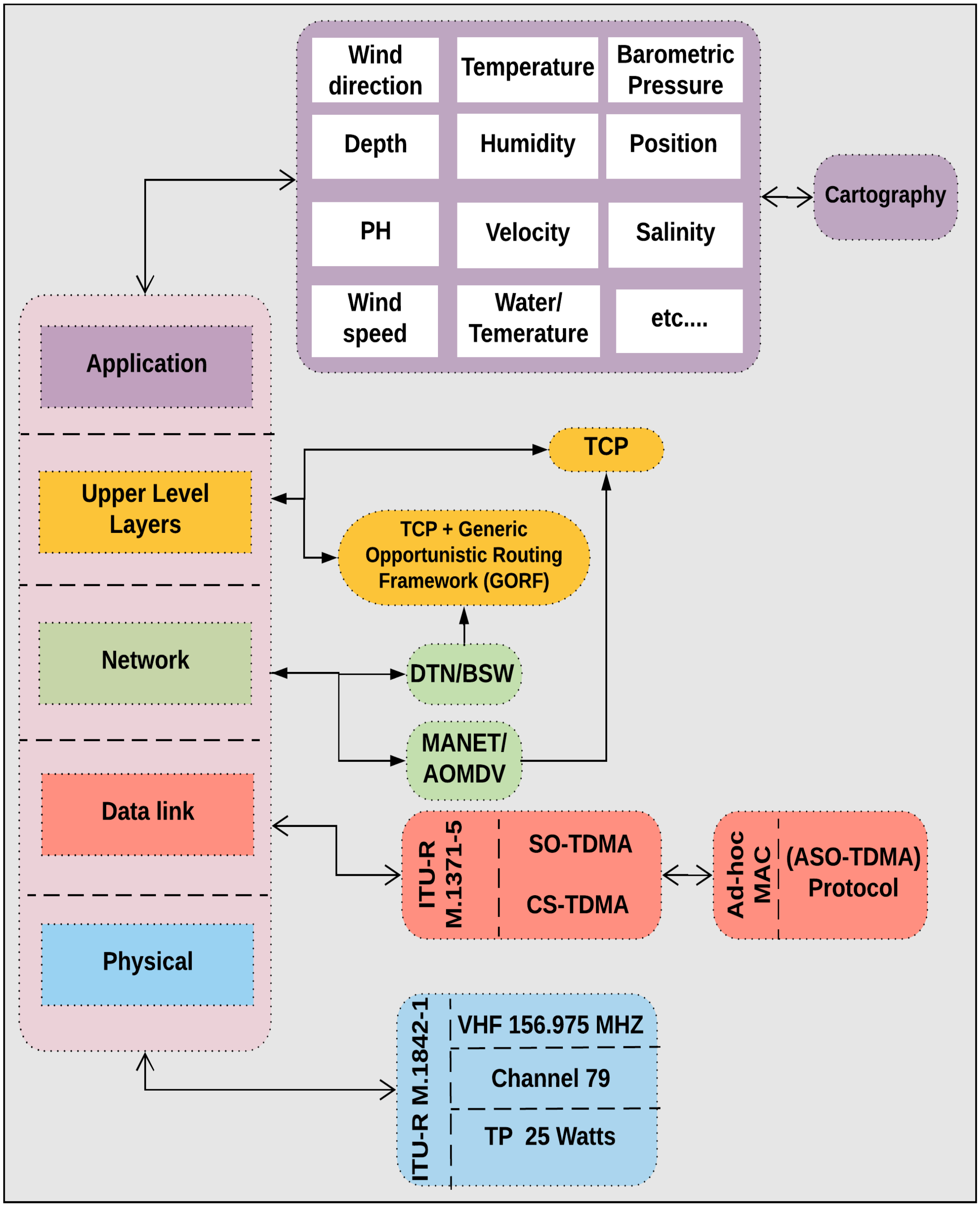


Figure 5.2: Proposed Protocol Stack in SANET's

Table 5.1: MADNET Simulation Parameters

Parameter	Sparse Area (North Sea)	Moderate Area(Clacton (UK)-Middleburg (Netherlands))	Dense Area (English Channel)
Simulation Time(s)	43200	43200	43200
Simulation Area	350 x 400 (km)	175 x 255 (km)	200 x 200 (km)
Average Number of nodes	53	79	100
Routing Protocol	AOMDV,Epidemic, Binary(SW),MADNET	AOMDV,Epidemic, Binary(SW),MADNET	AOMDV, Epidemic, Binary(SW),MADNET
Transmission Range	30(km) - 40(km)	30(km) - 40(km)	30(km) - 40(km)
TTL (s)	14400	14400	14400
Buffer size (MB)	25	25	25
Message size(bits)	223	223	223
Movement Model	Real mobility from live AIS website	Real mobility from live AIS website	Real mobility from live AIS website

5.2.1 MADNET Routing Protocol

The algorithm below shows the basic switching function between AOMDV and Binary (SW) in MADNET routing. The routing process at each ship always starts with AOMDV route discovery, where a Route Request message (RREQ) is propagated towards the destination. Any intermediate node (ship) receiving the RREQ packet sets up a reverse path to the source using the previous hop that sent the RREQ as a next hop for the reverse path. If the same route request is received through more than one neighbour, two reverse paths are created accordingly. Also if the intermediate node has a route to the destination, it immediately generates and sends a route reply message (RREP) back through every neighbour that sent a RREQ for the identified destination. Else, the RREQ message is re-broadcast and the same process is repeated until the RREQ reaches a ship with a route to the destination or the destination itself.

Thereafter, a RREP message is forwarded back to the source via the reverse path(s). As the RREP message proceeds towards the source, the forward path(s) are set up towards the destination.

Algorithm 2 AOMDV, Binary (SW) switching algorithm

```
while DST has not received the message do
    if RREP exists then
        Send by AOMDV
    else
        Send by Binary(SW)
    end if
end while
```

In case no route exists towards the destination and the timeout for receiving a RREP expires, the source node does not regenerate a RREQ packet as would happen in normal AOMDV operation. Instead, the source node initiates Binary (SW) DTN routing that does not rely on end-to-end communication. The DTN bundles are transmitted hop-by-hop and have a certain lifetime, also known as time-to-live (TTL). After the TTL expires, the bundle is deleted from intermediate nodes. Bundle retransmission can take place either at the originating node or at an intermediate node that has obtained custody of the bundle in a process called custody transfer. Several DTN protocols exist and differ in the way they propagate DTN bundles throughout the network. In this work, we use Binary (SW) routing that limits the number of bundle copies to a fixed value. Basically in Binary (SW), any node x that has more than a single bundle token, and encounters another node y (with no tokens), hands over half of the tokens. When node x or y is eventually left with only a single token, the corresponding bundle can be forwarded to the destination node only using AOMDV routing and no more copies are sent to any relay nodes.

MADNET routing protocol provides an optimum routing solution for the marine environment and achieves the performance trade-off between ship density, sparsity and low available data bandwidth. In sparse areas, the MADNET initially jump-starts spreading message copies, in a manner similar to epidemic routing. When enough copies have been spread to guarantee that at least one of them will find the destination through end-to-end AOMDV routing, it stops generating copies. Otherwise, if a direct AOMDV routing path exists from the beginning, then DTN routing will not be triggered for the specified message. This approach limits the propagation of DTN message copies to sparse areas only, since dense areas that are more exposed to link congestions, usually encounter AOMDV routing.

5.2.2 MADNET Simulation and Performance Evaluation

To evaluate the performance of the proposed marine network, we use a model of the VHF radio that complies with the ITU standards to setup a Physical layer in the Network Simulator Version 2 (NS2). VHF transmission ranges were calculated using the Free Space Propagation model as in equation 3.1.

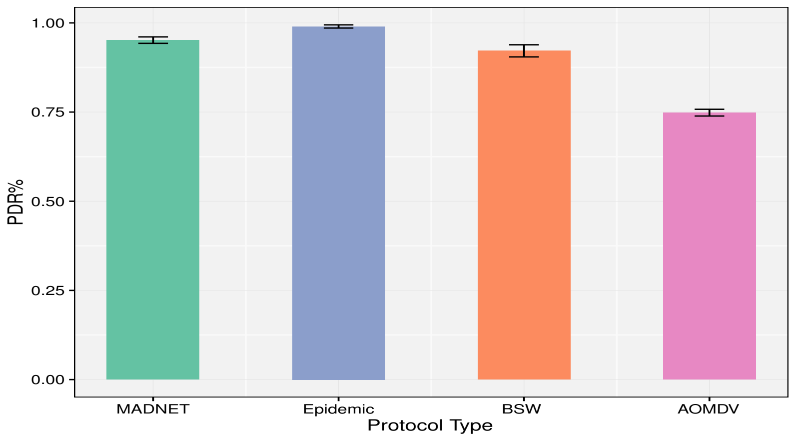
The simulation was performed using four routing protocols: Epidemic, Binary (SW), AOMDV and MADNET. The traffic source type used in the simulation is CBR (Constant Bit Rate) traffic generated corresponding to the size of the collected and processed sensory data as in section (4.10) and published in [91]. We have chosen to use 25 MB bundle buffer space, which does not become a bottleneck in the simulations. Bundle lifetime is set to 14400 seconds, after which all copies of the bundle will be deleted.

Total simulation time was set to 43200 seconds and three simulation scenarios were evaluated. The first is a sparse scenario in the North Sea with simulation area of 350 x 400 km, and the second is a dense scenario in the English Channel with simulation area of 200 x 200 km and the third is a moderate scenario between Clacton (UK) and Middleburg (Netherlands) with simulation area of 175 x 255 km. Each scenario was simulated 6 times corresponding to 6 consecutive days from 9:00 am to 9:00 pm in order to show the variation in performance. We have used real ship trajectories and speed extracted from the real AIS data website in [9]. Table 5.1 shows a summary of the simulation parameters used in our simulation.

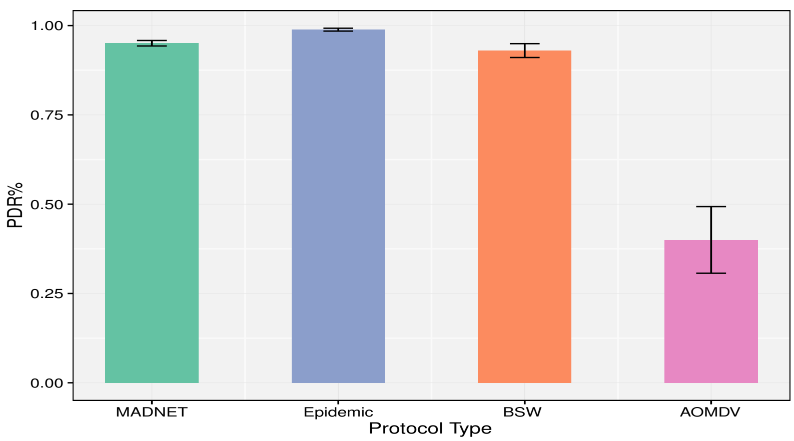
5.2.3 Results and Comparison

The performance of MADNET routing is evaluated in terms of Packet Delivery Ratio (PDR) and hop count. PDR is the ratio of data packets that arrive at the destination successfully. And a hop count is the number of intermediate hops (ships) where a packet traverses from its source to the 5G Base Station (sink).

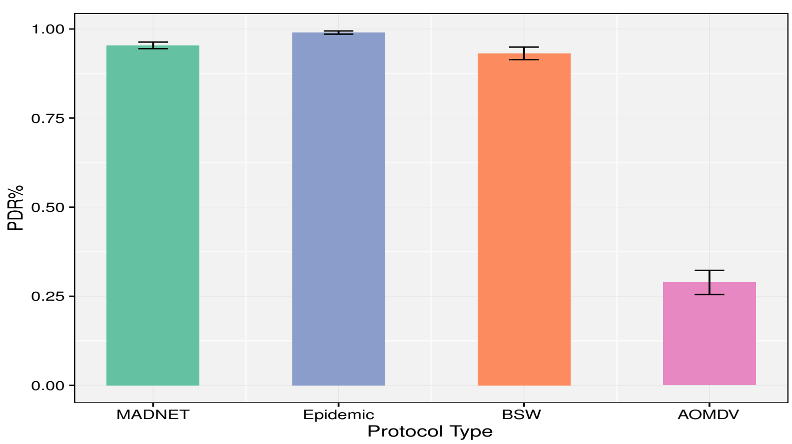
Figures 5.3a, 5.3b and 5.3c compare the simulated routing protocols in terms of PDR rates in dense, moderate and sparse scenarios respectively. It is clear that the MADNET routing protocol achieves the PDR rate of approximately 97% in all scenarios, which is close to the Epidemic routing protocol that has the PDR rate of about 99%. However, the MADNET routing protocol outperforms Epidemic routing in terms of packet delivery cost. This is due to the vast amount of message replication required for the Epidemic routing. MADNET protocol does not require as much replication due to the effective switching between DTN and MANET routing according to the network connectivity. On the other hand, Binary (SW) and AOMDV achieved 94% and 75% respectively in the dense scenario vs. 95% and 40% respectively in the moderate scenario and 95% and 30% respectively in the sparse scenario. The difference in performance is obviously due to the frequent dis-connectivity in the sparse scenario as compared to the dense scenario which mostly impacts the AOMDV protocol that depends on end to end route establishment.



