An adaptive medium access control protocol using $m$-ary tree algorithms for quality-of-service support in single-cell ad hoc networks

Orestis Tsigkas *,1, Fotini-Niovi Pavlidou

Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Received 18 November 2005; received in revised form 8 December 2006; accepted 28 December 2006

Available online 9 January 2007

Abstract

Concurrent with the rapid expansion of wireless networks is an increasing interest in providing Quality-of-Service (QoS) support to them. As a consequence, a number of medium access control protocols has been proposed which aims at providing service differentiation at the distributed wireless medium access layer. However, most of them provide only average performance assurances. We argue that average performance guarantees will be inadequate for a wide range of emerging multimedia applications and “per-flow” service assurances must be provided instead. Based on $m$-ary tree algorithms, we propose an adaptive and distributed medium access algorithm for single-cell ad hoc networks to provide “per-flow” service assurances to flows whose QoS requirement can be expressed as a delay requirement. Both analytical and simulation experiments are conducted to assess the performance of the proposed scheme.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Quality-of-Service; Medium access; Ad hoc networks

1. Introduction

The concept of wireless networking has gained popularity at an unprecedented rate over the last few years. Ad hoc networks have made spontaneous, on demand networking possible, while infrastructure-based Wireless Local Area Networks (WLANs) have enabled hassle-free, high perfor-

mance networking without the need for additional wires or retrofitting. Moreover, the recent advances that have been achieved in the physical layer have led to bitrates that were until recently achieved only in wireline networks, rendering the Medium Access Control (MAC) sublayer one of the most critical building blocks in any network which employs a common medium. By arbitrating access to the shared channel, medium access protocols have a direct impact on the efficient use of the available raw bandwidth. Low collision rate, low overhead and robustness are among the characteristics that a medium access protocol should possess in order to achieve high efficiency.
Concurrent with the expansion of wireless networks is a rapidly increasing demand for multimedia applications with very stringent and diverse Quality-of-Service (QoS) requirements. Providing QoS requires the network to guarantee hard bounds on a set of measurable prespecified attributes, such as delay, bandwidth, probability of packet loss, and delay variance (jitter). The necessity of providing QoS support in wireless networks is reinforced in the case that wireless access is just another hop in the communication path. Wireless networks are only a part, even though significant, of a global infrastructure that is strongly characterized by heterogeneity of available telecommunications platforms. To efficiently support multimedia applications in such a heterogeneous environment, it is of critical importance, not only to design the interconnection of the various platforms but also to provide QoS guarantees on an end-to-end basis. End-to-end QoS assurances can only be provided if each telecommunication system provides the means for offering QoS guarantees. However, the different characteristics of wireless networks compared to those of their wired counterparts pose a number of difficulties and unique challenges in guaranteeing bounds on the QoS metrics.

It is, thus, evident that in order to honour the specific needs of traffic classes while achieving high efficiency in terms of throughput, support from both medium access and scheduling algorithms is required. Indeed, in the literature there is a number of studies that aim at providing service differentiation at the distributed wireless MAC layer. The common feature of these distributed medium access algorithms is their attempt to provide QoS support by applying the ideas behind scheduling algorithms proposed for wired networks.

Following the paradigm of Weighted Fair Queuing (WFQ), [2–4] propose distributed algorithms for rate-based differentiation and throughput fairness. However, the tight coupling between rate and delay under WFQ, renders it inappropriate for providing delay guarantees, especially in the case of low-rate traffic requiring low delay bounds. Indeed, none of the schemes proposed in the above papers, addresses the problem of delay differentiation.

The objective of another class of QoS capable medium access protocols proposed in the literature is to provide service differentiation by allowing faster access to the channel to traffic classes with higher priority. Refs. [5,9,10] propose modifications to the IEEE 802.11 Distributed Coordinated Function (DCF) [18] to incorporate differentiated service by supporting two or more priority levels.

In [6], the authors propose a MAC protocol that provides multiple priority levels and adopts the black-burst mechanism [7] to guarantee that higher-priority packets will always be transmitted earlier than lower-priority ones. Packets with the same priority are then transmitted in a round robin manner. The Busy Tone Priority Scheduling (BTPS) [8] protocol makes use of two busy tone signals to create two priority levels and adds an additional interframe space to the 802.11 DCF to ensure channel access of high priority packets.

The authors in [11] present the performance evaluation of IEEE 802.11e Enhanced DCF and observe that, although EDCF provides service differentiation, the delay cannot be guaranteed even for the highest priority traffic. In [15], a distributed priority scheme is proposed which piggybacks the priority index of a head-of-line packet onto existing handshake messages of the 802.11 DCF. Neighbors monitor these transmissions and keep a table of the times in order to assess the relative priority of their own head-of-line packet. A station defers from contention as long as a time on its table precedes the arrival time of its own head-of-line packet. It is shown that this scheme achieves a closer approximation to an ideal deadline-based schedule than IEEE 802.11 does. A similar approach is followed in [16], where the priority index of a packet is given by its laxity budget, the packet delivery ratio of the flow which the packet belongs to and the packet delivery ratio desired by the user.

