Underwater Acoustic Networks

Ethem M. Sozer, Milica Stojanovic, and John G. Proakis, *Life Fellow, IEEE*

***Abstract*—With the advances in acoustic modem technology that enabled high-rate reliable communications, current research focuses on communication between various remote instruments within a network environment. Underwater acoustic (UWA) networks are generally formed by acoustically connected ocean-bottom sensors, autonomous underwater vehicles, and a surface station, which provides a link to an on-shore control center. While many applications require long-term monitoring of the deployment area, the battery-powered network nodes limit the lifetime of UWA networks. In addition, shallow-water acoustic channel characteristics, such as low available bandwidth, highly varying multipath, and large propagation delays, restrict the efficiency of UWA networks. Within such an environment, designing an UWA network that maximizes throughput and reliability while minimizing the power consumption becomes a very difficult task. The goal of this paper is to survey the existing network technology and its applicability to underwater acoustic channels. In addition, we present a shallow-water acoustic network example and outline some future research directions.**

***Index Terms*—Acoustic networks, telemetry, underwater communications.**

I. INTRODUCTION

UNDERWATER communication applications mostly involve long-term monitoring of selected ocean areas [1]. The traditional approach for ocean-bottom or ocean-column monitoring is to deploy oceanographic sensors, record the data, and recover the instruments. This approach creates long lags in receiving the recorded information. In addition, if a failure occurs before recovery, all the data is lost. The ideal solution for these applications (such as data collection from hot vent communities, offshore mining or drilling, etc.) is to establish real-time communication between the underwater instruments and a control center within a network configuration.

Basic underwater acoustic (UWA) networks are formed by establishing two-way acoustic links between various instruments such as autonomous underwater vehicles (AUV's) and sensors. The network is then connected to a surface station which can further be connected to a backbone, such as the Internet, through an RF link. This configuration creates an interactive environment where scientists can extract real-time data from multiple distant underwater instruments. After evaluating the obtained data, control messages can be sent to individual instruments and the network can be adapted to changing situations. Since data is transferred to the control station when it is available, data loss is prevented until a failure occurs.

Underwater networks can also be used to increase the operation range of AUV's. The feasible wireless communication range of an AUV is limited by the acoustic range of a single modem which varies from 10 to 90 km [1]. By hopping the control and data messages through a network that covers large areas, the range of AUV's can be considerably increased. However, as noted in [2], the effects of propagation time delays should be considered.

The shallow-water acoustic channel differs from radio channels in many respects. The available bandwidth of the UWA channel is limited and depends on both range and frequency [3], [4]. Within this limited bandwidth, the acoustic signals are subject to time-varying multipath [5], which may result in severe intersymbol interference (ISI) and large Doppler shifts and spreads, relative to radio channels. These characteristics restrict the range and bandwidth for reliable communications.

The propagation speed in the UWA channel is five orders of magnitude lower than that of the radio channel. Large delays between transactions can reduce the throughput of the system considerably if it is not taken into account. Also, since the ocean-bottom instruments are battery-powered, power efficiency is a desirable characteristic for underwater networks. Special attention should be given to these facts when designing a network protocol.

The paper is divided into two parts. The first part is an overview of network design principles. We survey multiple access, media access, automatic repeat request (ARQ), and routing methods for communication networks. An overview of the work carried out to date on the development of UWA networks is also included in this section. The second part presents a design example of a shallow-water network. Network topology is determined to provide minimum energy consumption, which is an important constraint for the battery-powered underwater modems. A media access method based on the multiple access with collision avoidance (MACA) protocol [22] is proposed and an Opnet simulation is conducted. Network performance is assessed in terms of delay and throughput analysis. Finally, Section IV summarizes the conclusions and gives an outline of future research directions in this area.

II. NETWORK DESIGN PRINCIPLES

The design of an information network is commonly carried out in the form of a layered architecture [7]. The first three layers of this hierarchical structure are the physical layer, the data link layer, and the network layer.

The physical layer is in charge of converting the logical information (bits 0 and 1) into signals which are transmitted over the communication channel. At the receiving end, it is in charge of detecting the signal corrupted by noise and other channel distortions and converting it back into logical bits.

The bits are often grouped into packets, which are processed at the data link layer. This layer has two major functions:

Manuscript received July 15, 1999; revised August 10, 1999.
The authors are with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA.
Publisher Item Identifier S 0364-9059(00)00294-6.

framing and error correction control. Framing refers to defining a packet, which includes the information sequence, synchronization preambles, source, and destination addresses, and other control information. Error correction control at the data link layer is most commonly implemented through a cyclic redundancy check (CRC). Redundant bits are formed from the bits in the packet and appended to it. At the receiver, a check is performed to detect errors in a packet. If the CRC fails, a node may ask for the packet to be retransmitted. This procedure is known as the automatic repeat request (ARQ). In order to implement an ARQ procedure, the nodes in a network must follow a data link protocol. Commonly used protocols are Stop & Wait, Go Back N, and Selective Repeat Protocol. These protocols control the logical sequence of exchanging the packets between two nodes and acknowledging the correctly received packets. They form the logical link control (LLC) sublayer of the data link layer. If there are more than two nodes communicating on the same channel, i.e., if the channel is broad-cast, additional measures must be taken to orchestrate the access of multiple sources to the same medium. These measures are known as the media access control (MAC). Examples of MAC protocols are the Aloha protocol, the carrier sense media access (CSMA) protocols (Ethernet is such a protocol), and the Token protocols. These protocols form the MAC sublayer of the data link layer.

The layer above the data link layer is the network layer. As every layer in a network architecture does, this layer assumes perfect operation of the layer below it. Thus, the network layer manipulates the packets without questioning their correctness. Its major function is that of routing. Routing involves finding a path through the network and forwarding the packet(s) from the source to the destination along that path. If all the packets of the same message follow the same path, established by the routing algorithm at the beginning of a message exchange, the network is said to employ virtual circuit switching. If a new path is determined for every packet, the network employs datagram switching. Finding a path through the network is often subject to an optimality criterion, such as minimizing the path distance (the delay, the number of hops, or some other "distance" measure). Static routing algorithms that compute the shortest path include the Dijkstra algorithm and the Bellman–Ford algorithm. Static algorithms provide the basis for the design of many dynamic routing algorithms used in practical networks that have to adapt to the changing operating conditions.

In this section, we survey the methods and protocols used in the data link layer and the network layer. A treatment of the physical layer is beyond the scope of this paper. We also discuss possible network topologies, which are an important constraint in designing a network protocol.

A. Network Topologies

There are three basic topologies that can be used to interconnect network nodes: centralized, distributed, and multihop topology [9]. In a centralized network scenario, the communication between nodes takes place through a central station, which is sometimes called the *hub* of the network. The network is connected to a backbone at this central station. This configuration is suitable for deep-water UWA networks, where a surface buoy with both an acoustic and an RF modem acts as the hub and controls the communication to and from ocean-bottom instruments. A major disadvantage of this configuration is the presence of a single failure point [9]. If the hub fails, the entire network shuts down. Also, due to the limited range of a single modem, the network cannot cover large areas.

The next two topologies are classified as peer-to-peer networks. A fully connected peer-to-peer topology provides point-to-point links between every node of the network. Such topology eliminates the need for routing. However, the output power needed for communicating with widely separated nodes is excessive. Also, a node that is trying to send packets to a far-end node can block the signals to a neighboring node, which is called the *near–far* problem [9].

