



HAL
open science

Data muling: an innovative solution to extend ranges and data rate

Nathan Fourniol, Nathan Fourniol, Andreas Arnold, Pierre-Jean Bouvet, Luc
Jaulin

► **To cite this version:**

Nathan Fourniol, Nathan Fourniol, Andreas Arnold, Pierre-Jean Bouvet, Luc Jaulin. Data muling: an innovative solution to extend ranges and data rate. mastMED 2022 - Maritime/Air Systems & Technologies, Nov 2022, Athens, Greece. hal-04707448

HAL Id: hal-04707448

<https://hal.science/hal-04707448v1>

Submitted on 24 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data muling: an innovative solution to extend ranges and data rate

Nathan Fourniol¹, Andreas Arnold², Pierre-Jean Bouvet³, Luc Jaulin¹

¹ ENSTA Bretagne, 2, Rue François Verny, 29200 Brest, France

² Thales DMS France SAS, 10 avenue de la 1^{ère} DFL, 29200 Brest, France

³ ISEN Yncréa Ouest, 20 Rue du Cuirassé Bretagne, 29200 Brest, France

Abstract—The emergence of new maritime activities, such as the search for mines with AUVs or wind turbine monitoring, requires more and more long-range, high-speed communications. However, in the underwater environment, no conventional wireless communication means allow this. We propose a data muling solution, *i.e.* the use of mobile agents spread between the data source and the data destination. These agents communicate using short range wireless communication and carry the data in internal memory from one point to another. By doing this, we show we can significantly increase the throughput of long-range communication. In this article, we will mainly focus in a 2D plan and the study of theoretical gain of this method and expose a example where data muling is interesting. In addition, we introduce a metric to choose the characteristic quantities for designing and implementing such a solution.

I. INTRODUCTION

The exploration of the seabed has become a strategic issue due to military, scientific or commercial use. With technological progress, these missions are increasingly carried out using Autonomous Underwater Vehicles (AUVs). This represents advantages in terms of area coverage rate, discretion, but also in terms of human resources, thus reducing costs and human risks.

However, the use of AUVs for such missions means that the data is not accessible immediately. It is common to carry out long campaigns with long periods of time below the surface, meaning a large amount of data is generated, but not immediately sent to the surface. Radio waves propagate poorly underwater and acoustic communication generally has a fairly low data rate. Thus quickly communicating a large amount of data, over a long distance is a real problem.

On the one hand, with a range of 100 m up to several hundred of km, acoustic waves are the most energy efficient communication mode, however the data rate is of the order of 100 kb/s.km [4] that is not suitable to transfer of a large amount of data. On the other hand, Underwater Wireless Optical Communications (UWOCs) can reach 1000 Mb/s on a typical range of the order of ten meters [5][3] depending on the turbidity of the water. Directional UWOC systems using

lasers can reach farther but they require to align the transmitter and the receiver, which adds an additional layer of complexity. A third approach is the use of Radio Frequency (RF) communications. Such systems have a performance figure of the order of 10 Mb/s over a range on dozens of centimeters [5] [4] due to strong attenuation of RF waves in salt water. As a result, wireless underwater transfer of large amount of data over a range of few kilometres is not feasible on a reasonable time with the existing communication technologies.

To tackle the problem of retrieving large amount of data from underwater sensor nodes or AUVs, interest in data muling solutions is rising. The key idea is the use of a mobile vector such as an AUV, hereafter called “agent” to physically carry data in its internal memory. An agent is characterised by its ability to move, relocate other agents or nodes, and to communicate with a high bandwidth, short range communication system [7] involving docking or not.

Several muling scenarios are studied in literature. A first one considers the case of Underwater Sensors Networks (USNs) where a series of sensing nodes (such as environment monitoring stations) gather data which is stored locally, and an agent is periodically sent to visit the nodes and collect the data to bring it to the surface. This is a specific instance of the Travelling Salesman Problem (TSP), solved using classical optimisation algorithms (simulated annealing, genetic algorithms, *etc.*) with constraints such as the node coverage or the energy used [2]. Energy constraints for this problem are handled either by putting a reloading station on the sea floor at the middle of the sensor network, or by using a larger ship at the surface having some means of energy production (solar panels, generators) and a surface docking station to reload the AUV [2].