All of the above schemes attempt to provide distributed service differentiation by assigning traffic to fixed priority classes. The channel access mechanism of the base MAC protocol is modified so that higher priority traffic classes have a higher possibility to gain access to the common medium and consequently perceive better performance. However, even though these schemes do provide service differentiation, the service assurances offered to a priority class are only on average better than those offered to lower-priority classes. The reason resides in the fact that most of them fail to ensure that nodes with lower-priority packets will always defer access in the presence of higher priority ones. We argue that average performance assurances will be inadequate for a wide range of emerging multimedia applications and “per-flow” service guarantees will have to be provided instead.
Based on \( m \)-ary tree algorithms, we propose an adaptive and distributed medium access algorithm for single-cell ad hoc networks to provide “per-flow” service assurances to flows whose QoS requirement can be expressed as a maximum allowable delay that its packets can suffer at each hop along the communication path. Packets whose delay exceeds this given threshold will be dropped, resulting in wasteful consumption of the scarce wireless resources. We aim at developing a medium access mechanism which is capable of meeting the delay requirements of data packets by serving first packets whose deadline is about to expire. This is similar to implementing an Earliest Deadline First (EDF) service discipline \([1]\). The EDF scheduler is a dynamic priority scheduler where the priority for each packet is assigned as it arrives and it increases accordingly to the amount of time it spends in the system.

While our proposed MAC scheme aims at providing QoS support to traffic with timeliness constraints, “per-flow” rate guarantees can be provided as well. Recent advances in core-stateless architectures allow for the provision of delay and rate guarantees by time-stamping packets without the need to maintain “per-flow” state in core nodes. In \([24]\), Stoica and Zhang introduced the notion of Dynamic Packet State to develop mechanisms for providing end-to-end assurance levels similar to those that can be provided with “per-flow” mechanisms. The Dynamic Packet State approach was further extended in \([25]\) to develop a unifying framework for providing scalable support for guaranteed services.

Network stations operate within a broadcast region, in which each station is within radio range of all other stations, and contend for channel access to the common medium in a completely distributed way. Based on overheard information, our proposed protocol can adapt to traffic conditions, rendering itself insensitive to traffic load variations; thus, the efficiency of the protocol in terms of achieved throughput remains quasi-constant regardless of the number of contending nodes, while the parameters that determine the protocol’s performance need not to be reconfigured when the length of transmitted packets changes. Both analytical and simulation results are derived to assess the performance of the protocol.

The rest of the paper is structured as follows. In Section 2, we provide an overview of EY-NPMA, the MAC protocol for HIPERLAN, and DP-TB, a protocol that is based on EY-NPMA to provide enhanced QoS support to multimedia applications. Section 3 provides the motivation of our work and describes the design of the proposed protocol. Section 4 presents the theoretical analysis of the proposed scheme and its performance evaluation through analytical experiments. Section 5 deals with the performance evaluation of the proposed scheme through simulation results, while Section 6 concludes the paper.

2. Background work

EY-NPMA \([17]\), the HIPERLAN MAC protocol, is a dynamic priority scheme, which provides hierarchical independence of performance by means of channel access priority. However, its ability to track an ideal EDF scheduler and, thus, provide service differentiation, degrades severely as traffic load increases and the number of contending nodes grows. This is mainly due to the fact that EY-NPMA supports only five priority levels. Based on EY-NPMA, we proposed a dynamic priority Medium Access Control protocol which modifies the channel access scheme of EY-NPMA to support a high number of priority levels \([22]\). Both of these MAC protocols are reviewed in this section.

2.1. EY-NPMA

EY-NPMA stands for Elimination-Yield Non-Preemptive Priority Multiple Access. Elimination-Yield describes the contention resolution scheme, while NPMA refers to the principle of the HIPERLAN medium access mechanism that provides hierarchical independence of performance by means of channel access priority. When a new packet arrives, its lifetime is set to a value that cannot exceed 500 ms. Depending on its residual lifetime, the packet is assigned one of the five priorities from 0 to 4, with 0 being the highest priority. Packets that cannot be delivered within the allocated lifetime are discarded. The synchronized channel access cycle comprises three phases: the prioritization, contention and transmission phase.

The prioritization phase ensures that only the data transmission attempts with the highest channel access priority will survive this phase. The contention phase consists of two subphases; namely, the elimination phase and the yield phase. During the elimination phase, a contending node transmits a channel access burst, whose length in slots is random between 1 and a predefined maximum
according to a truncated geometric distribution, and then listens to the channel. If the channel is sensed as idle, the node proceeds to the yield phase. During the yield phase, the contending nodes sense the channel for a random number of slots, and if the channel is sensed idle, they immediately enter the transmission phase by transmitting the packet stored in their buffer. All other stations sense the beginning of the transmission and refrain from transmitting. The parameters in the HIPERLAN standard were chosen so as to achieve a quasi-constant collision rate of 3.5% up to 256 simultaneous transmitting nodes. A performance study of EY-NPMA can be found in [12,14], where extended analytical and simulation results are presented. Further, it has been compared with DCF and EDCF in [19,13] respectively.

### 2.2. DP-TB

Based on EY-NPMA, we proposed a dynamic priority MAC protocol, DP-TB, to support time-bounded services in wireless networks [22]. The proposed medium access scheme provides support to traffic with delay requirements by closely approximating an ideal EDF schedule. The proposed scheme preserves all three phases of the synchronized access cycle of the EY-NPMA scheme; yet, it features a different structure for the prioritization phase. Instead of a maximum of five prioritization slots, we proposed a scheme that uses at most $N$ slots for this phase. The prioritization phase, in the proposed DP-TB scheme, is further subdivided into $n$ subphases, where subphase $j$ consists of at most $p_j$ slots, such that $\sum_{i=1}^{n} p_i = N$. We did not fix $N$ and $n$ to constant values, but rather let them be parameters of the system. Depending on the choice of $N$ and $n$, there is a trade-off between the extent that the ideal EDF scheduler can be approximated to and the throughput that can be achieved.

