RESEARCH ARTICLE

Providing throughput guarantees in heterogeneous wireless mesh networks

Antonio de la Oliva1*, Albert Banchs2, Pablo Serrano1 and Frank A. Zdarsky3

1 University Carlos III of Madrid, Madrid, Spain
2 Institute IMDEA Networks, Madrid, Spain
3 NEC Network Laboratories, Heidelberg, Germany

ABSTRACT

In this paper, we propose to provide throughput guarantees in heterogeneous wireless mesh networks by jointly optimizing routing and Medium Access Control configuration. Our solution is based on the notion of linearized capacity region, which provides a technology-independent way of representing the capacity of a wireless link (thereby hiding the technology specifics to the upper layers). From the available capacity of the underlying links as given by the linearized capacity region, we propose two routing algorithms (based on multipath and single path) that find optimal paths for all the flows in the network given their throughput requirements. The throughput allocation resulting from routing is then provided to each link, which uses this information to optimize its technology-specific Medium Access Control parameters. The proposed approach is evaluated in a heterogeneous scenario comprising Wireless Local Area Networks (WLAN) and Worldwide Interoperability for Microwave Access (WiMAX) technologies and is shown to outperform previous solutions by (at least) a factor of 2. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS
wireless; IEEE 802.11; capacity region

*Correspondence
Antonio de la Oliva, Avenida de la Universidad 30, 28911, Leganés, Madrid, Spain.
E-mail: aoliva@it.uc3m.es

1. INTRODUCTION

In the last few years, mesh networking has emerged as a cost-effective and efficient solution for realizing backhaul networks. The multi-hop wireless network architecture of mesh networks enables them to efficiently cover large areas without requiring many interconnections into a wired infrastructure. Furthermore, mesh networks are dynamically self-organized and self-configured, which ultimately results in reduced up-front cost and lower network maintenance costs for the operator. Along these lines, many major operators have already considered wireless mesh networks (WMNs) as a technology for their wireless cities initiatives [1].

Most existing WMN solutions are designed for a single specific radio technology. However, combining different technologies to realize a heterogeneous mesh solution allows for a more flexible designs that benefit from the complementary characteristics of different technologies, for example, extending the wireless mesh connectivity of wireless local area network (WLAN) access points, using a point-to-multipoint non-line-of-sight technology like Worldwide Interoperability for Microwave Access (WiMAX).

A critical concern for operators is to provide their customers with service guarantees. Although already challenging in WMNs with single radio technology, this becomes even more difficult in heterogeneous mesh networks. This is because different radio technologies typically exhibit vastly different link characteristics, in particular in the way capacity is shared between links, and this needs to be considered in admission control, routing, traffic engineering, etc.

In this paper, we aim at designing a solution to provide service guarantees in heterogeneous WMNs, while preserving flexibility and cost-efficiency through a technology-independent approach. The paper first analyzes the capacity region of each wireless link and then proposes a routing algorithm that optimizes performance within the capacity region of all wireless links of the mesh. Our key contributions are as follows: (i) we propose a novel technology-independent way to represent the capacity region of a wireless link, hereafter referred to as linearized capacity region; (ii) we provide a method to
map the capacity region of an 802.11 link to the proposed linearized capacity region; and (iii) we design a novel routing algorithm for heterogeneous WMNs that builds upon the linearized capacity region concept. We evaluate the performance of the resulting approach by means of simulations for heterogeneous WMNs comprising 802.11 and 802.16 technologies and show that it outperforms previous approaches by at least a factor of 2.

The rest of the paper is structured as follows. In Section 2, we propose the linearized capacity region concept, which is mapped to the 802.11 technology in Section 3. From this concept, in Section 4, we present a routing algorithm for heterogeneous WMNs. This algorithm is evaluated and compared against other routing approaches in Section 5. Finally, Section 6 reviews some related work and Section 7 closes the paper with some final remarks.

2. LINEARIZED CAPACITY REGION

In order to allocate resources in a network that comprises wireless links, one needs to know the capacity region of those links. However, these regions are typically very complex to compute and depend on the specific technology used for each wireless link. In this paper, we advocate for the need of a simple and technology-independent way of representing the capacity region of a wireless link, in order to reduce the complexity of resource allocation algorithms.†

Following the previous paragraph, in this section, we present a novel approach to represent the feasible allocations in a wireless link. The key advantages of the proposed approach are twofold:

(i) The approach is technology independent and can be mapped to different technologies as we do in Section 3. Thus, we can use it for resource allocation in heterogeneous mesh networks.

(ii) The proposed solution can be easily combined with efficient algorithms to optimize the mesh network performance, as we do in Section 4.

2.1. Capacity region of a wireless link

Whereas the available capacity of a wired link is a precise notion, it is a blurred notion for wireless links. Take for instance a WLAN link with two nodes. If all the link capacity is allocated to one of the nodes, the available capacity equals the WLAN nominal rate‡ as no time is wasted in collisions. However, if capacity is fairly shared between the two nodes, the available capacity is smaller because of the bandwidth wasted in contention.

We note that, although it may be theoretically possible to compute the exact capacity region of all the technologies and links of a heterogeneous WMN, designing an algorithm for resource allocation that relies on such complex and technology-dependent computations would be highly inefficient. Instead, in the following, we propose a novel notion to characterize the capacity region of a wireless link in a simple and technology-independent way. This notion allows us to easily determine the set of resource allocations that are feasible in the link.