(a) Dense Area



(b) Moderate Area



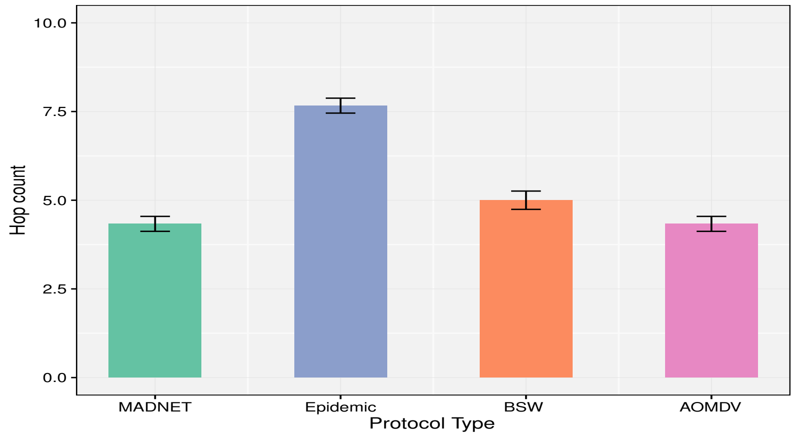
(c) Sparse Area

Figure 5.3: Packet Delivery Ratio vs. Protocol Type

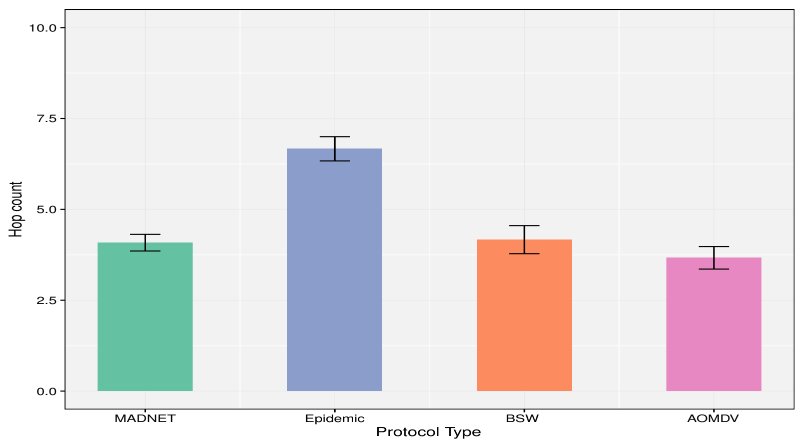
Figures 5.4a, 5.4b and 5.4c show the average hop stretch per routing protocol in dense, moderate and sparse scenarios respectively. Obviously, AOMDV achieves the lowest hop count in all scenarios due to its end to end path establishment property based on shortest path algorithms. On the other hand, MADNET imposes higher hop count than AOMDV (due to the use of Binary (SW) in sparse locations) in favour of better PDR rates while still maintaining lower hop count than Epidemic routing despite that similar PDR rates are observed.

Figures 5.5a, 5.5b and 5.5c show the average packet duplication ratio per routing protocol in dense, moderate and sparse scenarios respectively. The packet duplication rate is defined as the number of nodes in the network that hold a copy of a given packet over the total number of nodes in the network. It can be seen from the figures that Epidemic routing protocol has the highest duplication ratio in all scenarios due to its stochastic nature, where the packet is always duplicated to all nodes within range until it reaches the destination. However, variations in the protocols performance can be seen from the three scenarios where the duplication ratio increases from 70% in the dense scenario, to about 80% in the moderate scenario and all the way to 90% in the sparse scenario. This is due to the network dis-connectivity rate that increases in sparse scenarios and decreases the possibility of packet delivery to the sink and also the possibility that an anti vaccine packet would spread back through the network nodes and stop the duplication in case the packet has been successfully delivered. It can also be observed from the figures that although the packet duplication rate for MADNET routing protocol increases in a similar trend to Epidemic and Binary (SW) in the three evaluated scenarios; it outperforms both protocols substantially with duplication rates of 24%, 44% , and 57% in dense, moderate and sparse scenarios respectively. This is because of its efficient switching between AOMDV and Binary (SW) routing according to the route paths available. Finally, it can be seen that AOMDV routing protocol has 0% duplication rate in all three scenarios, as it does not duplicate packets under any circumstances.

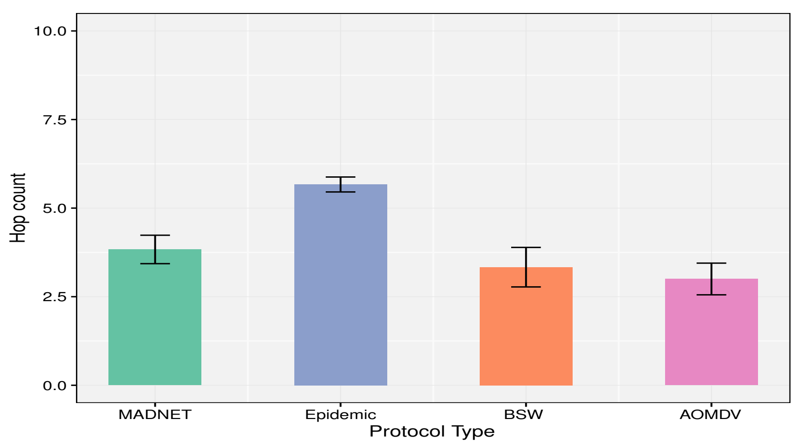
To conclude the results of the evaluated scenarios in Figures 5.3, 5.4 and 5.5, it can be clearly seen that MADNET routing protocol outperforms all three protocols in its overall performance with the objective of maximizing the packet delivery ratio and minimizing both hop count and packet duplication ratio. Therefore, MADNET routing protocol is considered the best option for the marine environment with its specific characteristics of low transmission bandwidth, random topology, no underlying infrastructure and mixed dense/sparse scenarios.



(a) Dense Area

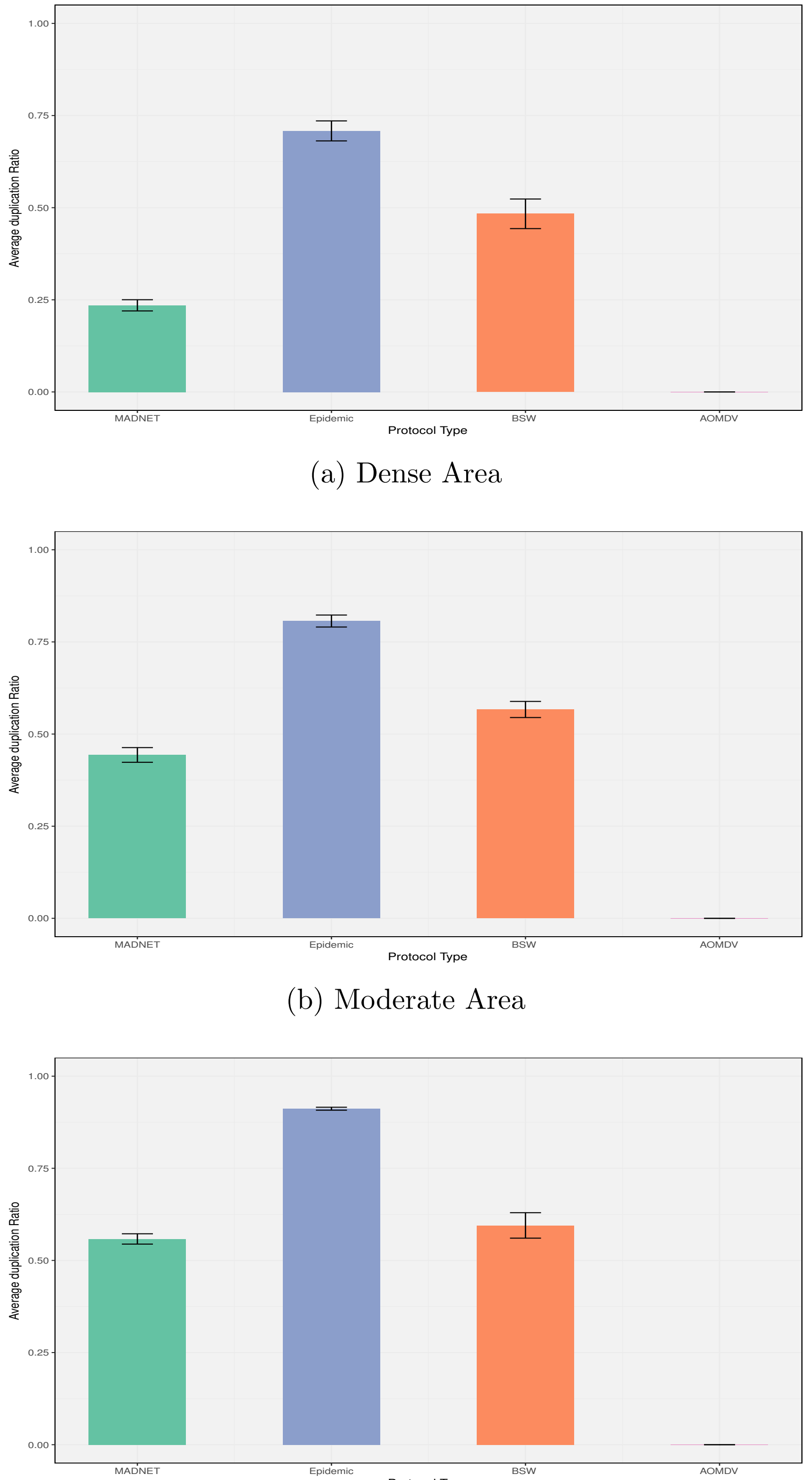


(b) Moderate Area



(c) Sparse Area

Figure 5.4: Hop Count vs. Protocol Type



(a) Dense Area

(b) Moderate Area

(c) Sparse Area

Figure 5.5: Average duplication Ratio vs. Protocol Type

5.3 Link Congestion and Bottlenecks Towards the Sink

The proposed SANET is a sink based network where traffic bottlenecks are likely to happen due to high competition of sensory traffic towards the network sink at shore. In order to analyse the traffic behaviour of SANETs, and get an idea of the throughput at the network bottleneck links; we have used the same simulation environment in Section 5.2.2 to identify the most congested link in the network and evaluate the throughput per ship, and thus the amount of sensory data per ship that can be sent over that congested link per unit time. Three simulation scenarios were evaluated, all at the English Channel at different time windows with a simulation area of 200 x 200 km. The first, second and third scenarios include 82, 100 and 123 ships respectively as found at the specific investigated time windows. The three scenarios are intended to show the effect of link bottlenecks and traffic congestion at three different network sizes over the same area. The proposed MADNET protocol was used to route the data in all three scenarios. Each of the scenarios was run three times, each time with different source CBR value at each ship which are unmodified, quantized and compressed sensory readings respectively as in Section 4.10. These readings are collected at 11 marine sensors on every ship each minute. All the rest of simulation parameters remain as in Section 5.2.2. Table 5.2 shows a summary of the simulation parameters used in this simulation.

