ADAPT: A DYNAMICALLY SELF-ADJUSTING MEDIA ACCESS CONTROL PROTOCOL FOR AD HOC NETWORKS*

I. Chlamtac, A. Faragó, A. D. Myers, V. R. Syrotiuk, and G. Záruba
Center for Advanced Telecommunications Systems and Services (CATSS)
The University of Texas at Dallas
{chlamtac, farago, amyers, syrotiuk, zaruba}@utdallas.edu

Abstract

This paper presents A Dynamically Adaptive Protocol for Transmission (ADAPT) for ad hoc networks that combines, in a novel way, a collision-free allocation based protocol and a contention based protocol while retaining the advantages of each. At low loads, ADAPT uses its contention mechanism to reclaim/reuse bandwidth that would otherwise be wasted by a pure allocation based protocol. At high loads, ADAPT provides bounded delay guarantees by dynamically changing its operation to that of its allocation based protocol, avoiding the fundamental problem of instability associated with pure contention based protocols. Thus, ADAPT self-adjusts its behavior according to the prevailing network conditions. Both analysis and simulation results demonstrate that the two protocols interact in a positive way, showing that it is possible to combine the advantages of two fundamentally different design philosophies without suffering from their drawbacks.

Introduction

A mobile ad hoc network is a self-organizing system of wireless nodes that requires no fixed communications infrastructure. In the event any two nodes cannot communicate directly, each node must act as a relay, forwarding packets on the behalf of other nodes. Due to the broadcast nature of a radio channel, overlapping transmissions (collisions) may occur resulting in increased packet loss and delay due to retransmissions. Thus a key issue is determining when nodes are allowed to access the channel (i.e., transmit a packet), a decision made by a Media Access Control (MAC) protocol.

Generally, MAC protocols may be broadly classified into two groups based on their strategy for determining access rights. In *contention* protocols, such as Aloha, CSMA, MACA, MACAW, FAMA, and 802.11 [1, 3, 8, 9, 10, 11], nodes compete asynchronously to access the shared channel. Some use collision avoidance mechanisms [3, 8, 9, 11], and all ultimately use randomized retransmissions. The primary advantage of this group is that they are *mobility transparent*, i.e., the protocol does not change its operation as the topology changes. While contention based protocols cannot provide deterministic delay bounds, they are effective at low load when few collisions have to be resolved. Their primary disadvantage surfaces at high load, when these protocols spend most of their time resolving collisions. As a result, the throughput approaches zero resulting in an unstable network.

In order to avoid instability, deterministic *allocation* protocols were introduced. These protocols, which include TDMA, variations on spatial reuse TDMA [6], and TSMA[4], assign each node a transmission schedule indicating in which of the synchronized slots the node may transmit. Since there is a guarantee that at least one slot in the schedule will be successful (i.e., collision-free), these protocols have bounded delay.

Simple TDMA assigns a permanent, unique transmission slot to each node in the network. While TDMA is mobility transparent, its throughput is very low since there is no *spatial reuse*, i.e., no multiple simultaneous transmissions are allowed even when the transmitting nodes are sufficiently far enough apart such that no collision would occur. Variants of TDMA attempt to increase the spatial reuse factor by dynamically computing the transmission schedules. However, such protocols are no longer mobility transparent as the transmission schedules must be recomputed as the network topology changes. Furthermore if the network is highly mobile, these protocols potentially become unstable as the nodes can spend virtually all of their time maintaining their transmission schedules.

The Time-Spread Multiple-Access (TSMA) family of protocols are mobility transparent and have a relatively high degree of spatial reuse. However, the a priori computation of the schedule assumes a fixed upper bound on the *maximum degree* of the network, i.e., the maximum number of nodes that are in the transmission range of a node (its neighbors). If the degree constraint is violated, the guarantees on delay are lost and, consequently, these protocols may also become unstable. This constraint was overcome in threaded-TSMA[5], however, the resulting schedules, hence delay, can be prohibitively long.