2.2. Proposed concept

Let us consider a wireless link $L$ shared by $N$ flows, where a flow corresponds to the traffic from a given source node to another destination node in the link. Let $R_i$ denote the throughput allocated to flow $i$. The key problem of allocating resources in link $L$ is to determine whether a given allocation $\{R_1, \ldots, R_N\}$ of flows in this link is feasible or not. In wired links, this is straightforward: as long as the total resources in the link do not exceed the link’s capacity $C$, the allocation is feasible; otherwise, it is not feasible, that is,

$$\sum_{i \in L} R_i \leq C \iff \{R_1, \ldots, R_N\} \text{ is feasible}$$  \hspace{1cm} (1)

Determining the feasible allocations in a wireless link is much more difficult because, in contrast to wired links, the total amount of resources allowed is not constant but depends on a number of factors including, for example, the wireless technology used in the link (contention-based technologies waste some resources in collisions, which centralized technologies do not) or the modulation and coding scheme used by each of the nodes of the link.§

In order to provide a way of expressing the capacity region of the wireless link as accurately as possible while avoiding the complexity involved in considering all the above aspects, our key proposal is to linearize the capacity region of a wireless link. In particular, with our proposal, any allocation that satisfies

$$\sum_{i \in L} c_i R_i \leq C$$  \hspace{1cm} (2)

is guaranteed to be feasible, where $c_i$ is defined as the cost of flow $i$ and $C$ as the wireless link capacity. Hereafter,

†The need for a technology-independent representation of the available capacity of a link has already been detected by the 802.21 standard [2], which includes a primitive for this purpose. However, the 802.21 does not address the actual meaning of available capacity in wireless systems and leaves the computation and interpretation of this parameter up to the implementation.

‡By the nominal rate of a technology, we mean the highest data rate provided by this technology. For instance, for the 802.11b technology, we mean a data rate of 11 Mbps.

§Hereafter, by modulation rate, we refer to the rate provided by the modulation and coding scheme used. For example, the modulation rates available with 802.11b are 11, 5.5, 2, and 1 Mbps.
we refer to the capacity region resulting from this linearization as the "linearized capacity region." Note that this way the problem of allocating resources in a network with wireless links becomes as easy as with wired links, as we can determine it with a simple linear function. This concept is shown in Figure 1.

The shortcoming of this approach is that because we are using a lower bound of the actual capacity region, there may be some feasible zones in this region that are not allowed, and this may lead to a suboptimal allocation that does not take full advantage of the wireless link feasible allocations. However, the results presented in Section 5 show that there is no significant performance loss because of this reason.

2.3. Mapping to wireless technologies

The challenge of the proposed model is the computation of the $c_i$ and $C$ values for the different wireless technologies. For technologies with centrally coordinated medium access such as IEEE 802.16, this computation is rather direct. Indeed, as centralized approaches do not waste any bandwidth in contention, channel time is fully used, and as a result, the total capacity depends only on how time is shared among the different stations and the modulation rate that these are using. Therefore, in this case, the boundary of the capacity region can be computed as

$$\sum_{i \in L} c_i R_i = C$$

(3)

where $C$ is the nominal bit rate of the technology and $c_i$ is the ratio between this nominal bit rate and the rate of the modulation scheme used by flow $i$.$^4$

The obtention of the linearized capacity region parameters for contention-based wireless technologies is

$$2 \frac{1}{CW_j}$$

(4)

Let $\rho_{i,j}$ denote the probability that a packet transmitted by station $j$ belongs to flow $i$. Let us further denote by $\tau_i$ the probability that flow $i$ makes a transmission attempt in a randomly chosen slot time. Then,

$$\tau_i = \rho_{i,j} \frac{2}{CW_j + 1}$$

(5)

The throughput provided to a flow can be computed as a function of the $\tau_i$'s as follows [5]$:^5$:

$$r_i = \frac{p(s_i)l}{p(s)T_s + p(c)T_c + p(e)T_e}$$

(6)

$^4$The mapping to contention-based technologies other than IEEE 802.11 can be performed following a similar method to the one outlined in the following section.

$^5$Note that our definitions of slot time duration and probability follow the notion of a slot time given in [5], which defines a slot time as the period elapsed between two backoff counter decrements.

$^6$We note that, like in [5], our throughput analysis assumes saturation conditions, that is, stations always have packets ready for transmission. We argue that this assumption is appropriate to compute the bandwidth allocated to a flow because it provides the maximum throughput that can be obtained by the flow.

3. LINEARIZED CAPACITY REGION MAPPING TO IEEE 802.11

In this section, we first present a model for 802.11 that, given the configuration of the contention parameters of the nodes of a link, computes the throughput obtained by each node. Then, we propose a bandwidth allocation strategy that, given a desired throughput allocation, finds (if it exists) the optimal configuration that satisfies this allocation. Finally, we compute the linearized capacity region resulting from our throughput allocation strategy.

3.1. Bandwidth allocation model

The resulting bandwidth allocation in 802.11 depends on the $CW_{\text{min}}$ and $CW_{\text{max}}$ parameters of each node, which in the latest release of the standard are configurable parameters [3]. Following our previous results in [4], in this paper, we take that $CW_{\text{min}} = CW_{\text{max}}$ as in [4], it was shown (both analytically and via simulation) that no other configuration provides a better throughput performance. We denote by $CW_j$ the CW configuration of node $j$.

Following the analysis in [5], if the backoff window size of a station is constant, the probability that it transmits a packet in a slot time is given by $^*$

$$2 \frac{1}{CW_j + 1}$$

(4)

Let $\rho_{i,j}$ denote the probability that a packet transmitted by station $j$ belongs to flow $i$. Let us further denote by $\tau_i$ the probability that flow $i$ makes a transmission attempt in a randomly chosen slot time. Then,

$$\tau_i = \rho_{i,j} \frac{2}{CW_j + 1}$$

(5)

The throughput provided to a flow can be computed as a function of the $\tau_i$'s as follows [5]$:^5$:

$$r_i = \frac{p(s_i)l}{p(s)T_s + p(c)T_c + p(e)T_e}$$

(6)
where \( I \) is the average packet length; \( T_s \), \( T_c \), and \( T_e \) are the average durations of a successful slot time, an empty one, and a collision, respectively; and \( p(s_i) \), \( p(s) \), \( p(c) \), and \( p(e) \) are the probabilities that a slot time contains a success of flow \( i \), a success of any flow, and a collision and is empty, respectively.

The probabilities are computed as

\[
p(s_i) = \tau_i \left( \prod_{j \in S \setminus S_i} \left( 1 - \sum_{k \in F_j} \tau_k \right) \right)
\]

\[
p(s) = \sum_{i \in F} p(s_i)
\]

\[
p(e) = \prod_{j \in S} \left( 1 - \sum_{k \in F_j} \tau_k \right)
\]

\[
p(c) = 1 - p(e) - p(s)
\]

where \( S \) denotes the set of stations in the link, \( S_i \) is the station to which flow \( i \) belongs, \( F \) is the set of flows in the link, and \( F_j \) is the set of flows of station \( j \).