A second use case, is the long range communication in the vertical plane (bottom-surface) using one or more mobile agents. Two experimental results presented in [6] and [9], tackle the problem of developing a solution where a few AUVs go down to the node, download its data and come back to the surface. Teixeira *et al.* [9] developed a generic protocol called Underwater Data Muling Protocol (UDMP) protocol to retrieve data from a node or an AUV over a long range. In this protocol, they considered that the node always has a secondary acoustic communication link with the destination point in order to request data; this is feasible in the vertical plane when the channel depth is not too important (a few hundred meters).

This work was in part sponsored by the regional council of Brittany and by Thales Group. Corresponding author: Andreas Arnold
andreas.arnold@fr.thalesgroup.com.

They provided a comparison of the impact of the size of the batch of data transported, the number of mules and the distance travelled by the mule on the overall throughput of the whole system. Then they showed experimentally that their solution solves their problem. A limitation of their approach is that mules cannot communicate with each other but only with the sensor node and the surface ship.

In this paper, we focus on a study where moving agents are used as relays. Each agent in the relay system is assigned to a specific area, which can be fairly large. At the edge of its area it can communicate with the agent of the neighbouring area. The agent can travel within its assigned area to physically transport the data and act as the moving relay. To put things into perspective, this is nothing more but a transposition of the short-lived Pony Express mail service (April 3, 1860 – October 26, 1861) to the marine case! This has not been studied in the literature to the best of our knowledge. Our study analyses the pros and cons of such a system and compare it to an other one detailed in [9]. We will show that for different parameters this solution is suitable and adapted to extend the range and data rate of a communication link.

The rest of the paper is organised as follows: in section II we present the data muling system model. In section III, simulation results are analysed and compared with the state of the art. Section IV concludes the paper.

II. SYSTEM MODEL

We use a similar model to what is exposed in [1]. Let the set of agents be designated as $\mathcal{N} = \{n_0, n_1, \dots, n_{N+1}\}$, where agent n_0 and n_{N+1} are the data source (\mathcal{S}) and destination (\mathcal{D}) respectively. The others will be referred to as “mules”. The source is the agent producing the data while the destination is the agent waiting for the data. Furthermore, all agents have the ability to communicate over a maximum range d with a data rate R . Furthermore, mules have a limited memory m . Figure 1 is a graphical representation of the scenario.

We assume that the agents are distributed along a polygonal line, thus the position of agent n_i is solely described by a curvilinear abscissa p_i . Without loss of generality we may consider this polygonal line to be a simple line. We further assume that the source and destination are static, with $p_0 = 0$ and $p_{N+1} = D$. We shall denote by vector $\mathbf{p} = \{p_0, p_1, \dots, p_{N+1}\}$ the positions of all agents at a given time; the dependence on time is implicit and left out for the sake of brevity. We assume that the other agents have the ability to move at speed of modulus v_{mule} and that they are evenly distributed between p_0 and p_{N+1} in non-overlapping areas of length L . This length L shall be characterised in section II-A. For this paper, navigation range and positioning accuracy are not considered to be an issue, meaning navigation range is infinite and positioning is accurate. This is a way to simplify the study. Also, we consider that agents never fail and the environment disturbs neither communication nor navigation.

A. Distribution of the mules in the area

In order to create a communication link between agents n_0 and n_{N+1} , mules are distributed in non overlapping areas

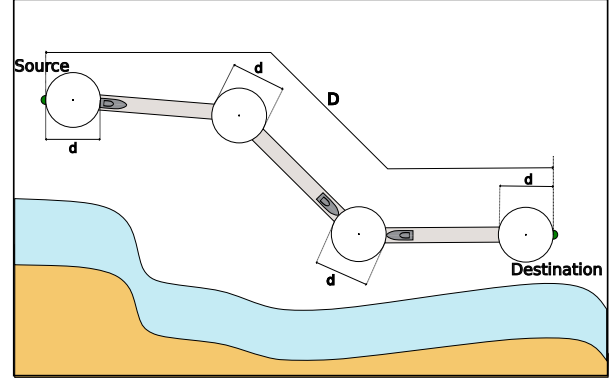


Fig. 1. Linear data muling scenario to transmit data from agent n_0 to n_{N+1} over the distance D

of length L . The space between the areas is equal to the communication range d . By dividing distance D in N equal chunks separated by the communication range d , we end up with:

$$L = \frac{D - (N + 1)d}{N} \quad (1)$$

So, for $i \in [1, N]$, the i -th mule lies in interval $[id + (i - 1)L, id + iL]$.