In this paper, we propose a new MAC protocol that combines, in a novel way, an allocation and contention protocol. Moreover, the protocol has a simple method of dynamically

*This work was supported in part by the DOD USARO (Army Research Office) under contract No. DAAG55-97-1-0312.

0-7803-5796-5/99/$10.00 © 1999 IEEE

adapting its behavior according to the prevailing traffic loads and node densities, hence the name: *A Dynamically Adaptive Protocol for Transmission* (ADAPT). Thus, ADAPT is a stable, mobility transparent MAC protocol that provides a deterministically bounded delay while maintaining a high spatial reuse factor. In the next section, we describe the ADAPT protocol in detail. We then present an analysis of ADAPT demonstrating that the combined protocols interact in a positive way. Next, we offer numerical results gathered from simulation that confirms our analysis. Finally, we end the paper with a summary of our conclusions.

ADAPT: Combining Allocation and Contention Protocols

In this section, we propose an approach to combine an allocation protocol with a contention protocol in order to obtain the combined advantages of each. Although in principle any allocation protocol can be combined with any contention protocol, in ADAPT we use simple TDMA (allocation) as the *base* protocol, and combine it with CSMA/CA (contention) [7].

Our choice for using TDMA as the base protocol in this approach is motivated by the fact that it provides the shortest possible transmission schedule for the situation when every node is in the transmission range of every other (i.e., a fully connected network).

Our choice for using CSMA/CA is motivated by the *hidden terminal problem*, the conflict situation that may arise when two or more nodes, sharing a common intermediate neighbor i, attempt to transmit a packet to i at the same time. In a radio network, collisions cannot be detected since generally a node may not both transmit and receive simultaneously. Instead, CSMA/CA attempts to avoid such collisions by preceding data transmissions with a RTS/CTS (Request-To-Send/Clear-To-Send) *control packet* exchange. These control packets are much shorter than data packets, so collisions among them have less of an impact on the protocol's performance.

Briefly, CSMA/CA works as follows. A source node s has a data packet p to send to a neighboring node d. Before sending p, s sends a RTS packet r to d. Upon receiving r, node d responds with a CTS packet c. Once s receives c it finally transmits p. All other nodes that receive either r or c realize that nodes s and d are communicating, and defer any transmissions until after s has sent p to d. If the transmission of p is unsuccessful, s schedules the packet for retransmission at some randomly chosen time.

Figure 1 shows how CSMA/CA is combined with TDMA in ADAPT. For a network of N nodes, we construct a TDMA schedule of N *slots* for each node. A slot is large enough to accommodate the following. We establish a *sensing period* in which all nodes j determine whether or not a node i is using its assigned slot, s_i. Determination is made by listening for any transmissions within the specified sensing period. If node i has a data packet to send in s_i, it immediately contends for the slot using the RTS/CTS exchange of CSMA/CA.

If node i does not have a packet to transmit, then after the *sensing period* all other nodes will have determined that i is not using s_i. At this time, any node with a packet to transmit will contend for use of this slot using a RTS/CTS exchange. If any node j successfully performs this exchange, then it is allowed to transmit its data packet in the remaining portion of the slot. Notice that even though the base protocol is full time division, we obtain spatial reuse of any unused dedicated slots.

If there is a collision of nodes contending for use of a slot we manage the contention by using a backoff interval b, initialized to zero, at each node. If node i does not use its dedicated slot then other nodes compete for slot i. Whether or not the contention is successful we increment b (up to some maximum value). This reflects active contention for the slot. Consistent with binary exponential backoff techniques [2], if a node j contending for a slot $i, i \neq j$, experiences a collision (i.e., its RTS/CTS exchange is unsuccessful), it will wait a random number r, $1 \leq r \leq 2^b$, of slots before contending for a slot again. (Of course, node j is always allowed to use its own dedicated slot j.) The only time we reduce the backoff interval is when a slot i is unused. In this case, b is decremented, reflecting the decrease in contention for the slot.

