

A Multi-Band Random Access Messaging Protocol

Nick Van Stralen, Orhan Imer,
Robert J. Mitchell, Scott C. Evans
GE Global Research

One Research Circle, Niskayuna, NY 12309
{vanstran, imer, evans, rjmitchell}@research.ge.com

Suresh Iyer

Lockheed Martin Integrated System Solutions**
700 N. Frederick Avenue
Gaithersburg, MD 20879
suresh.iyer@lmco.com

ABSTRACT

Media access protocols can be classified as either demand assigned or random access. Demand assigned protocols are often inflexible, inefficient under diverse traffic, and do not scale well to large networks. Random access schemes usually suffer from collisions and cannot manage the multi-frequency capabilities of many radios. We present new results on our pseudorandom message scheduling approach. The protocol has been extended to manage multiple frequency channels. In dense networks, sub-networking can be performed. Nodes are partitioned a priori and are assigned different reference frequencies for their discovery/broadcast segments. We demonstrate in networks where nodes use transmission based start times along with conventional carrier sense channel access, collision free communications.

I. INTRODUCTION

Different media access protocols have emerged over the years. The goal of all media access protocols is to share access to a limited resource in as fair a manner as possible which achieving good performance under different traffic conditions. Two types of media access protocols are demand assigned and random access. Demand assigned protocols are often inflexible, and inefficient under diverse traffic conditions, and do not scale well to large networks. Random access schemes usually suffer from collisions and cannot manage the multi-frequency capabilities of many radios.

In this paper, we present new results on our pseudorandom message scheduling approach known as the Connectionless Scheduling Protocol (CSP) [1]. With this protocol, a number of global message schedules are established. Each schedule instance denotes a message start opportunity in time. Each node is assigned unique message start schedules, which can be either transmission or reception based.

This paradigm has been extended to manage multiple frequency channels. This is done by assigning receivers a receive frequency allocation based on their unique identifier. If several frequency bands are available, then any individual node will change its receive band according to a known and predetermined algorithm.

An optional RTS/CTS along with our boomerang acknowledgement concept improves link reliability in situations where hidden nodes are present. In dense networks, sub-networking can be performed as a means of improving overall network performance. Nodes are partitioned into subnets a priori and are assigned different reference frequencies for their discovery/broadcast segments. We demonstrate the subnet concept in networks where nodes use transmission based start times along with conventional carrier sense channel access, collision free communications.

A. Previous Work

The Adaptive Connectionless Scheduling Protocol (CSP) is a random access media access protocol developed by GE for DARPA. As described in [1] it provides an approach to minimizing energy costs in wireless sensor networks through a protocol that implements near perfect scheduling based on a near perfect clock and some discriminating node feature. The Adaptive CSP pseudorandom scheduling paradigm provides flexibility to serve data rate requirements over a wide range of traffic scenarios and automatically adapts to load and latency requirements, providing a high performance scheduling protocol for sensor networks and homeland security applications

II. TRANSMISSION-BASED CSP

The Adaptive CSP was developed initially in order to enable low power communications for flexible topologies and traffic patterns. However, key innovations and advantages of the Adaptive CSP protocol, including Message Start Time (MST) based channel access, pseudo-random schedule of MST based on node ID or location, and the ability to designate based on protocol mode whether the pre-determined schedules are start times for message transmission or reception makes this paradigm attractive for non energy sensitive cases as well. The channel access schedules that can be readily understood with minimal overhead – basically only node discovery is required – promote low collision and optimal, fair channel utilization at low latency as we analyze in this document. Additionally, several extensions to the original Adaptive CSP design are proposed to tune it effectively for tactical networks of software-defined radios.

A. Throughput of transmission based CSP

The performance of transmission based CSP was assessed through simulation and analysis. A discrete event

Lockheed Martin funded this work under a dual use shared vision agreement between GE Research and Lockheed Martin.

simulation of several MAC protocols was implemented. The simulation generated packet arrivals according to a Poisson arrival process. The packets were delivered to one of several media access protocols under test. The MAC protocol under test decides when the packet is to be queued and when it is to be transmitted. When the packet is transmitted, the channel is analyzed and a collision determination is made. A collision is recorded if another node was transmitting. Collision checks were made at both the beginning and the end of the packet. With the known Poisson arrival rate and the collision statistics, a channel throughput calculation can be made. For some of the known MAC protocols, the throughput results are compared to prior published results. Throughput performance of CSP is compared to several classic random access media access schemes in Figure 1.

