Focused Beam Routing Protocol for Underwater Acoustic Networks

Josep Miquel Jornet
Massachusetts Institute of Technology
jmjornet@mit.edu

Milica Stojanovic
Massachusetts Institute of Technology
millitsa@mit.edu

Michele Zorzi
University of Padova, Italy
zorzi@dei.unipd.it

ABSTRACT

Multi-hop transmission is considered for large coverage areas in bandwidth-limited underwater acoustic networks. In this paper, we present a scalable routing technique based on location information, and optimized for minimum energy per bit consumption. The proposed Focused Beam Routing (FBR) protocol is suitable for networks containing both static and mobile nodes, which are not necessarily synchronized to a global clock. A source node must be aware of its own location and the location of its final destination, but not those of other nodes.

The FBR protocol can be defined as a cross-layer approach, in which the routing protocol, the medium access control and the physical layer functionalities are tightly coupled by power control. It can be described as a distributed algorithm, in which a route is dynamically established as the data packet traverses the network towards its final destination. The selection of the next relay is made at each step of the path after suitable candidates have proposed themselves.

The system performance is measured in terms of energy per bit consumption and average packet end-to-end delay. The results are compared to those obtained using pre-established routes, defined via Dijkstra’s algorithm for minimal power consumption. It is shown that the protocol’s performance is close to the ideal case, as the additional burden of dynamic route discovery is minimal.

Categories and Subject Descriptors: A.1. [General]: Introductory and survey; H.1.1. [Information systems]: Systems and information theory.

General Terms: Theory.

Keywords: Underwater acoustic networks, routing, power control.

1. INTRODUCTION

Underwater networking is attracting the attention for an increasing number of applications, in which nodes may be required to cooperate, minimizing their energy consumption without compromising the network connectivity and the ability to deliver data to a final destination.

In underwater acoustic networks, the transmission power depends on the distance that needs to be covered, and this may be many times greater than the power required for reception (e.g., tens of Watts as opposed to a Watt or less). Because the power is battery-supplied, conservation of energy is an important aspect in the design of underwater acoustic networks. Several approaches can be pursued to this end. One approach is to reduce the transmission power by routing a message over multiple short hops instead of sending it directly over one long link. Another is to increase the transmission bandwidth, which effectively reduces the bit duration and, consequently, the energy per bit. Note that the two approaches go hand-in-hand, since higher bandwidth is available over shorter acoustic links [1].

The benefits of multi-hopping in the underwater acoustic channel have been recently addressed in the literature. In [2], the capacity of an acoustic relay link is analyzed for a noise-limited scenario, showing that it increases with the number of hops used to span a given distance. This serves as an upper bound on all practical systems in which the channel access must be regulated, in either a deterministic or a random fashion. The effects of interference on the system capacity in a contention-based acoustic network have been assessed in [3], showing similar results.

In [4], discrete power control is introduced as a practical means of enabling multi-hop communications in underwater acoustic networks. In this case, nodes are able to switch their transmission power $P$ over a finite set of power levels. The performance of various MAC protocols was measured when following pre-established routes, designed in light of minimum power consumption. This can only be done when each node has a complete knowledge of the network topology.

In this paper, we propose a routing methodology, and evaluate its performance when coupled with power control [4]. This routing technique assumes that nodes know their own locations. Such assumption is justified in underwater systems where fixed bottom-mounted nodes have location information upon deployment, while the mobile nodes, i.e.,

---

Note also that transmitting at a higher bit rate means shorter packet durations (assuming a fixed number of bits per packet) which in turn implies fewer chances of packet collisions when random access is employed, because propagation delay is not negligible in an acoustic channel (on the contrary, it equals about one second over 1500 meters). This fact affects the performance of a MAC protocol, which, in turn, will affect the performance of a routing protocol.
autonomous underwater vehicles (AUVs), are equipped with internal navigation systems. In addition, a source node also knows the location of its desired final destination, but not the locations of other nodes. Due to drifting, the location of each node may change in time, and, hence, it may not be realistic to assume that each node knows exactly the location of all other nodes. This case is representative of an underwater network in which the distributed nodes are required to transmit to a common sink, or a set of sinks.

Without location information, a large number of broadcast or multicast queries may cause unnecessary network flooding, thus reducing the user perceived throughput. This is one of the main limitations in non-geographical ad hoc routing protocols. In proactive protocols (e.g., DSDV [5], OLSR [6]), or reactive protocols (e.g., AODV [7], DSR [8]), large signaling overhead and high latency may compromise the network performance.

