
Downloaded from: https://e-space.mmu.ac.uk/628052/

Version: Accepted Version

Publisher: Institute of Electrical and Electronics Engineers

DOI: https://doi.org/10.1109/JSEN.2018.2791487

Please cite the published version

https://e-space.mmu.ac.uk
Building Novel VHF-based Wireless Sensor Networks for the Internet of Marine Things

Rabab Al-Zaidi, Member, IEEE, John C. Woods, Mohammed Al-Khalidi, Member, IEEE and Huosheng Hu, Senior Member, IEEE

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 paper proposes a novel Internet of Marine Things (IoMaT) 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 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. Several observations and obstacles unique to the marine environment have been discussed and feed into the solutions presented. The impact of marine sparsity on the network is examined and a novel hybrid Mobile Ad-hoc/Delay Tolerant 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 data synchronization and transmission approach has also been proposed at the 5G network core using Information Centric Networks (ICN).

Index Terms—Internet of Marine Things, Very High Frequency, Mobile Ad-hoc Networks, Delay Tolerant Networks, Information Centric Networks.

I. INTRODUCTION

A. Background

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 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 [1]. Therefore, other solutions are required in order to provide more accurate readings and future estimates of the earth’s climate. One favorable 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, etc... 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 [2]. 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 [3]. 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 [4].

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. Extending MANET communication to sea applications requires consideration of the characteristics of radio signal propagation at sea, and the need for repeaters and multi hop relays (MMR) to achieve Ad-hoc networking objectives. The system design relies on the accurate modelling of the radio propagation. However, 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/disruption tolerant networks (DTNs). MANET routing protocols do not work properly in DTNs since under these protocols, when packets arrive and no contemporaneous end-to-end paths for their destinations can be found, these packets are simply dropped [5].

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 156MHz to 163MHz [6]. In preparation for the revolutionary 5G networks, 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 [7].

B. Wireless Sensor Networks

There has been a large number of research activities at using wireless sensor networks (WSNs) to monitor marine environments. However, the main drawback of the these research efforts is the low coverage area of the proposed wireless transmission. Most of the developed WSN technologies cannot effectively cover the vast areas of oceans and seas in scales of thousands of kilometres. Summarizing the available proposed technologies for marine wireless sensor networks, various IEEE 802 standards have been proposed in the literature to build wireless sensor networks at sea. The highest communications range proposed within the IEEE 802 family is the WiMAX standard which barely exceeds 10 km. Other proposals involve GSM or GPRS that provide a higher communication coverage of about 35 km. But they are impractical for marine environments that are 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. Below, we elaborate some of the most mature research efforts in the field of marine sensing and communication networks and applications.

- An environment monitoring framework was proposed in [8], 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 [9]. 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 in 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 [10]. The coverage extension was achieved through a multi-hop WiMAX wireless mesh network where ships, maritime beacons and buoys were the nodes. Ship to ship communication transmission range was specified as 20 km at maximum.

- A project was presented in [11] to test methods and technology in a network of autonomous vehicles to increase efficiency of marine data acquisition operations. To provide communication, the Marine Broadband Radio system (MBR) was implemented. The MBR operates in the 5 GHz frequency band and uses highly dynamic beamforming and adaptive power control to secure stable communication in maritime operations with signal obstructions, fading and ranges in excess of 50 km. However, this project (as with the previous efforts discussed) does not have an IoT perspective and does not provide any IoT solution.

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 [12]. 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 [13]. 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 [14]. 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.

C. The proposed Approach

Based on our previous work in [15], we present a novel marine data acquisition and cartography system in this paper, which could be used in ships containing multiple sensors such as sea depth, temperature, wind speed and direction, etc. It is a combination of multiple sensors and Ship Ad-hoc Networks (SANETs) that communicate over the marine VHF radio channels. The data collected by multiple sensors is sent to a central server to produce an information map for public view. The whole system could be seen as the realization of 5G IoMaT connectivity.