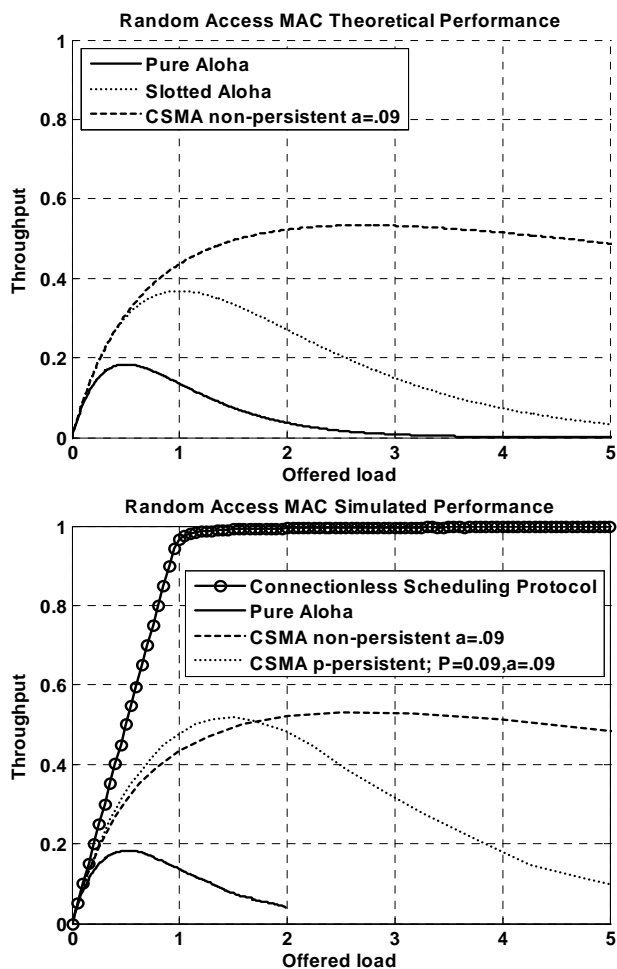


Figure 1: Throughput performance of CSP.

In both plots the axes are normalized to the maximum channel data rate. The figures show that the classic media access schemes offer throughput of only up to 50% of available channel throughput. The Connectionless Scheduling Protocol achieves near ideal performance. When the offered load is less than the available channel throughput, all offered load is communicated without

collision. When the offered load is more than the available channel throughput, then nearly 100 percent of channel throughput is utilized, again without collision.

Demand assigned reservation schemes such as TDMA and FDMA and others [2,3,4] are capable of achieving high channel utilization. These schemes achieve high channel utilization by reserving specific portions of the channel bandwidth to individual nodes. Under heavy load conditions, (when each node in the network always has data to send,) these schemes also exhibit an offered load/throughput relationship similar to the Connectionless Scheduling Protocol. All demand assigned reservation schemes generally suffer either high latency or high overhead. These problems are exacerbated when the number of nodes is high or when the channel traffic is dynamic.

B. Latency of transmission based CSP

We also performed a preliminary latency analysis of the Connectionless Scheduling Protocol. This analysis was performed through both analysis and simulation. The simulation, which generated the throughput performance results, was modified to record the packet arrival time and the packet delivery time. At each packet arrival, the latency was calculated. Figure 2 shows the average simulated packet latency of the Connectionless Scheduling Protocol.

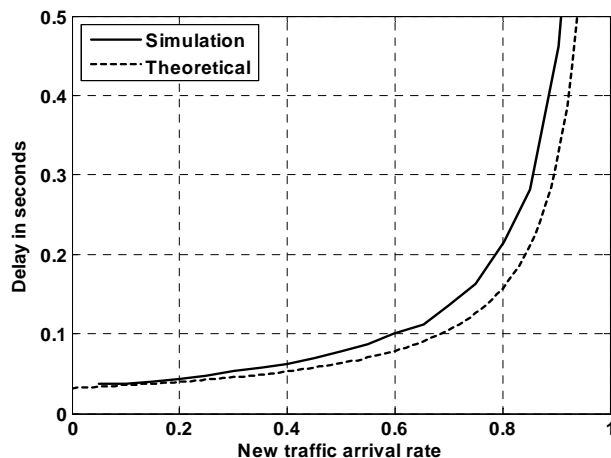


Figure 2: Latency performance of CSP

The minimum latency (at zero channel load) is due to the scheduling. A new arrival is scheduled for transmission in the first available transmit slot for that node. The average wait time for the slot is the inverse of the number of transmits slots per second. In the simulation, we allocated 32 slots per second, resulting in approximate no load latency of 0.033 seconds. The latency can be further reduced by increasing the number of transmit start times available to each node.

