IEEE 802.16j RELAY-BASED WIRELESS ACCESS NETWORKS: AN OVERVIEW

VASKEN GENC, SEAN MURPHY, YANG YU, AND JOHN MURPHY, UNIVERSITY COLLEGE DUBLIN
SCHOOL OF COMPUTER SCIENCE AND INFORMATICS

ABSTRACT

Multihop wireless systems have the potential to offer improved coverage and capacity over single-hop radio access systems. Standards development organizations are considering how to incorporate such techniques into new standards. One such initiative is the IEEE 802.16j standardization activity, adding relay capabilities to IEEE 802.16 systems. This article provides an overview of this relay-based technology, focusing on some of the most pertinent aspects. In particular, the different modes of operation (transparent and non-transparent), framing structures, and network entry procedures are described. Some consideration of the issues in designing such systems is then given, which highlights when different features within the standard are most appropriate. As these systems are very new, many open issues remain to be resolved.

INTRODUCTION

Multihop wireless access networks are the subject of much interest at present. In the last few years they have moved from the domain of an interesting research topic to one having a significant impact on the commercial world. Two of the most visible manifestations of this are the increased standardization activities relating to multihop networking and the growth of companies with product portfolios based on such technology (e.g., Tropos, Belair).

There are a number of different types of multihop wireless networks, notably ad hoc networks, sensor networks, and wireless mesh networks. Each of these network types has different characteristics—mobility rates, power constraints, scale, form factor, and so on—which result in different system design (e.g., routing protocols, medium access control mechanisms). Another type of multihop wireless network that is the subject of some attention is based on a relay architecture. This is the focus of this article.

Relay-based systems typically comprise small form factor low-cost relays, which are associated with specific base stations (BSs). The relays can be used to extend the coverage area of a BS and/or increase the capacity of a wireless access system [1, 2]. Typically, it is envisaged that they could be used in the early stages of network rollout to provide coverage to a large area at lower cost than a BS only solution; they can also be used to provide increased capacity in more developed networks, as coverage holes such as areas in the shadows of buildings.

Most consider relay systems to be most applicable in network operator contexts, where the operator plans and deploys a wireless access network operating in some licensed spectrum. Typically, they are characterized by tree-based routing in which end terminals connect to the BS over short routes (2–3 hops). However, even within this scope, many different designs are possible, and there is much work to be done to understand the most appropriate use cases for different designs.

One of the most important initiatives within the context of relay networking architectures is the IEEE 802.16j standardization initiative.1 This initiative, which is currently in its final stages, is developing solutions for relay-based networks that can operate with legacy 802.16e-2005 end terminals: the group is working on developing new BS and relay station (RS) capabilities to enable relay networks to be realized and provide support for access by legacy devices. A key driving force for this work is the fact that there are many parts of the world in which IEEE 802.16 technologies based on [3, 4] have not been rolled out (for a number of reasons): an initial rollout based on relay architectures would potentially facilitate much greater coverage at lower cost.

This article, then, provides an overview of the most important aspects of the current version of the 802.16j standard. As with most standards, there are a large number of possible options, and it remains unclear which are most appropriate for which contexts. Furthermore, the draft is still evolving, although most of the major deci-
sions have been made at this point. Hence, the
description provided here attempts to capture
the most essential aspects of the current draft,
which should see little modification in the final
standard.

The article is structured as follows. In the
next section the IEEE 802.16j standardization
activity is presented, including some context
relating to other 802.16 work, the scope of the
standardization activity, and the timelines. This
is followed by a description of the key aspects of
the current draft of the standard. The following
section discusses some of the different design
decisions that must be made when realizing
802.16j systems. Finally, the article is concluded.

The IEEE 802.16j Standardization Activity

Before describing the current draft of the stan-
dard in detail, it is useful to give some context.
Here, the 802.16j standardization activity is dis-
cussed: first, there is a brief description of relat-
ed 802.16 work; the scope of the 802.16j
standardization activity is presented next; and
finally, the 802.16j timelines are discussed.

Context of the 802.16j Initiative

The original 802.16 standard was published in
2001 as a solution for fixed broadband wireless
access (FBWA) [5]; the standard described an
orthogonal frequency-division multiplexing (OFDM)-
based point-to-multipoint (PMP) solution. It was
superseded by 802.16-2004, which incorporated
lower-frequency operation, and most systems shipping
today that do not have mobility support are compliant
with this standard. Mobility support added with the 802.16e-
2005 standard has now been implemented, and
802.16e systems are currently in rollout (e.g., the
Sprint XOHM service).