A model of the VHF radio is used to setup a Physical layer for ship Ad-hoc networks, which complies with the
International Telecommunication Union (ITU) standards for data communication in the marine environment. The impact of marine sparsity on the network is examined and the performance of MANET and DTN routing protocols is evaluated according to the network connectivity within the evaluated area. A novel hybrid Mobile Ad-hoc/Delay Tolerant routing protocol (MADNET) is also proposed to switch automatically between MANET and DTN routing according to the network connectivity.

The data transmission cost of the proposed system is analysed with respect to the problem of low rate data transmission offered by VHF radio and the bottleneck of the links closest to the sink. A quantization and compression method specific to marine sensory data has been proposed in our previous work [16] and partially alleviates this problem. 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 performance of the proposed data acquisition network is evaluated through simulation with respect to the minimum amount of sensory data successfully delivered to the sink for each ship when different sizes of SANETs are simulated. Finally, we design a data synchronisation and transmission approach at the 5G network core using Information Centric Networking (ICN). This approach is aimed at providing efficiency and removing any duplication of marine sensory transmission readings from the base station/sink nodes towards the central cloud. At this point, the central cloud is ready to serve any user demand.

Allocating data computation to local servers is well studied in the context of IoT since low latency is a key to success in 5G networks. However, how to retrieve local data from IoT devices is not well studied. An efficient synchronization and transmission method is needed between the edge and central clouds in order to preserve network resources and guarantee efficient delivery. ICN provides means for information dissemination that identifies information at the Network layer. This allows for efficient data synchronization between edge and central clouds and elevates any possibility of duplicated data transmission.

The rest of the paper is structured as follows. Section II outlines the proposed system architecture, which 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.

II. OVERVIEW OF THE PROPOSED IoMaT SYSTEM ARCHITECTURE

This section describes the individual layers of the proposed IoMaT system architecture as shown in Fig. 1. The ship Ad-hoc network is used to collect different marine sensory data from ships and vessels and send this data back to onshore sinks collocated with 5G base stations. The 5G base station connects to a 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 paper 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 underwater 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.

Fig. 1: The proposed IoMaT Cartography Network

Fig. 2 shows the four service layers in the proposed cartography network. As can be seen, there are four service layers in the system, namely the marine sensor layer, the marine Network layer, the cloud data management layer and the data analysis and cartography application layer.

A. 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. These sensors can sample numerical values reflecting the states of the monitored objects. The lower and upper limits of individual sensor readings are set according to 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) proposed in [16] has given effective data compression rates in comparison to the main compression methods.

B. The marine Network layer:

This layer consists of 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 Section IV.

C. 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 synchronize sensory data delivery between edge clouds and ensure duplicate-less arrival of data at the central cloud.

D. 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 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.

This layer is concerned with analysing and managing the bulk of collected sensory data at a central repository. It is also responsible for extracting the most important information in the collected data regarding the behaviour of the monitored environmental factors. This will facilitate not only up to date cartography systems, but more advanced applications such as future predictions of weather forecasts and climate changes through the use of techniques such as big data, data mining, neural networks and machine learning.

III. MARINE SENSORS LAYER

For the proposed marine application, the data gathered from sensors is predictable, therefore it is essential to quantize and compress the data locally on ships for optimum transmission. This reduces the amount of bits needed to represent each sensory reading in the binary representation.

A. Sensory Reading Quantization

We use linear quantization where the range of readings for each sensor and the required steps within that range are used 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 are represented so as to reduce even more the bit representation required. Ships latitude is represented in degrees and tenths of a degree, measured in terms of degrees north or south of the equator. Ships Longitude 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 [17].

Table I shows the most important sensors applied in the proposed sensor data acquisition network.
and upper limits of the individual sensor readings likely to be found in the marine environment have been set. Also, the level of accuracy required to represent each reading is chosen so as 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 [15].

B. Average Marine Data Compression algorithm (AMDC)

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, the first sensor reading in every five readings block is considered as a reference for the block. The AMDC algorithm is shown in Fig. 3 and 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 eq. 1 below:

\[
AR = \sum_{j=1}^{4} R_{i+j}/4
\]  

(1)

Then the deviation from the 1st reference reading \( D_v \) is calculated as in eq. 2 below:

\[
D_v = R_i - AR
\]  

(2)

- 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 \([0,1]\), 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. For a more comprehensive explanation of arithmetic code compression, the reader is referred to [18].