III. MULTIPLE FREQUENCY BANDS

Extending CSP to manage multiple frequency channels is done by assigning receivers a receive frequency allocation based on their unique identifier. If several frequency bands are available, then any individual node will change its receive band according to a known and predetermined algorithm. Other nodes that wish to pass a message to it will begin transmitting the message at their unique message start time using the frequency channel of the target receiver. The assigned receive channels of each node will be changed in a pseudo-random way, so that the effect of high bandwidth flows (such as video) will be distributed among all nodes in the network and not only on the nodes assigned to a particular receiver frequency allocation.

The receive frequency channel assignment is changed at frame boundaries. A separate pseudo-random scheduler is called to select the receive channel assignment from the possible list of channel assignments. All nodes in the network or sub-net (see section on sub-netting) use the same pseudo-random function. When a source node has a message to transmit to some destination node, it schedules the transmission according to its own message start time. The source node will set a timer at this scheduled transmission time. At the appointed transmission time the source node determines the receive channel to which the target receiver is assigned. The source node will change its channel assignment to match the channel that the target is scheduled to listen. It will then begin a transaction on the target's assigned listening channel. Normally, the transaction begins by sensing activity on the channel and only initiating a transaction if the channel is free.

If the target's assigned listening channel is sensed free, the target will initiate a message transaction. The transaction may begin with an RTS (c.f. next section), or the data message may be transmitted. When the transaction is completed, the source transmitter returns to its designated listening channel. Of course, when the target node changes its channel, it stops listening to its designated listening channel for the duration of the transaction. Transactions targeted to it during that time would not be successful and would be retried.

Under this scheme, the source transmitter must be aware of the target node's identifier so that it can determine the target's listening channel. This is accomplished by maintaining a small lookup table, which associates potential target nodes IP addresses with their identifier. When a new message from the Network layer is received by the CSP protocol, the identifier is retrieved and used to determine the target node's listening channel.

IV. RTS/CTS FOR HIDDEN NODES

Hidden node resolution is addressed by a combination the RTS/CTS paradigm used in 802.11 ad hoc networks

and the boomerang acknowledgement concept discussed in [1]. When the RTS/CTS handshaking is enabled, the start of a transaction begins with an RTS message. The reception of RTS messages by nodes near to the source node will clear local transmitters. When handshaking is enabled, the start of a transaction would also require a CTS message, simultaneously informing the source node that a collision did not occur and clearing all nodes in its neighborhood. Both the RTS and the CTS message contain information relating the target and source node identifiers and the duration of the message. Nodes receiving either the RTS or CTS can then set an internal channel reservation timer, expiring at the end of the transaction. Any scheduled transmission event during this channel reservation would be rescheduled.

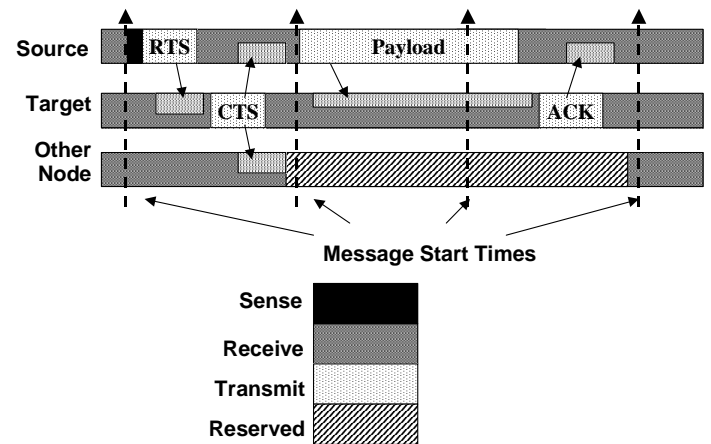


Figure 3: Exemplary Message Transaction using RTS/CTS/ DATA/ACK.

RTS packets initiate the transaction at the scheduled transmission start time of the target node. All subsequent packets involved in the transaction are sent immediately, i.e. they are not re-scheduled. Both the RTS and CTS messages contain information regarding the data packet length. All nodes hearing an RTS or CTS message set an internal channel reservation that expires at the end of the expected transmission. As shown in the Figure 3, the CTS packet is returned immediately after reception and processing. When RTS/CTS handshaking is enabled, the message start time intervals are sufficiently separated in time to allow for the handshaking. Since the transmitter's start time is uniquely determined by the CSP scheduling algorithm, and another transaction cannot begin until the next message start time, we guarantee that multiple RTS/CTS handshakes will not collide.