At present, the IEEE 802.16 Maintenance
task group is working on an updated version of
these two standards that will be merged into one
standard, probably known as IEEE 802.16-2008.
One key modification is the removal of the mesh
mode defined in IEEE 802.16-2004 to allow
multihop communication among the BS and sub-
scriber stations (SSs). Its removal is due to its
incomplete specifications, which were unlikely to
be completed, and its lack of support for PMP
and mobility.

The other major effort within 802.16 is de-
veloping an 802.16-based solution for input to the
International Telecommunication Union’s
(ITU’s) IMT-Advanced initiative. This is the
mandate of the 802.16m group2 that is working
on new radio interfaces and system architectures
to support highly mobile high-data-rate commu-
nications — up to 100 Mb/s mobile and 1 Gb/s
fixed. A key differentiator between 802.16m and
802.16j is that the former is not constrained by
legacy issues and hence is at liberty to design an
entirely new radio access system.

Motivation for 802.16j

802.16j is developing a standard to support relay
mode operation for 802.16 systems. A number of
different use cases are envisaged, as depicted in

- Figure 1. Use cases for 802.16j.

Fig. 1. The two main use cases can be broadly
divided into increased coverage and capacity
enhancement. The former can be further subdi-
vided into two slightly different use cases:
extending the coverage range of a BS using mul-
tihop techniques and addressing coverage hole
problems (e.g., shadows of buildings). Increased
system capacity can be achieved through use of
multiple links with greater efficiency, as opposed
to single-hop links over poor-quality channels;
also, multihop communications can support spa-
tial reuse, which can result in increased overall
system capacity.

The 802.16j task group is expected to define
RSs with significantly lower complexity than
802.16e-2005 BSs. Hence, by using such RSs an
operator could deploy a network at a lower cost
than using only (more) expensive BSs to provide
wide coverage while delivering a required level
of service to users. Low cost and ease of deploy-
mation are the key to the business case for 802.16j,
and are issues that require further investigation
to understand where the viable markets lie. In
order to appreciate more clearly the differences
between an 802.16e-2005 deployment and an
802.16j deployment, Table 1 provides the key
differences between the two technologies.

Timelines

The 802.16j task group commenced its activities
in March 2006. The fifth draft [6] of the 802.16j
specification was released in May 2008, and the
standard is expected to be ratified in late 2008
or early 2009. It is important to note that this is
the first step in a longer process: after ratifica-
tion it is necessary to define appropriate profiles
for interoperability testing within the WiMAX
Forum, and then there will be laboratory trials.
Rollout of 802.16j systems will not happen for
another two to three years.

IEEE 802.16j Specifications

The current 802.16j draft [6] is a very substantive
document, detailing many aspects of operation
of an 802.16j system. It is difficult to capture all
The standard defines two different relay modes of operation: transparent mode and non-transparent mode. The key difference between these two relay modes of operation is how framing information is transmitted.

<table>
<thead>
<tr>
<th>802.16e-2005</th>
<th>802.16j</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology</td>
<td>PMP only</td>
</tr>
<tr>
<td>Hops</td>
<td>Single hop</td>
</tr>
<tr>
<td>Traffic aggregation</td>
<td>No</td>
</tr>
<tr>
<td>System capacity</td>
<td>Lower</td>
</tr>
<tr>
<td>Coverage</td>
<td>Lower</td>
</tr>
<tr>
<td>Cost</td>
<td>Higher</td>
</tr>
<tr>
<td>Legacy 802.16e-2005 stations</td>
<td>—</td>
</tr>
<tr>
<td>Mobility support</td>
<td>Yes</td>
</tr>
<tr>
<td>PHY support</td>
<td>OFDMA</td>
</tr>
</tbody>
</table>

Table 1. Comparison of 802.16j and 802.16e-2005 capabilities.

Aspects of such a comprehensive document succinctly, the focus here is on those aspects of the standard that differ significantly from previous standards. The specific aspects of the standard detailed here are the different relay modes, the differences in the PHY, focusing only on the frame structure, and a number of differences at the medium access control (MAC) layer. These are described in the following subsections.

**RELAY MODES**

The standard defines two different relay modes of operation: transparent and non-transparent. The key difference between these two relay modes of operation is how framing information is transmitted: in transparent mode operation, the relays do not transmit frame header information; but in non-transparent mode, the relays do transmit frame header information. The frame header contains essential scheduling information the nodes use to determine when they can transmit and receive information. This difference in treatment of framing information gives rise to very different characteristics between the two operating modes.

Related to the relay modes are two different options for scheduling: centralized and distributed. In the former, all scheduling for all nodes in the system takes place in the BS; in the latter, the RSs have some autonomy and can make scheduling decisions for the nodes with which they communicate.