EY-NPMA uses $N$ prioritization slots to support $N$ priority levels. By subdividing the prioritization phase in $n$ subphases, DP-TB provides a maximum of $P = \prod_{i=1}^{n} p_i$ priority levels, with 0 being the highest and $P - 1$ the lowest one. The lifetime of a packet that has just arrived is set to a value that cannot exceed $500$ ms. Then, the interval of $500$ ms is divided into $P$ time intervals, each of which has a duration of $t_p = 0.5/P$ s. The priority index $PI$ of a packet with residual lifetime $RL$ can be computed as

$$ PI = \{ k : k \cdot t_p \leq RL < (k + 1) \cdot t_p \} = \left[ \frac{RL}{t_p} \right]. \quad (1) $$

The priority index of a packet can be computed as a function of the maximum number $p_i$ of slots allocated to each subphase $i$ and the number $p_j$ of slots for which a node should sense the channel in each subphase $j$, in order to determine whether it has the highest priority packet for transmission.

$$ PI = \sum_{i=1}^{n-1} \left( ps_i \cdot \prod_{j=i+1}^{n} p_j \right) + ps_n. \quad (2) $$

Then, the algorithm presented below can be used to compute the set of parameters \{ps$_1$, ..., ps$_n$\}.

```plaintext
for (i = 1; i < n; i++) do
    \{ ps$_i$ := $\prod_{j=i+1}^{n} p_j$ \}
    PI := PI - ps$_i$ \cdot $\prod_{j=i+1}^{n} p_j$
\}
ps$_n$ := PI
```

As soon as the set of parameters \{ps$_1$, ..., ps$_n$\} has been computed, a packet can contend for channel access in the prioritization phase. The prioritization phase of DP-TB works as follows. At the beginning of the first subphase, a station that has a packet eligible for transmission senses the channel for as many as $ps_1$ slots. If the channel is idle for the whole sensing interval, the station transmits a burst of one slot and proceeds to the second subphase. Otherwise, the station exits contention and will have another chance for accessing the channel at the next cycle. In the same manner, during the second subphase the station senses the channel for $ps_2$ slots, and if the channel is sensed idle it transmits a burst slot. The procedure is repeated until the last subphase, where the node transmits a burst of random length, instead of just one burst slot. The length of this burst is between 1 and a predefined maximum number of slots. The contention phase in DP-TB works as in EY-NPMA.

However, the computation of the backoff interval in the yield phase is modified to ensure that if there is a successful transmission, the transmitting station is the one with the lowest residual lifetime among those who survived the elimination phase. The proposed medium access protocol allowed us to define a large number of priority levels by using a relatively small number of prioritization slots. The added
overhead of the prioritization phase in DP-TB can be alleviated by its lower collision rate. Simulation results showed that DP-TB can closely approximate an ideal EDF scheduler while achieving high medium utilization.

3. Proposed protocol

3.1. Motivation

In [22], the ability of DP-TB and Priority Broadcast for IEEE 802.11 DCF [15] to track an ideal EDF scheduler was evaluated. These two schemes correspond to two different approaches in resolving the priorities (deadlines) of contending packets. In [15], nodes exchange information about their queued packets so that a node can deduce when it should defer to nodes that have a higher priority packet. On the other hand, DP-TB, as EY-NPMA, performs priority resolution by using signalling; highest priority nodes signal first to preempt lower priority ones from gaining access to the channel. The comparison made it evident that DP-TB can closely approximate an ideal EDF schedule while 802.11 DCF fails to resolve the priority of contending packets even for a small network population. Moreover, it was observed that the large overhead introduced by the signalling of DP-TB does not come at the expense of reduced efficiency in terms of achieved throughput. By closely approximating an EDF scheduler, DP-TB ensures that packets whose deadline is about to expire will not be preempted by packets that have enough delay budget (residual lifetime) to be transmitted in succeeding channel access cycles. Consequently, DP-TB minimizes the number of lost packets due to incorrect scheduling, achieving high medium utilization.

However, the number of priority levels provided by DP-TB is fixed and prevents it from adapting to traffic conditions. This number corresponds to a fair balancing between the extent that an ideal EDF scheduler can be approximated to and the throughput that can be achieved. Better performance could be achieved if the number of priority levels offered by DP-TB was dynamically varied, allowing it to adapt to traffic conditions. Admittedly, DP-TB possesses certain inherently adaptive qualities, since it reduces its overhead as traffic load increases so that its efficiency is not degraded; nevertheless, this is not enough to characterise the protocol as adaptive. Furthermore, DP-TB suffers from the hidden terminal problem and this shortcoming renders it inappropriate for ad hoc networks.

The hidden terminal problem can be alleviated if the MAC protocol incorporates the Request To Send/Clear To Send (RTS/CTS) mechanism, as in the case of IEEE 802.11 DCF. Moreover, the RTS/CTS mechanism turns out to be very beneficial when the length of transmitted packets is large. For large packet lengths, the dominant factor that affects most of the MAC protocols’ performance is the collision rate rather than the overhead. The RTS/CTS mechanism constrains collisions just to the exchange of RTS and CTS messages, preventing the collision of data packets. Indeed, simulation results in [22] showed that for large packet lengths and large network populations, Priority Broadcast for DCF achieves slightly better throughput than DP-TB. However, the large overhead introduced by the RTS/CTS mechanism prevents DCF from achieving high efficiency for small payloads.

Based on the above, it becomes obvious that using signalling is a more promising solution towards providing QoS support at the distributed wireless MAC layer than merely overhearing the common medium to share information. Moreover, the RTS/CTS mechanism can be very beneficial to a protocol’s efficiency, apart from protecting its performance against the hidden terminal problem. Based on m-ary tree algorithms, we propose an adaptive medium access scheme to provide “per-flow” service assurances to flows whose QoS requirement can be expressed as a delay requirement. As we discussed earlier, recent advances in scalable core-stateless architectures can be applied to allow the proposed scheme to provide both delay and rate guarantees. The proposed scheme incorporates the RTS/CTS mechanism to protect itself against the hidden terminal problem. Moreover, we derive analytical results to show that our protocol adapts to traffic conditions and is, thus, insensitive to traffic load variations; that is, the efficiency of the protocol in terms of achieved throughput remains quasi-constant regardless of the number of contending nodes, while the parameters that determine the protocol’s performance need not to be reconfigured when the length of transmitted packets changes.