The end of the transaction would similarly have an end of packet/acknowledgement. The end of the transaction would then inform nodes at both the source and target destination that the channel is again free. When multiple frequency channels are employed, all message transmissions occur on the channel of the target receiver. That is, after the source transmitter sends the RTS, it listens for the corresponding CTS on the same channel. The final

acknowledgement would also be transmitted on the target receivers listening frequency.

V. BROADCAST MESSAGE SEPARATION

Many application layer packets and many other packets types in the network are intended for only a single target node. However, some applications and many network layer function require that packets be broadcast to several nodes or any available node. These two classes of packets present are difficult to treat under a single packet-handling framework at the MAC layer due to their different requirements.

In the previous section, we discussed the implementation of an RTS/CTS handshake in the CSP MAC protocol. The RTS/CTS paradigm will not work well with broadcast messages. If an RTS packet were transmitted for a broadcast packet, a CTS packet would be potentially generated for each node that received the RTS. These multiple packets could collide when received at the source node, causing confusion.

When a node has access to multiple channel frequencies, it is desired that several unicast messages be transmitted simultaneously on separate frequencies to increase the throughput performance. However for broadcast messages, it is most efficient to have all nodes on the same frequency when the message is transmitted. Transmitting on only one of the available frequencies would only be received by a subset of all nodes. Radiating on all frequencies simultaneously is wasteful of transmission power, and would often require hardware changes to enable transmission on multiple channels.

For these reasons, we have chosen to divide the message frame into a unicast message portion and a broadcast message segment. When a message is delivered to the CSP MAC, the message is checked to determine whether it is a unicast packet or a broadcast packet. The packet is handled differentially based on packet type. During the unicast portion of the frame, only unicast messages are communicated. The multi-frequency management and handshaking are performed as previously discussed.

Broadcast messages are scheduled in a separate portion of the frame denoted as the broadcast segment. Again, to promote fairness, messages are scheduled according to a pseudo random scheduling algorithm. The number of message start times during the broadcast segment and the number of start times available to each node during each broadcast segment is configurable. During the broadcast segment of the frame, all nodes have their receivers tuned to a common channel. By this means all nodes within range of a transmission receive the transacted broadcast messages. With broadcast message service differentiated and scheduled, we also limit the volume of messaging taken by overhead functions such as route determination.

VI. SUBNETS

Densely populated wireless adhoc networks create a challenging environment for achieving good network throughput when there is only a single radio channel available. In a demand-assigned protocol the amount of potentially wasted bandwidth increases while the opportunities to transmit decreases if slots are statically assigned. If slots are dynamically assigned the complexity of the assignment algorithm increases while the opportunities for a node to transmit decreases per unit time.

Using a random access scheme, as node density increases the number of nodes competing for the same bandwidth also increases thus limiting the average amount of data any one node can send. Using the RTS/CTS feature to solve the hidden node problem also exacerbates the problem by increasing the effective area within which nodes compete for the channel. Also, as the density of nodes increases the amount of broadcast traffic per unit area is also likely to increase. Multi-band systems are created to help mitigate this problem. These systems allow nodes to simultaneously transmit and/or receive on two or more non-interfering channels.

In a single channel environment, the CSP achieves fairness by providing multiple opportunities (i.e. schedules) for a node to transmit and to order these opportunities in a random fashion from frame to frame. Further, the CSP also partitions the time allocated within each frame between broadcast and unicast schedules. As node density increases there is a need to allocate more of the frame to broadcast traffic which ultimately limits the performance of unicast traffic.

In a multi-band or multi-channel system CSP can take full advantage of these features to improve overall network throughput. The CSP can take advantage of multiple bands using different forms of subnetting.

The objective of subnetting is to use the available channels to mitigate the need of increasing the size of the broadcast portion of a CSP frame and to provide a second dimension for increasing the opportunities for each node to transmit its data. Two methods of subnetting are considered in this paper.

A. Unicast And Broadcast Partitioning

The first method of subnetting is coupled with subnetting at the IP layer. With this method, nodes are divided into CSP subnets according to their IP subnet. All of the nodes in a given IP subnet are statically assigned the same set of channels. Router nodes will have multiple interfaces where each interface is assigned channels corresponding to its subnet. One of the channels assigned to each subnet can be configured in CSP as the broadcast channel for that subnet.