According to the AMDC phases described above, the AMDC algorithm has been formulated as follows:

The AMDC algorithm is particularly suitable for the limited transmission resources of the proposed ship Ad-hoc network. According to the results obtained in our previous work [16], the number of bits required to represent the readings of the deployed sensors is decreased by approximately 90% from 695 bits to approximately 70 bits every five minutes when the AMDC compression algorithm is used.

IV. MARINE NETWORK LAYER

The marine Network layer shown in Fig. 2 consists of sensor nodes (ships) and sink nodes that communicate in an Ad-hoc (point-to-point) manner.

---

**Algorithm 1** The proposed AMDC algorithm

<table>
<thead>
<tr>
<th>Input</th>
<th>Sensory Data Readings ( R_t ) → ( R_{t+5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>Compressed Sensory Data</td>
</tr>
<tr>
<td>Variables</td>
<td>( R_t, AR, DV, i, H, L, R_i, C(x), P(x), t )</td>
</tr>
</tbody>
</table>

---

Step 1: Compute Average Reading

\( AR \leftarrow \sum (R_{i+2}:R_{i+5})/4 \)

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

\( R \leftarrow H - L \)

\( L \leftarrow R \times C(x_i) \)

\( H \leftarrow L + R \times P(x_i) \)

end for

\( t \leftarrow (H + L)/2 \)

---

A. Radio Specifications

The international telecommunication union (ITU) has defined marine VHF radio to operate on internationally agreed frequencies in the band from 156MHz to 163MHz. The VHF 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 in eq. 3 below:

\[
d = \sqrt{12.746 \times hb}
\]

(3)

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 [19].
In the marine environment, VHF radio communication is achieved via the Automatic Identification System (AIS). AIS is a VHF radio broadcasting system that transfers packets of data over VHF Data Links (VDL’s) and enables vessels and shore-based stations to send and receive identification information that can be displayed on an electronic chart, or compatible radar.

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” [20]. They also provide a guideline on the use of digital technologies by VHF systems of different bandwidths. In this paper, our ship Ad-hoc network proposal over VHF is completely based on ITU recommendations for VHF data communication.

B. 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 Ship Ad-hoc Networks as proposed by ITU Recommendation ITU-R M.1371-4 [21], 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) [22].

As a consequence we choose to use the most relevant: the Ad-hoc Self-Organizing TDMA (ASOTDMA) protocol for this Ad-hoc network. 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 [22].

C. Routing Protocols

Mobile 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 proactive and reactive MANET approaches. With low density of nodes, 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 zones [23]. This is typically the domain of DTN [24]. One can characterize the relevant routing paradigms in mobile wireless networks along the two main parameters of node density and mobility. MANET routing has a positive relationship with node density, and a negative relationship with node mobility, in contrast with DTN routing which envisages node mobility as a positive relationship and node density as a negative one.

1) Mobile Ad Hoc Networking (MANET): The primary goal of any MANET routing protocol is to cope with the dynamically changing topology; i.e., to establish an efficient route between any two nodes with minimum routing overhead and bandwidth consumption. Routing protocols in MANETs are usually divided into two main groups depending on the method by which the routes are preserved inside the network [25]. Proactive (Table-Driven) protocols preserve information on routes in the network in advance. Every node in the network preserves information on routes to every other node. The protocols that fall under this category don’t perform efficiently when applied in large networks, because eventually they end up with very large numbers of routing table entries. An example that has been used in our simulation and evaluation of SANETs in [15] is the Destination-Sequenced Distance-Vector (DSDV) routing protocol. In reactive (on demand protocols), nodes start discovering and establishing routes to the destination of packets only when needed through a route discovery process [26] [27]. Ad hoc On-demand Multihop Distance Vector (AOMDV) routing protocol is an example that has been used in our simulation.