We can use the following approximations for the probabilities in the previous paragraph:

\[
p(s_i) = \tau_i \left( 1 - \sum_{j \in S \setminus S_i} \sum_{k \in F_j} \tau_k \right)
\]

\[
p(e) = 1 - \sum_{i \in F} \tau_i + \sum_{i \in F} \sum_{j \in F \setminus \{F_1, \ldots, F_{S_i} \}} \tau_i \tau_j
\]

\[
p(c) = \sum_{i \in F} \sum_{j \in F \setminus \{F_1, \ldots, F_{S_i} \}} \tau_i \tau_j
\]

The average slot time durations are computed as

\[
T_s = \frac{1}{p(s)} \sum_{i \in F} p(s_i) T_{s,i}
\]

\[
T_c = \frac{1}{p(c)} \sum_{i \in F} \sum_{j \in F \setminus \{F_1, \ldots, F_{S_i} \}} \tau_j \tau_i T_{c,i,j}
\]

where \( T_{s,i} \) is the duration of a successful transmission of flow \( i \) and \( T_{c,i,j} \) is the duration of a collision between flows \( i \) and \( j \).

\[
T_{s,i} = T_{\text{PLCP}} + \frac{H + l}{C_i} + \text{SIFS} + T_{\text{PLCP}} + \frac{\text{ACK}}{C_i} + \text{DIFS}
\]

where \( T_{\text{PLCP}} \) is the Physical Layer Convergence Protocol preamble and header transmission time, \( H \) is the Medium Access Layer (MAC) overhead (header and frame check sequence), \( \text{ACK} \) is the size of the acknowledgment frame, \( l \) is the packet length, \( \text{SIFS} \) and \( \text{DIFS} \) are time constants defined by the standard [3], and \( C_i \) is the bit rate of the modulation scheme used for flow \( i \).

Finally, \( T_{c,i,j} \) is computed as

\[
T_{c,i,j} = T_{\text{PLCP}} + \frac{H + l}{\min(C_i, C_j)} + \text{EIFS}
\]

where \( \text{EIFS} \) is another time constant defined by the standard.

### 3.2. Bandwidth allocation strategy

On the basis of the model presented previously, in the following, we address the problem of finding the optimal 802.11 configuration to meet a given set of bandwidth requirements. In particular, given a set of desired throughputs for each flow in the link, \( \{R_1, \ldots, R_N\} \), our goal is to find the set of \( CW \)'s that meets these throughput requirements, that is, the configuration that provides each flow \( i \) with an allocated throughput \( r_i \) no smaller than its desired throughput:

\[
r_i \geq R_i
\]

The first step towards finding the \( CW \) configuration is to find the \( \tau_i \) values that provide each flow with the desired throughput. These \( \tau_i \) values are found with the algorithm described next. We first distribute throughput among the competing flows proportionally to their allocated rate,

\[
\frac{r_i}{r_j} = \frac{R_i}{R_j}
\]

which, combined with Equation (6), yields

\[
\frac{\tau_i (1 - \sum_{k \in S_j} \tau_k)}{\tau_j (1 - \sum_{k \in S_i} \tau_k)} = \frac{R_i}{R_j}
\]

The preceding equation can be approximated by \( \tau_i / \tau_j \approx R_i / R_j \).

From the previous equation, we can express all \( \tau_i \)'s as a function of a reference \( \tau_1 \) as

\[
\tau_i = w_i \tau_1
\]

where \( w_i = R_i / R_1 \).
With the preceding equation, the total throughput in the WLAN \( r \) can be computed as a function of \( \tau_1 \) as

\[
    r = \sum_{i \in F} r_i = \frac{l (a \tau_1 + b \tau_1^2)}{c + d \tau_1 + e \tau_1^2}
\]  

(22)

where

\[
    a = \sum_{i \in F} w_i \\
    b = -\sum_{i \in F} \sum_{j \in F \setminus F_s} w_i w_j \\
    c = T_e \\
    d = \sum_{i \in F} w_i T_{s,i} - \sum_{i \in F} w_i T_e \\
    e = -\sum_{i \in F} \sum_{j \in F \setminus F_s} w_i w_j T_{s,j} \\
    \quad + \sum_{i \in F} \sum_{j \in F \setminus F_s} w_i w_j T_{c,j,j} \\
    \quad + T_e \sum_{i \in F} \sum_{j \in F \setminus F_s} w_i w_j
\]  

(23)

In order to find the optimal configuration, we look for the \( \tau_1 \) value that maximizes the \( r_i \)'s. Considering that throughput is distributed among flows following Equation (19), this \( \tau_1 \) value can be found by maximizing \( r \); hence,

\[
\frac{\partial r}{\partial \tau_1} = 0
\]  

(24)

which yields

\[
(a + 2b \tau_1)(c + d \tau_1 + e \tau_1^2) - (a \tau_1 + b \tau_1^2)(d + 2e \tau_1) = 0
\]  

(25)

The preceding equation can be expressed as

\[
A \tau_1^2 + B \tau_1 + C = 0
\]  

(26)

where

\[
A = bd - ea \\
B = 2bc = -2 \sum_{i \in F} \sum_{j \in F \setminus F_s} w_j T_e \\
C = ca = T_e \sum_{i \in F} w_i
\]  

(27)

From solving the preceding second-order equation, we can isolate \( \tau_1 \) and then compute the \( \tau_i \)'s from Equation (20). Once the \( \tau_i \) values have been obtained, we can compute the \( CW_j \) values as follows. From Equation (5),

\[
\sum_{i \in F_j} \tau_i = \sum_{i \in F_j} \frac{\rho_i j 2}{CW_j + 1} = \frac{2}{CW_j + 1}
\]  

(28)

Finally, by isolating \( CW_j \) from the preceding equation, we obtain

\[
CW_j = \frac{2}{\sum_{i \in F_j} \tau_i} - 1
\]  

(29)

which terminates the configuration of the \( CW \) parameters for bandwidth allocation.