Nodes wishing to transmit are only competing with nodes in the same subnet. Thus data exchanges within the same subnet can take place in parallel with similar data exchanges in other subnets. Data exchanges between subnets must pass through the routers between the two subnets.

Since each subnet is assigned a unique broadcast channel, only the nodes in that subnet share the common broadcast segment of the CSP frame. Therefore, the amount of the CSP frame time allocated to broadcast traffic can be smaller because it is serving fewer nodes. Thus, more time can be allocated to the unicast part of the frame.

The available number of unicast and broadcast schedules per node can be larger per frame since fewer nodes compete in the same frame. This method of partitioning therefore increases the effective number of opportunities a node has to transmit by the number of subnets.

This type of a subnetting approach would work well if the traffic load was not well balanced across the network. That is, the channels could be assigned in such a manner that groups of nodes having higher demands could be assigned more channels and the overall performance of the network would be improved. Alternatively, more channels could be assigned to nodes with more sensitivity to jitter to achieve a more predictable latency.

B. Broadcast Only Partitioning

Another subnetting strategy that can be considered with CSP is to make every channel available to every node for unicast transmission and assign unique channels to different groups of nodes for broadcast transmission. In this configuration each node has a different set of unicast schedules on each different channel. If one channel is busy, the node will have another opportunity on a different channel before the next opportunity on the same channel.

The ability to assign a unique broadcast channel to each subnet also mitigates the problem of allocating more of the frame time to broadcast traffic as before and this increasing the number of opportunities available to transmit unicast traffic. Furthermore, since broadcast traffic in different subnets uses different channels it is possible for different subnets to transmit broadcast traffic simultaneously. Care would need to be taken in how bridge nodes between subnets are chosen to minimize the number of hops between any two nodes.

This approach would provide better performance if the network load were highly dynamic since it offers every node the maximum number of opportunities for sending either unicast or broadcast traffic while at the same time ensuring a great deal of fairness in terms of transmission opportunities across all of the nodes.

VII. RESULTS

To investigate the performance of CSP in a realistic network environment, we have developed an implementation of the CSP protocol in QualNet, a discrete-event based network simulator from SNT (Scalable Network Technologies.) In this section, we present results from CSP simulation studies conducted in QualNet.

We consider a network scenario where 51 nodes are placed uniformly on a 90m x 90m terrain. The nodes communicate over a single channel at 2.4Mhz with PHY data rate 2Mbps. Every node is assumed to be in the radio range of every other node, hence there is no need for RTS/CTS handshaking. CSP frame length is set to 200ms with 20% of the frame allocated for broadcast traffic. There is a 10ms guard time between the unicast and broadcast portions of the frame resulting in a $150/200 \times 2\text{Mbps} = 1.5\text{Mbps}$ raw data rate for the network. Now with approximately 10-15% overhead from Network, MAC, and PHY layers, the application layer data rate is about 1.3Mbps.

Nodes generate application layer flows at a uniform rate of 2Kbits/sec by generating 250 byte-long packets every second. We vary the number of flows generated by all 51 nodes from 100 to 1300, and measure network throughput, latency, and jitter as a function of the offered load. Note that 650 flows correspond to an application data rate of 1.3Mbps.

As can be seen from Figure 4, CSP achieves collision-free communication up to an offered load of about 1.28Mbps. This is very close to the network bandwidth of 1.3Mbps.

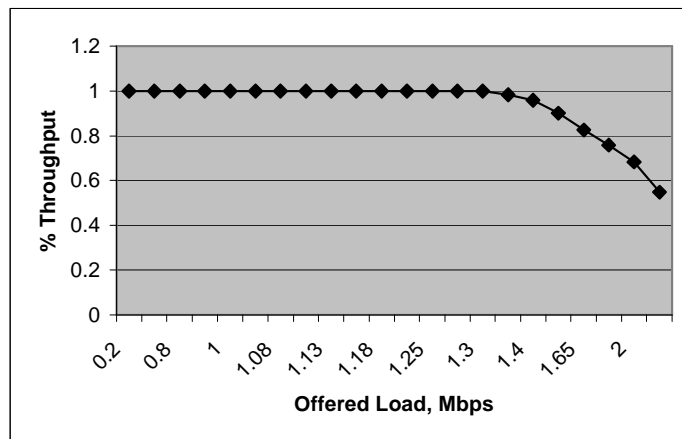


Figure 4: % Network throughput as a function of the aggregate offered network load.