Table 5.2: Link Congestion Simulation Parameters

Parameter	Value
Simulation Time (s)	43200
Simulation Location	English Channel
Simulation Area	200 x 200 (km)
Number of nodes	82, 100, 123
Routing Protocol	MADNET
Transmission Range	30(km) - 40(km)
TTL (s)	14400
Buffer size (MB)	25
Message size (bits)	2085, 1560 and 223
Movement Model	Real mobility from live AIS website

5.3.1 Results Analysis

The average period of captured sensory data successfully transmitted to the sink per minute over the bottleneck link is the metric used in this simulation to evaluate the performance of the proposed data acquisition system. Since on each ship, only one reading from every sensor is generated and transmitted every minute, therefore, the average period of captured sensory data per ship successfully transmitted to the sink refers to the bandwidth gain factor for each ship over the investigated link. In general, the bandwidth gain factor for each ship over a bottleneck link is the ratio of available bandwidth per ship (at the bottleneck link) divided by the actual sensory data transmission cost per ship. The gain factor depends on the number of ships sending sensory data and also the amount of control traffic passing over the bottleneck link.

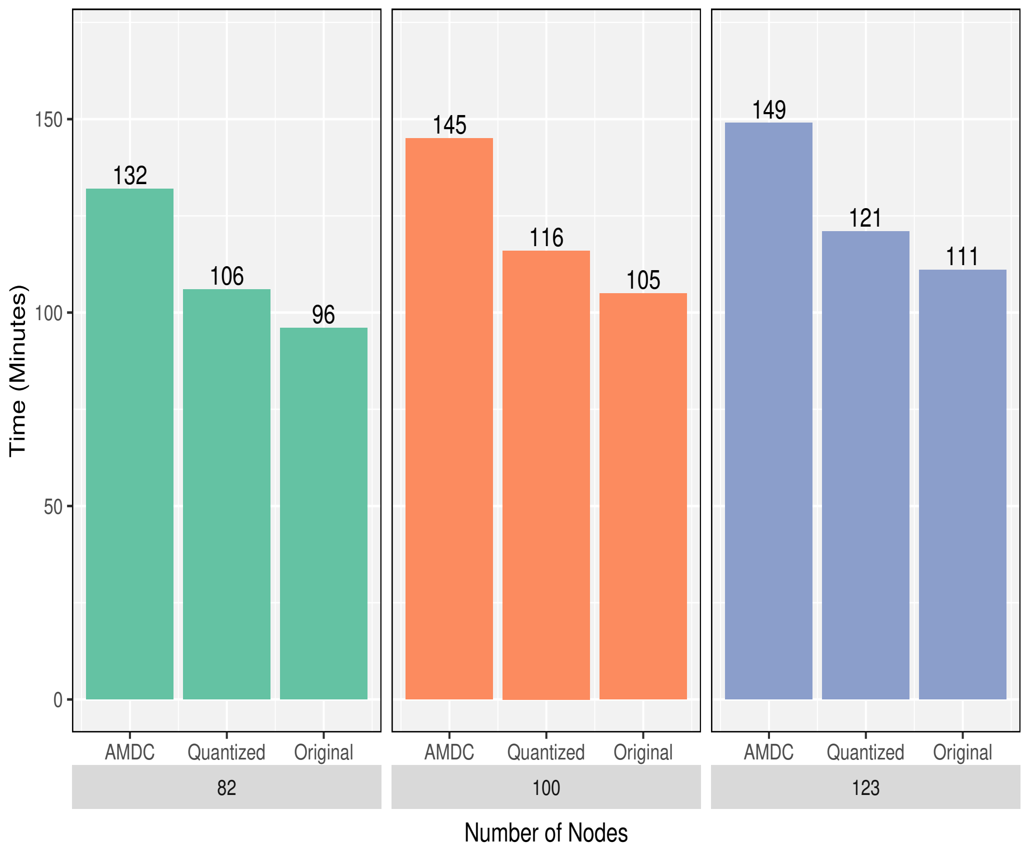


Figure 5.6: The average period of captured sensory data successfully transmitted to the sink per minute

Figure 5.6 shows the average period of collected sensory data transmitted to the sink every minute for each ship according to the limitations of the network bottleneck created over the simulation runs. The figure compares the results in terms of different network sizes ranging from 82 to 123 ships versus the three different types of sensory data discussed earlier (unmodified, quantized and compressed data). It can be clearly seen from the results that AMDC compressed data always provides the highest time period due to the high compression rate which necessitates lower bits over the link to represent the sensory data. The figures show that about an average of 132 minutes of collected data for ship on the link are transmitted to the sink each minute with a network of 82 ships. This number increases to 149 minutes of readings per minute when a larger network of 123 ships is simulated. The figures for quantized data transmission follow with an average of 106 minutes of readings when 82 ships exist in the network and increases to about 121 minutes for 123 ships. Also the figures for unmodified data fall below quantized data reaching an average of 96, 105 and 111 minutes of readings for 82, 100 and 123 nodes respectively. It can be seen from the figures that a higher number of ships in the same simulation area does not necessarily result in a lower number of minutes of transmitted sensory readings over the bottleneck link, as this is also governed by topological factors that influence the severity of congestions created. In the example, it could be the case that the network with a lower number of nodes has only one VHF link (ship) at the edge of the network towards the sink while the network with higher number of nodes has several VHF links (ships) towards the sink. It is also obvious from the results that even though the proposed SANET operates over VHF radio that offers low rate data transmission at 28.8 kbps; sensory data delivery rates of up to 149 minutes of collected data per ship per minute have been observed with a network of 123 ships.

5.4 Summary

Due to increased shipping and the high cost of other available technologies, the demand for data networks in the marine environment for cartography, safety and convenience shows an increasing trend. In this chapter, the behaviour of SANETs has been discussed high-lighting the problem of sparsity which necessitates the inclusion of selective opportunistic networking and MANET protocols. We have evaluated the deployment of the proposed hybrid MANET/DTN routing protocol (MADNET) in the marine environment with both dense and sparse scenarios. Obtained results verify the effectiveness of the proposed protocol in all evaluated scenarios. The drawbacks of low rate data transmission offered by VHF radio has also been investigated in terms of the network bottlenecks near the sink and the achievable data rate at the sink for each ship. Results show that even though the proposed SANET operates over VHF radio that offers low rate data transmission at 28.8 kbps; high sensory data delivery rates of up to 149 minutes of collected data per ship per minute have been observed.

Chapter 6

IoMaT Cloud Data Management and System Overview

6.1 Introduction

The SANET discussed and evaluated in the previous chapters is proposed to collect different marine sensory data from ships and vessels and send this data back to onshore sinks collocated with 5G base stations as part of a cartography IoT application. The base station/sink nodes send the collected sensory readings to edge clouds that include dedicated storage as part of the Mobile Edge Computing (MEC) services. Mobile Edge Computing usually relates to mobile network applications and data stream acceleration through caching and/or compressing of relevant (mainly localized) data at the edge of the mobile Network, as near as possible to the end user location. A new application of MEC is proposed in this chapter where part of the edge computing resources is exploited as edge repositories (clouds) of the sensory data delivered to the shore. The edge clouds eventually connect to a central cloud in the internet where all the sensory data is aggregated, filtered and analysed to produce real-time maps of surface and under water environmental information that produces accumulative maps for beneficiary customers. The proposed cartography system can collect data including but not limited to: sea state, depth, temperature, wind speed/direction, humidity, salinity, etc.

6.2 Information Centric Networking for Cloud Data Management

The applications and usage of the Internet of Things (IoT) often entails information centric usage patterns, where users consume IoT generated content from the network instead of communicating with specific hosts or devices. Information Centric Networking (ICN) provides means for information dissemination by identifying information at the Network layer. Such identification can be realized, for instance, through some form of naming scheme. This enables the decoupling of sender and receiver and naming data independently from its location. Therefore, using an ICN architecture for IoT data potentially provides advantages such as more efficient service provisioning and content management compared to using traditional host-centric networks. A wide range of ICN implementations have been proposed in various research projects, of which Publish Subscribe Internet Technology (PURSUIT) [92] architecture is used as an exemplary reference model. PURSUIT provides an architecture that completely replaces the IP protocol stack with a publish-subscribe protocol stack.

6.2.1 Conceptual Architecture

PURSUIT employs a Publish-Subscribe paradigm for a path-based information dissemination that names information at the Network layer decoupling request resolution from data transfer in both time and space. The asynchronous nature of the Publish/Subscribe architecture simplifies data management and greatly facilitates IoT service provisioning. Individual information items are arranged into a context named scoping. Scopes allow information items to be grouped according to application requirements, for example different categories of information. Relationships between information items and scopes are represented as a directed acyclic graph Figure 6.1 of which leaves represent pieces of information and inner nodes represent scopes. Each node in the graph is identified with its full path starting from a root scope [93].

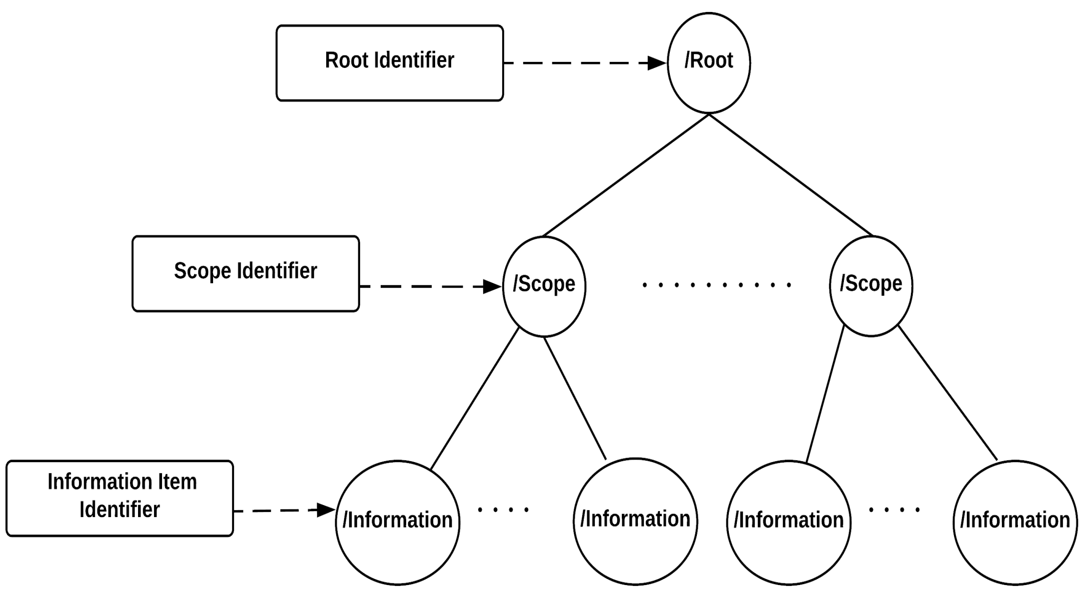


Figure 6.1: ICN Namespace

There are two main functional entities in PURSUIT to compose the core of the architecture, namely, the Rendezvous (RV) and the Topology Manager (TM). The RV s responsible for matching publications and subscriptions of information items; and the TM is responsible for constructing a delivery tree for the information object. Once the RV has matched a publication and one or more subscriptions, the delivery tree is created using a Bloom-filter-based approach, where the TM constructs a Forwarding IDentifier (FID) by encoding the link identifiers in a source routing manner. The FID is then sent to the publisher to forward the packets containing the information object to the subscriber. Therefore, ICN does not use end-to-end addresses in the network for data delivery, and instead of naming end hosts, it names links [94]. Figure 6.2 shows a basic ICN reference architecture.