Multihop peer-to-peer networks are formed by establishing communication links only between neighboring nodes. Messages are transferred from source to destination by hopping packets from node to node. Routing of the messages is handled by intelligent algorithms that can adapt to changing conditions. Multihop networks can cover relatively larger areas since the range of the network is determined by the number of nodes rather than the modem range. However, as the number of hops increases, the packet delay also increases. Special attention should be given to applications that are sensitive to delays. In Section III-A, we compare fully connected and multihop network topologies for a specific example.

B. Multiple Access Methods

In many information networks, communication is bursty, and the amount of time that a user spends transmitting over the channel is usually smaller than the amount of time it stays idle. Thus, network users should share the available frequency and time in an efficient manner. In this section, we discuss three major multiple access methods: frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA). We compare these methods from the point of view of a shallow-water acoustic communication network.

1) Frequency Division Multiple Access (FDMA): FDMA divides the available frequency band into subbands and assigns each subband to an individual user. The channel is used only by that user until it is released. The bandwidth of FDMA channels may be smaller than the coherence bandwidth of the transmission channel. In such a case, FDMA users are vulnerable to fading. The severe fading present in UWA channels creates a difficult environment for FDMA-based systems. This problem is usually solved by coding [32]. Also, due to the fixed channel bandwidths, this method is not flexible and can be very inefficient in bursty traffic [6], [9], [10].

2) Time Division Multiple Access (TDMA): Instead of dividing the frequency band, TDMA divides a time interval, called a frame, into time slots. Each time slot is assigned to an individual user. Time slots and overhead bits are combined into frames. Collision of packets from adjacent time slots are prevented by including guard times that are proportional to the propagation delay present in the channel. As a consequence, TDMA requires more overhead than FDMA [6].

In cellular environments, where many centralized subnets are combined together, TDMA is used with FDMA to reduce the effects of intercell interference. For example, in GSM, the European mobile cellular system, a seven-cell reuse pattern is employed by dividing the frequency band into seven subbands. TDMA users access the entire available frequency band that is assigned to a cell during their time slot. The higher transmission bandwidth relative to FDMA increases the resistance of the TDMA system to frequency-selective fading [10].

Since the data is buffered until the assigned time slot, TDMA transmission occurs in bursts. The transmitter can be turned off during idle periods and battery consumption can be reduced. Bursty transmission also requires a high bit rate relative to FDMA and increases ISI. Therefore, adaptive equalizers are usually required in TDMA systems [6], [10].

A major advantage of TDMA as compared to FDMA is its flexibility. Since the modem hardware of each user is the same, the number of time slots assigned to a user can be changed without the need for additional hardware. In this way, the data rate of users can be increased on demand [9].

The major disadvantage of TDMA in underwater acoustic networks is that it requires strict synchronization. A way of establishing a common timing reference in centralized topologies is to broadcast periodic probe signals. However, the time slots formed in this manner should be kept long enough to avoid collisions due to differences in propagation delays. The excessive delays present in the underwater channel causes large idle times, which results in low throughput.

3) Code Division Multiple Access (CDMA): CDMA allows multiple users to operate simultaneously over the entire frequency band. Signals from different users are distinguished by means of pseudonoise (PN) codes that are used for spreading the user messages. There are two basic spreading techniques: direct sequence (DS) and frequency hopped (FH) spread spectrum. In the DS-CDMA case, information signals are linearly modulated using wide-band PN codes, whereas in FH-CDMA the carrier frequencies of the users are changed according to a pattern obtained from the PN codes. A unique PN code is assigned to each user in the system.

The large bandwidth of CDMA channels provides resistance to frequency-selective fading. In a FH system, frequency bands separated by more than the coherence bandwidth of the channel (the inverse of the multipath spread) fade independently; hence, a coding scheme can be employed to extract a diversity gain from the fading channel. In a DS system, fine time resolution of the spreading codes provides the possibility to coherently combine the multipath arrivals by employing Rake filters at the receiver. If the resolvable multipath components fade independently, it is possible to extract a time diversity gain present in the channel [6], [11].

The CDMA system capacity is mostly restricted by multiple access interference (MAI). This interference is caused to the desired user's signal by the signals of other users, which, due to the nature of spread-spectrum signaling, may overlap with the desired signal in both time and frequency. In a DS-CDMA system, every active user contributes a wide-band signal to the overall MAI, effectively increasing the level of noise. In a FH-CDMA system, MAI is caused when a hopping pattern of an interfering user coincides with the hopping pattern of the desired user. The effects of MAI can be controlled by careful design of codes with large spreading gains and by employing multiuser receivers. In a network environment, where the data traffic is in the form of bursts, if the transmission is stopped during idle periods, such as silence in speech, the system capacity (the number of users that can be supported) can be increased. If a user transmits only 50% of the time, the capacity of the system can be doubled [11]. Since each additional user only increases the overall interference, the system performance degrades, but there is no hard limit in the capacity of a CDMA system.

CDMA systems are vulnerable to the *near–far* problem [11]. A power control algorithm is used to reduce the output power level of each node such that it can establish reliable packet transfer without creating excessive interference. This minimization of output power is also essential in underwater networks to reduce the battery consumption. It is shown in [12] that the power requirement of CDMA systems can be less than that of TDMA systems.

Spread-spectrum signals can be used for resolving collisions at the receiver by using multiuser detectors [13]–[15]. Theoretically, if codes of length N are used in the system, collisions involving at most N users can be resolved. By this way, the receiving node can respond to all N messages and the number of retransmissions is decreased. This property both reduces battery consumption and increases the throughput of the network. Additional increase in system capacity can be achieved by use of directional antennae.

In conclusion, CDMA and spread-spectrum signaling appear to be a promising multiple access technique for shallow-water acoustic networks.

C. Media Access Protocols

The UWA channel has very limited resources, as we have discussed in Section I. These scarce resources should be shared in a fair and efficient manner by means of a media access protocol. In this section, we review the existing media access protocols and discuss their effectiveness in an underwater network.

1) ALOHA: The original ALOHA protocol is based on random access of users to the medium. Whenever a user has information to send, it transmits it immediately. An acknowledgement (ACK) is sent back by the receiver if the packet is received without errors. Due to the arbitrary transmission times, collisions occur and packets are lost. If this happens, the receiver does not issue an ACK, and, after randomly selected times, the senders retransmit their packets. Because the random times of retransmissions are selected independently, the chances of a repeated collision are low. Due to the retransmissions, the average time required to successfully transmit a packet through the channel is longer than the minimum required for a single packet transmission. The ratio of the useful time (time required for information transmission) to the total average time that the channel is occupied in transmitting this information defines the throughput of a media access scheme. With the simple ALOHA method, the maximum achievable throughput is 18% [8].

An enhanced version of the ALOHA protocol, the slotted ALOHA method, was proposed by Roberts [16]. In this method, the time is divided into time slots, and the local clock of each

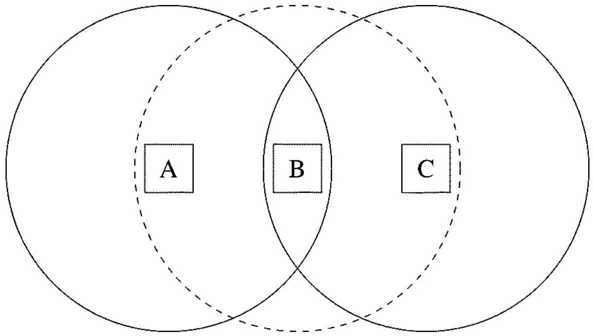


Fig. 1. Node A can communicate with node B, but not with node C. Node B can communicate with both A and C.

node in the network is synchronized according to these slots. When a node wants to send a packet, it waits until the next time slot and then begins transmission. Restricting packet transmission to predetermined time slots decreases the vulnerable time during which a collision can occur. With this method, the maximum throughput is increased to 36%.