The knowledge of location can eliminate this effect. In wireless sensor networks, location awareness has been previously considered leading to geographical routing protocols such as GeRaF [9, 10], a forwarding technique based on location information of the nodes involved, and random selection of the relaying nodes via contention among the receivers. An integrated MAC/Routing protocol based on geographical information and which makes use of power control is introduced in [11]. In this case, a competition is triggered at each hop in order to select the next relay node, in such a way that energy efficiency is maximized.

Routing protocols based on location information have been also designed explicitly for the underwater channel. In [12], a location aware variation of DSR in which link quality measurements are considered in the relay selection process is shown to reduce the system latency. In [13], the authors propose a vector-based forwarding protocol for sensor networks, in which a virtual transmission pipe is defined at each hop of the transmission path. In [14], the design of minimum energy routes is assessed, showing that in dense networks there is an optimal number of hops over which the system performance does not improve.

In Sec. 2, the proposed Focused Beam Routing (FBR) protocol is defined and illustrated. In Sec. 3, the coupling of the routing functionalities and the MAC protocol is addressed. Sec. 4 is devoted to analyzing the system performance through simulation examples, and studying the impact of different system parameters. Conclusions are summarized in Sec. 5.

2. FOCUSED BEAM ROUTING PROTOCOL

To illustrate the routing protocol, let us refer to the example of Fig. 1. Shown in this figure is a network of nodes, distributed in an arbitrary manner across some area. A simple two-dimensional scenario can be envisioned without loss of generality, while the extension to a three-dimensional case is straightforward.

Referring to Fig. 1, let us assume that node A wants to transmit to node B. To do so, node A will issue a request to send (RTS) to its neighbors. This request is a short control packet that contains the location of the source node (A) and of the final destination (B). Note that this is in fact a multicast request.

The initial transaction is performed at the lowest power level and the power is increased only if necessary. Power control is performed as an integral part of routing and medium access control. We assume open loop power control, in which the transmitting node decides which power level to use, rather than being instructed explicitly by a receiving node. In a practical system, power control will be implemented by choosing from several discrete levels. At the moment, we are not concerned with the exact way in which the power levels are determined, but simply assume that there is a finite number of increasing power levels, \( P_1 \) through \( P_N \). A detailed analysis of distributed discrete power control can be found in [4].

Corresponding to each power level \( P_n \), there is a transmission radius \( d_n \). Only the nodes that are within this radius are assumed to receive the signal at a level sufficient for detection. The signal of course propagates beyond this distance and can be overheard, but because of attenuation it cannot be detected. As such, it causes interference to other nodes, which will be taken into account when evaluating the system performance.

Returning to our example, let us draw an imaginary line between nodes A and B. All the nodes that receive A’s multicast RTS first calculate their location relative to the AB line. The objective in doing so is to determine whether they are candidates for relaying. Candidate nodes are those that lie within a cone of angle \( \pm \theta/2 \) emanating from the transmitter towards the final destination. If a node determines that it is within the transmitter’s cone, it will respond to the RTS. Those nodes that are outside the cone will not respond.

In our example, there are no nodes within the transmission cone that can be reached at the power level \( P_1 \). Hence, after an expected round-trip time (2\(d_1/c \) for the power \( P_1 \), where \( c=1500 \) m/s is the nominal speed of sound underwater), node A receives no responses. It now increases the transmission power to \( P_2 \), and sends a new RTS. In general, a transmitting node will keep increasing the power until it reaches someone, or until all power levels have been exhausted. If it cannot reach anyone at the maximal level \( P_N \), the transmitter will shift its cone and start looking for candidate relays left and right of the main cone. This strategy favors paths with minimal amount of zigzagging, while guaranteeing that all possible paths will eventually be searched. Alternatively, a node may first search in the \( d_1 \) vicinity by shifting its cone and then decide to increase the power to advance in distance.

Other strategies are also possible; for example, the strat-
energy used in GeRaF[9] involves defining relay zones as the intersections of concentric circles around the receiver and the coverage area centered in the transmitter. A relay is then chosen from the region that provides the greatest advancement. Specifically, the ring farthest from the transmitter is queried first, then the closer rings. The objective in doing so is not to conserve the power (transmission power is set to reach the farthest circle), but to find the shortest path (given the finite transmission power, equal for all nodes).

