Adaptive EY-NPMA: A Medium Access Protocol for Wireless LANs

Gerasimos Dimitriadis and Foteini-Niovi Pavlidou

Abstract: **Wireless local area networks have known an increasing popularity during the past few years. However, as new user applications diverge from the traditional data-centric model, the introduction of efficient, QoS aware medium access methods becomes of utmost importance. EY-NPMA is a medium access protocol belonging to the contention paradigm that provides support for service differentiation and low collision rates. In this paper, we address a shortcoming of EY-NPMA as indicated by previous studies, namely the insensitivity of the protocol to different working conditions. In this work, we study and evaluate a mechanism that allows a network employing EY-NPMA to adapt its operating parameters according to the offered load. Simulation studies further document and confirm the positive characteristics of the proposed mechanism.**

Index Terms: **EY-NPMA, medium access, wireless LANs.**

I. INTRODUCTION

Wireless local area networks (wireless LANs) have known an increasing popularity during the past few years. The recent advances that have been achieved in the physical layer have led to bitrates that are similar to those achieved in their wired counterparts, while simultaneously a clear trend towards mobile computing has become visible, as indicated by the penetration of portable computers, personal digital assistants (PDAs), and smart mobile phones in every day activities. The combined effect of the above has been the main drive behind this phenomenal growth in the deployment of wireless LANs. Ad hoc networks have made spontaneous, on-demand networking possible, while infrastructure based wireless LANs have enabled hassle-free, high performance networking without the need for additional wires or retrofitting.

One of the most critical building blocks in any network which employs a common medium, as is the case with wireless LANs, is the medium access control (MAC) sublayer. Medium access protocols define when and how network nodes gain access to the shared medium, thus having a major impact on how efficiently the available raw bandwidth is used. In order to ensure that as little of this finite resource is wasted, designers of medium access protocols work with a multitude of goals in mind. Low collision rates, low overhead, fairness, and robustness are among the most important characteristics that a medium access protocol should possess. Recently, another attribute has been added to this list, namely that of service differentiation. As user applications diverge from the traditional data-centric model, a network may need to simultaneously serve traffic corresponding to diverse content, such as audio, video, and data. In such a case, the medium access mechanism should try to honor the specific needs of each traffic class, without however degrading its performance as far as the other attributes are concerned.

In the field of wireless LANs, medium access protocols may be classified to two major families, according to the design philosophy they follow. The first one includes those that are based on the paradigm of dynamic assignment, while the other consists of contention based protocols. In the case of protocols belonging to the former family, transmissions are coordinated either by a central management entity or by a predefined distributed process that allows a station to transmit without interference from its peers. Techniques such as polling, making reservations, and token passing frequently form the basis for protocols belonging in the dynamic assignment family. These protocols generally result in a collisions-free environment for user data, yet simultaneous transmissions are still possible in many cases, although restricted to small, signaling packets. Furthermore, even though completely distributed dynamic assignment schemes are possible, in most cases a management entity is essential, undertaking roles such as polling, issuing tokens, and validating reservations. In infrastructure based wireless LANs, the obvious candidate for such tasks is the access point (AP), but in ad hoc networks a station is dynamically elected to become channel coordinator (in certain contexts called a clusterhead), usually through clustering techniques. Well known protocols that follow the dynamic assignment paradigm are packet reservation multiple access (PRMA) [1], the medium access layers of the Bluetooth [2], and HIPERLAN/2 [3] standards for wireless LANs, as well as the PCF (point coordination function) access mode of the IEEE 802.11 standard [4].

According to protocols belonging to the latter family—contention based schemes—there is little or no coordination between the stations that are willing to transmit. Under this paradigm, network stations constantly compete with each other in order to gain access to the common medium. One of the major goals in the design of such protocols is the minimization of packet collisions, i.e., simultaneous transmission of two or more stations, usually achieved by dictating that stations receive feedback by the common medium and reducing the aggressiveness with which stations compete for access. One of the widely employed concepts is that of carrier sensing (CSMA), since it prevents network stations from transmitting, when they sense that another transmission is already in progress. Contrary to wired networks, in wireless LANs a station may be in either receive or transmit mode, but not in both. Consequently, in such a case a station may not realize that its transmission is failing and thus enhancements such as collision detection (CSMA/CD) are not possible. For this reason, alternative techniques such as collision avoidance (CSMA/CA) have been devised and are

Manuscript received July 31, 2004.

The authors are with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece, email: {gedimitr, niovi}@auth.gr.

1229-2370/04/$10.00 © 2004 KICS

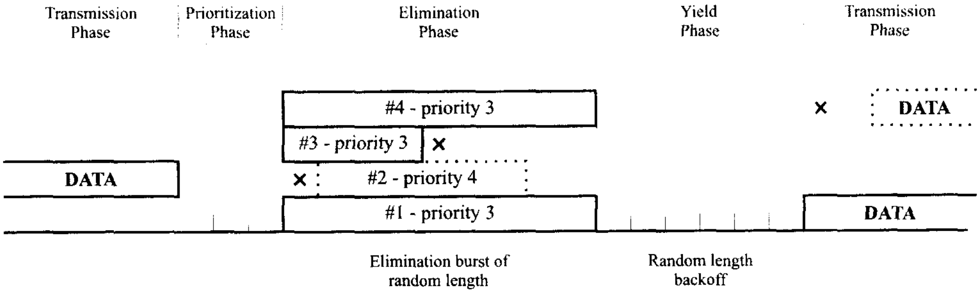


Fig. 1. EY-NPMA's synchronized channel access cycle.

widely employed in wireless networks. Further, the above problem, combined with other wireless LAN specific problems (hidden/exposed terminal [5]) have given rise to handshake employing protocols, such as MACA proposed by Karn in [6]. According to this scheme, packet collisions are restricted between signaling frames (RTS/CTS), while the actual data packets are transmitted collisions-free. Even though contention based protocols have generally an unpredictable behavior because of their stochastic nature, their simplicity and robustness have urged designers to embed QoS capabilities in them. This effort has led to protocols such as Blackburst [7], DFS (distributed fair scheduling) [8], IEEE 802.11e EDCF [9], and HIPERLAN EY-NPMA [10]. On the other hand, the most employed protocol today, DCF (distributed coordination function) of IEEE 802.11 [4] does not support service differentiation. In response to that, there have been many attempts to infuse QoS capabilities in 802.11 based networks [11], [12].