In this way, each node dynamically self-adjusts its contention for slots based on load. At low loads or density, ADAPT behaves as CSMA/CA with similar performance. As the load or density increases, ADAPT changes its operation into TDMA, where each node uses its dedicated slot. In fact, there is still the opportunity for spatial reuse, even at high load. There is very little overhead associated with the combined protocol, namely, the sensing period and the RTS/CTS exchanges, which are both very short in duration relative to the time to transmit a data packet. Furthermore, these adaptations occur independently at each node, according to its mobility, the density of its neighborhood, and the offered load. Thus, ADAPT is not only mobility transparent, but also density and load transparent.

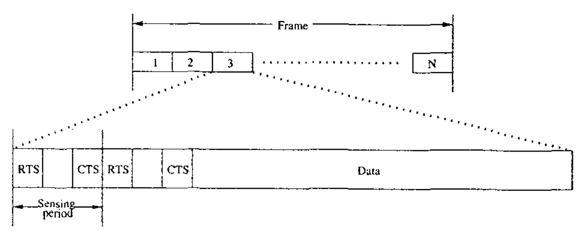

Figure 1: ADAPT: Combining allocation and contention.

Analysis of Generic Combined Protocols

In this section we first provide a simple approximation of the throughput of a generic combined protocol, then we investigate how performance can be optimized under certain simplifying assumptions. We do not consider the overhead caused by the exchange of RTS and CTS messages.

Let M be the number of nodes in the network. Let s_i be the average fraction of data slots that are used by the protocol at node i, under given load conditions. Let S be the average of s_i for all nodes in the network:

$$S = \sum_{0<i\leq M} s_i/M. \tag{1}$$

That is, S is regarded as the probability that a data slot is used by the protocol. Let r_j be the number of packets that are successfully transmitted in slot j in the whole network. Let R be the average of r_j's that are greater than zero. In other words, R is the average spatial reuse factor. Then the throughput can be characterized as $T = RS$, giving the average number of successfully transmitted packets in the network per slot. For the two component protocols we use the same quantities with subscripts that distinguish the protocols: T_a, R_a, S_a for the allocation and T_c, R_c, S_c for the (slotted) contention protocol, respectively. For calculating the values above we separate the two protocols but use the same network with the same load.

Now, we can roughly approximate the throughput of the combined protocol as follows:

$$T \approx T_a + (1-S_a)T_c = R_a S_a + (1-S_a)R_c S_c. \tag{2}$$

The formula follows from the reasoning that the allocation protocol has its own throughput T_a, while it leaves a fraction $1 - S_a$ of slots unused. In these leftover slots the contention protocol is running, producing its own throughput.

An interesting observation is that for any reasonable value of S_a (i.e., $0 < S_a < 1$), there is a strict increase in throughput in the combined protocol as compared to the pure allocation protocol. This is explained by the fact that the combined protocol never takes away a slot from the allocation protocol when it wants to use it, but the leftover slots are still further utilized for transmission by the contention protocol. We can formulate this fact in the following, surprising, statement: in the considered network scenario, *if a protocol guarantees optimum throughput, then it cannot be a pure allocation protocol.*

Now let us take a closer look at the behavior of the network at a randomly chosen time slot t. Assume that node i has λ_{ij} packets to send to a given neighbor j per slot, on the average. Let N_i be the set of neighbors of node i. Then i has altogether $\lambda_i = \sum_{j\in N_i} \lambda_{ij}$ packets per slot to its neighborhood. To avoid delay instability we assume $\lambda_i < 1$, which, of course, implies $\lambda_{ij} < 1$, too.