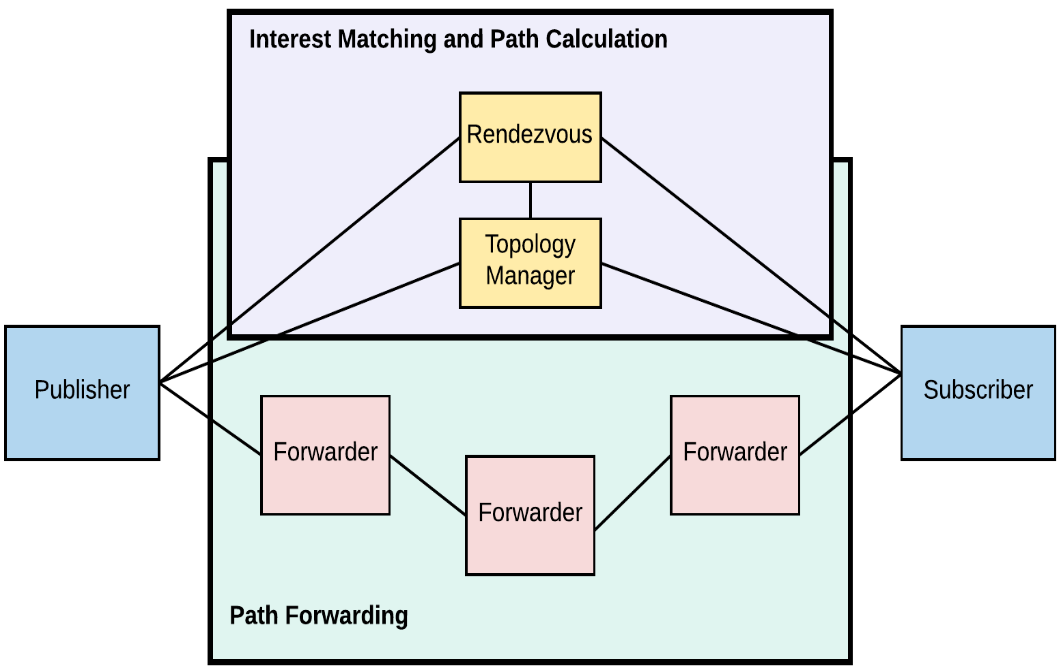


Figure 6.2: Basic ICN Architecture

In the network, there is also a set of Forwarders that simply forward the information object to the subscriber using the specific FID generated for this transmission. Each Forwarder acts on packets by executing a simple set membership test which can be efficiently implemented. For each link, the outgoing Link ID is ANDed with the FID in the packet. If the result matches with the Link ID, this confirms that the Link ID has been encoded into the FID and the packet is then forwarded along that link. Else, if the result does'nt match with the link ID, it is assumed that the Link ID has'nt been encoded into the FID and that the packet is not meant to pass over the link in question [95].

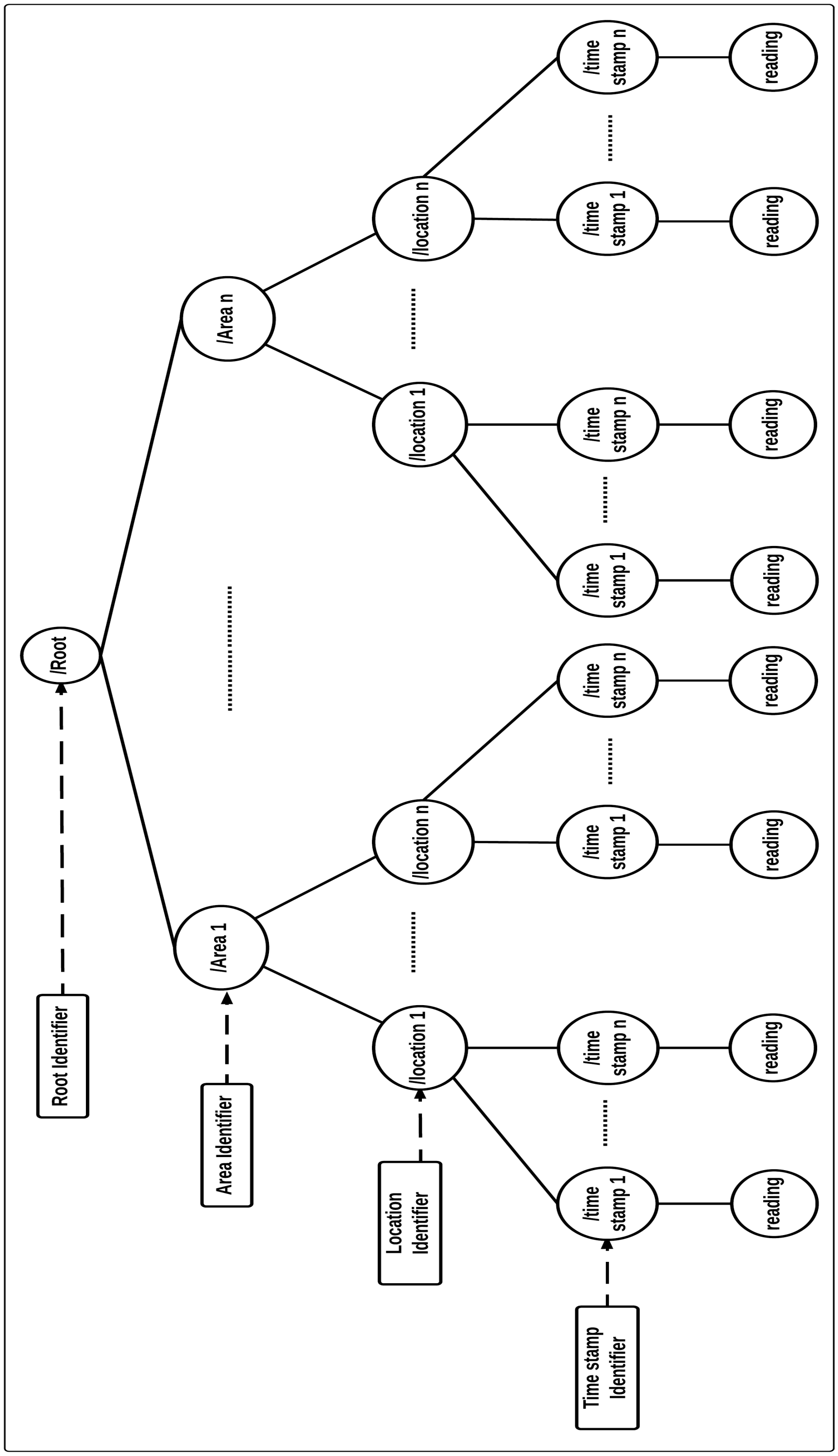


Figure 6.3: Namespace for Marine Data Management and Service Provisioning

6.2.2 Realizing an Information Centric IoMaT

Fig. 6.3 shows the ICN namespace proposed for managing marine sensory data at the 5G network core. The namespace proposed includes a root identifier that represents the root scope allocated to serve the subject network domain. Under this root scope, there exists a so-called area identifier that identifies a specific geographical area of the sea. This area is usually covered by several cellular base stations. The next level scope is the location identifier that identifies every specific location existing within the specified area (to GPS coordinates granularity). Then, the next level scope represents the recorded time stamps for all readings taken within the specified location. Finally, the leaves of the tree represent the individual sensory readings.

Fig. 6.4 shows a sequence diagram of the messages exchanged to establish the communication between the central cloud and edge clouds in the proposed cartography system over ICN architecture. In this scenario, edge cloud 1 sends a publish message towards the RV including the scope /root/area1. The RV then matches the publication with a previous subscription initiated by the central cloud for the whole /root scope of the domain which enables the central cloud to receive all data sent by edge clouds carrying marine sensory information. Otherwise, the central cloud can selectively subscribe to certain areas/locations only by sending subscription messages to the RV with the scopes of the requested areas/locations. i.e. scope /root/area1, /root/area1/location1, /root/area1/location2 or /root/area3. Even more specifically, the central cloud can subscribe to certain time stamps within the selected location, only by modifying the depth of the subscription scope, i.e. /root/area1/location1/time stamp1. Despite these flexibilities, it is always assumed that the central cloud is subscribed to the whole domain scope /root which facilitates

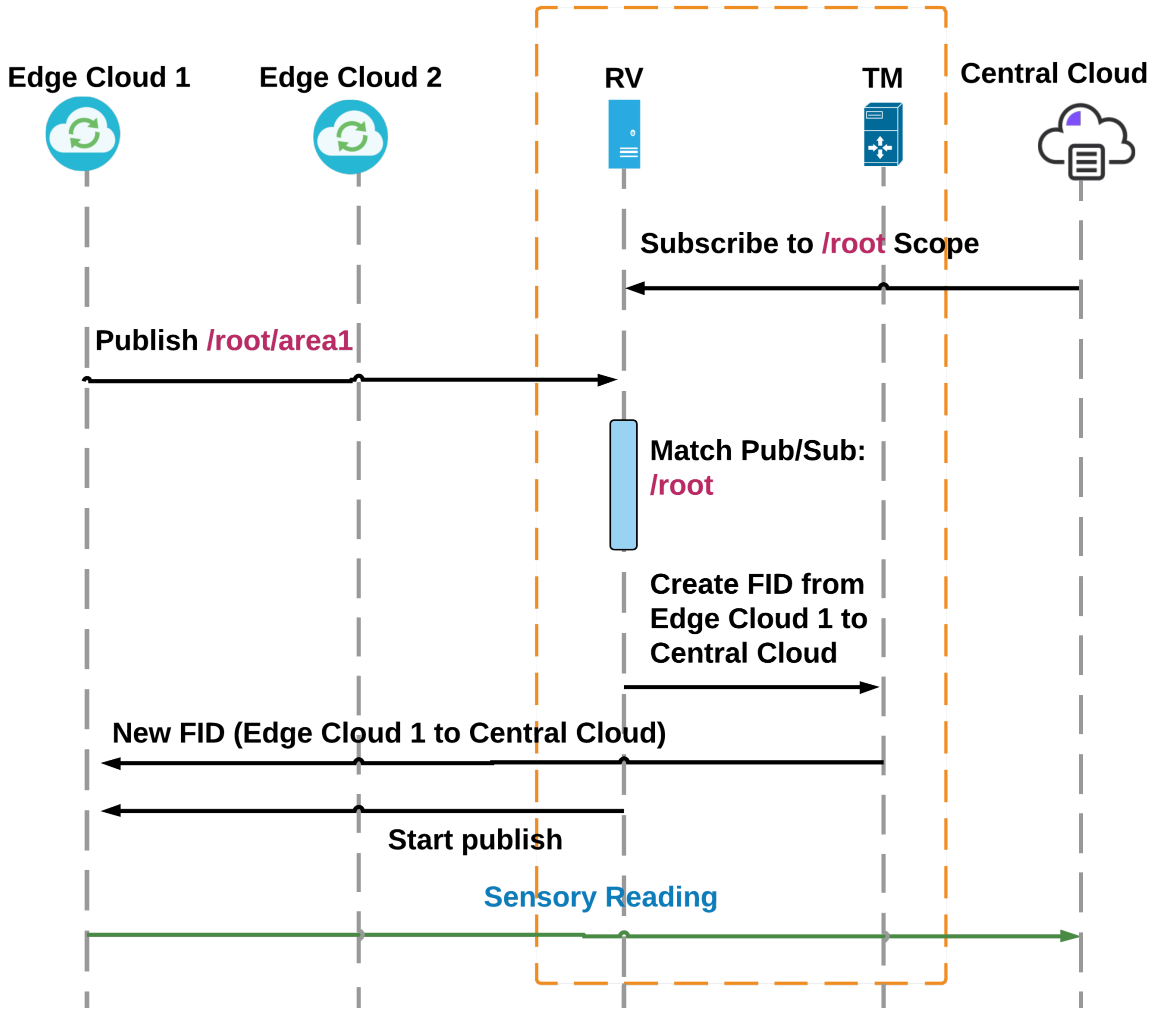
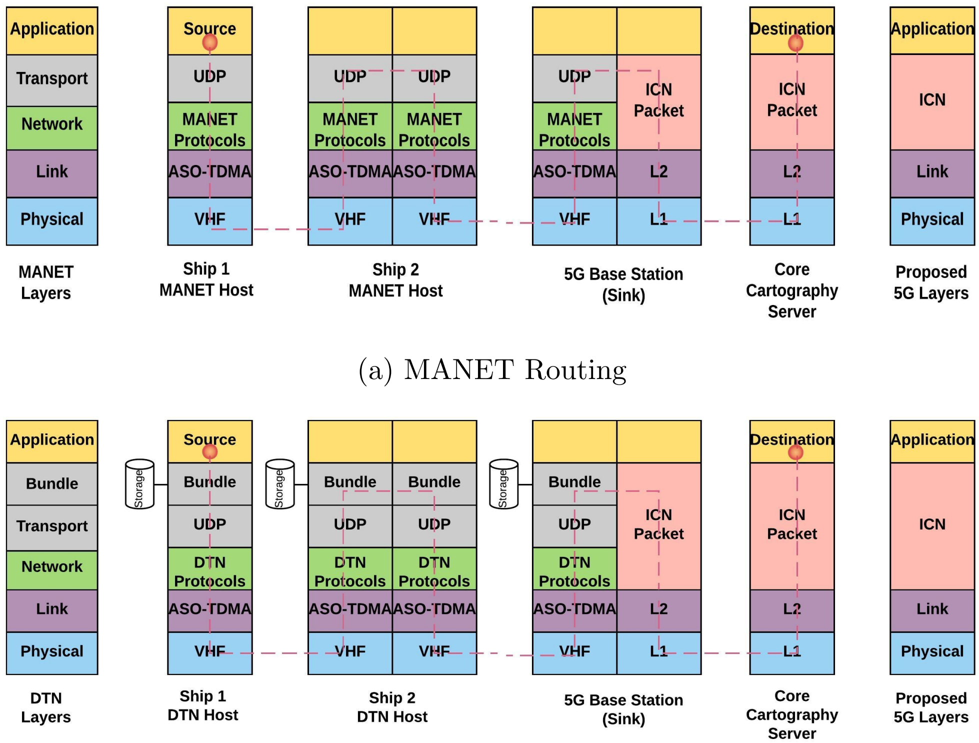


Figure 6.4: Sequence Diagram for Marine Data Management and Service Provisioning

the aggregation and provisioning of the data received from edge clouds. Other use cases may necessitate other subscription scenarios, but are out of scope of this work.