The two different relay modes are:

**Transparent mode:** In this mode the RSs do not forward framing information, and hence do not increase the coverage area of the wireless access system; consequently, the main use case for transparent mode relays is to facilitate capacity increases within the BS coverage area. This type of relay is of lower complexity, and only operates in a centralized scheduling mode and for topology up to two hops.

**Non-transparent mode:** The RSs generate their own framing information or forward those provided by the BS depending on the scheduling approach (i.e., distributed or centralized). They can support larger coverage areas and hence are mainly used to provide increased coverage. On the other hand, the transmission of the framing information can result in high interference between neighboring RSs; hence, the capacity enhancement that can be achieved using these relays is limited. Non-transparent relays can operate in topologies larger than two hops in either centralized or distributed scheduling mode, leading to different levels of complexity at the RS.

As transparent relays (T_RS) and non-transparent relays (NT_RS) have different advantages, there can be some scenarios in which it makes sense to associate both with a single BS. The standard, however, gives little detail on how this can be realized. Table 2 illustrates the difference between both relay modes.

**PHY LAYER SPECIFICATIONS: FRAME STRUCTURE**

As the frame structure defined in the earlier 802.16 standards was designed for single-hop wireless networks, modifications were required to support relay network architectures. As with the former 802.16 frame structure, the frame is divided into two subframes: downlink (DL) and uplink (UL). However, unlike the previous frame structure, these subframes are further divided into zones to support BS-RS communications: different zones facilitate communications between different parts of the system. In both transparent and non-transparent modes, so-called access zones are defined that support BS/NT_RS communications with the MS/T_RS. In the transparent mode a so-called transparent zone is defined for T_RS communications with the MS; in non-transparent mode relay zones are defined for BS/NT_RS communications with NT_RS. As there are differences between the frame structures in the different relay modes, these are described separately below.

**Frame structure in transparent relay mode**
Centralized scheduling only
Non-transparent RS

<table>
<thead>
<tr>
<th>Coverage extension</th>
<th>Transparent RS</th>
<th>Non-transparent RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Number of hops</td>
<td>2</td>
<td>2 or more</td>
</tr>
<tr>
<td>Inter RS cell interference</td>
<td>None</td>
<td>High</td>
</tr>
<tr>
<td>Performance</td>
<td>In BS coverage: high</td>
<td>In BS coverage: same as 16e</td>
</tr>
<tr>
<td></td>
<td>Outer BS coverage: —</td>
<td>Outer BS coverage: medium</td>
</tr>
<tr>
<td>RS Cost</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Scheduling</td>
<td>Centralized scheduling only</td>
<td>Centralized/distributed scheduling</td>
</tr>
</tbody>
</table>

Table 2. Comparison between transparent and non-transparent modes of operation.

As with the former 802.16f frame structure, the frame is divided into two subframes: a downlink subframe and an uplink subframe. However, unlike the previous frame structure, these subframes are further divided into “zones” to support BS-RS communications and RS-Mobile Station communications.

**MAC Layer Specifications**

Three different aspects of the MAC layer are presented here: the forwarding scheme, routing and path management mechanisms, and initial ranging and network entry mechanisms.

**Forwarding Scheme** — Two different forwarding schemes are defined; both are intended to maximize system efficiency by aggregating traffic where possible: the tunnel-based scheme and the connection ID (CID)-based scheme. Such traffic aggregation has two main benefits: it can result in gains in system efficiency as less signaling information is sent, and it results in simpler management as groups of flows can be handled together. The tunnel-based scheme provides support for explicit tunnels characterized by a unique CID, two specific endpoints, and quality of service (QoS) requirements. The CID-based scheme has no such tunnels and does not explicitly support traffic aggregation, but requires less complexity. In the tunnel-based approach tunnels are used to aggregate traffic from disparate MSs on the BS-RS connection for either management or transport connections with similar QoS requirements. The CID-based scheme, on the other hand, supports only the legacy management and transport connections as defined in the 802.16f-2005 standard [4]. These schemes can be differentiated in terms of QoS management, error handling, and overhead.

**The tunnel-based scheme**: In this mode, the station at the ingress of the tunnel adds a so-called relay MAC header to a packet or group of packets indicating the CID of the tunnel the packet(s) should traverse. This header is removed when the packet arrives at the destination RS. Use of the tunnel requires some intelli-
gence at the RSs along with the distribution of the different service flows’ parameters to the RSs in both centralized and distributed scheduling modes.