As mentioned above, EY-NPMA is a contention-based medium access protocol that supports service differentiation. EY-NPMA employs contention resolution in multiple levels, achieving this way remarkably low collision rates. In this paper, we address an issue that was reported in previous studies of EY-NPMA, namely the lack of adaptive elements in the mechanism of this medium access scheme. Specifically, we propose a simple yet robust and backwards compatible method that allows EY-NPMA to dynamically alter its working parameters, in order to enhance its performance under variable traffic loads. The rest of the paper is structured as follows. In Section II, we describe the mechanism of the base EY-NPMA scheme. Section III contains the description of the proposed scheme, dealing also with implementation and interworking issues. Section IV deals with the simulation results of the two examined schemes, while finally Section V concludes the paper.

II. EY-NPMA

EY-NPMA stands for elimination-yield non-pre-emptive priority multiple access and is a contention-based protocol. EY-NPMA is part of the HIPERLAN standard for wireless LANs and is designed for completely distributed operation. Even though it belongs to a design paradigm that generally does not scale well to heavy offered loads, EY-NPMA manages to achieve very good performance, even when large numbers of stations simultaneously try to gain channel access. Essentially, EY-NPMA was designed to be used when heavy loads are present; according to the HIPERLAN standard, simple carrier sensing is employed in cases of light traffic. Network stations switch to EY-NPMA as soon as contention begins to build up. The performance of this medium access scheme has been studied both analytically and via simulations in [13] and [14]. Further, it has been compared with DCF and EDCF in [15] and [16], respectively.

According to the HIPERLAN standard, when EY-NPMA is used, network time is divided into cycles. Each cycle begins immediately after the ACK frame of the previous data transmission and ends with the ACK frame of the data transmission that will take place in this cycle. Of course, if an ACK frame is missing—either because of a multicast/broadcast packet or a collision—a timer is used to announce the end of the cycle. Each cycle is further divided into four distinct phases, namely prioritization, elimination, yield, and data transmission. Each of the first three phases aims at reducing the number of stations that are contending for channel access. The first phase, prioritization, rules out all stations, but those that have the currently higher priority packets for transmission. EY-NPMA recognizes five levels of priority ranging from 0 to 4, with 0 being the highest priority. At the beginning of each cycle, each station senses the common medium for as many slots as the priority of the packet in its buffer. If the channel is idle for the whole sensing time, a station proceeds to the next phase, elimination, transmitting an energy burst. Naturally, the stations of highest priority enter elimination sooner than others of lower priority, thus preempting them from further trying to gain channel access at this cycle. During elimination, the stations of the highest priority at the time transmit energy bursts, whose length in slots is random between 0 and a predefined maximum, according to a truncated geometric distribution. As soon as a station finishes bursting, it immediately checks the common medium. If other bursts are still in progress, the station leaves the contention resolution process. On the other hand, the station(s) that has burst the longest survive this phase and proceed to the next one, yield. This phase is the equivalent of a simple backoff, where each station senses the channel for a random number of slots, which lies between 0 and a predefined maximum. In order to minimize the correlation between the backoff periods of different stations, this random number of slots is selected according to a uniform distribution. During yield, the station(s) that senses the channel as idle for the whole backoff period, commence transmitting a data frame. Those that do not, exit contention and will have another chance for accessing the channel at the next cycle. In Fig. 1, an example of an EY-NPMA access cycle is presented. Solid line boxes represent actual transmissions, while dashed line boxes represent projected transmissions that did not take place because the station left the contention process. The **x** marks show when stations left the cycle.

The multi-phase structure of the access cycles employed by EY-NPMA enables this protocol to scale well to very large numbers of simultaneously contending nodes and still provide acceptable rates of collisions. The three parameters of EY-NPMA (i.e., maximum number of slots for bursting and backing off, probability for bursting for one more slot), whose values are defined in the HIPERLAN standard, lead to very low collision rates, 3.5% for 256 simultaneously contending nodes. However, even though this performance is impressive, the cost of this virtual elimination of collisions is the increased overhead that accompanies the protocol, an effect that becomes more intense as shorter data packets are served and higher data speeds are used. This was the main drive behind a number of proposals that aim at reducing the number of slots that are experienced as overhead, without however increasing the very good collision rates of EY-NPMA. The common denominator between these proposals is the addition of memory attributes to the base protocol; according to EY-NPMA, each access cycle is completely independent from the previous ones. The modifications proposed in [17], [14], and [18] reduce protocol overhead by letting network stations to alter their behavior according to the outcome of previous access cycles.

III. ADAPTIVE EY-NPMA

A. Motivation

According to the HIPERLAN standard, the working parameters of EY-NPMA are fixed and are common for all 5 traffic priorities. The value of these parameters are defined in the standard and are the following. The maximum number of slots for which a station may burst during elimination (m_{es}) was set to 9, while the maximum number of slots that a station may backoff during the yield phase (m_y) was set to 12. Finally, the probability p_e that defines the truncated geometric distribution used for deciding for how many slots a station should burst was set to 0.5. The working parameters of EY-NPMA may be expressed as a triplet of values (m_{es}, m_y, p_e), which obviously the HIPERLAN standard defines as $(9, 12, 0.5)$.

The choice of these parameters mirrors the decision of the HIPERLAN designers to optimize medium access for extremely crowded networks. As mentioned in the previous subsection, the triplet that was chosen for EY-NPMA allows it to easily handle 256 simultaneously contending nodes, providing very low collision rates, namely 3.5%. However, even though this performance is truly impressive, the protocol operates in clearly suboptimal regions when the number of contending nodes is much lower, while in fact it is almost impossible that such a high density of network nodes may be observed in practice. Experience has shown that single wireless networks segments consist of a population that is well below 100. Furthermore, the fact that statistically only a fraction of these nodes will be active at a time, combined with the fact that the offered load will be spread on the 5 available priority classes dictates that for most of the time the number of contending nodes will be even lower. Consequently, the tuning of EY-NPMA for 256 simultaneously contending nodes automatically implies that for the greatest part of a network's operation, a significant fraction of each access cycle will be experienced as overhead, since the number of slots used for bursting and backing off will be dysproportional to those that are actually needed.