After a match publication/subscription happens at the RV, the later informs the TM to create a FID for the path from edge cloud 1 to the central cloud. The TM then creates the FID and sends it to the edge cloud. However, creating a new FID is only needed for the first sensory reading sent from the edge cloud towards the central cloud. Subsequent readings are then sent using the same FID. After receiving the



(a) MANET Routing

(b) DTN Routing

Figure 6.5: Ship Ad-hoc Networks/ICN Abstraction and Interfacing

FID, the edge cloud receives a start publish message from the RV, and has permission to start sending sensory readings towards the central cloud using the FID provided by the TM. Sensory data packets arrive at the 5G base stations at shore which act as sensory data sink nodes. Several base stations are proposed to cover a specific water surface area as described in the namespace for managing marine sensory data above. These base stations all connect to an edge cloud that serves the identified area and is responsible for aggregating and managing all the sensory data related to that area. The edge cloud includes a Network Abstraction Function (NAF) [96] [97] that occupies a key role in Ship Ad-hoc Networks/ICN abstraction and interfacing as shown in Fig. 6.5. When a new SANET packet is received at the base station, it is forwarded to the responsible edge cloud that inspects the packet for the sensory information described previously, such as the location identifier, time stamp and sensory reading. The edge clouds NAF constructs the scope /root/area/location/time stamp/reading from the received packet and compares it to a local database shared between all the base stations that cover the identified area. This local database is foreseen as part of the MEC resources available to the identified area. If the entry is found in the database, then it is dropped and considered a redundant reading. Otherwise, the scope including the reading itself is added to the local database, and also sent in full as an ICN packet towards the central cloud using the previously specified FID.

It is assumed that the domain RV maintains subscription state, and that duplicate publications from edge clouds are not permitted by the local area databases and therefore, duplicate sensor readings are not sent towards the central cloud even if they originated from different base stations within the area.

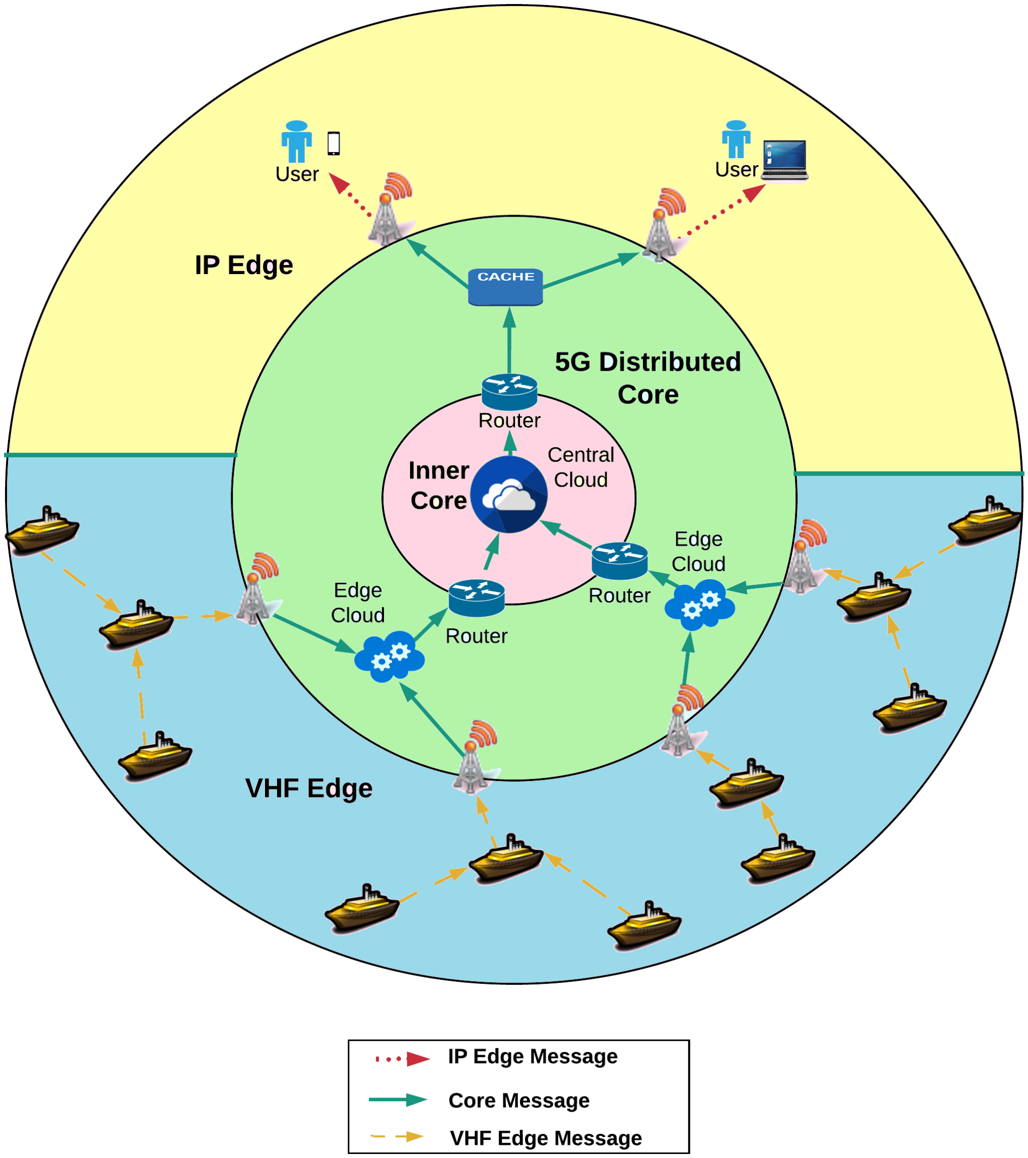


Figure 6.6: VHF over 5G

6.2.3 Benefits of an Information Centric Core

Utilizing an information centric core to manage IoMaT data offers significant benefits over traditional IP networks beyond sensory data aggregation. For instance, an information centric core that follows a publish/subscribe model can natively provide an efficient content retrieval mechanism that facilitates data retrieval from the closest edge cloud rather than the central one. With traditional IP networks, this would require overlay solutions such as DNS redirect that only provide a sub-optimal solution to the scenario above [98]. Another benefit is ICNs native support for multicast which facilitates aggregating requests coming from several consumers which reduces the overall traffic. In ICN, a single multicast FID created by the TM can identify a shortest delivery path for several users located at different locations. However, with traditional IP networks, supporting IP multicast is not so straightforward and comes with substantial delay and signalling cost [99]. A common capability of ICN is data replication and in-network caching. Delivering replicated data objects from caches decouples content consumption from data sources. Therefore, in-network caching can improve network performance by fetching content from nodes that are geographically placed closer to the end user. In current IP networks, caching and content distribution is approached by employing complex application-layer overlays such as Content Distribution Networks (CDNs); where data traffic typically follows sub-optimal paths as it is effectively routed, depending on the overlay topology instead of the Internet-layer topology. In ICN, caching is inherently possible at any network element, including routers and even end-user devices [100].

6.3 Overview of the Cartography System Architecture

The cartography system proposed in this work and illustrated in fig. 6.6 can collect data including but not limited to: sea state, depth, temperature, wind speed/direction, humidity, salinity .. etc. The proposed system consists of four layers: The marine sensors layer, the marine network layer, the cloud data management layer and the data analysis and cartography application layer. Figure 6.7 shows the overall architecture of the proposed cartography system.

1. The marine sensors layer:

   This layer consists of the various deployed marine sensors such as sea state, depth, temperature, wind speed/direction, humidity, salinity sensors .. etc. Sensors can sample numerical values reflecting the states of the monitored objects. A quantization and compression method specific to marine sensory data has been published in [91]. For each of the sensors mentioned previously we have set the extreme lower and upper limits of the sensors readings likely to be found in the marine environment as well as the level of accuracy required to represent each reading. The predictability of gathered sensor data makes it beneficial to quantize the data to reduce the amount of bits needed to represent each reading in the binary representation. Applying this quantization in conjunction with the compression algorithm (AMDC) has given effective data compression rates in comparison to the main compression methods as shown in Chapter 4.

2. The marine network layer:

This layer consists the ship connectivity layer with each ship hosting several sensors. The main task of each ship is to receive, store, and process the raw sampling data from the connected sensors, and then send the processed data (sampled data) to the cloud data management layer. The marine network layer possesses distinctive characteristics that are unique to the marine environment. Ship connectivity can be very dense in certain busy locations across main shipping channels and very sparse in others such as deep ocean and shallows. In addition, there is no infrastructure at sea that can be readily used for network connectivity. These characteristics impose the use of specific network protocols as discussed in detail in Chapters 3 and 5 .

3. The cloud data management layer:

   The cloud data management layer is responsible for the efficient synchronization and transmission of sensory data at the network core between the edge and central clouds in order to preserve network resources and guarantee efficient delivery. The cloud data management layer uses an Information Centric Networking (ICN) approach that provides an efficient way of information dissemination by identifying information at the network layer. This approach helps to synchronise sensory data delivery between edge clouds and ensure duplicateless arrival of data at the central cloud as discussed earlier in this Chapter.

4. The cartography data analysis and application layer:

   This layer consists of the data analysis, database management, data mining and cartography plotting at the central cloud data repository based on the massive sensor sampling data collected at the cloud data management layer. Through data analysis, we can get more useful information about the monitored objects and about the marine environment in general. Also further pattern behaviours can be extracted and predicted through the use of data mining.

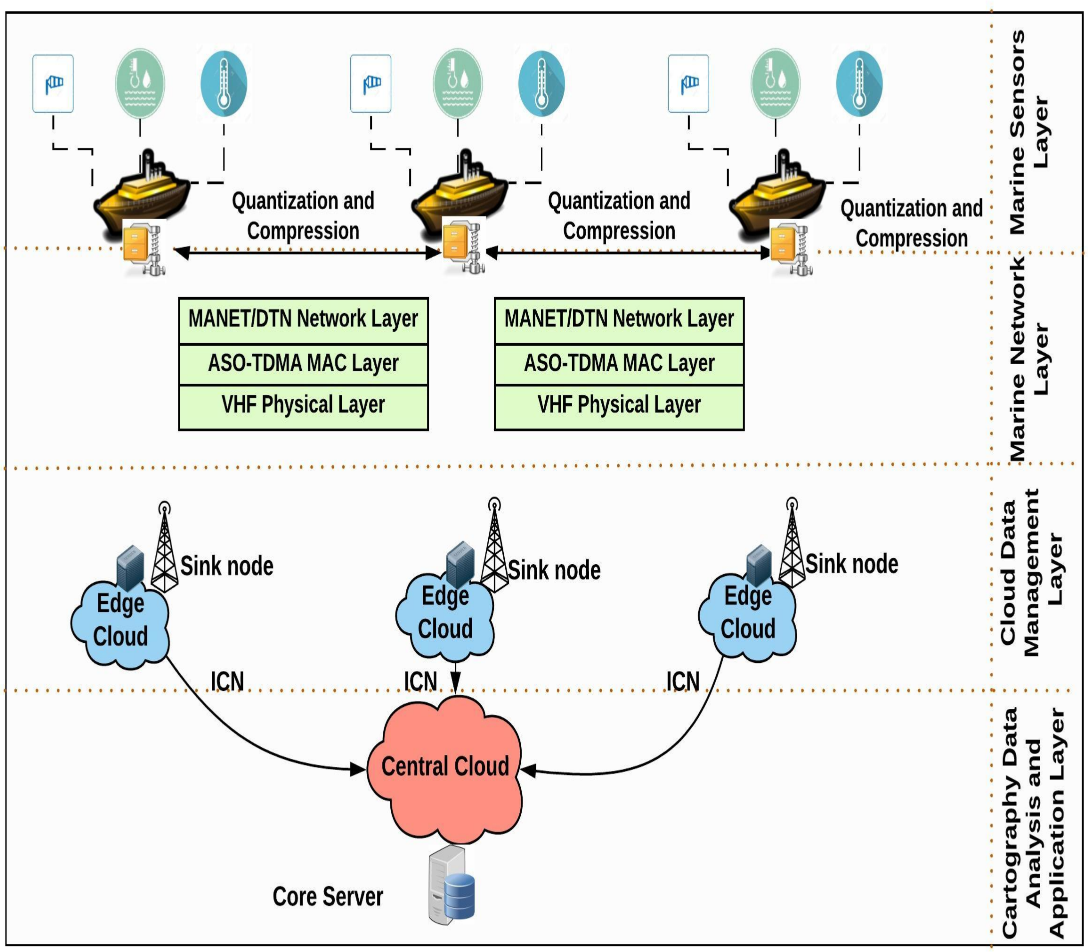


Figure 6.7: Cartography System Service Layers

6.4 Proposed System Prototype and IoMaT Software Components

The Automatic Identification System (AIS) is a worldwide automatic positioning system based on fitting small transponders to vessels that continuously transmit a signal. This alerts other vessels and shore stations with AIS receivers to the presence of that vessel. The position information is supplemented with additional information about the vessel. The signals and accompanying information can then be received by any vessel, land station or satellite fitted with an AIS receiver, and is then typically displayed on a screen using interactive chart-plotting software [101].