B. Communication time

As we modelled our agents with the ability to communicate, the communication time will be taken as:

$$T_{\text{com}} = \frac{m}{R} + \frac{d}{c} \quad (2)$$

where c is the celerity of the communication channel: c is about 1500 m.s^{-1} for acoustic communication and about $2.25 \times 10^8 \text{ m.s}^{-1}$ for UWOC, which we focus on. In the following, we will neglect the channel latency term d/c for convenience and because, for most of the cases, it is negligible for short range communication, being small in front of the time to transmit all data.

C. Equivalent throughput

In order to characterise the performance of the model, we use the metric introduced in [9]: the equivalent throughput $R_{b,eq}$ defined as the total amount M of data to transfer over the time, T_{total} , it took to be delivered:

$$R_{b,eq} = \frac{M}{T_{\text{total}}} \quad (3)$$

The expression of T_{total} will be developed in the next section.

D. Data muling scenario

For our study, we consider the following generic operational scenario where mules are distributed over a straight line joining the source to the destination:

1) *Problem*: There is an amount M of data in agent n_0 to send with our data muling protocol. This amount of data is divided in $P \in \mathbb{N}^*$ packets, in order to be adapted to the memory capacity m of the mules.

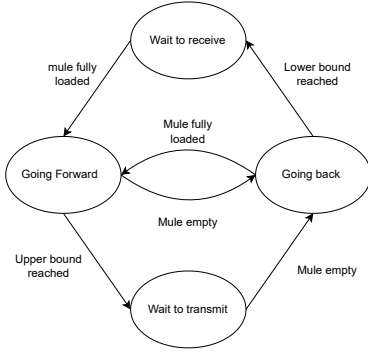


Fig. 2. State machine describing the behaviour of a mule

2) *Initialisation*: The agent n_0 is set at the abscissa 0 and agent n_{N+1} at D . Then we compute the position and the bounds for the roaming area of each mule, as defined in II-A. After that, mules are initialised at their lower bound:

$$\forall i \in 1..N, p_i(t=0) = id + (i-1)L$$

3) *Behaviour of a mule*: The behaviour of each mule follows the state machine shown at figure 2. After initialisation, each mule is in the state “Wait to receive”, on the lower bound of its area. Once a mule is fully loaded with data, the mule switches to state “Going forward” and starts moving towards the other bound of its area. If the mule sends all its data loaded to the next mule, its state changes to “Going back”, to go back to the lower bounds, to wait for an other batch of data. Otherwise, if the mule, being in the “Going forward” state, reaches the upper bound, the state switches to “Waiting to send”, until all data has been sent, before switching to “Going back”. We are sure we are not locked in state “Wait to transmit” because of our hypotheses of perfect navigation and communication and by the definition of the areas.

With the scenario presented above we can express a latency as the delay between the moment when the first bit of a packet is emitted by the source, and the moment when the last bit of this packet has been received at the destination:

$$T_{latency} = (N+1) \times (T_{com} + T_{rdv}) + N \times T_{travel} \quad (4)$$

...with T_{com} referring to the communication time, T_{travel} the travel time of a mule over its zone L , and T_{rdv} an arbitrary time modelling the rendezvous delay, *i.e.* the delay for two mules to meet and be ready to transfer data.

In most of the cases we have more than one packet to transmit; if P is the total number of packets, the total transmission time becomes:

$$T_{total} = T_{latency} + (P-1)(2 \times (T_{com} + T_{rdv}) + 2 \times T_{travel}) \quad (5)$$

Here T_{total} represents well the total time of data transmission but fails to express the total number of trips and short range data transfers. Those two quantities are equal to:

$$N_{com} = P \times (N+1) \quad (6)$$

$$N_{travel} = 2N \quad (7)$$

III. RESULTS

According to the model above, we have the possibility to calculate the time to send a quantity of data divided in packets for various configuration of mule speed or communication interface.