The HIPERLAN standard does not allow the working triplet (m_{es}, m_y, p_e) to be changed, since these three values are defined in the standard and should be hardcoded during the fabrication process of the network adapter. This inflexibility of EY-NPMA results in a medium access scheme that completely disregards the working environment, thus being confined to a suboptimal region of its operating space. Clearly, significant gains could be obtained if the working parameters of the protocol responded to the offered load and were updated on the fly. Admittedly, EY-NPMA possesses certain inherently adaptive qualities, yet they are not enough to characterise the protocol as adaptive. Specifically, the length of the elimination phase is highly correlated to the number of contending nodes. High levels of contention lead to long elimination phases, while the opposite holds true for low levels of contention. The truncated geometric distribution used for determining the length of the burst for each node is responsible for this behavior, since longer elimination phases appear with increased probabilities, when the level of contention is high. Consequently, the number of slots that are expended for bursting is lower, when fewer nodes simultaneously try to access the shared medium, thus leading to shorter access cycles and less overhead. However, this effect has a limited impact on protocol performance, compared to a scheme that could change its working parameters on the fly, in order to be better tuned to the specific conditions at the time.

Apart from the dependence of the optimal working parameters on the number of simultaneously contending nodes, there is also another characteristic of the offered load that has an effect on the performance of this medium access scheme. This characteristic is the average length of the data payload. For EY-NPMA, the triplet which defines the working parameters essentially balances two different qualities that are both desirable—low collision rates and short access cycles. Arbitrarily low collision rates may be achieved at the cost of long access cycles and vice versa. The efficiency of EY-NPMA is governed on a great degree by the equilibrium of these two conflicting qualities. By intuition, one could justify longer access cycles in the case of large data payloads, since packet collisions are particularly destructive as the payload size increases. On the other hand, the opposite holds true in the case of short payloads. In such scenarios, the time needed for a slot during prioritization, elimination, or yielding is a significant fraction of the time needed for the transmission of the actual payload. Consequently, it is inefficient to guarantee low collision rates, when most of the available capacity is expended as overhead. By introducing a controlled degree of collisions, the average cycle becomes shorter and the medium utilization increases. This relationship between payload size and protocol efficiency is studied and presented in [14].

Based on the above, it becomes evident that the optimal operating point of EY-NPMA depends on two characteristics of the offered load; number of simultaneously contending nodes and payload size. By allowing EY-NPMA to continuously monitor the offered load and periodically reconfigure itself based on these measurements, significant gains could be achieved in terms of throughput, access delay, etc. Furthermore, since the offered load is not spread uniformly throughout the available priority classes, by allowing each priority class to employ its own, custom-calculated operating parameters, further improvement in the protocol's efficiency is anticipated. In this paper, we propose, describe, and present an enhanced version of EY-NPMA, which takes the offered load into account in order to update in real time the parameters of the medium access scheme.

B. Design

In this subsection, we will describe the mechanics of the proposed protocol, especially insisting on the way the offered load is estimated, as well as the calculation of the optimal triplets. Generally, the process that is followed in order to enable a network employing EY-NPMA to dynamically update its working parameters, is provided below.

- Determination of the number of contending nodes and payload size per priority class.
- Calculation of the optimal triplets based on the values obtained from the previous step.
- Diffusion of the optimal triplets to the network population.

Out of the three steps described above, the first one—the estimation of the offered load for each priority class—is the most critical for the operation of the proposed scheme. During this phase, the average payload size, as well as the number of contending nodes must be estimated, in order to provide the input needed to calculate the optimal operating triplets later. The main problem during this phase is that a station is not allowed to actively make an inquiry about these measures; rather a non-intrusive approach must be followed. Out of the two desired metrics, the average payload size is the easiest to obtain. In order to estimate this value, a station has only to record the length of all successfully delivered frames for each traffic class. Subsequently, the average payload size may be calculated, simply by averaging these measurements. At this point it should be noted that it is very important that only the successfully transmitted frames are used for the estimation of the average payload size. Since the length of the transmission phase is governed by the longest frame in the case of collision of two or more packets, the inclusion of these samples would cause a bias towards longer estimated payload lengths.

On the other hand, the estimation of the average number of contending nodes is harder to be extracted by monitoring the access cycles, since there is no direct information about this measure provided in them. However, a successful estimation may be obtained by taking samples of other measures that are linked to the number of contending nodes. As stated in a previous subsection, the length of the elimination phase is highly correlated to the number of the nodes that entered the access cycle. By taking samples of the elimination phase length of subsequent access cycles, the number of contending nodes may be estimated, with an accuracy that increases with the number of available samples. The first step towards deducing the number of contending nodes is the construction of a vector that records the frequency with which the different elimination phase lengths occur. Specifically, the station(s) that undertake this task create and preserve a vector for each priority class L_i where i defines the priority examined and $i \in [0,4]$. Each L_i vector consists of $m_{es,i}+1$ elements $L_i(j)$, where j represents the length of the elimination phase in slots, thus $j \in [0, m_{es,i}]$. Since we allow different operating parameters for each traffic class, $m_{es,i}$ defines the maximum number of slots allowed for bursting for stations contending for access at the i-th priority. Correspondingly, $m_{y,i}$ defines the maximum number of slots allowed for backing off during yielding at the i-th priority, while $p_{e,i}$ the probability with which a station may burst for one more slot at the i-th priority.

Initially, the elements of the vectors L_i are set to 0. The contents of the $L_i(j)$ cell are increased by one for each occurrence at the i-th priority of an elimination phase of j slots. By monitoring many subsequent access cycles, these vectors will soon mirror the frequency with which the various elimination phase lengths occur. By comparing these values with those calculated using an analytical model, an estimation of the number of contending nodes may be deduced. According to the analytical model, the length of the elimination phase depends on the following three values: Number of contending nodes (N_i), maximum number of slots allowed for bursting ($m_{es,i}$), and the probability defining the truncated geometric distribution ($p_{e,i}$). According to [14], the probability that a station bursts for k slots at the i-th priority is given by the following relation

$$P_{E,i}(k) = \begin{cases} p_{e,i}^{k}(1-p_{e,i}), & \text{if } 0 \le k < m_{es,i} \\ p_{e,i}^{m_{es,i}}, & \text{if } k = m_{es,i}. \end{cases} \quad (1)$$

From (1), we can easily calculate the probability that an individual burst lasts k or less slots

$$P'_{E,i}(k) = \sum_{j=0}^{k} P_{E,i}(j) = \begin{cases} 1 - p_{e,i}^{k+1}, & \text{if } 0 \le k < m_{es,i} \\ 1, & \text{if } k = m_{es,i}. \end{cases} \quad (2)$$