ALOHA becomes inefficient in bursty traffic that occurs in information networks. Also, trying to resolve collisions by retransmissions increases the power consumption of the network nodes and reduces the lifetime of the network.

2) Carrier Sense Media Access (CSMA): In ALOHA, users do not take the channel state into account. This results in a high rate of collisions. The scarce resources of the channel can be utilized much better if users listen to the channel before transmitting a packet. The media access methods based on this idea are called carrier sense media access (CSMA) [9]. Details and various forms of this method can be found in [17]–[20].

The CSMA method tries to avoid the collisions by listening for a carrier in the vicinity of the transmitter. This approach does not avoid collisions at the receiver [21]. Let us consider a network formed by three users, as shown in Fig. 1. The circles around each node show the communication range of that node. Assume that node A is sending a packet to node B. At the same time, node C listens to the channel and, because it is out of the range of A and does not detect the carrier of A, it begins transmission. This creates a collision at B, which is the receiver node. Node A was hidden from node C. This situation is called the *hidden node* scenario [21]. To enable B to hear both messages, node C should defer its transmission. However, if the destination of the packet of C is not B, there is no reason to defer the transmission, provided that node B has the capability to deal with the interference generated by the signal from node C [21]. In case of B sending a packet to A, C detects a carrier. This creates an *exposed node* situation, where C is exposed to node B.

CSMA cannot solve these problems without adding a guard time between transmissions that is proportional to the maximum propagation time present in the network. The extensive propagation delays in underwater channels can cause this method to become very inefficient. If we consider an underwater acoustic network with a maximum range of 10 km, data rate of 1 kbit/s, and a packet size of 1000 bit, the transmission delay and the maximum propagation delay become 1 s and 6.7 s, respectively. In this situation, most of the time the channel will be idle, which results in low throughput. In [31], CSMA is employed, but, as also noted by the authors, the most important issue of that design was reliability, not efficiency.

3) Multiple Access with Collision Avoidance (MACA): The MACA protocol was proposed by Karn [22] to detect collisions at the receiver as an alternative to CSMA. This protocol uses two signaling packets called request-to-send (RTS) and clear-to-send (CTS). When A wants to send a message to B, it first issues an RTS command which contains the length of the message that is to be sent. If B receives the RTS, it sends back a CTS command which also contains the length of the message. As soon as A receives the CTS, it begins transmission of the data packet. Any node that overhears a CTS (C in this case) defers its transmission for the length of the data packet to avoid collision. If a node overhears an RTS but not a CTS, it decides that it is out of range of the receiver and transmits its own packet. Therefore, this protocol can solve both the hidden and the exposed node problems.

Automatic power control can be incorporated into the MACA protocol in a simple manner [22]. During the RTS-CTS exchange, nodes learn the minimum power level needed for reliable communication by trial and error. With power control, a node that hears a CTS need not stop its transmissions. If it restricts its output power to a level that will not be heard by the node that sent the CTS, it can continue its transactions with the nodes that require less output power [22].

This protocol highly relies on the symmetry of the channel, that is, a CTS should be heard by all the nodes within the range of the receiver node. However, in case of power control, this symmetry can be lost. Therefore, it is better to send CTS signals at a higher level to ensure that all the nodes within the range can hear it.

This protocol can be used as a basis of a media access protocol for UWA networks. It provides information for power control algorithms and has the ability to avoid collisions before they occur. Extra RTS-CTS exchange adds overhead, but the reduction of retransmissions can compensate for this increase. However, many aspects of the protocol, such as ACK's and scheduling, need to be considered by the designer. This protocol and its extensions that are suitable for UWA channels of interest are discussed further in Section III-B.

4) MACAW: The MACAW protocol is a modified version of MACA. It is proposed by Bharghavan [21] to improve the performance and reliability of MACA.

In MACA, the transaction between two nodes occurs as an exchange of RTS-CTS-DATA packets. In other words, packet reception is not acknowledged. Acknowledging the correctly received packets, normally a function performed at the data link layer, is sometimes deferred to a later stage. Namely, on highly reliable links, where packets rarely contain errors, sending acknowledgment on every hop of a link wastes time. To increase the throughput, messages, rather than packets, can be acknowledged in such networks. If there is an erroneous packet (or more) in a message, the destination node will ask the source node to retransmit. Thus, the ARQ procedure is not implemented between every two nodes along the communication route, but between the source and the destination nodes only. Consequently, it cannot be implemented at the data link layer, but at a layer higher than the network layer (after the route is found). This layer is the transport layer.

If the communication channel is of poor quality, a message will likely contain an erroneous packet. Recovering the errors in the data packet at the transport layer will now require excessive delay. Generally, error correction is better performed at the data link layer for channels of low reliability such as radio or shallow-water acoustic channels. For this purpose, an ACK packet is transmitted after each successful transaction. Inclusion of an extra packet in the transaction increases the overhead, which decreases the throughput. However, it is shown in [21] that, for radio channels, the gain in throughput exceeds the increase in overhead.

The MACAW protocol ignores power control and asymmetries that can occur. Its performance under power control needs to be investigated. Also, the effect of adding more overhead to the protocol in an environment where propagation delays are excessive should be addressed.

D. Automatic Repeat Request (ARQ)

ARQ is used to detect errors in the data link control layer and then to request the retransmission of erroneous packets. There are three basic ARQ schemes: Stop & Wait, Go Back N, and Selective Repeat Protocol.

The Stop & Wait ARQ is the simplest of all three. The source of the packet waits for an ACK from the destination before sending the next data packet. If after a preset time-out the ACK is not received, the source retransmits the packet. The channel stays idle during the round-trip propagation time. In full duplex links, the throughput of the system can be increased if the source continues transmitting the same packet while it is waiting for the ACK. However, this method has excessive power usage.

In the Go Back N ARQ method, the source transmits packets in a window of size N without waiting for any ACK. The data packets are numbered and the destination sends back ACK packets containing the number of the acknowledged packet. ACK for a packet acknowledges all the previous packets. If the source does not receive an ACK for a packet, it begins retransmissions starting from the unacknowledged packet. The destination accepts packets in order, which eliminates the need for buffering.

If a buffer is employed at the destination, packets can be accepted out of order. The source then has to transmit only the unacknowledged packets. This method is called Selective Repeat ARQ. This ARQ scheme is the most efficient of all three. The details of these ARQ schemes can be found in [8].

Go Back N and Selective Repeat Protocol require full duplex links. In [29], a Go Back N ARQ method is employed by assigning dedicated frequency bands for receiving and transmitting. Dividing the limited bandwidth in such a manner for frequency division duplexing decreases the data rate of the system.

E. Routing

There are two basic methods used for routing packets through an information network: *virtual circuit* routing and *datagram* routing. Networks using virtual circuits decide on the path of the communication at the beginning of a transaction. In datagram switching, each node that is involved in the transaction makes a routing decision, which is to determine the next hop of the packet.