### 3.3. Capacity region computation

In the following, we compute the linearized capacity region of a WLAN link that is using the strategy described previously to allocate bandwidth to its flows. In particular, we compute the costs \( c_i \) and the wireless link capacity \( C \) of Equation (2), which provides a mapping of the linearized capacity concept to the 802.11 wireless technology. In order to linearize the capacity region of a WLAN, we first have to choose the tangent point of the linearized region. Indeed, we can choose different tangent points \( \{R_1, \ldots, R_N\} \), which results in different linearized regions. Note that if we chose a tangent point in which \( R_i > R_j \), we are favoring flow \( i \) over flow \( j \), as the linearized capacity region is closer to the actual capacity region for those allocations where flow \( i \) takes larger throughputs.

Figure 2 illustrates three different tangent points that can be chosen to build the linearized region. Tangent point \((b)\), for which \( R_1 > R_2 \), favors flow 1 over flow 2 as it covers well the area where the throughput of flow 1 is high whereas it leaves out the area where flow 2 has a large throughput. In contrast, tangent point \((c)\) favors flow 2 as it covers well the area where \( R_2 > R_1 \).

In order to choose the tangent point for the linearized capacity region, we would like to find an appropriate compromise between favoring those flows that are using higher modulation rates, as allocating throughput to these flows yields a more efficient utilization of the overall wireless resources, and starving those flows that are using lower modulation rates because, in case one of these flows is needed (for instance, because it belongs to a critical path for routing), we would like to be able to use it.
The proportional fairness criterion has been defined precisely to satisfy the above compromise. In this paper, we take the proportional fair allocation as the tangent point for the linearized capacity region. In particular, the proportional fair allocation is the one that satisfies

$$\frac{R_i}{R_j} = \frac{C_i}{C_j}$$

(30)

where $C_i$ and $C_j$ are the modulation rates of flows $i$ and $j$, respectively [6].

Note that in the graph of Figure 2, the proportional fair allocation corresponds to tangent point (a), which is indeed a compromise between allocating more throughput to the better-off flows and not starving the worst-off ones (note that the graph uses different scales for the axes).

The above tangent point can be easily obtained by using the model of Section 3.2 together with Equation (30). We proceed as follows:

(i) Let us consider the function $\sum_i c_i r_i$, where $\{r_1, \ldots, r_N\}$ are the boundaries of the capacity region as computed in Section 3.2. With this function, we proceed as follows to compute the parameters of the linearized capacity region:

1. Any allocation $\{R_1, \ldots, R_N\}$ that satisfies $\sum_i c_i R_i \leq C$ falls within the capacity region.
2. The proportional fair allocation satisfies equality, that is, $\sum_i c_i R_i = C$.

We take the following approach.

According to Equations (2) and (30), our goal is to find the $c_i$ and $C$ parameters such that

1. Any allocation $\{R_1, \ldots, R_N\}$ that satisfies $\sum_i c_i R_i \leq C$ falls within the capacity region.
2. The proportional fair allocation satisfies equality, that is, $\sum_i c_i R_i = C$.

Let us consider the function $\sum_i c_i r_i$, where $\{r_1, \ldots, r_N\}$ are the boundaries of the capacity region as computed in Section 3.2. With this function, we proceed as follows to compute the parameters of the linearized capacity region:

(i) We first compute the parameters $c_i$ by imposing that $\sum_i c_i r_i$ is minimized for the $\{r_1, \ldots, r_N\}$ point that corresponds to the proportional fair allocation.

(ii) Next, we obtain the value of $C$ by evaluating the function $\sum_i c_i r_i$ at this point.

We next show that with this procedure we satisfy the two objectives stated previously. Indeed, objective 2 is clearly satisfied by (ii). On the other hand, (i) imposes that all the points in the boundary of the capacity region satisfy $\sum_i c_i R_i \leq C$. This implies that any point that satisfies $\sum_i c_i R_i \leq C$ will fall within the capacity region, and therefore, objective 1 is also met.

Following the preceding paragraph, the remaining challenge is to compute the costs $c_i$ for which function $\sum_i c_i r_i$ takes a minimum at $R_i = w_i R_1$, where flow 1 is taken as reference and

$$w_i = \frac{C_i}{C_i}, \ i \in \{2, \ldots, N\}$$

(31)

The preceding equation yields the following system of equations:

$$\frac{\partial \sum_j c_j r_j}{\partial w_i} \bigg|_{w_i = C_i/C_i} = 0, \ i \in \{2, \ldots, N\}$$

(32)

Substituting $r_i$ by the expression given in Equation (6) and operating on the equation, we obtain

$$\frac{\partial}{\partial w_i} \sum_j c_j p(s_j) (p(s) T_s + p(c) T_c + p(e) T_e)$$

$$- \sum_j c_j p(s_j) \left( \frac{\partial p(s) T_s}{\partial w_i} + \frac{\partial p(c) T_c}{\partial w_i} + \frac{\partial p(e) T_e}{\partial w_i} \right) = 0$$

(33)

We compute the first term of the preceding equation as follows

$$\frac{\partial}{\partial w_i} \sum_j c_j p(s_j) p\left( \frac{\partial p(s) T_s}{\partial w_i} + \frac{\partial p(c) T_c}{\partial w_i} + \frac{\partial p(e) T_e}{\partial w_i} \right)$$

$$= c_i \frac{\partial p(s_i)}{\partial w_i} + \sum_j c_j p(s_j) \left( \frac{\partial p(s) T_s}{\partial w_i} + \frac{\partial p(c) T_c}{\partial w_i} + \frac{\partial p(e) T_e}{\partial w_i} \right)$$

(34)

where

$$\frac{\partial p(s_i)}{\partial w_i} = \tau_1 + w_i \frac{\partial \tau_1}{\partial w_i} - \sum_j w_j \tau_1^2$$

(35)

and

$$\frac{\partial p(s_j)}{\partial w_i} = w_j \frac{\partial \tau_1}{\partial w_i} - w_j \tau_1^2 - \sum_k w_k w_j \tau_1 \frac{\partial \tau_1}{\partial w_i}$$

(36)

for $F_i \neq F_j$ and

$$\frac{\partial p(s_j)}{\partial w_i} = w_j \frac{\partial \tau_1}{\partial w_i} - \sum_k w_k w_j 2 \tau_1 \frac{\partial \tau_1}{\partial w_i}$$

(37)

for $F_i = F_j$.