Finally, the probability that the elimination phase lasts exactly k slots may be calculated as

$$P_{ED,i}(k) = \begin{cases} (1-p_{e,i})^{N_i}, & \text{if } k = 0 \\ P'_{E,i}(k)^{N_i} - P'_{E,i}(k-1)^{N_i}, & \text{if } 0 < k \le m_{es,i}. \end{cases} \quad (3)$$

Having this information at hand, we can proceed now into designing an estimator that will take as input the vector L_i and provide at its output the estimation of the number of contending nodes at priority i, $\hat{N}_i$. The decision making criterion is to minimize the probability of error in mapping each given observation vector L_i into a decision. Consequently, the optimum decision rule states that

$$P(n \text{ contending nodes} | L_i) \text{ is maximum for } n = \hat{N}_i, \quad (4)$$

where $P(n \text{ contending nodes}|L_i)$ is the probability of having n simultaneously contending nodes, provided that we observed the vector L_i. This rule is able to provide the optimal estimation regarding the number of contending nodes. However, even though this formula is simple and compact, the *a posteriori* probabilities are hard to evaluate. On the other hand, the *a priori* probabilities can be handled much more easily. By employing Bayes' rule, the above decision rule may be expressed as

$$\frac{p_n P(L_i|n \text{ contending nodes})}{P(L_i)} \text{ is maximum for } n = \hat{N}_i, \quad (5)$$

where p_n is the *a priori* probability of having n contending nodes, $P(L_i)$ is the probability of observing this L_i vector, while $P(L_i|n$ contending nodes) is the probability of observing the L_i vector, provided that there are n simultaneously contending nodes. This rule is quite more complex compared to the one provided in (4), yet it can be simplified so that it includes only one factor. By examining (5), we note that the denominator does not depend on the number of simultaneously contending nodes. Further, we can assume that the different numbers of contending nodes are equally probable, thus making p_n also independent of n, an assumption which must be initially made since the estimating station is not aware of the operating environment; optionally, this assumption may be lifted later on, as the station increases its knowledge regarding the distribution of p_n. Consequently, these two probabilities may be omitted and the decision rule may be expressed as

$$P(L_i|n \text{ contending nodes}) \text{ is maximum for } n = \hat{N}_i. \quad (6)$$

Essentially, (6) describes a *maximum-likelihood rule* and the estimator that employs it is referred to as a *maximum-likelihood estimator*. According to this rule, the estimator calculates $P(L_i|n$ contending nodes) for all possible numbers of contending nodes and decides in favor of the maximum. To calculate these probabilities, however, it is necessary that we define the distribution that the vector L_i follows. L_i is essentially a histogram of the elimination phase lengths for priority i, with each element of this vector containing the number of occurrences of a specific elimination phase length. At each i-priority access cycle, one of the $m_{es,i} + 1$ elements of this vector is increased by one, while the probability that element j is chosen to be increased is equal to $P_{ED,i}(j)$, which was defined in (3). From the above, it is evident that L_i follows a multinomial distribution and the probability of having a specific instance of this vector is equal to

$$P(L_i|n \text{ contending nodes}) = \frac{N_s!}{L_i(0)! \cdots L_i(m_{es,i})!} p_0^{L_i(0)} \cdots p_{m_{es,i}}^{L_i(m_{es,i})}, \quad (7)$$

where p_k is the probability of having a k slots long elimination phase at priority i, when there are n simultaneously contending nodes. So,

$$p_k = P_{ED,i}(k)\Big|_{N_i=n}. \quad (8)$$

With the availability of the estimation of simultaneously contending nodes N_i, as well as the average packet length for the i-th priority class, it is easy now to calculate the optimal working parameters of EY-NPMA. Based on the analytical model described in [14], the triplet that leads to the best throughput is derived, corresponding to the offered load, summarized by two values—average payload size and estimated number of simultaneously contending nodes. These two values, however, do not correspond to unique offered load scenarios. Specifically, various payload sizes may appear with different probabilities and still lead to the same average. Likewise, variations in the actual number of simultaneously contending nodes may still lead to the same estimations. Consequently, it is not evident whether the calculation of the EY-NPMA parameters based on these two scalar values truly lead to best throughput. Regarding the former issue, it can be easily proven that the average payload size—in contrast with a full histogram of the appearing payload sizes—is enough to make an optimal decision regarding the EY-NPMA parameters. The model used for the calculation of these parameters aims at maximizing the medium utilization, expressed by

$$mu = \frac{P_{NC} T_{pck}}{T_{cycle}}, \quad (9)$$

where P_{NC} is the probability of not having a collision in a given cycle, T_{pck} is the time needed for the transmission of the payload only, while T_{cycle} is the average duration of a cycle, that is T_{pck} plus all the overhead. Using the fact that the overhead does not change with different payload sizes, it can be proven that the average payload size may be used, instead of the full distribution of the various payload sizes that may appear.

Regarding the latter issue, unfortunately the use of the estimation of simultaneously contending nodes based on the distribution of subsequent elimination phases does not provide guarantees that overall the best performance will be attained. If the number of simultaneously contending nodes varies, the estimator will provide a value that lies between the minimum and maximum values. If the maximum value is much larger than the estimated one, the probability of collisions may increase significantly and intensify contention even more, since backlogged packets will attempt to be retransmitted. Consequently, the estimation of the number of contending nodes does not guarantee optimal performance, even though it is a useful first order approximation. On the other hand, even though it is possible, the probability of abrupt changes in a single estimation session is low. In order to gather 100 samples of successive elimination phase lengths in a 20 Mbps wireless LAN, less than 100 ms are needed. Consequently, if the stations act independently, there is little probability that many stations turn active in such a small time window. Nevertheless, in order to mitigate this problem, we propose two measures. First of all, we rule that the calculation of the optimal parameters are not conducted with the output of the maximum likelihood estimator, rather a more pessimistic value is used. Furthermore, an update containing new parameters may be issued, as soon as the rate of collisions increases above a predefined threshold.