● **AOMDV Routing Protocol**: AOMDV shares several characteristics with its predecessor, Ad hoc On-demand Distance Vector routing protocol (AOMDV). It is based on the distance vector concept and uses a hop-by-hop routing approach. Moreover, it also finds routes on demand using a route discovery procedure [28]. In AOMDV, whenever a traffic source needs a route to a destination, it initiates a route discovery by flooding a route request (RREQ) for the destination in the network and then waits for route replies (RREP). RREQ propagation from the source towards the destination establishes multiple reverse paths both at intermediate nodes as well as the destination. Multiple RREPs traverse these reverse paths back to form multiple forward paths to the destination at the source and intermediate nodes. Note that AOMDV also provides intermediate nodes with alternate paths as they are found to be useful in reducing route discovery frequency [17].

2) Delay Tolerant Networking (DTN): 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 [29]. 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 VN, or persistent, as in the Internet. 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.

Two of the most popular DTN routing protocols are the Epidemic and Spray and Wait protocol [30] [31].

1) 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.

2) 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, 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 [32].

D. Hybrid AOMDV/DTN (MADNET) Routing to overcome Disconnectivity

As an effective solution to the well-known sparsity problem that works against MANETs in real marine scenarios, 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 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 [33]. 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 routing protocol is proposed that switches automatically between AOMDV (the best performing routing protocol in dense locations as found in our previous simulations [15] [34]) and Binary (SW) DTN (the best performing routing protocol in sparse situations [35]). While our proposal is not the first to combine DTN functionality with other routing protocols [23], it is the first novel application of Binary (SW)/AOMDV hybrid routing over VHF Ad-hoc networks in a marine environment.

Fig. 4 shows the proposed protocol stack for SANETs and gives an overview of the solution paths for each of the corresponding layers of the novel protocol stack discussed in this paper. 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. The algorithm be-
low shows the basic switching function between AOMDV and Binary (SW) routing where Binary (SW) routing is triggered whenever there is no end to end route between the source and destination.

Algorithm 2 AOMDV, Binary (SW) switching algorithm

1: while DST has not received the message do
2:    if RREP exists then
3:        Send by AOMDV
4:    else
5:        Send by Binary(SW)
6:  end if
7: end while

1) 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 NS2 simulator (Network Simulator Version 2). VHF transmission ranges were calculated using the Free Space Propagation model as in eq. 4 below.

\[
pr = \frac{pt \times Gt \times Gr \times (\lambda)^2}{(4\pi)^2 \times d^2 \times L}
\]  

(4)

where \(pt\) is the transmitted power, \(pr\) 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 [36]. The Free-Space Path Loss (FSPL) was calculated from eq. 4 as:

\[
FSPL = \frac{\lambda^2}{(4\pi d)^2}
\]  

(5)

We use Self Organized Time division multiple access (SOTDMA) as the channel access method as proposed by the ITU for ship data communication over VHF channels. 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 at a rate of 14 bit/min corresponding to the size of the collected sensory data after compression. In the DTN case, the bundles are fragmented to 1500 bit packets before sending them to the MAC layer. We have chosen to use 25 MB bundle buffer space, which does not become a bottleneck in the simulations. Bundle lifetime is set to 10800 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 [37]. Table II shows a summary of the simulation parameters used in our simulation.

2) 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).

Fig. 5a, Fig. 5b and Fig. 5c 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.

Fig. 6a, Fig. 6b and Fig. 6c show the average hop stretch per routing protocol in dense, moderate and sparse scenarios

<table>
<thead>
<tr>
<th>TABLE II: MADNET Simulation Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td>Simulation Time (s)</td>
</tr>
<tr>
<td>Simulation Area</td>
</tr>
<tr>
<td>Average Number of nodes</td>
</tr>
<tr>
<td>Speed (km/h)</td>
</tr>
<tr>
<td>Routing Protocol</td>
</tr>
<tr>
<td>Transmission Range</td>
</tr>
<tr>
<td>TTL (s)</td>
</tr>
<tr>
<td>Buffer size (MB)</td>
</tr>
<tr>
<td>Message size (bits)</td>
</tr>
<tr>
<td>Movement Model</td>
</tr>
</tbody>
</table>

Fig. 6a, Fig. 6b and Fig. 6c 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.