Many of the routing methods are based on the *shortest path* algorithm. In this method, each link in the network is assigned a cost which is a function of the physical distance and congestion. The algorithm tries to find the shortest path, i.e., the path with the lowest cost, from a source node to a destination node. In a distributed implementation, each node determines the cost of sending a data packet to its neighbors and shares this information with the other nodes of the network. In this way, every node maintains a data base which reflects the cost of possible routes.

For routing, we consider the most general problem where network nodes are allowed to move. This situation can be viewed as an underwater network with both fixed ocean-bottom sensors and AUV's. The instruments temporarily form a network without the aid of any preexisting infrastructure. These types of networks are called *ad hoc* networks [23].

In *ad hoc* networks, the main problem is to obtain the most recent state of each individual link in the network, so as to decide on the best route for a packet. However, if the communication media is highly variable as in the shallow-water acoustic channel, the number of routing updates can be very high. Current research on routing focuses on reducing the overhead added by routing messages while at the same time finding the best path, which are two conflicting requirements. In a recent paper [27], the authors compared four *ad hoc* network routing protocols presented in the literature:

- destination sequence distance vector (DSDV) [24];
- temporally ordered routing algorithm (TORA) [25];
- dynamic source routing (DSR) [26];
- *ad hoc* on-demand distance vector (AODV) [28].

DSDV maintains a list of *next hops* for each destination node which belongs to the shortest distance route. The protocol requires each node to periodically broadcast routing updates to maintain routing tables.

TORA is a distributed routing algorithm. The routes are discovered on demand. This protocol can provide multiple routes to a destination very quickly. The route optimality is considered as a second priority and the routing overhead is reduced. DSR employs source routing, that is, the route of each packet is included in its header. Each intermediate node that receives the packet checks the header for the next hop and forwards the packet. This eliminates the need for intermediate nodes to maintain the best routing information to route the packets.

AODV uses the on-demand route discovery as in TORA and has the maintenance characteristic of DSR, and employs them in a hop-by-hop routing scheme instead of source routing. Also, periodic updates are used in this protocol.

In a mobile radio environment, DSR provides the best performance in terms of reliability, routing overhead, and path optimality [27]. The effect of long propagation delays and channel asymmetries caused by power control are issues that need to be addressed when considering application of these network routing protocols to UWA channels.

F. Development of UWA Networks

There are only a few UWA network applications reported in the literature. A deep-water acoustic local area network (ALAN) was deployed in Monterey Canyon, CA, for long-term data acquisition and ocean monitoring from multiple ocean-bottom sources [1]. The network was designed to transfer data from ocean-bottom nodes to a surface-deployed station and command and control signals in the reverse direction. Whenever a node collects a predetermined amount of data, it wakes up and sends a request to the surface receiver together with the size of the data it will transmit. Upon the reception of the request, the surface receiver schedules a transmission time that depends on the round-trip propagation delay and notifies the source node with an ACK. The source node immediately sends the data packet when it receives the ACK. Request, ACK, and data transmissions are carried out in different frequency bands. If a collision occurs on the request channel, the receiver tries to unscramble the collision by use of a multiuser structure. This increases the throughput of the network by eliminating the need for retransmissions. Because this protocol relies on correct estimation of the round-trip propagation times, any error in the estimation process decreases the throughput of the system by causing retransmissions or unnecessarily prolonged idle times. However, as the alternative to this solution is to fix the time between transmissions to the maximum round-trip propagation time, it is still advantageous.

A store-and-forward protocol was proposed in [29] to be employed in shallow-water ALAN's. The protocol is a modified version of the packet radio network (PRN) protocol [30] that matches the shallow-water acoustic channel characteristics. Different from PRN, each node in the network uses three separate channels (frequency bands) to transmit, to receive packets from its predecessor, and to receive ACK's from its successor. In this way, effective employment of a Go Back N ARQ scheme is enabled. The packets are routed using dynamically established paths, such that the nodes associated with a transaction cannot be used for other transactions simultaneously.

A peer-to-peer communication protocol has been developed to control AUV's [31]. The media access protocol of the network is CSMA. When a message becomes available for transmission at a node, the node begins to listen to the channel. When the channel becomes idle, it waits for an amount of time, called transmit delay, that is proportional to the maximum round-trip propagation time and sends its data packet. If it cannot receive an ACK from the destination, the source resends the packet after a time-out period calculated adaptively by using the maximum round-trip propagation time and priority of the source. Due to the extra idle times between transmissions, this protocol has a low throughput. On the other hand, it is highly reliable.

III. Design Example: A Shallow-Water Network

An underwater acoustic communication network is considered which consists of a large number of nodes operating in a shallow-water environment at a depth of approximately 50–100 m. The nodes are mounted on the bottom and separated by distances of up to 10 km. No special geometry of node placement is assumed, allowing for applications which require rapid deployment in unknown areas.

The nodes are equipped with sensors capable of detecting the presence of various objects or activity in the area. The information gathered by the sensors has to be transmitted to the base station, or the master node which is also located on the ocean bottom. The nodes have acoustic communication modems that can transmit and receive information. The total number of nodes is not restricted, with approximately several tens of nodes assigned per master node. The master node is connected to a surface buoy, from which a radio link is implemented to shore. The nodes are connected to the master node in a hierarchical manner. The number of hops that is required for a sensor node to communicate with the master node determines the level of the node. For example, a node that communicates directly to the master node is a first-level node. A node that must communicate via a first-level node in order to get to the master node is a second-level node.

The desired information transmission rate in the network is 100 bit/s from each node. The available frequency band is 8–15 kHz. The network is asynchronous, and the nodes transmit information when it becomes available. As in the majority of acoustic communication systems, the uncertainty about propagation delays prevents the design of a strictly synchronous system. The nodes transmit information in packets whose size is 256 bit. At most, a node is allowed to transmit five packets per hour; however, such a high duty cycle is unlikely. The transmission mode is half-duplex.

The network is constrained by energy consumption, because the acoustic modems are battery powered, as are the sensor processors. To conserve energy, the nodes must have two modes of operation: a sleep mode and an active mode. In the sleep mode, a node consumes much less energy than in the active mode. A node is awakened from sleep acoustically, that is, even when asleep, each node continually looks for a low-power wakeup signal. Upon detecting this signal, the node automatically powers up.

The goal of the network design is to achieve a good tradeoff between the quality of service and the energy consumption. In particular, a good quality of service includes maximum information throughput at minimum delay, as well as network reliability in the sense of its capability to adapt to channel outages and node failures.

A. Network Topology and Energy Consumption

One of the network design goals is to minimize the energy consumption while providing reliable connectivity between the nodes in the network and the master node. To assess the tradeoffs involved in such a design, we analyze the large-scale system and channel parameters that determine the energy consumption of the network.

Assuming that any number of connections can be supported by a chosen multiple-access strategy, we want to determine the energy consumption as a function of network topology, i.e., as a function of the connections established between the network nodes. As it was pointed out earlier, due to the nature of the application, the actual network geometry is such that node locations cannot be determined precisely. On the other hand, it is reasonable to assume in the first approximation that a uniform node distribution provides best area coverage.

To quantify the energy consumption of the network, we investigate a simplified scenario in which N nodes and a master node are arranged linearly along a stretch of length r. The nodes are uniformly distributed so that the distance between each two adjacent nodes, including the distance from the base station to the node closest to it, is r/N. Two extreme communication strategies are possible in this scenario. In the first strategy, each node has a direct access to the master node (fully connected topology). In the second strategy, each node transmits only to its nearest neighbor, who then relays the information toward the master node (multihop peer-to-peer topology). The energy consumption for both of these cases is determined below.