Now let k_i be the number of neighbors of i, i.e., $k_i = |N_i|$ and let L be the frame length of the allocation based protocol. Let us examine the average throughput (=successfully transmitted packets per slot) that node i can achieve to a neighbor j, denoted by T_{ij}. In a randomly chosen slot t two possibilities can occur. If the slot is owned by node i, then the node can transmit without any obstacle and this generates an average throughput of λ_{ij}/L (packets/slot) to j, since i has λ_{ij} packets per slot to j and the probability that the randomly chosen slot is owned by i is $1/L$. If the slot is a contention slot for i, then denote by $p_t^{(i)}$ the probability that i attempts transmission according to the contention protocol. The additional throughput from this can be estimated as follows. For successful contention transmission for i three events have to occur simultaneously: (1) the slot is a contention based slot for i, this occurs with probability $1-1/L$; (2) i contends for the slot, this occurs with probability $p_t^{(i)}$; (3) neither j nor its other $k_j - 1$ neighbors transmit to anybody (by owning the slot or by contention), this occurs with probability φ,

$$\varphi = \left(1-\frac{\lambda_j}{L}-(1-1/L)\lambda_j p_t^{(j)}\right)\times \prod_{r\in N_j-\{i\}}\left(1-\frac{\lambda_r}{L}-(1-1/L)\lambda_r p_t^{(r)}\right),$$

assuming the independence of events. Thus, the average throughput form i to j can be written as

$$T_{ij} = \left(\frac{\lambda_{ij}}{L} + (1-1/L)\lambda_{ij} p_t^{(i)}\right)\varphi. \tag{3}$$

Although the above equation looks quite complicated, it becomes much easier to manage if we assume the following uniformity conditions: $\lambda_{ij} = \lambda$, $p_t^{(i)} = p$ and $k_i = k$ for every

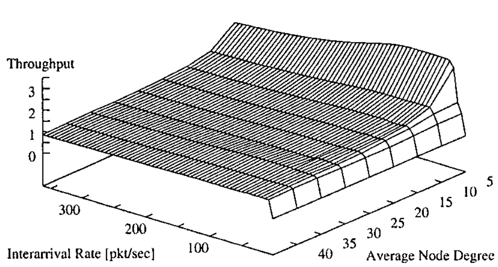

Figure 2: Throughput of the ADAPT protocol.

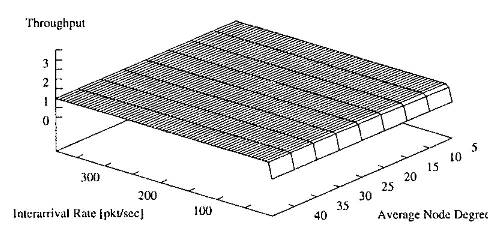

Figure 3: Throughput of the TDMA protocol.

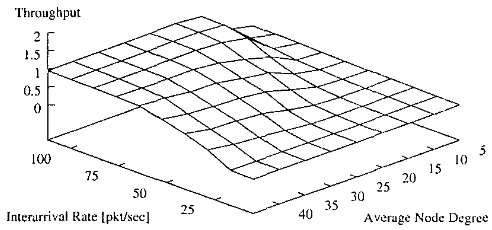

Figure 4: Throughput produced by allocation protocol.

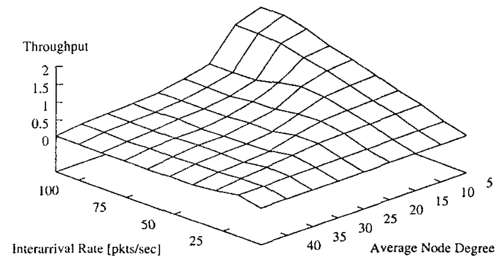

Figure 5: Throughput produced by contention protocol.

i, j, t. In other words, we replace the node and time dependent parameters by constant "typical" values, i.e., we assume that the topology and the load are homogeneous all over the network. This is, of course, only an approximation, but it greatly helps to follow the typical system behavior. With this, using that now $\lambda_j = k\lambda$, the throughput expression becomes

$$T_{ij} = \frac{\lambda}{L} + (1 - 1/L)\lambda p \left(1 - \frac{k\lambda}{L} - (1 - 1/L)k\lambda p\right)^k. \quad (4)$$

We introduce the variable $x = (1 - 1/L)\lambda p$, what can be seen as the probability that a node tries to use a leftover slot. The right hand side of (4) is the following function:

$$f(x) = \frac{\lambda}{L} + x\left(1 - \frac{k\lambda}{L} - kx\right)^k. \quad (5)$$