3.2. Proposed algorithm

In this subsection, we will describe the mechanics of the proposed protocol, especially focusing on the
way that priority resolution is performed. Tree algorithms were first proposed in [26] to resolve contentions in random access channels and were shown to achieve high throughput while reducing the mean delay required to gain channel access. We show that adaptive m-ary tree algorithms can efficiently be applied to resolve the priorities of contending nodes, where the priority index of packets is dynamically updated and the number of priority levels is infinite.

The proposed medium access scheme builds a tree whose degree of the root node is $k$ while the degree of any other node is $m$; that is, the number of branches that emanate from the root node is equal to $k$ while it is equal to $m$ for any other node in the tree. The tree algorithm is assumed to be dynamic in the sense that the tree is allowed to vary from epoch to epoch depending on the priority index of stations’ head-of-line packet. Within any single epoch, in which our correspondence to a channel access cycle, the tree is held fixed. Furthermore, the tree algorithm is adaptive in the sense that the degree $k$ of the root node is allowed to vary from epoch to epoch depending on traffic conditions. The degree $m$ of any other node in the tree is held fixed.

We denote by $n_{ij}$ a node $n$ whose depth is $i$ and index is $j$. The depth of a node is the number of branches between the node and the root node. The root node is assumed to be at depth zero. The index $j$ of a node lies in the range $0 \leq j \leq k - 1$, where $k$ is the degree of the root node, $m$ is the degree of any other node and $i$ is the depth of the node. Moreover, we denote by $T_{ij}$ the rooted subtree whose root node is node $n_{ij}$. We call source every node of the tree that has zero branches emanating from it.

At the beginning of each access cycle, the tree is built according to the following procedure. When a new packet arrives, its lifetime is set to a value from 0 to $S$ s. The Residual Lifetime (RL) of a packet is updated at the beginning of each access cycle. Packets that cannot be delivered within the allocated lifetime are discarded. A node belongs to subtree $T_{ij}$ if and only if its residual lifetime $RL$ satisfies the following inequality:

$$j \cdot \left(\frac{S}{k \cdot m^{i-1}}\right) < RL \leq (j + 1) \cdot \left(\frac{S}{k \cdot m^{i-1}}\right).$$

Moreover, a node $n_{ij}$ is a source if and only if it is the only node whose residual lifetime satisfies the previous inequality. It should be stressed that, although the number of contending nodes is finite, the tree can extend to infinity. Indeed, the structure of the tree depends on the priority index of sources’ head-of-line packets. Since the priority indices of packets are given by their residual lifetimes whose values are infinite, the number of priority levels is infinite as well.

Then, the channel access cycle of the proposed scheme works as follows. At the beginning of the channel access cycle, all sources that intend to transmit a packet wait for a time period equal to a Channel Synchronization (CS) interval and then priority resolution starts at depth 1. Each source determines the subtree $T_{ij}$ of node $n_{ij}$ that it belongs to, where $0 \leq j \leq k - 1$, and senses the channel for as many Priority Resolution Slots (PRS) as the priority index $j$ of its subtree. If the channel is idle for the whole sensing interval, the station transmits a Request To Send (RTS) message. Otherwise, the source exits contention and will have another chance for accessing the channel at the next cycle. When the receiving station detects an RTS message, it waits for a time period equal to a Verification Interval (VI) and responds with a Clear To Send (CTS) message. If the CTS packet is correctly received, the source transmits its data after waiting for a VI period.

Otherwise, all sources that transmitted an RTS message continue the priority resolution procedure at depth 2. Assume that all sources which did not receive a CTS packet transmitted their RTS message after sensing the channel for $l$ slots; thus, they belong to subtree $T_{ij}$. Then, each source determines the subtree $T_{2ij}$ of node $n_{2ij}$ that it belongs to, where $l \cdot m \leq j \leq (l + 1) \cdot m - 1$, and after waiting for a time period equal to VI senses the channel for as many as $(j - l \cdot m)$ slots. If the channel is idle for the whole sensing interval, the station transmits a new RTS message. In the case of a missing CTS, the priority resolution procedure is continued in the same manner until a source is granted clear transmission of its data packet. The receiving node acknowledges the successful reception of a packet with an Acknowledgement (ACK) message after waiting for a time period equal to VI.

From the above, it is evident that the tree is built in a dynamic way since the set of sources that belong to a given subtree $T_{ij}$ vary from epoch to epoch depending on the residual lifetime of the sources’ head-of-line packet. Furthermore, we allow

---

When priority resolution is performed at depth $i > 0$, a source which belongs to the subtree of node $n_{ij}$ senses the channel for as many as $(j - \frac{|L|}{m}) \cdot m$ slots.
the degree $k$ of the root node to vary from epoch to epoch so that the protocol can adapt to traffic conditions. The computation of $k$ is performed as follows. All stations that overhear the transmitted information during a channel access cycle can deduce an upper limit $RL^+$ on the residual lifetime of the packet that is being transmitted. Let $RL^+_r$ denote the average upper limit of the residual lifetime of packets that were transmitted in the preceding $r$ access cycles. Then, the degree $k$ of the root node is calculated as

$$k = \left\lfloor \frac{S}{RL^+_r} \right\rfloor.$$

The number $r$ of access cycles over which $k$ is calculated determines how fast the protocol responds to traffic variations. The value of $k$ is copied in the ACK frame so that hidden terminals, which cannot deduce its value, are aware of it. We require that the degree $k$ of the root node is at least equal to the degree $m$ of any other node in the tree, $k \geq m$.