To compute $\frac{\partial \tau_1}{\partial w_i}$, we take the expression of $\tau_1$ given by Equation (26) and proceed as follows,

$$\frac{\partial \tau_1}{\partial w_i} = \frac{-B' + 2 B B' - 4 A'(A' C + A' C')}{4 A^2}$$

(38)

where $A'$, $B'$, and $C'$ are the derivatives of parameters $A$, $B$, and $C$, respectively, given in Equation (26):

$$A' = \frac{\partial b}{\partial w_i} d + \frac{\partial d}{\partial w_i} a - \frac{\partial a}{\partial w_i} e$$

$$B' = -4 T_e \sum_{j \in F \setminus F_i} w_j$$

$$C' = T_e$$

(39)

The partial derivatives included in the preceding expressions are computed as follows:

$$\frac{\partial a}{\partial w_i} = 1$$

(40)
A. de la Oliva et al.  Providing throughput guarantees in heterogeneous wireless mesh

\[
\frac{\partial b}{\partial w_i} = -2 \sum_{j \in F \setminus S_i} w_j \quad (41)
\]

\[
\frac{\partial c}{\partial w_i} = 0 \quad (42)
\]

\[
\frac{\partial d}{\partial w_i} = T_{s,i} - T_e \quad (43)
\]

\[
\frac{\partial e}{\partial w_i} = - \sum_{j \in F \setminus F_i} w_j T_{s,i} - \sum_{j \in F \setminus F_i} w_j T_{s,j} + \sum_{j \in F \setminus F_i} w_j T_{c,i,j} \quad (44)
\]

Finally, the rest of the terms of Equation (33) are computed as follows:

\[
\frac{\partial p(s)T_s}{\partial w_i} = \sum_{j \in F \setminus F_i} \frac{\partial p(s_j)}{\partial w_i} T_{s,j} + \frac{\partial p(s_i)}{\partial w_i} T_{s,i} \quad (45)
\]

\[
\frac{\partial p(c)T_c}{\partial w_i} = \sum_{k \in F \setminus F \setminus F} \sum_{j \in F \setminus F_k} w_k w_j T_{c,k,j} 2 \tau_1 \frac{\partial \tau_1}{\partial w_i} \quad (46)
\]

\[
\frac{\partial p(e)}{\partial w_i} = - \tau_1 - \sum_{j \in F \setminus F_i} w_j \frac{\partial \tau_1}{\partial w_i} + \sum_{j \in F \setminus F_i} w_j \tau_1^2 + \sum_{k \in F \setminus F \setminus F} \sum_{j \in F \setminus F_k} w_k w_j 2 \tau_1 \frac{\partial \tau_1}{\partial w_i} \quad (47)
\]

By substituting Equations (33)–(47) into Equation (32), we obtain a system of \( N - 1 \) equations on the \( c_i \)'s. Note from Equation (2) that we have one degree of freedom when fixing the \( c_i \)'s and \( C \).

Therefore, we can set without loss of generality \( c_1 = 1 \), and as a result, we have a system of \( N - 1 \) equations with \( N - 1 \) unknowns. As this system of equations is linear, it can be easily solved by means of standard techniques like Cramer.

From the preceding discussion, we can compute all the \( c_i \)'s by resolving the system of equations, and then, we can compute \( C \) from

\[
C = \sum_{i} c_i f_i \bigg|_{w_i = c_i/c_1} \quad (48)
\]

which terminates the computation of the mapping to the linearized capacity region.

4. ROUTING AND BANDWIDTH ALLOCATION

The key objective of the proposed linearized capacity region is to aid the design of efficient algorithms to optimize network performance. In particular, the proposed model allows us to easily determine if a given flow allocation is feasible, and therefore, it is very useful for the design of efficient optimization algorithms. To illustrate this, in this section, we present a routing algorithm for mesh networks that relies on the proposed linearized capacity region in order to provide throughput guarantees.\(^{86}\)

The specific optimization problem that we address in this section is stated as follows: given a mesh network consisting of a set of wireless links with corresponding linearized capacity region, a set of gateways, and a set of flows, with each flow \( i \) originating at source node \( N_i \) and having a throughput requirement \( R_i \), we want to find a route for each flow to any of the gateways of the mesh network such that the throughput requirement of each flow is met and the number of admitted flows is maximized.

We assume that the mesh network has a proper radio resource management, such that packet transmissions do not fail because of interference from neighboring links. This assumption is supported by our measurements in [7], which show that as long as channel separation is large enough, different channels do not interfere with each other. Moreover, the use of rate adaptation techniques further mitigates the impact of interference.

We consider two scenarios: (i) a multipath (MP) scenario in which each flow can be split and distributed over different paths and (ii) a single-path (SP) scenario in which each flow is treated as an atomic, unsplittable entity and therefore it can be routed through only one SP.

Let us start with the MP routing problem. This problem can be viewed as a multicommodity flow problem [8], which has been widely studied in the literature, and can thus be formulated as the following linear programming problem.

Let \( r_{i,l} \) be the rate allocated to flow \( i \) on link \( l \), where here a link denotes a pair of directly connected nodes, as opposed to the notion of wireless link that we used in the previous sections to refer to a set of nodes that can communicate with each other by sharing a common wireless capacity (see Figure 3 for an example of a wireless link with three nodes and the corresponding links). Let \( s_i \) denote the set of links originating at the source node of flow \( i \). Furthermore, let \( N_{in} \) denote the set of links that terminate at node \( N \) and \( N_{out} \) the set of links that leave from this node. We further denote by \( I \) the set of nodes that are neither sources nor gateways. Following the

\[86\] It is important to note that the presented algorithm is only an example to show the potential of the proposed concept. Indeed, the linearized capacity region can be used to solve other optimization problems such as network planning, traffic engineering, or admission control.
subject to

\[ \sum_{i \in S_j} r_{i,l} = R_i, \forall i \]  

(50)

\[ \sum_{l \in N_{in}} r_{i,l} = \sum_{l \in N_{out}} r_{i,l}, \forall i, N \in I \]  

(51)

\[ \sum_{l \in L} c_l \sum_{i \in l} r_{i,l} \leq C_L, \forall L \]  

(52)

The above definitions, our objective is to find the allocation that satisfies

\[ \min \sum_{i,l} r_{i,l} \]  

(49)

subject to

\[ x_i = \sum_{l \in S_j} y_{i,l}, \forall i \]  

(54)

\[ \sum_{l \in N_{in}} y_{i,l} = \sum_{l \in N_{out}} y_{i,l}, \forall N, i \]  

(55)

\[ \sum_{l \in L} c_l \sum_{i \in l} y_{i,l} R_i \leq C_L, \forall L \]  

(56)

\[ x_i \in \{0, 1\}, \forall i \]  

(57)

\[ y_{i,l} \in \{0, 1\}, \forall i, l \]  

(58)

The interpretation of the above problem formulation is as follows. We aim at maximizing the number of routed flows \( \sum x_i \) subject to flow \( i \) being originated at node \( s_i \) (Equation (54)) while meeting the flow conservation and capacity constraints (Equations (55) and (56)) and imposing that flows cannot be split among different paths (Equations (57) and (58)).