Average network delay and jitter are shown in Figures 5 and 6, respectively. As expected, latency increases with

increased load, and jitter shows a similar behavior.

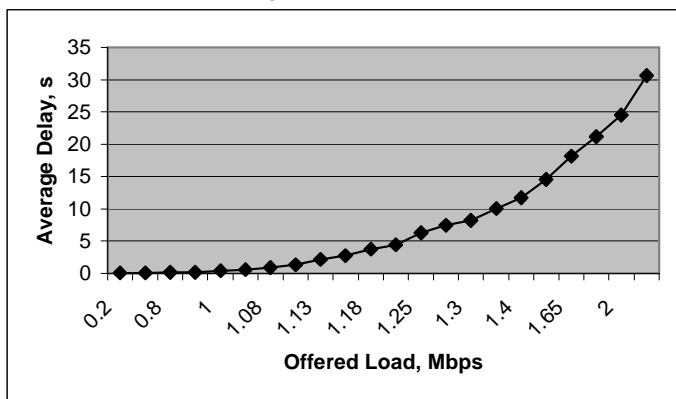


Figure 5: Average network latency as a function of the aggregate offered network load.

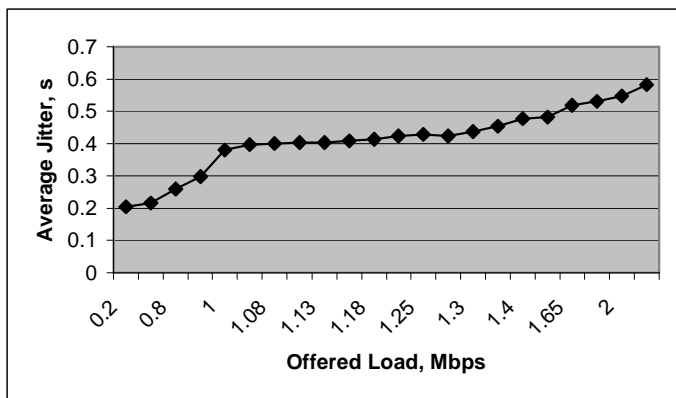


Figure 5: Average network jitter as a function of the aggregate offered network load.

VIII. CONCLUSIONS AND FUTURE WORK

We have extended the pseudorandom scheduling algorithm to manage multiple frequency channels, and included handshaking and link layer acknowledgements to enhance quality of service in networks where hidden nodes are of concern. Multiple channel support enhances network throughput by allowing simultaneous unicast messaging. Comprehensive RTS/CTS handshaking further enhances throughput by allowing widely separated nodes to simultaneously transmit on the same channel. Multi-channel operation and RTS/CTS handshaking presented several problems with respect to handling of differential packet types. Segmenting the CSP scheduling frame into a broadcast segment and a unicast segment solved these problems. The segmenting of the CSP frame offers flexibility to configure unicast and the broadcast segments and offers further performance enhancements.

Taking advantage of multi-channel configurations for subnetting at the MAC layer with CSP offers a significant amount of flexibility and potential for overall network performance improvement. Using subnets, a larger part of the total frame time can be dedicated to unicast traffic

rather than expanding the broadcast portion to compensate for increasing node density and a corresponding increase in broadcast traffic. Assignment of unicast and broadcast channels for each subnet could also be performed dynamically. Dynamic assignment would allow CSP to continually optimize the mapping of nodes to subnets perhaps based upon node position to achieve a better overall network performance. This is an area for future work with CSP.

REFERENCES

- [1] Energy Minimization in Wireless Sensor Networks through an Adaptive Connectionless Scheduling Protocol (ACSP), Scott C. Evans, Nick Van Stralen, John Hershey and Harold Tomlinson, In Proceedings MILCOM2005, Atlantic City, NJ, October 2005.
- [2] USAP Multiple Broadcast Access: Transmitter- and Receiver-Directed Dynamic Resource Allocation for Mobile, Multihop, Multichannel, Wireless Networking, C. David Young, Proc. IEEE MILCOM '99, October 2000.
- [3] C. David Young, "USAP Multiple Access: Dynamic Resource Allocation for Mobile Multihop Multichannel Wireless Networkmg", Proc. IEEE MILCOM '99, October 1999
- [4] David Young, "A Unifying Dynamic Distributed Multichannel TDMA Slot Assignment Protocol", Proc. IEEE MILCOM '96, Vol 1, October 1996.