To determine the energy needed for transmission of a single data packet of duration T_P from a node at distance r to the master node, over N relay hops, we assume that a required quality of reception is achieved if the received power level is P_0. To achieve a power level P_0 at the input to the receiver at distance x, the transmitter power needs to be $P_0 A(x)$, where $A(x)$ is the attenuation. The attenuation is given as [3]

$$A(x) = x^k a^x \tag{1}$$

where k is the energy spreading factor (k is 1 for cylindrical, 1.5 for practical, and 2 for spherical spreading), and

$$a = 10^{\alpha(f)/10} \tag{2}$$

is a frequency-dependent term obtained from the absorption coefficient $\alpha(f)$. The absorption coefficient for the frequency range of interest is calculated according to Thorp's expression [3] as

$$\alpha(f) = 0.11\frac{f^2}{1+f^2} + 44\frac{f^2}{4100+f^2} + 2.75\cdot 10^{-4} f^2 + 0.003 \tag{3}$$

in dB/km for f in kHz.

To transmit a data packet from one node to another over a distance r/N, each node needs to transmit at a power level $P_1 = P_0 A(r/N)$, and each needs to transmit for the duration of a packet T_P. Hence, the total consumed energy for transmission over N hops is

$$E = NP_1 T_P = P_0 T_P N A(r/N). \tag{4}$$

When each of the N nodes has a packet to transmit, the total consumed energy for packet relaying is

$$\begin{aligned} E_{\text{rel}} &= P_0 A(r/N) T_P + P_0 A(r/N) 2T_P + \cdots \\ &\quad + P_0 A(r/N) N T_P \\ &= P_0 A(r/N) T_P N(N+1)/2. \end{aligned} \tag{5}$$

For a direct-access strategy, the total consumed energy is

$$\begin{aligned} E_{\text{dir}} &= P_0 A(r/N) T_P + P_0 A(2r/N) T_P + \cdots \\ &\quad + P_0 A(Nr/N) T_P \\ &= P_0 T_P \sum_{i=1}^{N} A(ir/N). \end{aligned} \tag{6}$$

Fig. 2 shows total consumed energy for both strategies. Dashed curves represent the case of direct access, which obviously requires more energy. For direct access, inclusion of each additional node results in an increase in total energy. For relaying, the situation is reversed: inclusion of each additional node decreases the total energy consumption because the additional node serves as an additional relay along the same distance r. The price to pay, of course, is the need for a sophisticated communications protocol and the increased packet delay.

Hence, the strategy which minimizes energy consumption is relaying. The savings in energy are greater for a greater number of relay hops and become more pronounced at higher distances. Specifically, for the distances on the order of several tens of kilometers which are of interest to the present design, the savings are significant. Consequently, a network topology which implements connections only between the neighboring nodes is considered.

B. Media Access Protocol

The multiple access strategy is chosen as either FDMA or CDMA. In the case of FDMA, acoustic modems that employ Hadamard coded MFSK modulation are used. Details of this modem design can be found in [32]. As an alternative to FDMA, direct sequence spread-spectrum signaling and CDMA is also considered for the network. A Rake receiver is employed at the receiver to make use of the multipath propagation of the shallow-water acoustic channel. The modem design is given in [33].

The media access protocol for the shallow-water network is based on the MACA protocol, which uses RTS-CTS-DATA exchange. The network employs the Stop & Wait ARQ scheme. If the source cannot receive a CTS from the destination after a predetermined time interval, it repeats RTS. If, after K trials of RTS, the source cannot receive a CTS, it decides that the link is no longer available and returns to a low-power state. If the source receives a CTS, it immediately transmits the data packet. An ACK signal is sent by the destination upon the receipt of a correct data packet to provide positive acknowledgment to the source in the data link layer.

During a transaction, all other transmission requests are declined. If the source of the transmission that has been declined is not informed, it repeats its request. In doing so, the source increases its power as dictated by the power control algorithm. This situation results in increased battery consumption and increased probability of collisions. To prevent these undesirable effects, a WAIT command is added. This command informs the source node that the destination is busy and will send a CTS as soon as possible.

If two nodes send an RTS to each other, a deadlock may occur because both nodes will issue the WAIT command when they receive RTS's from each other. Each node will then wait forever for the other node to send a CTS. This problem is solved by assigning priority to the packets that are directed toward the master node, as explained below.

Assume that node A is a lower level node than node B, that is, A is closer to the master and, thus, the parent of B. Node A and node B both send RTS's to each other. Due to transmission

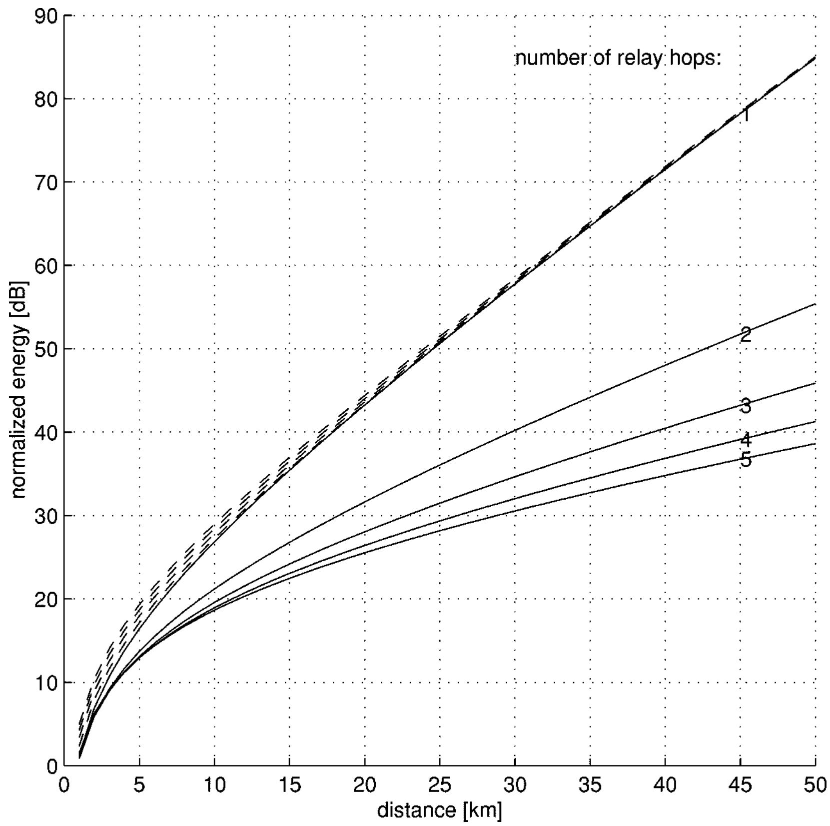


Fig. 2. Total (normalized) energy needed to transmit a packet from each of N nodes to the master node. Solid curves represent relaying $(E_{\rm rel}/P_0T_P)$; dashed curves represent direct access $(E_{\rm dir}/P_0T_P)$. The parameter on the curves is the number of hops. For relaying, the number of hops increases from the top curve downwards: although more packets are sent when there are more hops, total energy consumption is lower. For direct access, there is little difference between the curves, and the situation is reversed: the number of hops increases from 1 for the lowest energy consumption to 5 for the highest.

delays, packets arrive at their destinations while both nodes are waiting for a CTS packet. When node B receives the RTS, it notices that the packet is from its own destination, node A. Node B checks whether node A is its parent or child. By that time, node A receives the RTS of node B and does the same check. Because node B is its child, node A decides that it should wait for node B to complete its data transmission. Therefore, node A sends a CTS immediately, puts its own data packet into a queue, and waits for the CTS packet of node B.