The calculation of the optimal triplet employing this model is computationally intensive; thus it is highly impractical or even infeasible to have a mobile, battery powered device to execute this optimization process, which in order to have a beneficial effect must be repeated frequently. In order to avoid this resource consuming process, we propose the usage of precomputed EY-NPMA parameters for various instances of this optimization problem. According to this scheme, instead of calculating the optimal triplet from scratch, it is much more economical in terms of computational and power resources to find one of the precomputed scenarios that is most similar to the offered load sampled and then use the corresponding triplet. This way, the derivation of the EY-NPMA parameters becomes as easy as searching in a look up table, at the cost of a certain degree of suboptimality. However, by precomputing a sufficient large number of offered load scenarios, the introduced error may be kept to a minimum, being overshadowed in fact by the error introduced while estimating the number of simultaneously contending nodes and the average payload size.

From the above, it becomes evident that the mechanism proposed may be employed in both distributed and centralized ways. However, while it is possible to allow each station in the network to find the optimal EY-NPMA parameters for itself, a number of issues associated with this approach dictate otherwise. First of all, a completely distributed approach demands that all nodes constantly monitor the common medium, an undertaking which is costly in resources, as for example computational power and battery energy. Also, for the estimation of the number of contending nodes, the mechanism presented assumes that all stations employ the same working parameters at all times. Consequently, in a distributed scenario all stations must have synchronized clocks and switch parameters at the same instant. Further, if each network node executes the above outlined process periodically based on the locally stored samples, there is a significant probability that diverse results are computed. Since, networks nodes are randomly added, switched on and off, it is high probable that the calculation of the optimal EY-NPMA parameters is executed with each node having slightly different data. This diversity may lead to groups of stations in the same wireless LAN operating under different parameters. Since the method for finding the optimal working parameters assumes that all network nodes employ the same set of parameters, this error will not only propagate, but grow. On the other hand, the centralized approach negates all the objections provided above. If the estimation of the optimal working parameters is a task allocated to one and only node of the wireless LAN, it is possible to choose for this role the access point to the wired network, which is free of energy concerns. If an access point is not available, high end devices may undertake this role, each for small amounts of time, so that this role does not become too taxing for a single network station and thus, unfair. Furthermore, if one node estimates the optimal working parameters and subsequently diffuses them to the network, there is no possibility that a loss of synchronisation occurs. If a fraction of the network nodes miss the update message transmitted by the node that calculates the optimal parameters, this error will be short lived, since at the next update transmission, all network nodes will become synchronized again. Regarding the bandwidth requirements for the diffusion of the new parameters, the communications cost is negligible, since the contents of the broadcasted updates consist of a few numbers.

As far as different network architectures are concerned, it is obvious that in its current form the proposed medium access scheme poses several restrictions. One of the fundamental assumptions that were made during the design of the adaptive version of EY-NPMA is that at all times all stations are within one hop from each other. Restricting the diameter of the examined networks to unity allows all stations to share a common view of channel conditions, while it also eliminates potential occurrences of hidden or exposed terminals. Furthermore, this restriction enables a station to calculate the optimal working parameters of EY-NPMA by monitoring the common medium and easily diffuse them to all other network stations, since there exists a single, network-wide collision domain. The network may operate either in infrastructure or in ad hoc modes, where in the former case the optimal parameters are calculated by the access point, while in the latter case a station may be randomly elected to take up this role.

As the network diameter increases, a number of issues hindering the operation of the proposed scheme appear. In the case of a network of diameter two, most mechanisms remain unaltered, since a station that neighbors to all other nodes can still be found. Such a station de facto takes up the role of calculating and diffusing the working parameters, using however a different optimization procedure than the one proposed here. In this case, the managing station must take into account phenomena such as the issues of hidden/exposed terminals, as well as the possible space division multiplexing that may appear, since there is not a single collision domain. In generalized multihop scenarios, two different classes of networks may appear. In cases of multichannel environments, two hop wide clusters may be formed, each one using a different channel than its neighbors, with network wide connectivity being guaranteed by gateway nodes. In this case, each cluster internally operates independently and thus the methodology proposed above may be employed. On the other hand, in the case of single channel environments, there are many different, yet correlated collision domains. Since the estimator of the number of contending nodes assumes that all stations employ the same parameters for bursting, there must be a single, global set of EY-NPMA parameters in use. Of course, in this case these parameters should be chosen by taking into account the region with the highest level of contention and subsequently be flooded to all network stations.

C. Compatibility Issues

When performance enhancing modifications are presented, a major design issue is to ensure the proper and hassle free interoperability of the different protocol variants. Even though it is a very significant characteristic, perfect interworking is not guaranteed in previous modifications of EY-NPMA. In [14], the modifications introduced during the prioritization phase alter the chances that each network station has to access the channel, showing bias towards one class of stations. In the case of mixed populations consisting of EY-NPMA/TP and legacy EY-NPMA users, stations employing the base scheme access the channel with higher priority. Since with EY-NPMA/TP each priority class employs two slots during the prioritization phase, legacy users begin bursting much sooner, preempting thus those that employ the modified scheme. On the other hand, in the case of co-existing stations using EY-NPMA and EY-NPMA/ZP, the stations that employ the modified scheme have increased probabilities to gain channel access. According to EY-NPMA/ZP, stations that survive elimination but do not gain channel access are temporarily upgraded to priority 0, disallowing this way all stations of lower priority to contend for the channel. From the above, it becomes clear that legacy users—who do not get such

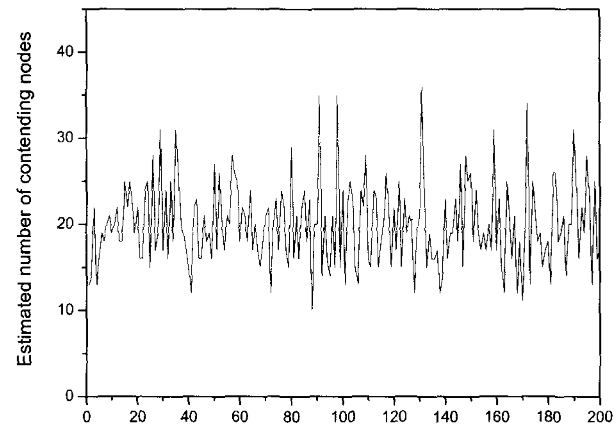


Fig. 2. Successive estimations for 25 analyzed samples.