Fig. 1 provides an example of an access cycle and the corresponding collision resolution tree built by the proposed scheme. We assume that the degree of the root node, as well as the degree of any other node in the tree, is 5. Moreover, there are five stations that have a packet eligible for transmission at the beginning of the channel access cycle. Sources are represented by hatched circles, while black-filled circles represent nodes of the tree. At the beginning of the channel access cycle, all stations which intend to transmit a packet determine the subtree $T_{1j}$ that they belong to and priority resolution is performed at depth 1. The source which has the packet with the lowest residual lifetime belongs to subtree $T_{1,2}$ of node $n_{1,2}$, so it senses the channel for as many as two slots and then transmits an RTS packet. However, there is at least one more source which belongs to the same subtree and, hence, transmits an RTS packet after sensing the channel for the same number of slots. Since there is a missing CTS message, all stations which belong to subtree $T_{1,2}$ continue contending for channel access by resolving their priorities at depth 2.

At that time, all contending stations are aware of the fact that there are no sources that belong to the subtrees of nodes $n_{2,j}$, where $j = 0, \ldots, 9$. Furthermore, sources $s_{1,3}$, $s_{2,20}$, and $s_{2,24}$ exit contention as they deduce that there is at least one packet with a lower residual lifetime than the residual lifetime of their own head-of-line packet. It should be noted that, there are zero branches emanating from node $s_{1,3}$ since this node is actually a source. The source that owns the packet with the lowest residual lifetime determines that it belongs to the subtree of node $n_{2,11}$, so it senses the channel for one slot and then transmits an RTS packet. Node $n_{2,11}$ is actually a leaf on the tree, so it receives a CTS message and proceeds by transmitting its data packet.

4. Performance evaluation

4.1. Theoretical analysis

The theoretical analysis of the proposed MAC protocol aims at developing a simple analytical model to calculate the average medium utilization.
The probability $P'(i,j,N,m)$ that a station belongs to any subtree $T_{ij}$, where $j \leq p \leq l(i,N,m) - 1$, can be derived from the previous equation:

$$P'(i,j,N,m) = \sum_{p=j}^{l(i,N,m)-1} P(i,p,N,m). \quad (8)$$

Given the previous probability, we can calculate the probability $P_D(i,j,N,m)$ that the first station to transmit an RTS packet belongs to the subtree of node $n_{ij}$. This probability is equal to the probability that all stations belong to the subtree of a node whose index is $j$ or higher and at least one station belongs to the subtree of node $n_{ij}$. That is, all stations should not belong to the subtree of a node whose index is $j + 1$ or higher.

$$P_D(i,j,N,m) = \begin{cases} P(i,j,N,m)^N - P(i,j + 1,N,m)^N, & 0 \leq j < I(i,N,m) - 1, \\ P(i,j,N,m)^N, & j = I(i,N,m) - 1. \end{cases} \quad (9)$$

The mean number $R_{PR}(i,N,m)$ of Priority Resolution Slots required to perform priority resolution at depth $i$ is equal to:

$$R_{PR}(i,N,m) = \begin{cases} \sum_{p=0}^{I(i,N,m)} p \cdot P_D(1,p,N,m), & i = 1, \\ \sum_{p=0}^{I(i,N,m)} (p - \lfloor \frac{p}{m} \rfloor \cdot m) \cdot P_D(i,p,N,m), & i > 1. \end{cases} \quad (10)$$

Eqs. (7) and (8) can be combined to derive the probability that $n$ stations, which belong to subtree $T_{ij}$, are the first to transmit an RTS packet when priority resolution is performed at depth $i$:

$$P_S(i,j,n,N,m) = \begin{cases} \binom{N}{n} P(i,j,N,m)^n P'(i,j,N,m)^{N-n}, & 0 \leq j < I(i,N,m) - 1, \\ P(i,j,N,m)^N, & j = I(i,N,m) - 1, \quad n = N, \\ 0, & j = I(i,N,m) - 1, \quad n < N. \end{cases} \quad (11)$$

By summing up $P_S(i,j,n,N,m)$ for all possible values of $j$, we can calculate the probability $P'_S(i,n,N,m)$ of having $n$ stations transmit an RTS message simultaneously regardless of the index of the subtree they belong to:

$$P'_S(i,n,N,m) = \sum_{p=0}^{I(i,N,m)-1} P_S(i,p,n,N,m). \quad (12)$$
We define as probability of correct scheduling the probability that the station which transmits an RTS message is unique. The probability of correct scheduling when priority resolution is performed at depth $i$ is equal to:

$$P_{CS}(i, N, m) = P'_S(i, 1, N, m). \quad (13)$$

Eqs. (10) and (13) can be used to calculate the mean length $L_{Cycle}(i, N, m)$ of an access cycle, when priority resolution is performed until depth $i$:

$$L_{Cycle}(i, N, m) = l_{CS} + R_{PR}(1, N, m) \cdot l_{PRS} + 2 \cdot l_{VI}$$

$$+ l_{RTS} + l_{CTS} + \sum_{q=1}^{i-1} (1 - P_{CS}(q, N, m))$$

$$\cdot (R_{PR}(q + 1, N, m) \cdot l_{PRS} + 2 \cdot l_{VI}$$

$$+ l_{RTS} + l_{CTS}) + P_{CS}(i, N, m)$$

$$\cdot (l_{Packet} + l_{VI} + l_{ACK}). \quad (14)$$

Then, the mean medium utilization can be derived as a function of the contending population $N$, the degree $m$ of all nodes, other than the root node, in the tree and the depth $i$ until which priority resolution is performed:

$$\text{MediumUtilization}(i, N, m) = P_{CS}(i, N, m) \cdot \frac{l_{Packet}}{L_{Cycle}(i, N, m)}. \quad (15)$$