Designed to prevent marine collisions, AIS is essentially a VHF marine band data network that operates primarily on two dedicated VHF channels (AIS1 - 161,975 MHz and AIS2 - 162,025 MHz). Where these channels are not available regionally, the AIS is capable of automatically switching to alternate designated channels [102]. Any AIS system requires two inputs and one output to function at its most basic level. The inputs are a GPS feed, so that it can locate its position, and a VHF feed, so that it can receive incoming AIS signals from other vessels. The output is also a VHF connection, so that it can transmit its position and core vessel information. All that is needed then is power. Thereafter, additional information from on board sensors can be fed in through a data feed as required from other systems [103]. AIS link layer communication is based on a self organizing TDMA access protocol. This means that a time period is divided into a number of time slots. When an AIS system switches on, it looks for a vacant time slot and reserves it. Other systems in range will avoid this slot and select another one. Precise timing is needed to ensure that all vessels are synchronised and this is derived from GPS; hence why an AIS system has its own GPS receiver [102]. In this section, we discuss the AIS system level considerations and how the IoMaT software components are designed to interact with the existing system. We also discuss the existing AIS message format and the proposed extensions for IoMaT support.

6.4.1 AIS System Level Considerations and IoMaT Integration

The AIS system in Figure 6.8 shows the principal parts of a ship borne mobile AIS station and the required components to support IoMaT integration. The AIS transceiver transmits and receives radio signals that form the VHF Data Link and interconnect the AIS stations to each other. The GPS receiver supplies the coordinated universal time (UTC) to the AIS station to correct its own clock in order to synchronize all transmissions such that there are no collisions or overlaps which would degrade the information being transmitted [103]. Most AIS models have a 9-pin NMEA port which can be connected to a computer serial port using a standard RS 232 serial cable [104] [105]. In our proposed system, this port is used for relaying the IoMaT sensory data captured on each ship. Therefore, the existing VHF infrastructure is completely utilized for data transmission and no additional equipment is needed. Although only one radio channel is necessary for communication, each AIS station transmits and receives over two radio channels to avoid interference problems, and to allow channels to be shifted without communications loss from other ships. Therefore, in the proposed system, one channel can be used for voice while the other can be used for data communication simultaneously. However, in emergency situations or when channel shifting is needed, both channels can be utilized for voice usage.

Figure 6.8 also shows the proposed IoMaT software module hosted on a laptop/PC that is connected to the AIS system through an RS 232 serial interface. The module consists of the IoMaT message parser/encapsulator, the AOMDV/DTN software router and the DTN buffer memory. The IoMaT message parser/encapsulator is responsible for parsing and processing the IoMaT input signals, processing messages into suitable transmission packets and sending the IoMaT output signals through the appropriate interface. The MADNET/DTN software router is responsible for routing the IoMaT messages through the SANET and making routing decisions based on IoMaT control messages. The DTN buffer memory provides a temporary storage unit for IoMaT DTN messages until a suitable target ship is available to receive the message or until the message is dropped on time-out. When a received IoMaT message is first detected at the IoMaT message parser/encapsulator, the message is processed and either sent to the AIS module for re-transmission using MADNET/DTN routing if a target ship is available and suitable, or else, sent to the DTN buffer memory. On the other hand, if the message is received locally through one of the IoMaT sensory units, the IoMaT message parser/encapsulator encapsulates the sensory information into a suitable AIS IoMaT message format, and proceeds to the routing phase where the message is either buffered at the DTN buffer memory or sent directly through the AIS module using AOMDV or DTN routing whichever suitable.

6.4.2 AIS IoMaT Message Format

AIS messages follow a specific format defined by the AIVDM/AIVDO protocol [106]. AIS receivers report ASCII data packets as a byte stream over serial or USB lines, using the NMEA 0183 or NMEA 2000 data formats. AIS packets have the introducer "!AIVDM" or "!AIVDO"; AIVDM packets are reports from other ships and AIVDO packets are reports from your own ship. To incorporate IoMaT messages, we propose the use of two more introducers that uniquely identify IoMaT packets in an AIS system, namely "!AIVDN" and "!AIVDP". In the same sense, AIVDN identifies IoMaT packets from other ships and AIVDP identifies IoMaT packets from your own ship. Figure 6.9 shows an example of a standard AIS packet Figure 6.9a and IoMaT packet Figure 6.9b in an AIS system.

Every field in the packet has a specific meaning and is translated according to the AIVDM/AIVDO protocol as follows. Field 1 identifies the packet type, i.e., in Fig. 6.9a, AIVDM indicates an AIS packet received from another ship and in Figure 6.9b, AIVDN indicates an IoMaT packet also received from another ship. Field 2 identifies the number of fragments in an accumulating message, where in the examples of Figure 6.9 it is 1, i.e., the message consists of one fragment. Field 3 identifies the sequence number of the current fragment, which is 1, meaning that this is the first fragment. And since the previous field is also 1, therefore the example packet is complete on its own. Field 4 is a sequential message ID for multi-sentence messages. Field 5 is a radio channel code, where AIS Channel A is 161.975Mhz and AIS Channel B is 162.025Mhz. Field 6 is the data payload, where in the case of standard AIS packet Figure 6.9a, every character of the ASCII code sequence, is interpreted into a specific message according to the AIVDM/AIVDO protocol (interested readers are referred to [106] for a complete mapping guide). While in the case of IoMaT packet (Fig. 6.9b), the IoMaT message parser/encapsulator translates the ASCII sequence of a IoMaT packet into a binary bit stream and the packet is processed at binary level.

Finally Field 7 is the number of fill bits requires to pad the data payload to a 6 bit boundary (0 in this example), and the *-separated suffix (*5C) is the NMEA 0183 data integrity checksum for the sentence.

6.5 Summary

In this Chapter, the proposed marine cartography system has been presented as the realization of 5G IoMaT connectivity. A data synchronisation and transmission approach has been proposed at the 5G network core using ICN. This is aimed at providing efficient and duplicate-less transmission of marine sensory readings from the base station/sink nodes towards the central cloud in order to preserve network resources and guarantee efficient delivery. Part of the Mobile Edge Computing resources is exploited as edge repositories (clouds) of the sensory data delivered to the shore. The edge clouds eventually connect to a central cloud in the internet where all the sensory data is aggregated, filtered and analysed to produce real-time maps of surface and under water environmental information that produces accumulative maps for beneficiary customers. An overview of the cartography system architecture has also been presented in this chapter in the form of four distinctive layers. namely, the marine sensor layer, the marine network layer, the cloud data management layer and the cartography data analysis and application layer. An IoMaT software module design has also been presented with a discussion of the AIS system level considerations and how the IoMaT software components are proposed to interact with the existing system. The existing AIS message format is also discussed with the proposed extensions for IoMaT support.

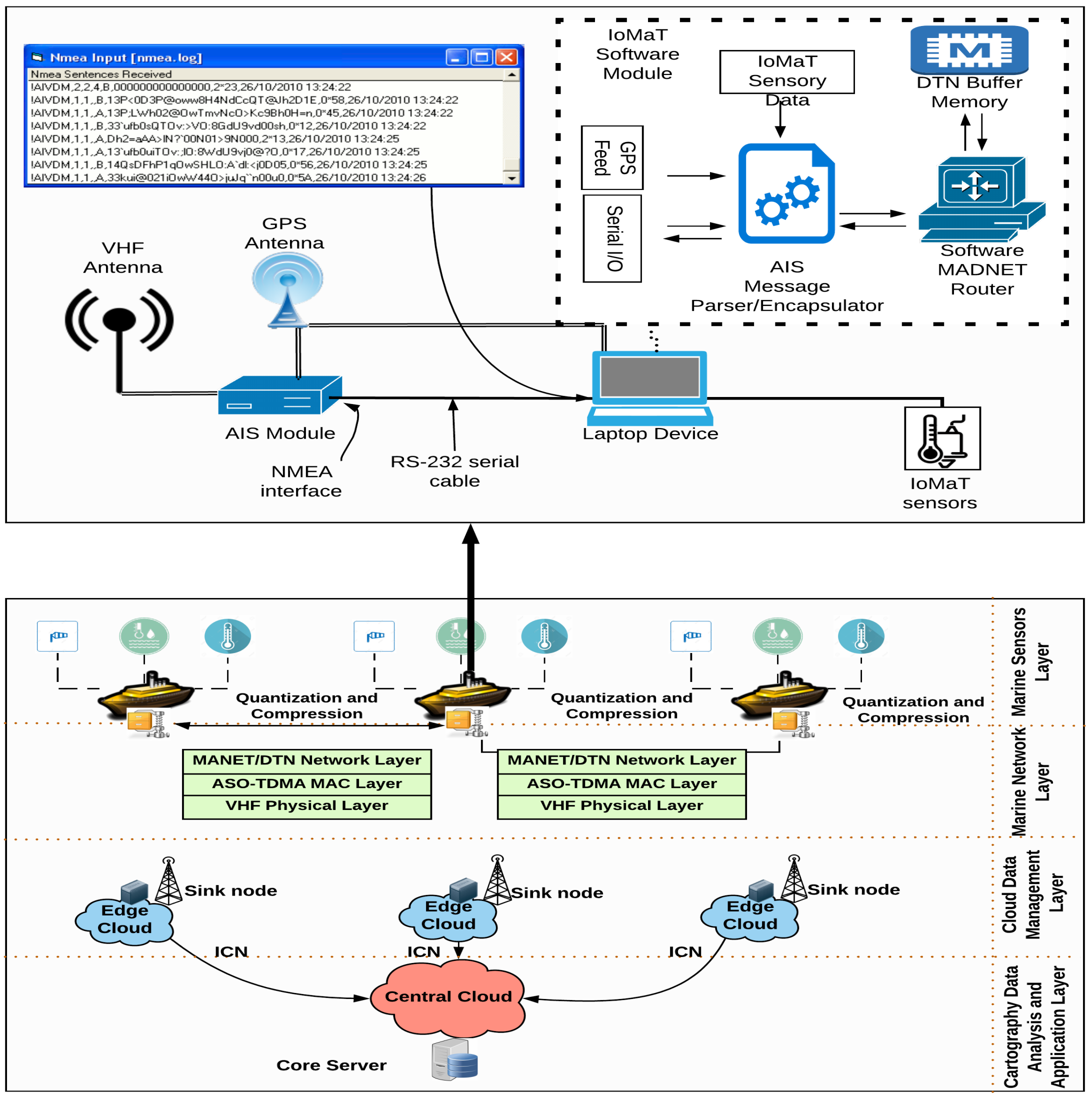


Figure 6.8: AIS with IoMaT Software Module

(a) AIS Standard Packet

(b) AIS IoMaT Packet

Figure 6.9: AIS ASCII Packet Format.

Chapter 7

Conclusions and Future Work

This chapter provides concluding remarks of the work presented in this thesis and guidelines for future work in light of the proposed solutions.