A. Latency

We recall that our latency above is the delay from the first bit of a packet sent by the source n_0 to the last bit of this packet received by the destination n_{N+1} . By developing and factorising equation 4, $T_{latency}$ evolves with the sign of $k = \frac{m}{R} + \frac{d}{c} - \frac{d}{v_{mule}} + T_{rdv}$. This criterion is composed by the time to transmit a packet, the data time of flight, the time taken by the mule if it has to travel over the communication range d and the time to gather before starting to exchange data. So depending of the sign of k it will be better to lower the number ($k > 0$) of mules or to maximize their number ($k < 0$). In the second case, the optimum is reached when the number of mules is large enough so that they do not have to move, *i.e.* $N = \frac{D}{d} - 1$, meaning we have evenly spaced static relays.

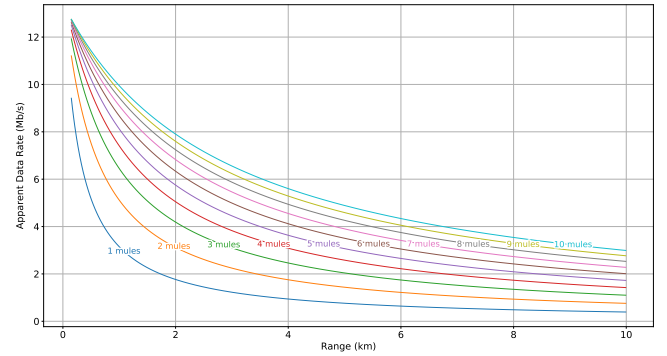


Fig. 3. Variation of apparent data rate $R_{b,eq}$ of the data muling approach, with the distance D between the source and the destination

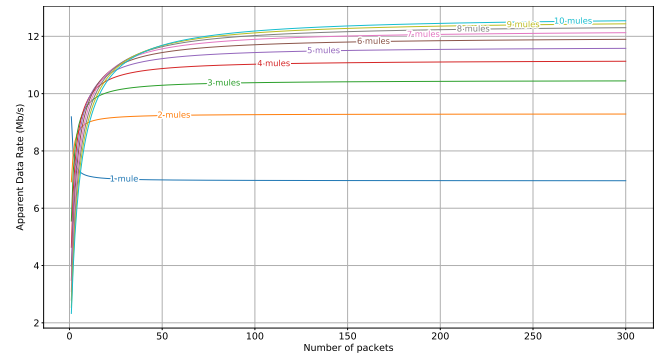


Fig. 4. Variation of apparent data rate $R_{b,eq}$ of the data muling approach as a function of the number P of packets, each packet being $m = 1000$ MB large.

B. Performance comparison in the underwater environment

To compare with other underwater solutions, we will consider the following scenario (see table I). Distance D between agent n_0 and agent n_{N+1} is 300 m and mules will be equipped

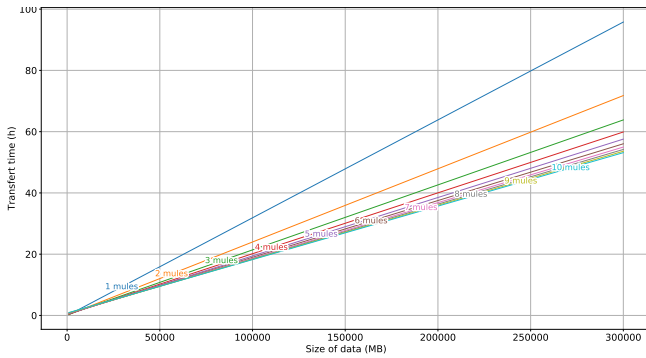


Fig. 5. Variation of the transfer time T_{total} of the data muling approach with the size of data divided in packets of 1000 MB

with an UWOC system. Those parameters are taken from [4] and [8], and are compatible with the value of what can be operated in the underwater environment.

Parameters	Value
D	300 m
$R_{optical}$	27.1 Mb/s
$R_{acoustic}$	300 kb/s
d	10 m
T_{rdv}	0 s

TABLE I

PARAMETERS FOR THE UNDERWATER COMMUNICATION SCENARIO

1) *Direct link*: Since we want to achieve a long range communication with mules, we can compare it with the equivalent setup that could be built at the same range without using mules. Currently, the only system that can achieve the 300 meters range is a pair of acoustic modems, which are known to not have a high data rate. Indeed, according to [4] an indicative figure for the throughput is rate \times range = 100 kb/s \times km. So $R_{acoustic} = R_{b,eq} \approx 300$ kb/s in our case. The data muling scenario with $N = 2$ mules, using UWOC at $R_{optical} = 27.1$ Mb/s yields an equivalent result of $R_{b,eq} = 9.26$ Mb/s. Thus, data muling is far more efficient in terms of data rate and latency (see table II).