E. 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 IV-D1 above to identify the most congested link in the network (closest to the shore base station) 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 139 bits/min, 104 bits/min and 14 bits/min corresponding to unmodified, quantized and compressed sensory readings respectively. These readings are collected at 11 marine sensors on every ship each minute according to [16]. All the rest of simulation parameters remain as in section IV-D1. Table III shows a summary of the simulation parameters used in this simulation.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Time (s)</td>
<td>43200</td>
</tr>
<tr>
<td>Simulation Location</td>
<td>English Channel</td>
</tr>
<tr>
<td>Simulation Area</td>
<td>200 x 200 (km)</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>82, 100, 123</td>
</tr>
<tr>
<td>Speed (km/h)</td>
<td>0-20</td>
</tr>
<tr>
<td>Routing Protocol</td>
<td>MADNET</td>
</tr>
<tr>
<td>Transmission Range</td>
<td>30(km) - 40(km)</td>
</tr>
<tr>
<td>TTL (s)</td>
<td>10800</td>
</tr>
<tr>
<td>Buffer size (MB)</td>
<td>25</td>
</tr>
<tr>
<td>Message size (bits)</td>
<td>250</td>
</tr>
<tr>
<td>Movement Model</td>
<td>Real mobility from live AIS website</td>
</tr>
</tbody>
</table>

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.

Fig. 7 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 also 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 138 minutes of collected data for each ship are transmitted to the sink with a network of 82 ships each minute. This number decreases to approximately 57 minutes of readings per minute when a larger network of 123 ships is simulated. The figures for quantized data transmission follow with a minimum of just above 40 minutes of readings when 82 ships exist in the network and decreases to about 16.7 minutes for 123 ships. Also the figures for unmodified data are below quantized data by approximately 20%. It is 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 138 minutes of collected data per ship per minute have been observed with a network of 82 ships.

\[ F. \text{ System Prototype and IoMaT Software Components Integration} \]

Fig. 8 shows the proposed IoMaT software module hosted on a laptop/PC that is connected to the AIS system through a serial interface. Most AIS models have a 9-pin National Marine Electronics Association (NMEA) port which can be connected to a computer serial port using a standard RS 232 serial cable [38] [39]. 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. The IoMaT module consists of the IoMaT message parser/encapsulator, the MADNET 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. 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 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. The MADNET 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. AIS messages are ASCII data packets that follow the NMEA 0183/NMEA 2000 data formats as defined by the AIVDM/AIVDO protocol [40]. IoMaT packets are proposed to have an identical format that is easily distinguishable by the IoMaT message parser/encapsulator through the introducer header. AIS packets have the introducer "/AIVDM" or "/AIVDO"; AIVDM identifies packets received from other ships and AIVDO identifies local packets. To incorporate IoMaT messages, we propose the use of two more introducers that uniquely identify IoMaT packets in an AIS system, namely "/AIVDN" and "/AIVDS". In the same sense, AIVDN identifies IoMaT packets received from other ships and AIVDS identifies local IoMaT packets.

Although only one radio channel is necessary for communication, each AIS station transmits and receives over two radio channels (maritime channels 87B (161.975 MHz) and 88B (162.025 MHz)) 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.

\[ V. \text{ Cloud Data Management Layer} \]

At the edge of the 5G core network, the marine sensory data collected on various ships is aggregated at base station sink nodes at shore. Several base stations that cover a designated water surface are connected to an Edge cloud. The Edge clouds process and synchronize the sensory data
with a central cloud in the network core. In this paper, we propose using an Information Centric Networking (ICN) approach, that provides means for information dissemination that identifies information at the Network layer. Such identification can be realized, for instance, through some form of naming scheme. A wide range of ICN implementations have been proposed in various research projects, of which Publish Subscribe Internet Technology (PURSUIT) [41] architecture is used as an exemplary reference model. There are two main functional entities in PURSUIT to compose the machinery of the architecture: (i) the Rendezvous (RV) that is responsible for matching publications and subscriptions of information items; and (ii) The Topology Manager (TM) that is responsible for constructing a delivery tree for the information object. This delivery tree is encoded in a forwarding identifier (FID) which is sent to the publisher to forward the packets containing the information object to the subscriber. 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 [42].