4.2. Analytical results

The experiments conducted in this study aim at evaluating the performance of the proposed medium access scheme, as well as comparing it to both the EY-NPMA protocol and the DP-TB medium access scheme. The performance metrics of interest are the average medium utilization and the probability of correct scheduling achieved by each scheme. As far as the proposed algorithm is concerned, we define as probability of correct scheduling the probability that, after contending stations perform priority resolution at depth $i$, there is only one station transmitting an RTS message. This station is always the one that had the packet with the lowest residual lifetime at the beginning of the channel access cycle. With respect to EY-NPMA and DP-TB, we define as probability of correct scheduling the probability that the transmitting station (if there are no collisions) is actually the one with the lowest residual lifetime. The performance evaluation of the proposed scheme is based on the analytical results derived in Section 4.1. The performance metrics of interest are examined with regard to the contending population $N$, the degree $m$ of all nodes, other than the root node, in the tree and the depth $i$ until which priority resolution is performed. The theoretical analysis of EY-NPMA and DP-TB can be found in [12] and [23], respectively.

Table 1 presents the values of the operating parameters of our protocol, while the values of the operating parameters of EY-NPMA were set according to the standard. DP-TB was designed to provide 3125 priority levels, by subdividing the prioritization phase in five subphases and allocating a maximum of five slots to each subphase. The triplet $\{m_{es}, m_{y}, p_{e}\}$, which defines the maximum number $m_{es}$ of slots allowed for bursting, the maximum number $m_{y}$ of slots allocated to the yield phase, and the probability $p_{e}$ that a station continues bursting for one more slot, was set to $\{2, 2, 0.3\}$. The values of the operating parameters of DP-TB were the same for both the analytical and the simulation experiments. The channel capacity was set to 23.5 Mbps.

In Section 4.1, we assumed that the distribution of the residual lifetime of packets is given. Many studies that consider a priority scheduler, such as [15], make the assumption that the priority indices of stations (in our case the residual lifetime of packets) are uniformly distributed. Under this assumption, the probability $P^{a}_{RL}(t)$ that the residual lifetime $RL$ of a station’s head-of-line packet at the beginning of an access cycle is less or equal to $t$ is given by the following equation:

$$P^{a}_{RL}(t) = \frac{t}{S},$$

where $S$ is the maximum value that the residual lifetime of packets can be set to. Therefore, we will refer to this case as Case a. Moreover, we make a more realistic assumption that the delay budget

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (bits)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l_{CS}$</td>
<td>705</td>
<td>Length of the channel synchronization interval</td>
</tr>
<tr>
<td>$l_{PRS}$</td>
<td>470</td>
<td>Size of priority resolution slots</td>
</tr>
<tr>
<td>$l_{VI}$</td>
<td>235</td>
<td>Length of the verification interval</td>
</tr>
<tr>
<td>$l_{RTS}$</td>
<td>160</td>
<td>Size of RTS</td>
</tr>
<tr>
<td>$l_{CTS}$</td>
<td>112</td>
<td>Size of CTS</td>
</tr>
<tr>
<td>$l_{ACK}$</td>
<td>112</td>
<td>Size of ACK</td>
</tr>
</tbody>
</table>
DB of flows that packets belong to are uniformly distributed in the range $[0, S]$. Then, the residual lifetime of packets can be considered to be uniformly distributed in the range $[0, DB]$. In this case, which we will refer to as Case b, the probability $P_{RL}^b(t)$ that the residual lifetime $RL$ of a station’s head-of-line packet at the beginning of an access cycle is less or equal to $t$ can be easily calculated to be equal to:

$$P_{RL}^b(t) = \frac{t + t \cdot (\ln S - \ln t)}{S}.$$ 

It should be noted that, the assumption that the residual lifetime of packets are uniformly distributed over the range $[0, DB]$ can be considered as realistic, only if packets are transmitted just before their residual lifetime expires. By applying Eq. (5), which gives the average residual lifetime left to packets before they are transmitted on the channel, when $N$ stations contend for channel access, we can observe that this average value approximates zero even for a low number of contending stations (e.g. $N = 25$). Thus, our assumption can be considered to be realistic.

The first analytical result that we obtain is the achieved throughput of our proposed scheme with respect to the degree $m$ of the underlying $m$-ary tree algorithm. Fig. 2 depicts the mean medium utilization for different values of $m$, $m = 2, \ldots, 6$. The packet size was set to 2383 bytes. The highest throughput is achieved for $m = 3$ and $m = 4$. However, it can be observed that the proposed scheme is relatively insensitive to the value of $m$. Even if $m$ is chosen equal to 6, which is the worst possible value, the protocol’s efficiency will be degraded by only 0.6% compared to the maximum feasible throughput. Henceforth, we assume that $m$ is set to 4 for all analytical and simulation experiments.