7.1 Conclusions

In this thesis, a novel Internet of Marine Things (IoMaT) data acquisition and cartography system has been proposed in the marine environment using Very High Frequency (VHF) available on the majority of ships. The proposed system is equipped with many sensors such as sea depth, temperature, wind speed and direction, and the collected data is sent to 5G edge cloudlets connected to sink/base station nodes on shore. The sensory data is ultimately aggregated at a central cloud on the internet to produce up to date cartography systems.

Implementing MANETs in a marine environment has been shown to be possible using the existing VHF communication infrastructure available on ships. This infrastructure is effective for transmission ranges of up to 40 Km carrying data traffic within the proposed network. The applied MANET protocols have proven to have limitations in the marine environment, which can be sparse in some locations and dense in others. It has been concluded that AOMDV is the most efficient MANET for Ship Ad-hoc Networks. This is due to its multipath route discovery process which takes into account frequent link breakages in networks with high mobility rates, maintaining alternative routes whenever needed.

The impact of marine traffic patterns on the performance of MANET routing protocols has been investigated. The results show that real-life marine density and mobility patterns show better (PDR) rates than the Random Waypoint Mobility model. This is because the navigational routes drawn by most vessels at sea tend to follow pathways which aid multi hop connectivity for data packets hopping towards the destination. This is not always the case as there are ships can be scattered sparsely so such pathways do not exist. It has been found that in high density and mobility environments, AOMDV is a better choice than AODV and DSDV for marine applications. It can be concluded that MANET protocol performance has a positive relationship with density and an inverse relationship with mobility and sparseness. PDR rates increase when density increases and on the contrary decrease when mobility or sparseness increases.

Due to bandwidth limitations of the VHF channel, minimizing transmission data redundancy overhead is essential for efficient use of the transmission channel. For our marine application, the predictability of gathered sensor data makes it beneficial to quantize the data to reduce the amount of bits needed to represent each reading in the binary representation. Applying this quantization in conjunction with the proposed compression algorithm (AMDC) has proved affective data compression rates in comparison with the major known compression methods.

The problem of sparsity in the marine environment has necessitated the inclusion of selective Opportunistic Networking and MANET protocols. We have evaluated the deployment of the proposed hybrid MANET/DTN routing protocol (MADNET) in the marine environment with both dense and sparse scenarios. Obtained results have verified the effectiveness of the proposed protocol in all evaluated scenarios.

The drawbacks of low rate data transmission offered by VHF radio has also been investigated in terms of the network bottlenecks near the sink and the achievable data rate collected at the sink for each ship within the network. Results show that even though the proposed SANET operates over VHF radio that offers low rate data transmission at 28.8 kbps; high sensory data delivery rates of up to 149 minutes of collected data per ship per minute have been observed.

The proposed marine cartography system has been presented as the realization of 5G IoMaT connectivity. A data synchronisation and transmission approach has been proposed at the 5G network core using ICN. This is aimed at providing efficient and duplicate-less transmission of marine sensory readings from the base station/sink nodes towards the central cloud in order to preserve network resources and guarantee efficient delivery. An overview of the cartography system architecture has also been presented in the form of four distinctive layers that describe the whole system functionality.

Finally, an IoMaT software module design has been presented with a discussion of the AIS system level considerations and how the IoMaT software components are proposed to interact with the existing system. The existing AIS message format is also discussed with the proposed extensions for IoMaT support.

In terms of research limitations, we have identified the need for security mechanisms such as encryption that can efficiently protect the proposed marine communication system and prevent unauthorized third parties or threat actors from accessing the data. In addition, since no mobility model specific for the marine environment exists, real marine mobility traces were used in the evaluation scenarios presented in this work. A mobility model specific to the marine environment would enable easier and more extensive evaluations. Finally, the VHF transmission used in this work is limited to over water communication and is not suitable for underwater networks.

7.2 Future Work

With the proposed IoMaT system, the potential exists for scenarios where the transmitted data can have commercial value and should not be accessible free of charge and protection. This was not seen necessary as part of the existing AIS system since its goals were limited to location awareness, navigation aid and safety. Therefore, efficient security mechanisms are needed to fulfil this requirement.

The need for a mobility model specific to the marine environment has also been identified, where an extensive study of ship trajectories can be exploited to formalize the model. Local parameters can also be used to identify mobility behaviours in various water surfaces around the globe. This would be very beneficial for marine network research.

Another useful addition to the proposed IoMaT system would be the involvement of underwater networks, i.e., submarines, underwater sensors, etc. This would enrich the proposed marine cartography maps with new possibilities of environmental data such as underwater data collection, pollution monitoring, offshore exploration and disaster prevention.

Finally, the development of a working prototype is proposed using a number of AIS transponders, marine VHF channels and PC based source, relay and sink nodes to replace the NS2 Simulation.

Bibliography

[1] V. Borrel, M. H. Ammar, and E. W. Zegura, "Understanding the wireless and mobile network space: a routing-centered classification," in *Proceedings of the second ACM workshop on Challenged networks*, pp. 11–18, ACM, 2007.

[2] F. Weng, X. Zou, and Z. Qin, "Uncertainty of amsu-a derived temperature trends in relationship with clouds and precipitation over ocean," *Climate dynamics*, vol. 43, no. 5-6, pp. 1439–1448, 2014.

[3] B. Hui, K. Jeon, K. Chang, S. Kim, J. Park, and Y. Lim, "Design of radio transmission technologies for vhf band ship ad-hoc network," in *ICT Convergence (ICTC), International Conference on*, pp. 626–629, IEEE, 2011.

[4] M. Manoufali, H. Alshaer, P.-Y. Kong, and S. Jimaa, "Technologies and networks supporting maritime wireless mesh communications," in *Wireless and Mobile Networking Conference (WMNC), 6th Joint IFIP*, pp. 1–8, IEEE, 2013.

[5] C. Yun and Y.-k. Lim, "Aso-tdma: ad-hoc self-organizing tdma protocol for shipborne ad-hoc networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–13, 2012.

[6] E. M. Royer and C.-K. Toh, “A review of current routing protocols for ad hoc mobile wireless networks,” *IEEE personal communications*, vol. 6, no. 2, pp. 46–55, 1999.

[7] Z. Zhang, “Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges,” *IEEE Communications Surveys & Tutorials*, vol. 8, no. 1, pp. 24–37, 2006.

[8] Ofcom, “More radio spectrum for the internet of things.”

[9] F. Inc., “Live marine traffic.”

[10] J. Pomfret, “Guidelines on use of marine-band vhf radio by inland pleasure craft on uk freight water ways,” March 2013.

[11] M. O. Ajewole, O. D. Oyedum, A. T. Adediji, A. S. Moses, and J. O. Eiche, “Spatial variability of vhf/uhf electric field strength in niger state, nigeria,” *International Journal of Digital Information and Wireless Communications (IJDIWC)*, vol. 3, no. 3, pp. 26–34, 2013.

[12] K. Schwehr and P. McGillivary, “Marine ship automatic identification system (ais) for enhanced coastal security capabilities: An oil spill tracking application,” in *OCEANS 2007*, pp. 1–9, IEEE, Sept 2007.

[13] F. Zhu and L. Lei, “Research on a method for analysis of ship traffic density in harbor water area based on gis,” in *Conference Anthology, IEEE*, pp. 1–4, IEEE, 2013.

[14] I. Recommendation, "Characteristics of vhf radio system and equipment for the exchange of data and electronic mail in the maritime mobile service rr appendix 18 channels," 2008.

[15] M. Series, "Technical characteristics for an automatic identification system using time-division multiple access in the vhf maritime mobile band," 2010.

[16] V. Del Duca Almeida, A. B. Oliveira, D. F. Macedo, and J. Nogueira, "Performance evaluation of manet and dtn routing protocols," in *Wireless Days (WD), 2012 IFIP*, pp. 1–6, IEEE, 2012.

[17] M. Patil and R. C. Biradar, "A survey on routing protocols in wireless sensor networks," in *Networks (ICON), 2012 18th IEEE International Conference on*, pp. 86–91, IEEE, 2012.

[18] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 27–34, ACM, 2003.

[19] A. Boukerche, *Algorithms and protocols for wireless, mobile Ad Hoc networks*, vol. 77. John Wiley & Sons, 2008.

[20] A. K. Gupta, H. Sadawarti, and A. K. Verma, "Performance analysis of aodv, dsr & tora routing protocols," *IACSIT international journal of Engineering and Technology*, vol. 2, no. 2, pp. 226–231, 2010.

[21] S. Mohseni, R. Hassan, A. Patel, and R. Razali, "Comparative review study of reactive and proactive routing protocols in manets," in *Digital Ecosystems and*

Technologies (DEST), 2010 4th IEEE International Conference on, pp. 304–309, IEEE, 2010.

[22] M. Abolhasan, T. Wysocki, and E. Dutkiewicz, "A review of routing protocols for mobile ad hoc networks," *Ad hoc networks*, vol. 2, no. 1, pp. 1–22, 2004.

[23] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Mobile Computing Systems and Applications, 1999. Proceedings. WMCSA'99. Second IEEE Workshop on*, pp. 90–100, IEEE, 1999.

[24] A. Moravejosharieh, H. Modares, R. Salleh, and E. Mostajeran, "Performance analysis of aodv, aomdv, dsr, dsdv routing protocols in vehicular ad hoc network," *Research Journal of Recent Sciences ISSN*, vol. 2277, p. 2502, 2013.

[25] M. K. Marina and S. R. Das, "Ad hoc on-demand multipath distance vector routing," *Wireless communications and mobile computing*, vol. 6, no. 7, pp. 969–988, 2006.

[26] S. Biradar, K. Majumder, S. K. Sarkar, and C. Puttamadappa, "Performance evaluation and comparison of aodv and aomdv.," *International Journal on Computer Science & Engineering*, vol. 2, no. 2, 2010.

[27] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, pp. 85–97, ACM, 1998.

[28] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, "Scenario-based performance analysis of routing protocols for mobile ad-hoc

networks," in *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pp. 195–206, ACM, 1999.

[29] P. Parvathi, "Comparative analysis of cbrp, aodv, dsdv routing protocols in mobile ad-hoc networks," in *Computing, Communication and Applications (ICCA), 2012 International Conference on*, pp. 1–4, IEEE, 2012.

[30] N. S. M. Usop, A. Abdullah, and A. F. A. Abidin, "Performance evaluation of aodv, dsdv & dsr routing protocol in grid environment," *IJCSNS International Journal of Computer Science and Network Security*, vol. 9, no. 7, pp. 261–268, 2009.

[31] X. Chen, J. Shen, and J. Wu, "Improving routing protocol performance in delay tolerant networks using extended information," *Journal of Systems and Software*, vol. 83, no. 8, pp. 1301–1309, 2010.

[32] K. Scott and S. Burleigh, "Bundle protocol specification, ietf rfc 5050, experimental," 2007.

[33] M. Demmer, J. Ott, and S. Perreault, "Delay-tolerant networking tcp convergence-layer protocol," tech. rep., 2014.

[34] L. Wan, F. Liu, J. Zhang, and H. Zhang, "Performance evaluation of routing protocols for delay tolerant networks," *Computer Science & Information Technology*, vol. 5, no. 11, 2015.

[35] T. Abdelkader, K. Naik, A. Nayak, N. Goel, and V. Srivastava, "A performance comparison of delay-tolerant network routing protocols," *IEEE Network*, vol. 30, no. 2, pp. 46–53, 2016.

[36] H. S. B. Shally and M. Garg, "Performance evaluation of rapid, and sppray-and-wait dtn routing protocols under backhole attack," *International Journal of Research in Engineering and Technology*, vol. 3, no. 1, pp. 35–40, 2014.

[37] C. Alippi, R. Camplani, C. Galperti, and M. Roveri, "A robust, adaptive, solar-powered wsn framework for aquatic environmental monitoring," *IEEE Sensors Journal*, vol. 11, no. 1, pp. 45–55, 2011.

[38] C. A. Pérez, M. Jimenez, F. Soto, R. Torres, J. López, and A. Iborra, "A system for monitoring marine environments based on wireless sensor networks," in *OCEANS 2011 IEEE-Spain*, pp. 1–6, IEEE, 2011.

[39] M.-T. Zhou, H. Harada, P.-Y. Kong, C.-W. Ang, Y. Ge, and J. Pathmasuntharama, "A method to deliver aodv routing messages using wimax mesh mac control messages in maritime wireless networks," in *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1537–1541, IEEE, 2009.