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 synchronization and greatly facilitates cloud services. 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 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.

Fig. 9 shows the ICN namespace proposed for collecting and synchronising the transmission of marine sensory data at the 5G network. 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. 10 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.
and /root/area3. Or 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/timeStamp1. Despite these flexibilities, it is always assumed that the central cloud is subscribed to the whole domain scope /root which facilitates the aggregation and synchronization of the data received from edge clouds. Other use cases may necessitate other subscription scenarios, but are out of scope of this paper.

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 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.

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 information. The edge cloud includes a Network Abstraction Function (NAF) [43] [44] that occupies a key role in Ship Ad-hoc Networks/ICN abstraction and interfacing as shown in Fig. 11. The edge clouds NAF constructs the scope /root/area/location/timeStamp/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.

VI. CONCLUSIONS

Due to increased shipping demands and the high cost of other available technologies, it is necessary to develop advanced and safe data networks used in the marine environment. This paper proposes the deployment of the existing VHF communication infrastructure available on all ships to build a new generation of wireless sensor networks for small and middle-size ships and vessels. Several observations and obstacles unique to the marine environment have been discussed.
and feed into the solutions presented. Novel data quantization and compression techniques specific to the marine sensor data collected have been investigated in order to reduce the burden on the channel links and achieve better transmission efficiency. The influence of marine traffic models on the behaviour of SANETs has also been discussed, which highlights the problem of sparsity and the necessary inclusion of selective opportunistic networking and MANET protocols.

The drawbacks of low rate data transmission offered by VHF radio limited to 28.8 kbps 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. A data synchronisation and transmission approach has also 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. Therefore, it is clear that low-cost SANETs could be used as part of a 5G infrastructure for marine environment monitoring.

Our future work will be focused on the development of a working prototype using a number of AIS transponders marine VHF channels and PC based source and sink nodes to replace the NS2 Simulation.

ACKNOWLEDGMENTS

This work was carried out through the support of EU Horizon 2020 project POINT under grant agreement No 643990, and the School of Computer Science and Electronic Engineering, University of Essex, UK. Rabab Al-Zaidi is financially supported by HCED Iraq Scholarships for her PhD study.

REFERENCES

Rabab Al-Zaidi received her B.Sc. degree in Computer and Software Engineering from Al-Mustansiriya University, M.Sc. degree in Software Engineering from University of Technology, in 2006 and 2009 respectively, she is currently working towards a Ph.D. degree in Electronic Systems Engineering at CSEE department, Essex University, United Kingdom. Her main research interests are in Mobil Ad-Hoc Networks, Internet of Things, Sensor Networks and data formatting in the marine environment.

John C. Woods was born in a small fishing village near Colchester, U.K., in 1964. He received the B.Eng. (honors) degree (first class) in 1996 and the Ph.D. degree in 1999 from the University of Essex, Colchester, UK. He has been a Lecturer in the Department of Computer Science and Electronic Systems Engineering, University of Essex, since 1999. Although his field of expertise is image processing, he has a wide range of interests including telecommunications, autonomous vehicles and robotics.

Mohammed Al-Khalidi Received his B.Sc. and M.Sc. degrees in Computer and Software Engineering, in 2004 and 2013 respectively. He is currently a Research Officer and working towards a PhD degree in Electronic Systems Engineering at the School of Computer Science and Electronic Engineering, University of Essex, United Kingdom. Main research interests are in ICN, IoT, Network Coding, Network Security, Mobile Computing and Mobile Ad-Hoc Networks.

Huosheng Hu (M94,SM01) received the MSc degree in industrial automation from Central South University, China, and the PhD degree in robotics from the University of Oxford, United Kingdom. He is currently a Professor with the School of Computer Science and Electronic Engineering, University of Essex, United Kingdom, where he is leading the Robotics Research Group. His current research interests include robotics, human-robot interaction, embedded systems, mechatronics, and pervasive computing. He has authored over 500 conference and journal papers in these areas. Prof. Hu is Editor-in-Chief for International Journal of Automation & Computing and the MDPI Robotics Journal, a fellow of the Institution of Engineering and Technology, and a Fellow of the Institute of Measurement Control.