If the transmitter, after increasing the power to some level, reaches a single neighbor, it passes the data packet on to that neighbor, who becomes a relay. A positive acknowledgement at each hop is expected. The relay now initiates an identical procedure, looking for candidate nodes within its cone. It has become an effective transmitter, searching for the next relay towards the final destination. If there is more than one candidate relay, the current sender will have to decide which one will become the next relay. In our example, A reaches two candidates, C and D, at power $P_t$. (The protocol does not change if there are more than two candidates.) When they receive the RTS from A, each one knows that it can help in relaying, and each replies to A’s request using a very short control packet, akin to the clear to send (CTS) signal. Note that there is a subtle difference between the traditional CTS, issued by the destination node, and this one, which is issued by a candidate relay. A candidate’s CTS contains the address (name and location) of the node issuing it (C or D) as well as the addresses of the source and destination (A and B). The two candidate relays are not (yet) aware of each other’s existence, so it is possible that their replies will collide. However, because the CTS is very short, and the distances CA and DA are unlikely to be exactly the same, the chances of the two CTSs colliding at A are minimal. For example, with 500 bits in a CTS packet, and a bit rate of 5 kbps, there will be no collision if the distances CA and DA differ by more than 75 m. Transmission times may also be randomized in order to avoid node synchronization effects.

If there is no collision, A receives both replies. A reply includes the sender’s location, and, hence, A knows which candidate is closer to the final destination – node D in this case. It may then choose D as the relay, and pass the data packet on to it. Node C will overhear the data packet transmission and deduce from its header that it has not been chosen as relay. Alternatively, more intelligence can be incorporated into making this decision. For example, A could know from overhearing previous transmissions that D is already engaged elsewhere and is thus becoming a bottleneck; it could therefore choose C as its relay. Alternatively, the CTS packet can include information about the network activity that each one of the candidates is measuring. In that case, routing is performed by exploiting first and second order neighborhood information for more efficient, integrated MAC/routing schemes [16]. As the authors show in [16], this information can be used as part of the relays’ decision of whether and when to respond a multicast RTS. However, such details are of no concern for the basic routing principle. An important observation to be made is that the (long) data packet is transmitted only after the relay has been chosen, i.e., the link is secured and there are no risks of data packet collisions. In other words, the only packets that can collide are the (short) control packets.

Although the chances of collision are small, it can still happen. If A detects a collision (e.g., by detecting signal energy without being able to decode a packet), it will send the RTS again, using the same power level. In this round, however, C and D may know of each other’s existence. This can only be guaranteed if they are inside a cone with an aperture less than or equal to $60^\circ$. In this case, they have also learned each other’s location, and only that node that knows to be closest to the final destination will reply. Hence, the next CTS collision will be avoided. In a more general case, C and D may not be aware of each other either because of the half-duplex operation of acoustic modems, or because the distance CD is greater than the transmission range associated with the power level in use. In this situation, they are able to detect that the previous query has not been completed successfully because they will have received exactly the same request as before. Then, they may delay their CTS retransmissions by $T_{\text{delay}} = N_{\text{res}} \cdot T_{\text{CTS}} \cdot x$ seconds, where $N_{\text{res}}$ is the number of retransmissions, $x$ is a uniformly distributed random variable between 0 and 1, and $T_{\text{CTS}}$ stands for the duration of the reply packet.

When the next relay has been chosen, the procedure continues. The cone emanating from node D is illustrated by dashed lines in Fig. 1. Note that the original sender does not need to know the location of the destination exactly. As the route discovery advances, the final relay will reach all the nodes in its own cone, and, so long as the destination has not moved outside of this region, it will be reached. Hence, there is a region over which a node can move without affecting the protocol performance. Normally, underwater nodes can be moving at a much slower rate than the speed of propagation (few m/s as compared to few km/s) and, hence, it is reasonable to assume that a node will not "escape" before it is reached.

As the algorithm progresses, and a cone is formed at each relay, the route will zoom in on the final destination so long as there are candidate relays within reach of one another. Fig. 2 illustrates the region of candidate relay locations for the case when a relay can be found in each hop within a single cone, i.e., no node needs to shift its cone and look outside of the angle $\theta$. Note that this region is bounded, as dictated by the definition of the transmitter’s cone. This is why the protocol is called focused beam routing.