C. Initialization and Routing

Since the network in consideration is an *ad hoc* network, an initialization algorithm is needed to establish preliminary connections autonomously. This algorithm is based on polling, and as such it guarantees connectivity to all the nodes that are acoustically reachable by at least one of their nearest neighbors. During initialization, the nodes create *neighbor tables.* These tables contain a list of each node's neighbors and a quality measure of their link, which can be the received SNR from the corresponding neighbor. The neighbor tables are then collected by the master node and a routing tree is formed. The master node decides on the primary (and secondary) routes to each destination. Initialization ends when the master node sends primary routes to the nodes. The initialization algorithm is detailed in [34]. It provides either a single set of connections or multiple connections between the nodes. Multiple connections are desirable to provide greater robustness to failures.

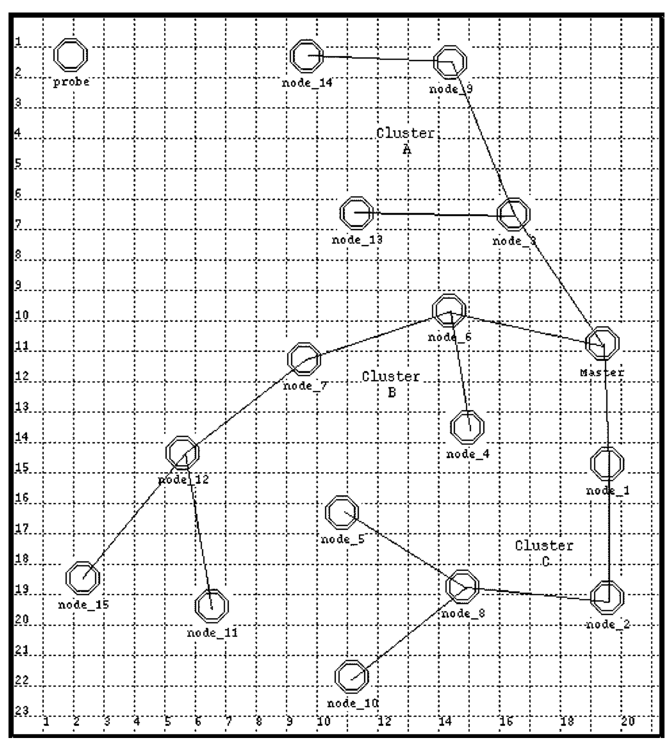


Fig. 3. The network consists of one master, fifteen sensors, and a probe node.

The performance of acoustic links between nodes can degrade, and even a link can be permanently lost due to a node

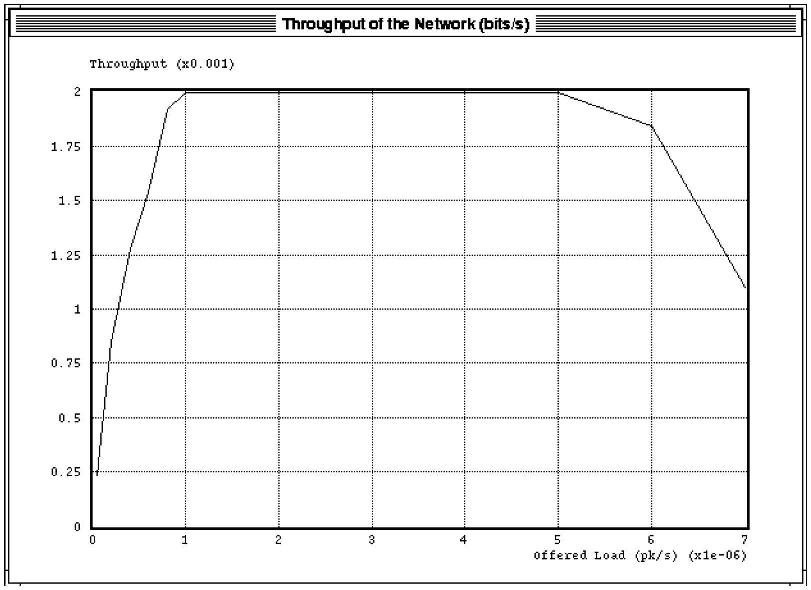


Fig. 4. Throughput of the network with respect to the offered load.

failure. In such cases, the network should be able to adapt itself to the changing conditions without interrupting the packet transfer. This robustness can be obtained by using an adaptive routing algorithm.

In the current design, the master node creates a routing tree depending on the neighbor tables reported by its nodes. If a node reports that a link's performance has degraded or it is no longer available, the master node selects new routes that take the place of the failed link. The changes in the routing tree are reported to all related nodes. This procedure ensures that nodes will not attempt to use a failed link. In this way, unnecessary transmissions that increase battery consumption are avoided.

D. Results

The network performance was tested using Opnet's Radio Modeler tool. We modified the standard radio channel model of Opnet to match the characteristics of the shallow-water acoustic channel. These characteristics include the transmission loss as a function of distance and frequency, and the probability of error as a function of the SNR. The data packet generation process is modeled as a Poisson process. Although this model does not reflect the real situation where data generation in different nodes can be dependent, it provides a way for performance evaluation.

Network nodes use FDMA-based modems. Although potentially inferior to CDMA in a number of UWA channels, this technique is considered for reasons of practical interest. Namely, the network design used in simulation is intended to support an ongoing experiment in which a number of acoustic modems that use an M-ary FSK modulation scheme are deployed in a shallow-water area. In this experiment, the available frequency band is divided into subbands and each subband is assigned to a cluster of nodes. A cluster of nodes is deployed in the same general geographical region. The neighboring clusters are assigned different frequency bands to assure low interference. Each cluster communicates with the master node through its first-level node.

Fig. 3 shows the network topology created using the network editor of Opnet. The network consists of a single master node and fifteen sensor nodes. Sensor nodes are divided into three clusters. The master node uses frequency group M. Sensor nodes are assigned frequency groups A, B, and C, as shown in the figure. The lines connecting the nodes represent the virtual communication paths created by the acoustic modems. There is also a probe node in the network that is used for simulation purposes. The probe node initializes the random number generators and records statistics at the end of the simulation. The details of the simulation model can be found in [35].

Using simulations, we have obtained the throughput and packet delay characteristics of the proposed network for data packets 256 bit in length. Throughput is defined as the number of successfully transmitted data packets per unit time. Packet delay is calculated by averaging time passed from the time a data packet is generated and when the packet is received by the final destination.

Fig. 4 shows the change in the average throughput of the system as a function of offered load (the number of packets per second that are offered to the network to transmit). The throughput of the system increases until the offered load reaches 10^{-6} packets per second. At this point, the system reaches saturation and throughput remains constant. If we further increase the offered load the throughput of the system begins to decrease. This is due to the fact that we do not use dynamic routing. If a source node cannot complete a transaction due to congestion, it drops the packet. The throughput of the system is so small because of the low data rate and high propagation delay of the underwater channel.