The above linear programming problem can be solved by using standard techniques. As a result of this, we obtain a routing strategy that admits as many flows as possible while meeting the desired throughput guarantees.

We next address the SP routing problem. This can be viewed as another very widely studied problem in the literature, which is the unsplittable flow problem, and can be formulated as the following integer programming problem.

Let \( x_i \) be 1 if flow \( i \) is routed and 0 otherwise. Furthermore, let \( y_{i,l} \) be 1 if the path chosen for flow \( i \) traverses link \( l \) and 0 otherwise. Then, we want to find the allocation that satisfies

\[ \max \sum_i x_i \]  

(53)

subject to

\[ x_i = \sum_{l \in S_j} y_{i,l}, \forall i \]  

(54)

\[ \sum_{l \in N_{in}} y_{i,l} = \sum_{l \in N_{out}} y_{i,l}, \forall N, i \]  

(55)

\[ \sum_{l \in L} c_l \sum_{i \in l} y_{i,l} R_i \leq C_L, \forall L \]  

(56)

\[ x_i \in \{0, 1\}, \forall i \]  

(57)

\[ y_{i,l} \in \{0, 1\}, \forall i, l \]  

(58)

The above integer programming problem can be solved by using standard relaxation techniques [9], which provides as a result an approximation to the optimal SP routing strategy, which solves the SP routing problem.

It is worth noting that, because each point of the capacity region of a link corresponds to a given configuration of the MAC parameters, when solving our routing problem, we are implicitly taking into account all possible MAC configurations, and therefore, we are looking for the optimal solution that provides the best performance considering all possible routing and MAC configurations. As a result, the obtained solution jointly optimizes routing and MAC. Once we obtain the routing solution, we then commit the MAC configuration that corresponds to the desired point on the capacity region.

The performance of this routing algorithm and the MP one are evaluated in the next section and compared against other routing algorithms for mesh networks.

---

**Figure 3.** Wireless link with three nodes.

---

Note that, although we impose the requirement that the entire flow’s throughput is originated at its source, we do not impose any restriction regarding its destination. The reason is that the flow does not have a specific destination and we just want to send it to any of the gateways in the mesh. This is a key difference between the classical multicommodity flow problem and the formulation presented here.
Table I. Model validation.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>N</th>
<th>Simulation (Mbps)</th>
<th>Analytical (Mbps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single rate (11 Mbps), $\alpha = 0.5$</td>
<td>8</td>
<td>7.26</td>
<td>7.22</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>7.22</td>
<td>7.18</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>7.20</td>
<td>7.16</td>
</tr>
<tr>
<td>Single rate (11 Mbps), $\alpha = 0.1$</td>
<td>8</td>
<td>8.25</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>8.25</td>
<td>8.20</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>8.25</td>
<td>8.20</td>
</tr>
<tr>
<td>Two rates (1/2 stations (11, 5.5 Mbps)), $\alpha = 0.5$</td>
<td>8</td>
<td>5.28</td>
<td>5.26</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>5.26</td>
<td>5.24</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>5.25</td>
<td>5.23</td>
</tr>
<tr>
<td>Multiple rates (1/4 stations (11, 5.5, 2, 1 Mbps)), $\alpha = 0.1$</td>
<td>8</td>
<td>7.09</td>
<td>7.10</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>7.06</td>
<td>7.09</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>7.02</td>
<td>7.04</td>
</tr>
</tbody>
</table>

5. VALIDATION AND PERFORMANCE EVALUATION

In this section, we validate and evaluate the performance of the schemes proposed in this paper. Unless otherwise stated, the experiments with WLAN are based on the 802.11b physical layer.

5.1. Model validation

In order to validate the model presented in Section 3, we performed several experiments and compared the results of our model against those obtained via simulation. The simulator we used was developed in OMNET++ and closely follows the details of the 802.11 protocol.

We simulated a scenario consisting of $N$ flows sharing the same WLAN link, where each flow $i$ sent a throughput proportional to its weight $w_i$. The weights were allocated according to $w_{i+1}/w_i = \alpha/1 - \alpha$, where $\alpha$ is a variable parameter that we used to set different throughput distributions ($\alpha = 0.5$ corresponds to equally distributing throughput among all flows whereas smaller values of $\alpha$ yield a more uneven distribution).

Table I gives the obtained simulation and analytical results. Results are given for different $\alpha$ values and number of stations for different scenarios. The results show that the analytical values closely follow the ones from simulation because the error is well below 1% in all cases. We conclude from these results that the analytical model is very accurate.

5.2. Linearized capacity region: two flows

In order to evaluate the accuracy of the linearized capacity region mapping to 802.11 proposed in Section 3, we performed the following experiment. We considered two flows sharing the WLAN. In the first scenario, both flows sent at a modulation rate of 11 Mbps (hereafter, we refer to this scenario as ‘homogeneous rates, 11 Mbps’). In the second scenario, one flow is sent at a modulation rate of 11 Mbps and the second one at 1 Mbps (‘heterogeneous rates’). The results for the three scenarios are shown in Figures 4 and 5, respectively. The exact capacity region as given by the exact throughput model presented in Section 3 is plotted with a continuous line. The approximate capacity region as given by the approximation resulting from Equations (11)–(13) is plotted with squared dots. Finally, the linearized capacity region is computed in Section 3.3, with a dotted line. As Figures 4 and 5 show, the actual capacity region and the approximate one coincide exactly for the homogeneous case and are quite close for the heterogeneous case. This confirms the accuracy of our approximation. Also, the linearized capacity region covers most of the area of the actual capacity region, which means that by linearizing the capacity region we do not waste significant resources. Indeed, the largest deviation between the linearized region and the actual one, which occurs in the axes, does not exceed 10%.

http://www.omnetpp.org/
5.3. Linearized capacity region: multiple flows

In the above experiments, only two flows are considered. In order to understand the impact of the number of flows into the behavior of the linearized capacity region, we evaluated the following additional scenarios.