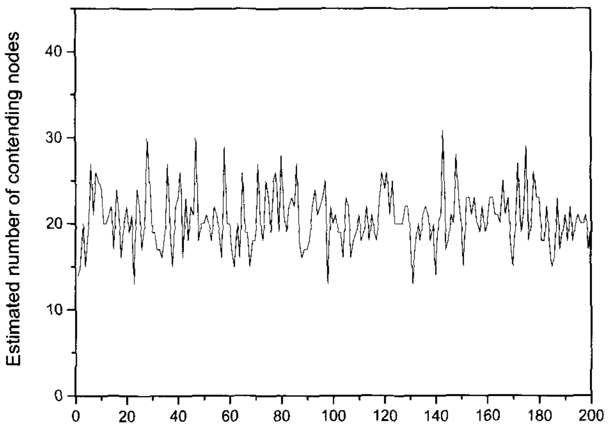


Fig. 3. Successive estimations for 50 analyzed samples.

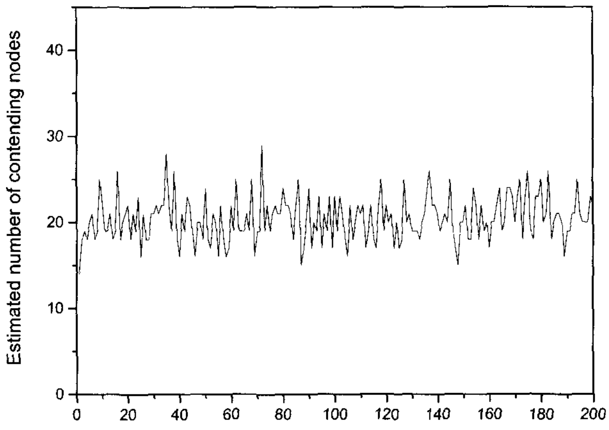


Fig. 4. Successive estimations for 100 analyzed samples.

priority upgrades—get a smaller share of the available bandwidth compared to EY-NPMA/ZP users. In [17], the authors modify the elimination phase, where stations send only the tail and not full length bursts. Consequently, stations employing this scheme on average burst for shorter periods of time compared to legacy EY-NPMA users, thus having a lower probability of gaining channel access.

In this work, no major modifications on the core mechanism of EY-NPMA were conducted. Indeed, the distinct phases of the synchronised cycle—prioritization, elimination, and yield—were not altered. The adaptive variant of EY-NPMA proposed is built upon the base protocol and can provide fair interworking in cases of mixed populations. Let us examine a network consisting of two set of stations, one employing the base EY-NPMA scheme, while the other uses the adaptive version. In this scenario, one of the enhanced stations takes up the role of monitoring the access cycles, computing the optimal working parameters and diffusing them throughout the network. Naturally, as soon as the new parameters are transmitted, a set of stations will adopt them, while the other will disregard them. If the level of contention is much lower than 256 simultaneously contending nodes—which is almost always the case—the computation of the optimal parameters will lead to fewer slots allowed for bursting (m_{es}). The station issuing these updates may consequently sense the existence of legacy users, when monitoring suspiciously frequent long burst, or even worse when bursts exceed the maximum number of slots for bursting which was issued previously. In this case, this station issues again as optimal working parameters the ones that are standardized under HIPERLAN, namely the triplet (9, 12, 0.5). This way, all stations, both legacy and modified, work with the same parameters and fairness between them is established.

IV. PROTOCOL EVALUATION

A. Simulation Environment

The experiments conducted in this work aim at evaluating the performance of the proposed medium access scheme, as well as comparing it with the base EY-NPMA protocol. The tool that was used for these experiments was custom coded by the authors in C++. Regarding the physical channel, the capacity of the common medium was set to 20 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). The duration of each slot employed during elimination was set to 212 bits, while the duration of each yield slot was set to 168 bits, as defined in the HIPERLAN standard. Further, all network stations were within one hop from each other, eliminating thus the appearance of hidden/exposed terminals.

In the first set of experiments, the precision of the mechanism that estimates the number of simultaneously contending stations is assessed. In this scenario, a population of 20 stations contend at every cycle for gaining channel access at the same priority—namely priority 1. The optimal parameter estimation mechanism is triggered periodically, with a period of 0.5 sec. We evaluate the precision of the maximum likelihood estimator by altering the number of available samples of the elimination phase lengths. Specifically, three values are tested, 25, 50, and 100 samples. For the second set of experiments, we evaluate the improvement in performance that the adaptive version of EY-NPMA offers, compared to the base medium access scheme. For this scenario, a variable number of network stations compete in every cycle for channel access at priority level 1. The two protocols are compared in terms of throughput and mean access delay. Under the term access delay, we measure the time needed between a packet reaching the head of its priority queue and its successful transmission. This metric was mainly chosen

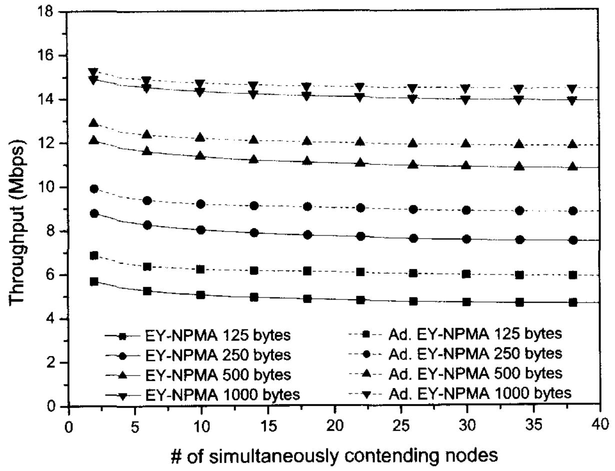


Fig. 5. Throughput vs. number of contending nodes.

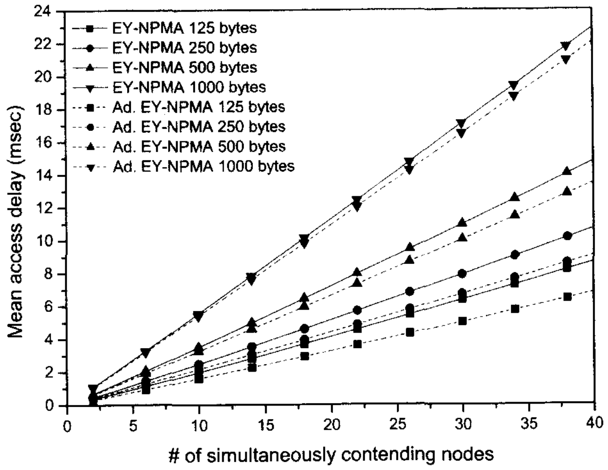


Fig. 6. Mean access delay vs. number of contending nodes.

because it decouples the delay measurements from queue length. The performance of both protocols was evaluated for different packet sizes, ranging from 125 to 1000 bytes long.