With some calculus one can show that $f(x)$ is maximized with the choice of $x_{opt} = (1 - k\lambda/L)/(k^2 + k)$. Substituting this into the expression $x = (1 - 1/L)\lambda p$ we obtain that the optimal choice for the "persistence probability" p is

$$p_{opt} = \frac{1 - \frac{k\lambda}{L}}{k(k+1)(1 - \frac{1}{L})\lambda}. \quad (6)$$

Re-substituting p_{opt} into (4) and taking into account that the total throughput to all neighbors of node i is kT_{ij}, then, using $(1 - \frac{1}{k})^k \approx \mathrm{e}^{-1}$, we obtain that the optimal throughput of the node is approximated as

$$T_{opt} \approx \frac{k\lambda}{L} + \frac{1}{\mathrm{e}k}\left(1 - \frac{k\lambda}{L}\right)^{k+1}. \quad (7)$$

Let us compare this with the initial estimation (2). If we take into account that $k\lambda$ is the offered load of the node to all neighbors (in terms of packets/slot), then $\frac{k\lambda}{L}$ is the average throughput T_a of the allocation based protocol. The term $\left(1 - \frac{k\lambda}{L}\right)^{k+1}$ can be interpreted as the probability that no node in a given neighborhood uses an allocated slot, so this corresponds to the fraction of leftover slots. Finally, $\frac{1}{\mathrm{e}k}$ equals to the estimated optimized throughput of a p-persistent slotted Aloha protocol in a packet radio network (see [2]), which is essentially the protocol we used in the modeling assumptions. With this interpretation we exactly get back the initial estimation (2), thus showing it is a valid approach.

Simulation Results

A discrete event simulator was used to evaluate the performance of the ADAPT protocol in a mobile ad hoc network environment. Our study was limited to MAC-layer details, thus no specific higher layer protocols were simulated.

Node mobility was simulated using a random graph model, where vertices represent mobile nodes and each edge (i, j) represents a bidirectional communication link between nodes i and j. To simulate the movement of nodes in our network we create new random graphs at specified intervals in time (the mobility update period). Moreover, with each such graph there were no isolated nodes and the node degree remained below a fixed threshold parameter.

Packet transmissions were considered to take place within a single radio channel operating at a rate of 1Mbps. (For simplification, we assumed a noiseless radio channel.) Control packets were 12 bytes in length, and were used to model the RTS and CTS packets used in the ADAPT protocol. Data packets, consisting of 2048 bytes, were introduced into the network according to a Poisson arrival process, and simulated traffic load.

Using the described model and assumptions, we simulated our ADAPT protocol in networks with varied traffic loads and average node densities. Each simulation consisted of 50 nodes, and data packets were distributed in a homogeneous fashion among them. To demonstrate that the contention mechanism of ADAPT would not adversely affect the operation of the underlying TDMA protocol, we simulated a TDMA protocol in order to compare their performance in similar environments.

In Figures 2 and 3, we show the normalized network throughput of both protocols with respect to both increasing traffic load, measured in (data) packets per second, and average node degree. The throughput is normalized with respect to bit rate of the channel, so a throughput of 1 indicates that the data rate (measured in bps) is equal to the channel rate. As we can see, at high loads and node densities ADAPT's performance is that of its underlying allocation based protocol. Moreover, at low loads and low node densities ADAPT outperforms the TDMA protocol due to its ability to reuse any available slots. This phenomenon is also evident in Figures 4 and 5. Here, we have separated the total ADAPT throughput into that portion produced by the allocation protocol (Figure 4) and the contention protocol (Figure 5). We can see that, at high loads and high density, the allocation protocol produces most (if not all) of the total throughput. At low loads and low density, ADAPT's contention protocol provides most of the total throughput.

Finally, in Figures 6 and 7, we present the average packet delay as a function of the traffic load for various average node densities. Clearly, at low loads the ADAPT protocol has lower average packet delays than the TDMA protocol. This is due to the higher spatial reuse gained by ADAPT's contention mechanism.