In the first scenario, we considered a varying number of flows sending at the same modulation rate of 11 Mbps (‘single rate’). In the second scenario, one fourth of the flows transmit at each of the following modulation rates: 11, 5.5, 2, and 1 Mbps (‘multiple rates’).

For each of the above scenarios, we consider a traffic pattern like the one described in Section 5.1, where the traffic distribution depends on a parameter $\alpha$. For each value of $\alpha$ and number of flows, we depict the total capacity allowed according to the exact capacity region (‘exact capacity’) and the linearized one (‘linearized capacity’).

The results for the single-rate and multiple-rate scenarios are given in Figures 6 and 7, respectively. Note that because in the single-rate scenario the linearized capacity does not depend on $\alpha$, in Figure 6, there is only one line drawn for the linearized capacity.

The main conclusion that we draw from the above results is that the number of flows has a fairly small impact on the linearized capacity region. Indeed, for the multiple-rate scenario the difference between the exact and linearized capacity does not change noticeably with the number of flows, whereas for the single-rate scenario, there is a noticeable change but it is not very significant. These results show that the conclusions given above for two flows also hold for multiple flows.

5.4. Routing: homogeneous scenario

The previous experiments have validated the mapping of the linearized capacity region to 802.11. In the following, we evaluate the performance of the proposed routing algorithm based on the linearized capacities. In this section, we focus on a homogeneous WMN consisting of WLAN links only, whereas in the next section, we consider a heterogeneous network consisting of WLAN and WiMAX.

In order to conduct a performance evaluation independent of the chosen topology, we generated multiple random topologies and evaluated the average performance (and its deviation) among all topologies. To generate these random topologies, we used the Hyacinth–Laca tool, available at http://www.ecsl.cs.sunysb.edu/multichannel/ used in several well-known works such as [10] and [11].

We configured this tool to create random topologies with a node count between 40 and 70 nodes (which yields a mean of 55 nodes) spread over an area of $400 \times 400$ m$^2$. Once a topology is available, before performing a routing experiment, we need to assign the channels used by each interface. For this purpose, we used a channel assignment policy that follows a Common Channel Set configuration [12–14]. In order to calculate the modulation rate at which each node is able to communicate with its neighbors, we further used the curves of throughput versus distance given in [15].
We evaluated routing performance for 10 gateways and a varying density of source nodes (25%, 50%, and 100% of the nodes). The metric that we used to evaluate the routing performance is the maximum amount of traffic that can be admitted to the network while providing all flows with the same throughput. For each experiment, we generated a set of 35 random topologies, and we provide the average throughput performance and confidence intervals over the throughput resulting from each topology. In order to show the performance improvement resulting from the proposed SP and MP schemes, we compared them with well-known routing approaches for mesh networks, namely the expected transmission count (ETX) [13], the expected transmission time (ETT) [16], and the shortest path (ShP).

The results on the routing performance of our approaches (SP and MP) against previous proposals (ETX, ETT, and ShP) are given in Figure 8. From these results, we observe that our two approaches, SP and MP, perform similarly, which means that the performance gain resulting from using MPs is limited. This is because SP can route two flows originating at the same access point (AP) through different paths and the additional benefit of splitting individual flows with MP is low. Even more, our two algorithms clearly outperform the other approaches, that is, they outperform ETX and ETT approximately by a factor of 2 and ShP by a factor of 3.

Among other reasons, the performance improvement of our approach over ETX and ETT is due to the fact that ETT and ETT are additive metrics and, as a result a path consisting of a few rather congested links, may be preferred over a longer and less congested path, which harms throughput performance. Furthermore, ETT and ETX do not take into account that the flows that belong to the same wireless link share the same resources; in contrast, our approach considers that allocating throughput to one flow harms the other ones in the same link. We must also consider that by setting the configuration of the MAC parameters according to the algorithm presented in Section 3, our algorithm jointly optimizes routing decisions and MAC configuration, which results in an improved performance. We conclude from the above results that our method is effective in optimizing throughput performance, making an efficient use of the linearized capacity region and clearly outperforming traditional approaches.

5.5. Routing: heterogeneous scenario

In the following, we evaluate the performance of our routing algorithm in a heterogeneous WMN based on WLAN and WiMAX technologies. The parameters for WiMAX used in our simulations and the parameters for the linearized capacity region are given in Table II, where $c_{\text{uplink}}$ and $c_{\text{downlink}}$ are the costs of the uplink and downlink flows and $C$ is the wireless capacity.

The scenario we simulated to evaluate the routing performance is the same one as in the previous experiment, with 50% of the nodes acting as sources and 10 as gateways. To this scenario, we added a variable number of nodes with WiMAX capabilities (5, 10, and 20 WiMAX nodes) and an additional node that can act as a WiMAX base station, depending on the actual experiment performed.

In order to compare our approach against other proposals, we consider ShP and the well-known ETX [13] and ETT [16] metrics, which we extend to account for the heterogeneity of our scenario. These extensions, denoted as ETX-H and ETT-H, respectively, are described as follows:

- ETX-H: The original ETX metric is based on the number of retransmissions required to transmit a packet over a given link. Hence, it favors paths with low utilization. We have extended this metric by assigning a value of 1 to all WiMAX links that have enough bandwidth to serve new flows, and a value of $\infty$ to those that do not have enough bandwidth. Indeed, as long as there is enough bandwidth to serve a flow, in WiMAX, all flows can be served with a single retransmission, whereas otherwise the number of retransmissions required is infinite.
- ETT-H: The original ETT uses as a metric of a link the time required to transmit a packet over the link and hence chooses the links with the smallest delays. In order to extend this metric for a heterogeneous scenario, we assign a value of $1/BW$ to all WiMAX links, where the $BW$ value corresponds to the available bandwidth in the link.