For the third set of experiments, the performance of both base and modified schemes was evaluated in terms of throughput and mean access delay, this time however in a multiple priority scenario. In this scenario, a population of 40 stations is offered variable traffic. Three poisson sources are attached to each station, one for each priorities 1, 2, and 3. All sources generate constant length packets that are different between the 3 priorities. Specifically, priority 1 sources generate 250 bytes long packets, while priorities 2 and 3 sources generate 500 bytes and 1000 bytes long packets, respectively. The mean interarrival time of these sources was set, so that the bandwidth ratio between the three priority sources is always 1:1:2. That is if each priority 1 source generates 32 kbps of traffic, priorities 2 and 3 sources generate traffic of 32 kbps and 64 kbps, respectively. At this point, It should be noted that no priority migration scheme was employed in these simulations. Since this task is undertaken by higher sublayer and the mechanism for doing so is not unique, we preferred to not allow packets getting their priority upgraded.

B. Simulation Results

In the first set of experiments, the accuracy of the degree of contention estimates is measured and evaluated. In Figs. 2–4, successive estimates of the maximum likelihood estimator are depicted, showing the relationship between estimation accuracy and number of processed samples. In the scenario examined, the level of contention is constant, with 20 stations willing to transmit at every cycle. With an ideal estimator the results should be flat lines with all estimates being equal to 20. However, because of the stochastic errors and the limited sample size, a certain degree of inaccuracy is present. As was expected, when more elimination phase lengths are sampled and analysed, the output of the estimator deviates less from the actual number of simultaneously contending nodes. When 25 successive samples are captured and analysed, the estimator approximates the real number of contending nodes, yet at times significant errors appear. As the number of samples increases, these errors naturally become smaller. With a sample pool equal to 100, the deviation is clearly less intense.

From these figures, it is clear that more measured elimination phase lengths enable more precise estimations. Even though accurate estimates are attractive, a major practical issue is connected with such an increase of the sample pool. As the number of samples that must be analysed increases, so does the computational cost of the estimation process. Also, the model employed for finding the number of contending nodes is based on the multinomial distribution, which makes heavy use of factorials. With large number of samples, the formula presented in (7) may not be used in a straight forward fashion, since overflows are almost certain to appear.

In the second set of experiments, the two medium access protocols are examined for different payload sizes and simultaneously contending nodes. In Figs. 5 and 6, the achieved throughput and the mean access delay are respectively presented versus the number of simultaneously contending nodes. In Fig. 5, we see that the adaptive version of EY-NPMA outperforms the base scheme for all combinations of payload sizes and contending nodes. However, these gains are more significant when smaller payload sizes are served, in contrast with larger payloads where the difference in performance between EY-NPMA and its adaptive variant is smaller. In the case of 125 bytes long packets, the adaptive version of EY-NPMA demonstrates an increase in throughput that can reach as high as 27%, while when 1000 bytes long packets are served the corresponding increase is only 4%. In the case of small payload sizes, the duration of an elimination or yield slot is a significant fraction of the total cycle time. Consequently, the gain that can be achieved by reducing the number of slots experienced as overhead is much more important in such a case, compared with a longer payload scenario, where the data transmission accounts for the biggest part of the access cycle. The same effect is also present in the results depicted in Fig. 6. As hinted by the previous figure, the adaptive version of EY-NPMA shows better gains in performance in the case of smaller payloads, while for longer packets, even though it is still better, the improvement is not as big. Furthermore, we notice that in both cases the mean access delay increases almost linearly, as more stations enter the contention process.

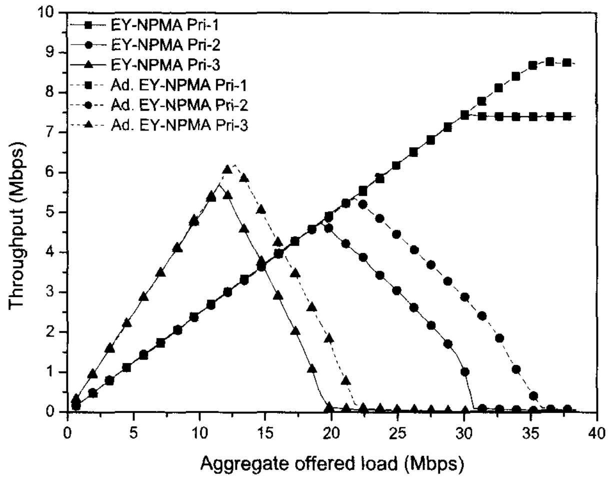


Fig. 7. Throughput per priority vs. offered load.

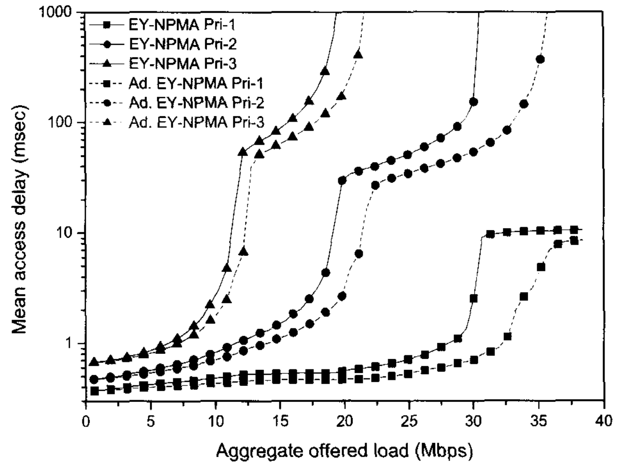


Fig. 8. Mean access delay per priority vs. offered load.

This characteristic is due to the very good scaling attributes of EY-NPMA and its variant, where only a slight degradation in terms of throughput is observed, as the number of contending nodes increases.