Fig. 5 gives the end-to-end packet delay of the system. As expected, the end-to-end delay experienced by the information sequences increases with increasing load. However, between 10^{-6} and $6 \cdot 10^{-6}$ end-to-end delay stays constant, and for offered

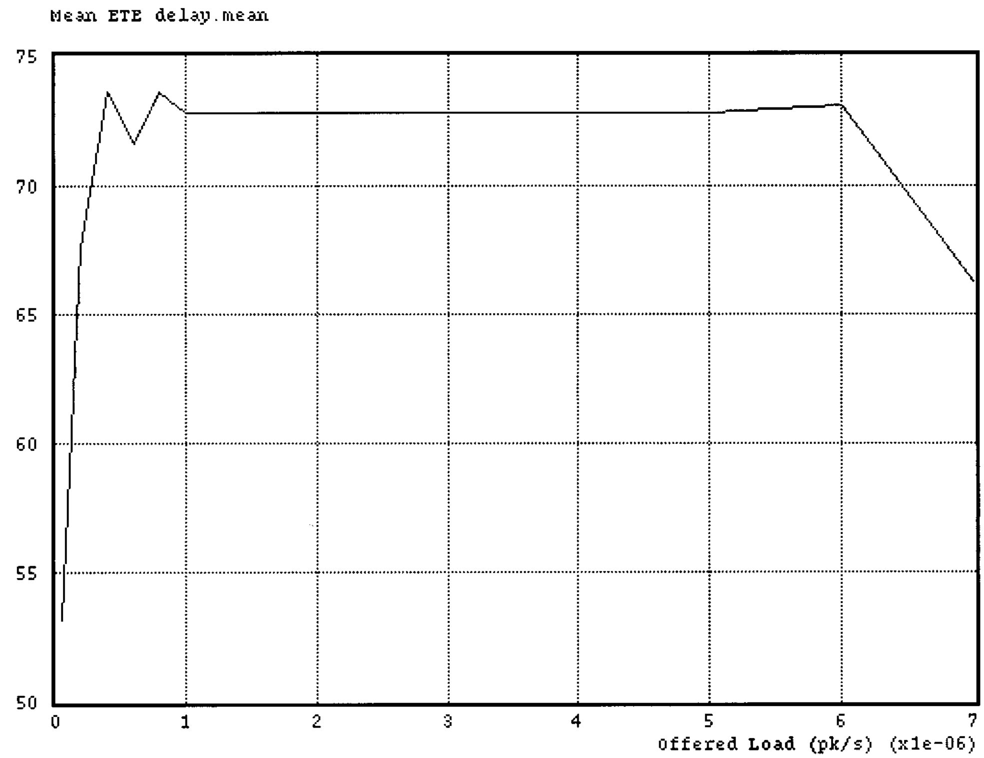


Fig. 5. Variation of end-to-end delay experienced by the layer 3 information sequences.

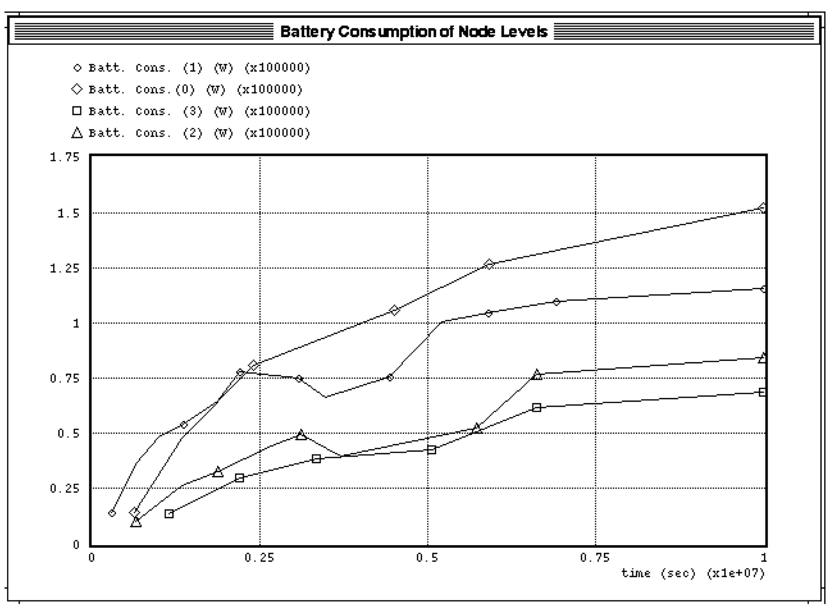


Fig. 6. Battery consumption of the master node and average battery consumption of layer 1, 2, and 3 nodes as a function of time.

load greater than $6 \cdot 10^{-6}$ it begins to decrease. This is because of the highly congested higher level nodes. Due to congestion, high-level nodes cannot complete their transaction. However, the first-level nodes continue to communicate with the master node, and the only completed transaction occurs between first-level nodes and the master node. Also, the end-to-end delay does not include the lost packets.

Fig. 6 presents the battery consumption of the master node, and the average battery consumption of level 1, level 2, and level 3 nodes as a function of time. The battery consumption decreases for higher level nodes, which is due to the fact that lower level nodes carry more traffic.

The network design described above is part of a larger network development study. The study focuses on various multiple-access techniques and various network protocols. The current stage in the network development is at-sea testing of the FDMA-based network. A series of experiments under the name of SEAWEB is scheduled for the next two years. These experiments will focus on network designs of increased complexity, with the goal of conducting a feasibility analysis of various networking techniques and determining those techniques that are best suited for use in the shallow-water channels.

IV. Concluding Remarks and Future Directions

In this paper, the focus was on the problem of designing reliable UWA networks that are capable of transferring data from a variety of sensors to on-shore facilities. Major impediments to the design of such networks were considered, which are: 1) severe power limitations imposed by battery power; 2) severe bandwidth limitations; and 3) channel characteristics including long propagation times, multipath, and signal fading. Various multiple-access methods, network protocols, and routing algorithms were also considered.

Of the multiple-access methods considered, it appears that CDMA, achieved either by frequency hopping or by direct sequence, provides the most robust method for the underwater network environment. Currently under development are modems that utilize these types of spread-spectrum signals to provide the multiple access capability to the various nodes in the network. Simultaneously with current modem development, there are several investigations on the design of routing algorithms and network protocols.

The design example of the shallow-water network served to illustrate the design methodology that embodies the power and bandwidth constraints. Experimental data that will be collected over the next two to three years will be used to assess the performance of the network and possibly validate a number of assumptions and tradeoffs included in the design. Over the next decade, significant improvements are anticipated in the design and performance of UWA networks as more experience is gained through at-sea experiments and network simulation tools.

References

[1] J. Catipovic, D. Brady, and S. Etchemendy, "Development of underwater acoustic modems and networks," *Oceanography*, vol. 6, pp. 112–119, Mar. 1993.

[2] J. Catipovic, L. Freitag, and S. Merriam, "Underwater acoustic local area network for ROV and instrument communications," in *Proc. AUVS'91*, Washington, D.C., Aug. 1991, pp. 447–460.

[3] L. Berkhovskikh and Y. Lysanov, *Fundamentals of Ocean Acoustics*. New York: Springer, 1982.

[4] A. Quazi and W. Konrad, "Underwater acoustic communications," *IEEE Commun. Mag.*, pp. 24–29, Mar. 1982.

[5] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol. 21, pp. 125–136, Apr. 1996.

[6] T. S. Rappaport, *Wireless Communications*. Englewood Cliffs, NJ: Prentice Hall, 1996.

[7] A. Tannenbaum, *Computer Networks*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1996.

[8] D. Bertsekas and R. Gallager, *Data Networks*, NJ: Prentice Hall, 1992.