![Figure 8. Routing, homogeneous scenario, 10 gateway nodes.](image)

**Table II.** Linearized region and physical parameters for Worldwide Interoperability for Microwave Access.

<table>
<thead>
<tr>
<th>Permutation mode</th>
<th>PUSC</th>
<th>UL</th>
<th>2.2 Mbps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame duration</td>
<td>5 ms</td>
<td>DL</td>
<td>8.4 Mbps</td>
</tr>
<tr>
<td>FFT size</td>
<td>512</td>
<td>C</td>
<td>UL + DL</td>
</tr>
<tr>
<td>Modulation</td>
<td>64QAM</td>
<td>$c_{\text{uplink}}$</td>
<td>C/UL</td>
</tr>
<tr>
<td>FEC code</td>
<td>3/4</td>
<td>$c_{\text{downlink}}$</td>
<td>C/UL</td>
</tr>
</tbody>
</table>

DL, downlink; FEC, forward error correction; FFT, fast Fourier transform; PUSC, partial usage of the subchannels; UL, uplink.
One of the first and major contributions is the seminal work of Gupta et al. [17], which, on the basis of a geometric analysis, provides an upper bound on the maximum capacity of a network where every node is able to share any portion of the channel it is using with any of its neighbors. Further extensions of this work, for example, [18,19], account for the geographic locations and the transmission power of the rate tuples at which a reliable communication is possible. The problem is studied for the case of ad hoc networks with infrastructure support in [20] and in [21] for the case of WMNs, where different regions of the theoretical maximum capacity are derived, depending on the relative ratio of mesh clients, routers, and gateways. These works are devoted to the computation of limiting upper bounds on the capacity and, as such, cannot be used to support optimal routing or perform admission control.

With respect to routing algorithms for WMNs, the first proposals were based on algorithms already available for mobile ad hoc networks (e.g., hop count); however, given that WMNs significantly differ from MANETS [22] (e.g., because of the infrastructure support), these are far from providing optimal performance. New metrics have been proposed for WMN routing [13,16,22,23], although typically tailored to a particular technology, for example, IEEE 802.11: ETX [13] is based on the number of attempts to send a frame using the lowest-modulation probes; ETT [16] extends it to account for the physical rate and frame length used; minimum loss [23] proposes to find the route with the minimum end-to-end loss probability, whereas modified ETX and effective number of transmissions [22] extend ETX to account not only for average values but also for standard deviations. Our approach significantly differs from these approaches because of two major reasons: (i) our proposal is technology agnostic as the routing algorithm is oblivious to the technology supporting a given set of weights and a capacity boundary and (ii) for the particular case of IEEE 802.11, our computation algorithm optimally configures the wireless network, this way minimizing the number of collisions and rendering metrics like ETX that are based on retransmissions useless.

Because of the evolution of wireless networks towards heterogeneity [24], researchers have recently started to address the definition of a quality-of-service-aware metric supporting different technologies [25,26]. Most of the solutions proposed so far are based on hop counts or nominal bandwidth [27], this way being unable to support smart routing. A first attempt to use a linear model to optimize the throughput allocation is the work of [28]. However, only the case of 802.11 is considered, and the performance of the MAC protocol is not taken into account—authors assume that the nominal capacity coincides with the achievable capacity of the WLAN.

7. CONCLUSIONS

In this paper, we have proposed a novel approach for throughput allocation in heterogeneous mesh networks comprising different wireless technologies.

A key feature of our approach is that it relies in a technology-independent interface between the routing layer and the underlying link layers. Such a design decision
allows hiding the technology specifics from the underlying links to the routing layer, which is essential for the support of heterogeneous technologies. We note that many previous routing solutions (e.g., ETT and ETX) rely on technology-specific parameters, which makes them unsuitable for heterogeneous mesh networks based on different technologies.

In order to represent the available capacity of a wireless link in a technology-independent way, we propose the notion of *linearized capacity region*. The key advantages of our proposal are as follows: (i) it relies on a few parameters, which yields a simple interface between link layer and routing; (ii) by properly adjusting its parameters, the linearized capacity region covers most of the actual capacity region, which provides a high level of efficiency; and (iii) by relying on a linear function, it allows for solving network-wide optimization problems at a low computational cost.

The proposed linearized capacity region concept relies on a number of parameters that need to be computed for each of the technologies present in the mesh network. Whereas such a mapping is relatively straightforward for centralized technologies, whose capacity region is already linear, it is more challenging for distributed technologies. In this paper, we have addressed the mapping of the linearized capacity region to a distributed technology like WLAN, by first proposing an algorithm that configures the WLAN to optimize throughput performance and then linearizing the capacity region resulting from this configuration.

In this paper, we have further proposed two routing algorithms on top of the linearized capacity region concept. The first algorithm assumes that flows have the ability of being split over MPs (following recent standardization work at the Internet Engineering Task Force) whereas the second algorithm is based on traditional SP flows. The proposed algorithms result from linear and integer programming formulations, in which the capacity available at each wireless link is given by a linear constraint derived from the linearized capacity region. The approaches proposed in this paper have been extensively evaluated by simulations and have been shown to vastly outperform previous solutions.

ACKNOWLEDGEMENTS

The work presented in this paper has been partially funded by the CARMEN (carrier-grade mesh networks) EU project, which involves major European operators. For more information, see http://www.ict-carmen.eu/. The work of Albert Banchs and Antonio de la Oliva has also been partially funded by the Spanish Government, MICINN, under research grant TIN2010-20136-C03.

REFERENCES

null
Dr. Frank A. Zdarsky holds a joint master’s degree in Electrical Engineering and Business Administration (Dipl.-Wirtsch.-Ing.) from the University of Technology Darmstadt, Germany, and a doctorate degree (with distinction) in Computer Science from the University of Kaiserslautern, Germany. Since 2008, he has been a project manager at NEC’s European Networking Laboratories. Among other responsibilities, he is/was a technical manager of the European Union projects BeFEMTO and CARMEN, respectively. Frank has published more than 40 research articles in the areas of femtocells, mobile backhaul networks, radio resource management, network security, and performance engineering.