![Figure 2: The region of candidate relay locations is contained in a cone emanating from each relay. The region of all reachable relays at the lowest power level is the shaded beam-like area.](image-url)
3. MEDIUM ACCESS CONTROL:
IMPLEMENTATION DETAILS

The algorithm we propose can be coupled with any MAC protocol. Since the exchange of short control packets is an inherent part of the proposed routing protocol, DACAP, a collision avoidance protocol based on virtual carrier sensing [15], seems a suitable choice. Below, we summarize the more important aspects of coupling the MAC and the routing layers.

Multicast Requests – When requesting a route, the transmitter sends a multicast RTS. Each control packet contains three \( \{ID, Position\} \) pairs: one for the current transmitter, one for the final destination and one for the next intermediate node, i.e., the relay. In a multicast RTS, this field is left empty. A node proposing itself as a relay overwrites it with its own ID and position. After sending a multicast RTS, the transmitter will wait twice the maximum propagation delay corresponding to the current transmission power level even if it has already received one or more CTSs (plus the corresponding additional delay if it is a retransmitted packet). The value of the cone aperture is also specified in the multicast RTS.

Silence Packets – After a multicast RTS, the requesting node may receive no answers. This will occur if there are no neighbors, or there are, but they are already engaged in another communication. In the latter case, if the transmitter is not aware of the situation, it will decide to increase the transmission power, increasing the chances of disturbing other ongoing transmissions. To prevent this situation, a node aware of a concurrent communication that overhears a multicast RTS will send a very short silence packet to the requesting node. A node receiving a silence packet will defer its transmission. The length of this type of packet minimizes the chances of interfering with the ongoing communication.

Implicit Acknowledgement – Apart from an end-to-end acknowledgement which may be generated at the transport or application layers, each intermediate node expects a positive acknowledgement from the current receiver. The concept of implicit acknowledgement, as outlined in [4] is used, i.e., if nodes use omnidirectional transducers, which is the case of mobile nodes, the transmitter can deduce that its last data transaction has been successfully completed if it overhears its own packet being transmitted to the next relay. However, this may not be always possible. If the power level used to reach the next node is lower than the one used for the previous transmission, the acknowledgement should be sent explicitly using a higher power level. The same should be done when the packet reaches its final destination. At the same time, if for any reason a node receives an RTS from the same transmitter for a packet that has been successfully transmitted (each packet has a unique ID), an acknowledgement is explicitly sent, avoiding the long data packet retransmission.

Dynamic Backoff and Waiting Times – The power level that is being used is specified in each control packet. By doing this, any node that overhears an ongoing communication can dynamically adjust its backoff and waiting times.

![Figure 3: Scenario for simulation: a variable number of active nodes randomly positioned within a grid over an area of 200 km² and 4 sinks at the edges of the geometry.](image)

4. PERFORMANCE ANALYSIS

To assess the protocol performance, we have used a discrete-event underwater acoustic network simulator implemented in standard Python [17].

The simulation scenario is shown in Fig. 3. The network is composed of a varying number of active nodes, randomly located over a square area of 200 km². There are four sinks, located at the corners. A Poisson distribution with an average packet generation rate \( \lambda \) in packets/second for each transmitter is assumed. An active node chooses the closest sink.

Each node makes use of discrete power control with four uniformly separated levels, redefined for each node density as introduced in [4]. The system frequency allocation (center frequency \( f_c \) and bandwidth \( B \)) is made so as to optimize the system performance in terms of energy per bit consumption, end-to-end delay and number of collisions [4]. The average energy per bit consumption takes into account the energy invested in transmission, listening and active reception of control and data packets, as well as their possible retransmissions.

4.1 Network Density

We first investigate the impact of node density on the protocol performance. A higher node density implies a shorter internode distance. For each link distance, there exists an optimal center frequency that minimizes the acoustic path loss [1]. At the same time, interference changes with the link distance and the frequency allocation scheme. This trade-off was assessed in [4] for different MAC protocols and different power allocation policies. Here, we compute the optimal frequency allocation scheme \( (f_c \) and \( B \) ) for every node density.

Fig. 4 shows the energy per bit consumption, the average packet end-to-end delay and the total number of collisions in the network, as functions of the network node density. The performance is evaluated for two choices of the cone aperture, \( \theta=30^\circ \) and \( \theta=150^\circ \), and is compared to the case in which routes are established using Dijkstra’s algorithm, with the cost between two nodes defined as the minimal power level required to guarantee connectivity. Two nodes are connected if they can reach each other with a reference SNR₀, 20dB in our simulations.
Figure 4: Energy per bit consumption, average packet end-to-end delay and number of collisions when using a transmission cone with $\theta=30^\circ$ and $\theta=150^\circ$ as functions of the number of active nodes over a fixed area of 200 km$^2$ ($\lambda=0.5 \cdot 10^{-3}$ packets/second, packet length=9600 bits).