Conclusions

In this paper we presented a new MAC protocol for mobile ad hoc networks. ADAPT combines a contention protocol with a collision-free allocation protocol, resulting in a hybrid that maintains the advantages of each while avoiding their individual drawbacks. Analysis and empirical evidence have demonstrated that the performance of ADAPT mirrors the contention protocol at low load (density), and the allocation protocol at higher load. Thus the two fundamentally different protocols interact in a positive way, yielding a combined protocol that remains stable at high loads and provides a high degree of spatial reuse.

References

[1] ABRAMSON, N. "The ALOHA System — Another Alternative for Computer Communications." *Proceedings of the Fall Joint Computer Conference*, 1970, 281–285.

[2] BERTSEKAS, D. AND GALLAGER, R. "Data Networks." Prentice Hall, 1992.

[3] BHARGHAVAN, V., DEMERS, A., SHENKER, S., AND ZHANG, L. "MACAW: A Media Access Protocol for Wireless LANS." *Proceedings of ACM SIGCOMM*, 1994, 212–225.

[4] CHLAMTAC, I. AND FARAGO, A. "Making Transmission Schedules Immune to Topology Changes in Multi-Hop Packet Radio Networks." *IEEE/ACM Transactions on Networking*, Vol. 2, No. 1, 1994, 23–29.

[5] CHLAMTAC, I., FARAGO, A. AND ZHANG, H. "Time-Spread Multiple-Access (TSMA) Protocols for Multihop Radio Networks." *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, 1997, 804–812.

[6] CHLAMTAC, I. AND PINTER, S.S. "Distributed Node Organization Algorithm for Channel Access in a Multihop Dynamic Radio Network." *IEEE Transactions on Computers*, Vol. 36, No. 6, June 1987, 728–737.

[7] COLVIN, A. "CSMA with Collision Avoidance." *Computer Communications*, Vol. 6, No. 5, 1983, 227–235.

[8] FULLMER, C.L. AND GARCIA-LUNA-ACEVES, J. "Floor Acquisition Multiple Access (FAMA) for Packet-Radio Networks." *Proceedings of ACM SIGCOMM*, 1995.

[9] KARN, P. "MACA — A New Channel Access Protocol for Packet Radio." *ARRL/CRRL Amateur Radio Ninth Computer Networking Conference*, 1990, 134–140.

[10] TOBAGI, F., AND KLEINROCK, L. "Packet Switching in Radio Channels: Part I — Carrier Sense Multiple Access Modes and Their Throughput Delay Characteristics." *IEEE Transactions on Communications*, Vol. 23, No. 12, 975, 1400–1416.

[11] Wireless Medium Access Control and Physical Layer WG, IEEE Draft Standard P802.11 Wireless LAN. *IEEE Standards Department*, D3, Jan. 1996.

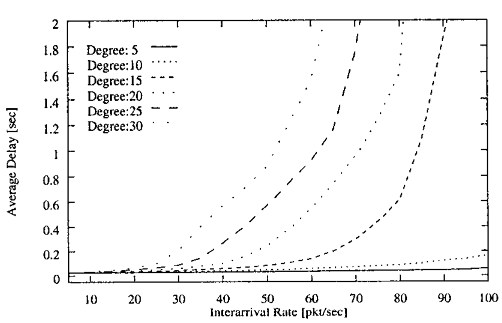

Figure 6: Average packet delay of ADAPT.

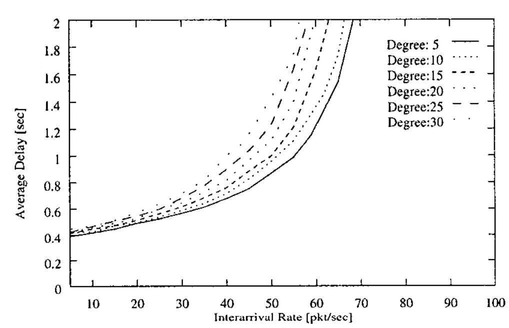

Figure 7: Average packet delay of TDMA.