For the third and final set of experiment, a scenario with multi-priority traffic is examined. In Fig. 7, the throughput for the three different priorities is presented, while in Fig. 8 the mean access delay is depicted. On a first observation, the results in this scenario confirm what has already been shown in the previous set of experiments, with the network working below its saturation point for wider ranges of offered load and increased throughputs when the working parameters are adaptively changed. However, under further scrutiny another important phenomenon becomes apparent. For the same aggregate offered load, we see that priority 3 traffic achieves better throughput with adaptive EY-NPMA rather with the base scheme and this increase is bigger than what one would expect having Fig. 5 as a guide. While on average in Fig. 5 adaptive EY-NPMA outperforms the base scheme by 0.4 Mbps, in this scenario priority 3 traffic—which also consists of 1000 bytes long payloads—achieves better throughput with the adaptive version by a margin that is well above 1 Mbps. This improvement cannot be accounted only on the better channel utilization done by the stations at this priority class. Rather, the increased efficiency in the higher priorities effects the performance in the lower ones. In this scenario, the more efficient "packing" of data transmissions in priorities 1 and 2, increases the network time that is available for priority 3 transmissions, thus further increasing the throughput at this priority class. The same effect can also be observed for priority 2 traffic. Regarding the mean access delay times, the adaptive variant of EY-NPMA outperforms the base scheme in all examined values of offered load, having a major impact on the network responsiveness, and the degree of quality of service that it can provide. Furthermore, since with adaptive EY-NPMA the network operates below its saturation point under a wider range of offered loads, extremely high mean access delay times appear later with the adaptive version of EY-NPMA, compared with the base scheme.

V. CONCLUSIONS

In this work, we have proposed and evaluated the performance of a medium access scheme based on EY-NPMA. The mechanisms behind the proposed protocol that allow it to achieve better performance rely on the calculation of the optimal working parameters for the currently offered load and its adoption by all network stations. In order to enable a sensing station to estimate the number of contending nodes and the average payload size, a non-intrusive method for doing so was devised, based on the analysis of the characteristics of successive access cycles. The good characteristics of the proposed scheme were confirmed via simulations, where significant gains in performance were witnessed for all scenarios examined.

REFERENCES

[1] D. Goodman *et al.*, "Packet reservation multiple access for local wireless communications," *IEEE Trans. Commun.*, vol. 37, no. 8, pp. 885–890, Aug. 1989.

[2] Bluetooth SIG, "Specification of the Bluetooth system - core v1.0b," 1999.

[3] ETSI, "Broadband radio access networks (BRAN): High performance local area network (HIPERLAN) type II: Data link control (DLC) layer: Part 1: Basic transport functions," *DTS BRAN-0020004-1 v0.m*, 1999.

[4] IEEE, "Wireless LAN medium access control (MAC) and physical layer specification," *IEEE 802.11*, 1997.

[5] F. Tobagi and L. Kleinrock, "Packet switching in radio channels: Part II - the hidden terminal problem in carrier sense multiple-access and the busy-tone solution," *IEEE Trans. Commun.*, vol. 23, no. 12, pp. 1417–1433, Dec. 1975.

[6] P. Karn, "MACA - a new channel access protocol for packet radio," in *Proc. ARRL/CRRL Amateur Radio 9-th Computer Networking Conf.*, 1990, pp. 134–140.

[7] J. Sobrinho and A. S. Krishnakumar, "Quality-of-service in ad hoc carrier sense multiple access wireless networks," *IEEE J. Select. Areas Commun.*, vol. 17, no. 8, pp. 1353–1368, Aug. 1999.

[8] N. Vaidya *et al.*, "Distributed fair scheduling in a wireless LAN," in *Proc. 6-th Annal Int. Conf. Mobile Computing and Networking*, 2000.

[9] M. Benveniste *et al.*, "EDCF proposed draft text," *IEEE Working Document 802.11-01/131r1*, 2001.

[10] ETSI, "Broadband radio access networks (BRAN): High performance radio access network (HIPERLAN) type I: Functional specification," *EN 300 652 v1.2.1*, 1998.

[11] X. Pallot and L. Miller, "Implementing message priority policies over an 802.11 based mobile ad hoc network," in *Proc. IEEE MILCOM 2001*, Oct. 2001.

[12] S.-T. Sheu and T.-F. Sheu, "A bandwidth allocation/sharing/extension protocol for multimedia over IEEE 802.11 ad hoc wireless LANs," *IEEE J. Select. Areas Commun.*, vol. 19, no. 10, pp. 2065–2080, Oct. 2001.

[13] G. Anastasi *et al.*, "HIPERLAN/1 MAC protocol: Stability and performance analysis," *IEEE J. Select. Areas Commun.*, vol. 18, no. 9, pp. 1787–1798, Sept. 2000.

[14] G. Dimitriadis and F.-N. Pavlidou, "Two alternative schemes to EY-NPMA for medium access in high bitrate wireless LANs," *Int. J. Wireless Pers. Commun.*, vol. 28, no. 2, pp. 121–142, Jan. 2004.

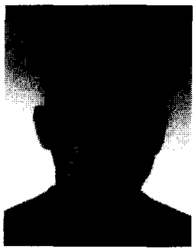
[15] J. Weinmiller *et al.*, "Performance study of access control in wireless LANs - IEEE 802.11 DFWMAC and ETSI RES 10 HIPERLAN," *ACM/BALTZER Mobile Networks and Applications*, vol. 2, no. 1, pp. 55–67, 1997.

[16] G. Dimitriadis and F.-N. Pavlidou, "Comparative performance evaluation of edcf and EY-NPMA protocols," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 42–44, Jan. 2004.

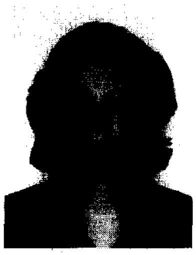
[17] T. Janczak and J. Wozniak, "Modified EY-NPMA channel access scheme," *IEE Electron. Lett.*, vol. 37, no. 10, pp. 655–656, May 2001.

[18] G. Dimitriadis and F.-N. Pavlidou, "On the performance of a medium access scheme based on EY-NPMA in a high bitrate wireless LAN," in *Proc. PWMC 2003, Venice, Italy*, 2003.

[19] A. Santamaría and F. J. López-Hernández, Eds., *Wireless LAN Standards and Applications*, Artech House, 2001.



Gerasimos Dimitriadis received his Diploma in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece in 2001. He is currently working towards his Ph.D. degree in the same department. His research interests include medium access, as well as routing in multihop wireless networks.



Foteini-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 an Associate Professor at the Dept. 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 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 6-th "International Symposium on Power Lines Communications-ISPLC 2002," the "International Conference on Communications-ICT'98," 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.