[40] A. J. Pullin, S. Presland, and C. Pattinson, "Using ship movement in the irish sea for manet evaluation," in *Computer Modeling and Simulation. EMS'08. Second UKSIM European Symposium on*, pp. 394–399, IEEE, 2008.

[41] Z. Haider and F. Shabbir, "Genetic based approach for optimized routing in maritime tactical manets," in *Applied Sciences and Technology (IBCAST), 2014 11th International Bhurban Conference on*, pp. 488–492, IEEE, 2014.

[42] D. Do, H. Nguyen, N. Tran, T. Ta, T. Tran, and Y. Vu, "Wireless ad hoc network based on global positioning system for marine monitoring, searching

and rescuing (msnr)," in *Microwave Conference Proceedings (APMC), 2011 Asia-Pacific*, pp. 1510–1513, IEEE, 2011.

[43] J.-C. Kuo, W. Liao, and T.-C. Hou, "Impact of node density on throughput and delay scaling in multi-hop wireless networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 10, pp. 5103–5111, 2009.

[44] S. S. Kulkarni, A. Iyer, C. Rosenberg, and D. Kofman, "Routing dependent node density requirements for connectivity in multi-hop wireless networks," in *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, vol. 5, pp. 2890–2896, IEEE, 2004.

[45] J.-C. Kuo and W. Liao, "Hop count distribution of multihop paths in wireless networks with arbitrary node density: Modeling and its applications," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 4, pp. 2321–2331, 2007.

[46] Y. Gu, R. V. Prasad, and I. Niemegeers, "Analyzing the effect of node mobility in clustered wireless ad hoc networks," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pp. 1–6, IEEE, 2009.

[47] M. Bhatt, R. Chokshi, S. Desai, S. Panichpapiboon, N. Wisitpongphan, and O. K. Tonguz, "Impact of mobility on the performance of ad hoc wireless networks," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, vol. 5, pp. 3025–3029, IEEE, 2003.

[48] J. Ott, D. Kutscher, and C. Dwertmann, "Integrating dtn and manet routing," in *Proceedings of the 2006 SIGCOMM workshop on Challenged networks*, pp. 221–228, ACM, 2006.

[49] J. Lakkakorpi, M. Pitkänen, and J. Ott, "Adaptive routing in mobile opportunistic networks," in *Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems*, pp. 101–109, ACM, 2010.

[50] C. Kretschmer, S. Ruhrup, and C. Schindelhauer, "Dt-dymo: delay-tolerant dynamic manet on-demand routing," in *Distributed Computing Systems Workshops, 2009. ICDCS Workshops' 09. 29th IEEE International Conference on*, pp. 493–498, IEEE, 2009.

[51] R. Pant, A. Tunpan, P. Mekbungwan, R. Virochpoka, and K. Kanchanasut, "Dtn overlay on olsr network," in *Proceedings of the Sixth Asian Internet Engineering Conference*, pp. 56–63, ACM, 2010.

[52] Q. Ma, K. Liu, T. Zhu, W. Gong, and Y. Liu, "Bond: Exploring hidden bottleneck nodes in large-scale wireless sensor networks," in *Distributed Computing Systems (ICDCS), 2014 IEEE 34th International Conference on*, pp. 399–408, IEEE, 2014.

[53] V. D. Hoang, Z. Shao, and M. Fujise, "A new solution to estimate the available bandwidth in manets," in *2006 IEEE 63rd Vehicular Technology Conference*, vol. 2, pp. 653–657, IEEE, 2006.

[54] C. Wu, K. Kumekawa, and T. Kato, "A manet protocol considering link stability and bandwidth efficiency," in *2009 International Conference on Ultra Modern Telecommunications & Workshops*, pp. 1–8, IEEE, 2009.

[55] N. Wright and H. Chan, "Low-cost internet of things ocean observation," in *OCEANS 2016 MTS/IEEE Monterey*, pp. 1–5, IEEE, 2016.

[56] Y. Zixuan, W. Zhifang, and L. Chang, "Research on marine environmental monitoring system based on the internet of things technology," in *Electronic Information and Communication Technology (ICEICT), IEEE International Conference on*, pp. 121–125, IEEE, 2016.

[57] C.-C. Kao, Y.-S. Lin, G.-D. Wu, and C.-J. Huang, "A comprehensive study on the internet of underwater things: Applications, challenges, and channel models," *Sensors*, vol. 17, no. 7, p. 1477, 2017.

[58] G. Xu, W. Shen, and X. Wang, "Applications of wireless sensor networks in marine environment monitoring: A survey," *Sensors*, vol. 14, no. 9, pp. 16932–16954, 2014.

[59] H. Thriveni, G. M. Kumar, and R. Sharma, "Performance evaluation of routing protocols in mobile ad-hoc networks with varying node density and node mobility," in *Communication Systems and Network Technologies (CSNT), 2013 International Conference on*, pp. 252–256, IEEE, 2013.

[60] E. Natsheh and K. Buragga, "Nodes density and broadcast management in heterogeneous environments of mobile ad-hoc networks," *Journal of Computer Science*, vol. 6, no. 3, p. 312, 2010.

[61] J. Zuo, Y. Wang, Y. Liu, and Y. Zhang, "Performance evaluation of routing protocol in vanet with vehicle-node density," in *Wireless Communications Net-*

working and Mobile Computing (WiCOM), 2010 6th International Conference on, pp. 1–4, IEEE, 2010.

[62] F. Bai and A. Helmy, “A survey of mobility models,” *Wireless Adhoc Networks. University of Southern California, USA*, vol. 206, 2004.

[63] D. Raymond, I. Burbey, Y. Zhao, S. Midkiff, and C. P. Koelling, “Impact of mobility models on simulated ad hoc network performance,” in *Proc. 9th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pp. 398–402, 2006.

[64] J. Yoon, M. Liu, and B. Noble, “Random waypoint considered harmful,” in *INFOCOM 2003. twenty-second annual joint conference of the IEEE computer and communications. IEEE societies*, vol. 2, pp. 1312–1321, IEEE, 2003.

[65] K. Amjad and A. J. Stocker, “Impact of node density and mobility on the performance of aodv and dsr in manets,” in *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*, pp. 61–65, IEEE, 2010.

[66] F. Bai, N. Sadagopan, and A. Helmy, “Important: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks,” in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 2, pp. 825–835, IEEE, 2003.

[67] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless communications and mobile computing*, vol. 2, no. 5, pp. 483–502, 2002.

[68] R. Mohsin and J. Woods, "Performance evaluation of manet routing protocols in a maritime environment," in *Computer Science and Electronic Engineering Conference (CEEC), 2014 6th*, pp. 1–5, IEEE, 2014.

[69] A. Mitra, "Lecture notes on mobile communication," *A Curriculum Development Cell project Under QIP, IIT Guwahati*, 2009.

[70] A. K. Maurya, D. Singh, and A. K. Sarje, "Median predictor based data compression algorithm for wireless sensor network," *International Journal of Smart Sensors and Ad Hoc Networks*, vol. 1, no. 1, pp. 62–65, 2011.

[71] "Digital signals - sampling and quantization."

[72] H. Johansson, "Sampling and quantization," *Academic Press Library in Signal Processing: Signal Processing Theory and Machine Learning*, vol. 1, p. 169, 2013.

[73] W. Sheppard, "On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant division of a scale," *Proceedings of the London Mathematical Society*, vol. 1, no. 1, pp. 353–380, 1897.

[74] R. M. Gray and D. L. Neuhoff, "Quantization," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2325–2383, 1998.

[75] A. Kondoz, "Sampling and quantization," *Digital Speech: Coding for Low Bit Rate Communication Systems, Second Edition*, pp. 23–55.

[76] T. Anttalainen, *Introduction to telecommunications network engineering.* Artech House, 2003.

[77] M. Mahoney, "Data compression programs," 2008.

[78] H. M. Sun, K. H. Wang, and W. C. Ting, "On the security of the secure arithmetic code," *IEEE Transactions on Information Forensics and Security*, vol. 4, pp. 781–789, Dec 2009.

[79] S. Kodituwakku and U. Amarasinghe, "Comparison of lossless data compression algorithms for text data," *Indian journal of computer science and engineering*, vol. 1, no. 4, pp. 416–425, 2010.

[80] D. P. Stern, "Latitude and longitude," *Web page, NASA, Goddard Space Flight Center, Greenbelt, Maryland*, vol. 17, 2004.

[81] U. D. O. COMMERCE.

[82] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, 1987.

[83] A. Gersho and R. M. Gray, *Vector quantization and signal compression*, vol. 159. Springer Science & Business Media, 2012.

[84] P. G. Howard and J. S. Vitter, "Arithmetic coding for data compression," *Proceedings of the IEEE*, vol. 82, no. 6, pp. 857–865, 1994.

[85] E. C. Ifeachor and B. W. Jervis, *Digital signal processing: a practical approach.* Pearson Education, 2002.

[86] "Tonbridge-weather."

[87] J. Whitbeck and V. Conan, "Hymad: Hybrid dtn-manet routing for dense and highly dynamic wireless networks," *Computer Communications*, vol. 33, no. 13, pp. 1483–1492, 2010.

[88] R. J. Mohsin, J. Woods, and M. Q. Shawkat, "Density and mobility impact on manet routing protocols in a maritime environment," in *Science and Information Conference (SAI), 2015*, pp. 1046–1051, IEEE, 2015.

[89] R. Mangrulkar and M. Atique, "Routing protocol for delay tolerant network: A survey and comparison," in *Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on*, pp. 210–215, IEEE, 2010.

[90] K. K. Sevimli and M. Soyturk, "Lifetime determination for delay tolerant communications in sparse vehicular networks," in *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on*, pp. 250–255, IEEE, 2010.

[91] M. Rabab J., W. John, and M. Q. Shawkat, "(amdc) algorithm for wireless sensor networks in the marine environment," *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 6, no. 6, 2015.

[92] "PURSUIT project."

[93] D. Trossen and G. Parisis, "Designing and realizing an information-centric internet," *IEEE Communications Magazine*, vol. 50, no. 7, 2012.

[94] M. Al-Khalidi, N. Thomos, M. J. Reed, M. F. Al-Naday, and D. Trossen, "Seamless handover in ip over icn networks: A coding approach," in *2017 IEEE International Conference on Communications (ICC)*, May 2017.

[95] P. Jokela, A. Zahemszky, C. Esteve Rothenberg, S. Arianfar, and P. Nikander, "Lipsin: line speed publish/subscribe inter-networking," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 4, pp. 195–206, 2009.

[96] "POINT project."

[97] M. Al-Khalidi, N. Thomos, M. J. Reed, M. F. AL-Naday, and D. Trossen, "Network Controlled Mobility Management using IP over ICN Architecture," *http://arxiv.org/abs/1610.09011*, 2016.

[98] F. Zhang, Y. Zhang, and D. Raychaudhuri, "Edge caching and nearest replica routing in information-centric networking," in *Sarnoff Symposium, 2016 IEEE 37th*, pp. 181–186, IEEE, 2016.

[99] C. Stais, G. Xylomenos, and A. Voulimeneas, "A reliable multicast transport protocol for information-centric networks," *Journal of Network and Computer Applications*, vol. 50, pp. 92–100, 2015.

[100] G. Xylomenos, X. Vasilakos, C. Tsilopoulos, V. A. Siris, and G. C. Polyzos, "Caching and mobility support in a publish-subscribe internet architecture," *IEEE Communications Magazine*, vol. 50, no. 7, 2012.

[101] B. O. Data, "The definitive ais handbook."

[102] F. E. P. Development, "Ais: A guide to system development."

[103] R. Essaadali, C. Jebali, K. Grati, and A. Kouki, “Ais data exchange protocol study and embedded software development for maritime navigation,” in *Electrical and Computer Engineering (CCECE), 2015 IEEE 28th Canadian Conference on*, pp. 1594–1599, IEEE, 2015.

[104] J. Deng, “Real-time navigation monitoring system research for lng-fuelled ship in inland water,” *Journal of Maritime Research*, vol. 12, no. 2, pp. 87–94, 2017.

[105] E. Alincourt, C. Ray, P.-M. Ricordel, D. Dare-Emzivat, and A. Boudraa, “Methodology for ais signature identification through magnitude and temporal characterization,” in *OCEANS 2016-Shanghai*, pp. 1–6, IEEE, 2016.

[106] E. S. Raymond, “Aivdm/aivdo protocol decoding,” *GPSD documentation Version*, vol. 1, 2014.