In Fig. 3, the mean medium utilization of each scheme for 2383 bytes packet length is presented. From the results, it is apparent that the proposed protocol outperforms DP-TB, as well as EY-NPMA, for any number of contending stations. For the maximum number of contending stations, the efficiency of our scheme is approximately 4.5% higher than that of DP-TB and 10% higher than the efficiency of EY-NPMA. In Case a, the increasing throughput of DP-TB is attributed to its inherently adaptive qualities; the added overhead of DP-TB is reduced as the number of contending stations increases. Nevertheless, the high number of priority levels, which is held fixed for any number of contending stations, prevents DP-TB from achieving high efficiency when the network population is small. In Case b, the number of priority levels of DP-TB can be considered well-chosen for small and medium network populations; yet, a higher number of priority levels would result in better performance for a high number of contending stations. As far as EY-NPMA is concerned, its medium utilization curve exhibits a decreasing trend, which is primarily accounted to its increasing collision rate. When the residual lifetimes of packets are distributed over a shorter range, e.g. Case b, the utilization curve of EY-NPMA decreases even more rapidly due to the low number of priority levels that it provides.

On the other hand, the achieved medium utilization of the proposed scheme is quasi-constant for
any number of contending stations. This positive feature of our protocol is accounted to its ability to adapt to traffic conditions. When the number of contending stations grows, the proposed scheme builds a tree whose degree $k$ of the root node is high enough, so that the number of sources that belong to a given subtree is reduced. In this way, the depth until which priority resolution must be performed in order to determine the source with the lowest residual lifetime is reduced as well. Furthermore, the fact that the performance of the proposed scheme is not affected by the distribution of the residual lifetimes of packets reinforces our observation that our scheme can effectively adapt to traffic conditions.

Fig. 4 illustrates the probability of correct scheduling of each scheme for different numbers of contending stations. The results derived from it make obvious the ability of the proposed scheme to provide QoS assurances, since it can closely approximate an ideal EDF scheduler. The probability of correct scheduling of the proposed protocol is close to 100% even when the priority resolution procedure is performed at a low depth. In this figure, we assumed that priority resolution is performed until depth 7. Moreover, the probability of correct scheduling is unaffected by the distribution of the residual lifetimes of packets. With regard to DP-TB, the high number of priority levels that it provides helps it to correctly select the packet, among those competing for channel access, with the lowest residual lifetime. DP-TB’s probability of correct scheduling is higher than 99% in Case a; yet, it is slightly decreased to 96% in Case b. On the other hand, EY-NPMA struggles to make the correct scheduling decision and its ability to grant access to the packet with the lowest residual lifetime deteriorates severely when the residual lifetimes of packets are distributed over a shorter range.

The probability of correct scheduling of the proposed protocol is higher than 99% when priority resolution is performed at depth 4 and higher than 99.9% when it is performed at depth 6. Consequently, the proposed scheme can quickly detect the packet with the lowest residual lifetime without spending much overhead. In this way, it can achieve high efficiency even in the case of short payloads, as illustrated in Fig. 5. This figure presents the achieved medium utilization of each scheme for 128 bytes packet length. As in the case of large payloads, the performance of our scheme is better than that of DP-TB and EY-NPMA and remains quasicontant regardless of the number of contending stations and the distribution of the residual lifetime of packets. In Case a, the mean medium utilization of the proposed scheme is 21% higher than the maximum achieved throughput of DP-TB and 18% higher than that of EY-NPMA, whereas in Case b, our proposed protocol outperforms DP-TB and EY-NPMA by 12% and 35%, respectively.

Finally, Table 2 presents the achieved medium utilization and probability of correct scheduling for different depths until which priority resolution is performed. The packet size is set to 2383 bytes and the number of contending stations to 250. The presented figures make it evident that the high probability of correct scheduling that our scheme exhibits does not come at the expense of reduced efficiency. Indeed, both the probability of correct
scheduling and the throughput of the protocol are increased as priority resolution is performed at a higher depth.

5. Simulation experiments

5.1. Simulation environment

The tool that was used for the simulation experiments was customly coded by the authors in C++. Regarding the physical channel, the capacity of the common medium was set to 23.5 Mbps and was considered to be ideal; that is, the only reason behind erroneous reception was the simultaneous transmission of more than one stations (packet collision). Furthermore, all network stations were within one hop from each other, thus, eliminating the appearance of hidden/exposed terminals.

Each station initiates a video stream with QoS requirements which are expressed in terms of a maximum allowable delay. Video applications exhibit high variability in the frame sizes and bitrates, especially for low-quality encodings, and as they require stringent delay guarantees, they are expected to benefit significantly from the proposed medium access scheme. Moreover, encoded video traffic is expected to account for large portions of the traffic in future wireline and wireless networks. Upon its arrival, each packet is assigned a lifetime which is equal to the delay requirement of the flow that it belongs to. In contrast, the lifetime assigned to it is equal to the delay requirement of the flow that it belongs to.

Table 2
Priority resolution depth | Probability of correct scheduling (%) | Mean medium utilization (%)
--- | --- | ---
1 | 58.72420 | 85.269
2 | 87.86240 | 87.283
3 | 96.84705 | 87.487
4 | 99.20448 | 87.531
5 | 99.80067 | 87.542
6 | 99.95014 | 87.544
7 | 99.98753 | 87.545

In Fig. 6, the mean medium utilization achieved by each scheme is presented. All of the three schemes exhibit the same throughput under low and medium load conditions. However, as the number of contending stations increases, the superiority of our proposed scheme becomes evident. For the maximum number of contending stations, the medium utilization achieved by the proposed protocol is 11% higher than the throughput of DP-TB and 16% higher than that of EY-NPMA. The increased medium utilization of the proposed mechanism is primarily due to its ability to grant access to the station who has the packet with the lowest residual lifetime. In this way, our protocol serves first those packets whose deadline is about to expire and would be otherwise lost. In contrast,

5.2. Simulation results

Each station is assumed to initiate one video stream. The performance metrics of interest were the achieved medium utilization, the probability of correct scheduling and the packet loss ratio. All of these metrics were examined for different node populations (1–256 stations). Each newly generated flow has a delay budget which is uniformly distributed in the interval [0.5 ms, 20 ms]. Upon an arrival of a new packet, the residual lifetime assigned to it is equal to the delay requirement of the flow that it belongs to.