[9] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*, New York: Wiley, 1995.

[10] D. D. Falconer, F. Adachi, and B. Gudmundson, "Time division multiple access methods for wireless personal communications," *IEEE Commun. Mag.*, pp. 33–36, Jan. 1995.

[11] R. Kohno, R. Meidan, and L. B. Milstein, "Spread spectrum methods for wireless communications," *IEEE Commun. Mag.*, pp. 33–36, Jan. 1995.

[12] A. C. Chen, "Overview of code division multiple access technology for wireless communications," in *Proc. IECON'98*, 1998, pp. T15–T24.

[13] D. Brady and J. A. Catipovic, "Adaptive multiuser detection for underwater acoustic channels," *IEEE J. Oceanic Eng.*, vol. 19, pp. 158–165, Feb. 1994.

[14] Z. Zvonar, D. Brady, and J. A. Catipovic, "Adaptive decentralized linear multiuser receiver for deep water acoustic telemetry," *J. Acoust. Soc. Amer.*, pp. 2384–2387, Apr. 1997.

[15] M. Stojanovic and Z. Zvonar, "Multichannel processing of broadband multiuser communication signals in shallow water acoustic channels," *IEEE J. Oceanic Eng.*, vol. 21, pp. 156–166, Apr. 1996.

[16] L. G. Roberts, "ALOHA packet system with and without slots and capture," *ACMSIGCOM Comput. Commun. Rev.*, pp. 28–42, 1975.

[17] L. Kleinrock and F. A. Tobagi, "Carrier sense multiple access for packet switched radio channels," in *Proc. ICC'74*, June 1974, pp. 21B-1–21B-7.

[18] ——, "Packet switching in radio channels, part I: Carrier sense multiple access modes and their throughput-delay characteristics," *IEEE Trans. Commun.*, pp. 1400–1416, 1975.

[19] F. A. Tobagi and L. Kleinrock, "Packet switching in radio channels, part II: The hidden terminal problem in carrier sense multiple access and busy tone solution," *IEEE Trans. Commun.*, pp. 1417–1433, 1975.

[20] H. Takagi and L. Kleinrock, "Correction to 'throughput analysis for persistent CSMA systems'," *IEEE Trans. Commun.*, pp. 243–245, 1987.

[21] V. Bharghavan, A. Deers, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," *ACMSIGCOMM*, pp. 212–225, Aug. 1994.

[22] P. Karn, "MACA—A new channel access method for packet radio," in *ARRL/CRRL Amateur Radio 9th Computer Network Conf.*, Sept. 1990.

[23] D. B. Johnson, "Routing in ad hoc networks of mobile hosts," in *Workshop on Mobile Computing and Applications*, Dec. 1994, pp. 159–163.

[24] C. E. Perkins and P. Bhagwat, "Highly dynamic destination sequence distance vector routing (DSDV) for mobile computers," in *Proc. SIGCOMM'94*, Aug. 1994, pp. 234–244.

[25] V. D. Park and M. S. Corson, "A highly adaptive distributed routing algorithm for mobile wireless networks," in *Proc. INFOCOM'97*, Apr. 1997, pp. 1405–1413.

[26] D. B. Johnson and D. A. Maltz, "Protocols for adaptive wireless and mobile networking," *IEEE Personal Commun.*, 1996.

[27] J. Broch, D. A. Maltz, D. B. Johnson, Y. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *ACM/IEEE Int. Conf. Mobile Computing and Networking*, Oct. 1998.

[28] C. E. Perkins, "Ad hoc on demand distance vector (AODV) routing,", Internet-Draft, draft-ietf-manet-aodv-00.txt, 1997.

[29] J. L. Talavage, T. E. Thiel, and D. Brady, "An efficient store-and-forward protocol for a shallow-water acoustic local area network," in *Proc. OCEANS'94*, Brest, France, Sept. 1994, pp. I883–I888.

[30] J. Jubin and J. D. Tornow, "The DARPA packet radio network protocols," in *Proc. IEEE*, vol. 75, Jan. 1987, pp. 21–32.

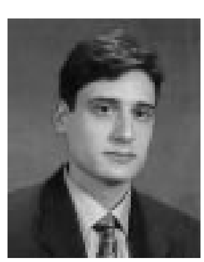
[31] S. M. Smith and J. C. Park, "A peer-to-peer communication protocol for underwater acoustic communication," in *OCEANS'97*, Oct. 1997, pp. 268–272.

[32] K. F. Scussel, J. A. Rice, and S. Merriam, "New MFSK acoustic modem for operation in adverse underwater channels," in *Proc. Oceans Conf.*, Halifax, NS, Canada, Oct. 1997, pp. 247–254.

[33] E. M. Sozer *et al.*, "Direct sequence spread spectrum based modem for under water acoustic communication and channel measurements," in Proc. OCEANS'99, Nov. 1999, to appear in, to be published.

[34] J. G. Proakis, M. Stojanovic, and J. A. Rice, "Design of a communication network for shallow water acoustic modems," in *Proc. Oceans Comm. Conf.*, MD, Nov. 1998.

[35] E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Design and simulation of an underwater acoustic local area network," in *OPNETWORK'99*, Washington, D.C., Aug. 1999.



Ethem M. Sozer received the B.S. and M.S. degrees from Middle East Technical University, Ankara, Turkey, in 1994 and 1997, respectively. He is currently working toward the Ph.D. degree in electrical engineering at Northeastern University, Boston, MA.

His research interests include underwater acoustic communications, data networks, and cellular communications.

Milica Stojanovic received the Dipl. Ing. degree in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Boston, MA, in 1991 and 1993, respectively.

She has been a Post-Doctoral Scholar at the Woods Hole Oceanographic Institution, Woods Hole, MA, where she is now a Guest Investigator. She has also been a Visiting Assistant Professor at Northeastern University. Her research interests include digital communications theory and adaptive signal processing, with related problems in mobile radio and underwater acoustic communications.

John G. Proakis (S'58–M'62–SM'82–F'84–LF'97) received the B.S.E.E. degree from the University of Cincinnati, Cincinnati, OH, in 1959, the S.M. degree from the Massachusetts Institute of Technology (MIT), Cambridge, in 1961, and the Ph.D. degree in engineering from Harvard University, Cambridge, MA, in 1966.

He was a Staff Member at the MIT Lincoln Laboratory, Lexington, MA, from 1961 to 1963, and a Member of the Technical Staff at GTE from 1966 to 1969. Since September 1969, he has been on the faculty of the Electrical and Computer Engineering Department, Northeastern University, Boston, MA, where he held the position of Department Chair (1984–1997), Associate Dean and Director of the Graduate School of Engineering (1982–1984), and Acting Dean (1992–1993). His professional experience and interests are in the general area of digital communications and digital signal processing. He is the author of the book *Digital Communications* (New York: McGraw-Hill, 3rd ed., 1995) and co-author of *Introduction to Digital Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall, 3rd ed., 1996), *Digital Signal Processing Laboratory* (Englewood Cliffs, NJ: Prentice-Hall, 1991), *Advanced Digital Signal Processing* (New York: Macmillan, 1992), *Digital Processing of Speech Signals* (New York: Macmillan, 1992), *Communication Systems Engineering* (Englewood Cliffs, NJ: Prentice-Hall, 1994), *Digital Signal Processing Using MATLAB* (Boston, MA: PWS, 1997), and *Contemporary Communication Systems Using MATLAB* (Boston, MA: PWS, 1998).