In [14], it was shown that, especially in very dense networks, paths following minimum power routes (maximum number of hops) are not optimal in terms of energy savings, but that instead there is a minimum distance that should be traversed in each hop. However, the authors also show that for the node densities that we are considering, both options are the same. For this reason, we use this minimum power routes as the gold standard. Alternatively, in very dense networks, this problem could be managed by sweeping the power levels from the highest one, instead of the lowest one.

First of all, we note from Fig. 4 that both the energy per bit and packet end-to-end delay are very close to those of the network with minimum-power pre-established routes. This shows that the method we are proposing is able to dynamically discover minimum energy routes with minimal network knowledge. In addition, by combining the channel reservation process with the route discovery phase, the extra delay introduced by routing is small. As described in Sec. 2, after a multicast RTS, the current transmitter should wait twice the maximum propagation delay corresponding to its current transmission power. For a given transmission range, the node offering the maximum advancement towards the destination is the best relay. Therefore, even when packets follow pre-established routes, the waiting time necessary to reserve the channel tends to this value. At the same time, the number of collisions can be less than for pre-established routes because bottleneck situations are dynamically resolved (this is not the case for Dijkstra’s algorithm in the way it is defined).

4.2 Optimal Cone Aperture

The cone aperture $\theta$ plays an important role in the system performance. In sparse networks, limiting the area with potential relays to a cone with an aperture of $30^\circ$ turns out to be less energy efficient than opening the cone. Indeed, in low density networks, rather than reducing the amount of zigzagging, too small $\theta$'s can force the protocol to switch to higher power levels than necessary. This reduces the end-to-end delay but increases the energy consumption. On the contrary, in dense networks, forwarding a packet over too many relays (large $\theta$) can overload the network, thus degrading the system performance, as illustrated in Fig. 4c. As we are using DACAP with power control, network congestion translates into higher delays, but not into a noticeable energy consumption increase.

For each node density, an optimal cone aperture can be determined for a given scenario. In Fig. 5, the cone aperture that minimizes the average energy per bit consumption for each node density is shown for our example network. The energy per bit consumption, the average packet end-to-end delay, and the total number of collisions when using this optimal cone aperture are shown as functions of the network node density in Fig. 6. The results confirm our earlier conjecture. When the network is composed of only a few nodes, opening the cone can reduce the energy per bit consumption on average (in this case, zigzagging is not detrimental). At higher network densities, focusing on the receiver by closing the cone avoids making more hops than necessary by preventing zigzagging.

4.3 Packet Generation Rate

For a specific network density, the protocol performance as a function of the packet generation rate is measured and compared to the case in which pre-established routes are followed. In this way, the actual load in the network due to the route discovery mechanism is obtained.

Fig. 7 shows the energy per bit consumption, the average packet end-to-end delay, and the total number of collis-
Figure 6: Energy per bit consumption, average packet end-to-end delay and number of collisions when using a transmission cone with the optimal aperture of Fig. 5, as functions of the number of active nodes over a fixed area of 200 km² (λ=0.5·10⁻³ packets/second, packet length=9600 bits).

Figure 7: Energy per bit consumption, average packet end-to-end delay and number of collisions when using a transmission cone with θ=90° and θ=150° as functions of the packet generation rate of 100 active nodes, over a fixed area of 200 km².

which packets follow pre-established routes.

Future work involves several refinements and extensions, as well as validation of the system performance in different scenarios (e.g., a scenario containing only mobile nodes and one static sink). In particular, issues which should be considered are additional cost metrics in the candidate selection process (now only the location information is considered), alternative MAC protocols, and the effect of different power allocations strategies (e.g., defining the power levels according to the average number of neighbors that can be reached).

6. ACKNOWLEDGMENT

This work was supported by the NSF grants 0520075 and 0427502, and the NOAA Sea Grant NA060AR4170019.

5. CONCLUSIONS AND FUTURE WORK

A routing technique, based on location information, was proposed for energy-efficient multi-hop communications in underwater acoustic networks. The system performance was evaluated for different node densities and network loads. It was shown that, by properly coupling routing and MAC functionalities with power control, routes can be established on demand with a minimum impact on the network performance. Energy per bit consumption and average packet end-to-end delay were compared to those of a network in
7. REFERENCES