In Fig. 6, the mean medium utilization achieved by each scheme is presented. All of the three schemes exhibit the same throughput under low and medium load conditions. However, as the number of contending stations increases, the superiority of our proposed scheme becomes evident. For the maximum number of contending stations, the medium utilization achieved by the proposed protocol is 11% higher than the throughput of DP-TB and 16% higher than that of EY-NPMA. The increased medium utilization of the proposed mechanism is primarily due to its ability to grant access to the station who has the packet with the lowest residual lifetime. In this way, our protocol serves first those packets whose deadline is about to expire and would be otherwise lost. In contrast,
EY-NPMA fails to make the correct scheduling decision most of the time. Thus, packets whose residual lifetime is about to expire are preempted by packets who have enough residual lifetime to be transmitted in succeeding access cycles. The higher throughput of DP-TB, compared to the throughput of EY-NPMA, is accounted to its high probability of correct scheduling. However, the slight decrease in the probability of correct scheduling of DP-TB, when the number of contending stations increases, combined with its larger overhead results in reduced efficiency compared to the proposed scheme.

The probability of correct scheduling achieved by each medium access scheme is depicted in Fig. 7. This probability is 100% for the proposed scheme. It should be stressed that, in order to protect the performance of our protocol from the unlikely event that two or more packets have exactly the same residual lifetime, we set the value of the maximum depth until which priority resolution can be performed equal to 15. In the event that the priority resolution procedure reaches that depth, the contending packets are discarded. However, in our simulation experiments there was not such an occurrence. The high number of priority levels that DP-TB provides allows it to make the correct scheduling decision most of the time. The probability of correct scheduling of DP-TB is higher than 93%, even for the maximum number of contending stations. On the contrary, the probability of correct scheduling of EY-NPMA degrades severely as the number of contending stations grows. For the maximum number of stations, the probability that EY-NPMA makes the correct scheduling decision is lower than 17%.

The third performance metric of interest was the packet loss ratio witnessed by each scheme which is illustrated in Fig. 8. When the number of contending stations is less than 25, each scheme exhibits approximately zero packet loss. However, for medium and large contending populations, both DP-TB and EY-NPMA experience an increased portion of lost packets. The number of discarded packets is much higher in the case of EY-NPMA. The reason stems from the inability of EY-NPMA to make the correct scheduling decision. When the offered load increases, packets are transmitted just before their deadlines expire. Consequently, if the medium access protocol fails to make the correct scheduling decision the packet with the lowest residual lifetime will be lost.

The packet loss ratio is a significant metric of the ability of each scheme to provide service assurances, since the QoS requirements of most applications are expressed in terms of both maximum allowable delay and packet loss probability. With respect to real-time video applications, 3GPP defines that their performance will not be harmed as long as their packet loss ratio remains below 1%, while video streaming applications can tolerate up to 2% of their packets being lost [21]. In the first case, the proposed protocol could admit 177% more sessions than EY-NPMA (169–61 admissible sessions) and 15% more sessions than DP-TB (169–147 admissible sessions), while in the second case the admissible region of DP-TB would be 70% better than that of EY-NPMA (184–108 admissible sessions) and 15%
better than the admissible region of DP-TB (184–160 admissible sessions).

6. Conclusions

In this paper we proposed, described, and presented an adaptive medium access control protocol for single-cell ad hoc networks, which is based on m-ary tree algorithms to provide "per-flow" service assurances to traffic flows that can express their QoS requirement as a delay requirement. Based on overheard information, our proposed protocol was shown to easily adapt to traffic conditions. In this way, the efficiency of the protocol in terms of achieved throughput remained quasi-constant regardless of the number of contending nodes. Both analytical and simulation studies were conducted to assess the performance of the protocol. The derived results documented and confirmed the positive characteristics of the proposed mechanism.

References


Orestis Tsigkas received his Diploma in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece in 2002. He is currently working towards his Ph.D. degree in the same department. His research interests include medium access, as well as quality-of-service provisioning.
Fotini-Niovi Pavlidou received the Ph.D. degree in electrical engineering from the Aristotle University of Thessaloniki, Greece, in 1988 and the Diploma in mechanical-electrical engineering in 1979 from the same institution. She is currently a Professor at the Department of Electrical and Computer Engineering at the Aristotle University engaged in teaching for the under- and post-graduate program in the areas of mobile communications and telecommunications networks. Her research interests are in the field of mobile and personal communications, satellite communications, multiple access systems, routing and traffic flow in networks and QoS studies for multimedia applications over the Internet. She is being involved with many national and international projects in these areas (Tempus, COST, Telematics, IST) and she has been chairing the European COST262 Action on “Spread Spectrum Systems and Techniques for Wired and Wireless Communications”. She has served as member of the TPC in many IEEE/IEE conferences and she has organized/chaired some conferences like, the “IST Mobile Summit 2002”, the 6th “International Symposium on Power Lines Communications-ISPLC2002”, the “International Conference on Communications-ICT1998”, etc. She is a permanent reviewer for many IEEE/IEE journals. She has published about 80 papers in refereed journals and conferences. She has served as guest-editor on special issues as: “Mobile Ad Hoc Networks (MANETs): Standards, Research, Applications” in the International Journal of Wireless Information Networks and “Power Line Communications and Applications” in the International Journal on Communications Systems. She is a senior member of IEEE, currently chairing the joint IEEE VT&AES Chapter in Greece.