Solution	Latency	100 GB	Data Rate
Acoustic	7.4 h	740 h	300 kb/s
2 mules	0.31 h	23.98 h	9.26 Mb/s

TABLE II

COMPARED RESULTS IN THE UNDERWATER COMMUNICATION SCENARIO

2) *With the data muling from [9]*: In literature we found a similar muling system theorised and experimented by Teixeira *et al.* [9]. Our model is pretty close to theirs but differs in the approach. Their solution considers a circular approach where each mule travels distance D , that is, all the way from the destination to the source and back. In the solution we expose here, distance D is divided into N zones each having one mule. Our results with perfect navigation and rendezvous show that we achieve better results in terms of achievable data rate. That said, we have only considered perfect navigation, infinite energy, and instantaneous rendezvous. Future work we plan, is to carry out an experiment to address this issue.

IV. CONCLUSIONS, FUTURE WORK

Extended ranges and data rate for underwater network is a crucial for underwater operations. Our results show that data muling is theoretically effective. We provide a simple model to predict performance: we provide a criterion to characterise latency, which allows to play on the variables affecting it. for the moment strong hypotheses were used, but our future work aims at relaxing these hypotheses to better match reality.

ACRONYMS

AUV	Autonomous Underwater Vehicle.
Gb/s	Gigabits per second (10^9 bits per second).
kb/s	kilobits per second (10^3 bits per second).
Mb	Megabits (10^9 bits).
Mb/s	Megabits per second (10^6 bits per second).
RF	Radio Frequency.
TSP	Travelling Salesman Problem.
UDMP	Underwater Data Muling Protocol.
USN	Underwater Sensors Network.
UWOC	Underwater Wireless Optical Communication.

REFERENCES

- [1] Cory Dixon and Eric W. Frew. Optimizing Cascaded Chains of Unmanned Aircraft Acting as Communication Relays. *IEEE Journal on Selected Areas in Communications*, 30(5):883–898, June 2012.
- [2] Marek Doniec, Iulian Topor, Mandar Chitre, and Daniela Rus. Autonomous, Localization-Free Underwater Data Muling Using Acoustic and Optical Communication. In Jaydev P. Desai, Gregory Dudek, Oussama Khatib, and Vijay Kumar, editors, *Experimental Robotics*, volume 88, pages 841–857. Springer International Publishing, Heidelberg, 2013. Series Title: Springer Tracts in Advanced Robotics.
- [3] Marek Doniec, Iuliu Vasilescu, Mandar Chitre, Carrick Detweiler, Matthias Hoffmann-Kuhnt, and Daniela Rus. Aquaoptical: A lightweight device for high-rate long-range underwater point-to-point communication. In *OCEANS 2009*, pages 1–6, 2009.
- [4] Matthew Fewell. Empirical Range–Data-Rate Relationships for Undersea Wireless Communications Systems. Technical report, Defence Science and Technology Organisation, Australia, October 2020.
- [5] Hemani Kaushal and Georges Kaddoum. Underwater Optical Wireless Communication. *IEEE Access*, 4:1518–1547, 2016.
- [6] Clifford Pontbriand, Norm Farr, Johanna Hansen, James C. Kinsey, Leo-Paul Pelletier, Jonathan Ware, and Dehann Fourie. Wireless data harvesting using the AUV Sentry and WHOI optical modem. In *OCEANS 2015 - MTS/IEEE Washington*, pages 1–6, Washington, DC, October 2015. IEEE.
- [7] Rahul C. Shah, Sumit Roy, Sushant Jain, and Waylon Brunette. Data MULEs: modeling and analysis of a three-tier architecture for sparse sensor networks. *Ad Hoc Networks*, 1(2-3):215–233, September 2003.
- [8] Filipe B. Teixeira, Nuno Moreira, Nuno Abreu, Bruno Ferreira, Manuel Ricardo, and Rui Campos. UDMSim: A Simulation Platform for Underwater Data Muling Communications. In *2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–6, October 2020. ISSN: 2160-4894.
- [9] Filipe B. Teixeira, Nuno Moreira, Rui Campos, and Manuel Ricardo. Data Muling Approach for Long-Range Broadband Underwater Communications. In *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–4, Barcelona, Spain, October 2019. IEEE.