The CID-based scheme: The packets are forwarded based on the CID of the destination station. In centralized scheduling the BS sends a message to the RSs describing the relay link channel characteristics, including an extra field specifying the delay associated with each packet in either the DL or UL. Thus, the RS knows in which frame each packet should be transmitted. This is necessary in order to meet the QoS requirements of each connection. In the distributed case the RS has knowledge of the QoS requirements of each connection and can therefore make its own scheduling decisions.

Routing and Path Management — As 802.16j networks comprise multihop paths between the BS and MS, routing and path management issues arise. Although routing in such systems is tree-based, there can be decisions to be made regarding with which RS a particular MS should be associated. Path management refers to issues relating to path establishment, maintenance, and release for which different path management techniques have been proposed.

Routing/path selection: The standard provides for routing decisions based on metrics such as radio resource availability, radio link quality, and traffic load at the RSs, but does not indicate how the decision should be made — the particulars of the path selection decision are left to vendors. The decisions are made at the BS based on information provided by the RSs. The path management mechanisms are then used to create the path in the appropriate way.

The standard defines two approaches to path management: embedded and explicit. The key difference between the two approaches is how signaling information to manage the path is distributed in the system.

The embedded path management approach: In this approach a hierarchical CID allocation scheme is used in the system. The BS allocates CIDs to its subordinate stations such that the CIDs allocated to all subordinate RSs of any given station are a subset of the allocated CIDs for that station. In this way there is no specific routing table in each RS, and a reduced need for signaling to update path information. This is a very simple approach to path management.

The explicit path management approach: This mode uses an end-to-end signaling mechanism to distribute the routing table along the path. The BS sends the necessary information to the RSs involved in a path when a path is created, removed, or updated. Each path is identified by a path ID to which the CIDs are bound. This leads to a smaller routing table at the RSs and a reduction of the overhead required to update these tables. Optionally, the BS may include the QoS requirement associated with each CID to allow the RSs to make an independent decision regarding how to schedule the packet in distributed scheduling mode.

Initial Ranging and Network Entry — There are two quite different aspects to network entry in 802.16j: differences related to network entry for the MS and network entry procedures for the RSs. These are considered here.

As 802.16j must maintain compatibility with legacy terminals, the network entry procedure as seen by the terminal must remain unchanged. However, there are some differences regarding how the BS and RSs deal with this procedure arising from the fact that the network needs to determine which node should be the access node for the MS.

The initial ranging process in 802.16j systems varies depending on the scheduling mode and relay mode: the different ranging process can be distinguished as follows:

MS initial ranging in transparent mode relay: The RSs monitor the ranging channel in the UL access zone and forward the ranging codes they receive to the BS. The BS waits a specified time for other messages with the same ranging code — from other RSs — and then determines the most appropriate path for the station (i.e., direct or via an RS). If the direct path is chosen, the
BS sends a response directly to the MS. Otherwise, the response is sent to the RS, which then forwards it to the MS.

**MS initial ranging in non-transparent mode relay** Due to legacy constraints, the MSs choose the BS or NT_RS with the strongest preamble detected. This means that there is essentially no path decision to be made in this case, as the MS is communicating with a single RS. As the BS ultimately makes the network entry decision, the RS must communicate with the BS to ensure that the MS is permitted network entry. In the centralized case this involves communicating all ranging information back to the BS, but in the distributed case the RS handles the ranging functions and simply makes a network entry query of the BS.

The network entry process for RSs incorporates additional steps and defines a specific ranging process. More specifically, network entry is augmented with a neighborhood discovery and measurement process followed by a path selection algorithm to determine the most suitable access station for the RS.

**RS initial ranging:** During the process the BS (or NT_RS) can determine if a node performing ranging is an RS based on the ranging code used — a specific set of ranging codes are reserved for RSs. In this way transparent mode RSs can easily ignore ranging performed by other RSs. Also, this enables priority to be given for RSs performing ranging. The rest of the initial ranging process for RSs is then similar to that for MSs in non-transparent mode relay.

**RS network entry:** After the initial ranging, authentication, and registration processes, the BS may request the RS to determine the signal strength of each of its neighboring RSs and forward it to the BS. The BS can then determine the most suitable access station with which to associate the RS based on traffic load, signal strength, and so on. The final stage of network entry is then the configuration of the RS parameters, including its operation mode (e.g., transparent or non-transparent) and scheduling mode.

### Design Considerations for 802.16J Systems

The 802.16j standard will specify a wide variety of options. For implementers, it is necessary to consider which subset of these options is most appropriate for different use cases. Some of these options are listed in Table 3. The WiMAX Forum will define specific profiles, and implementers will be guided by its work; however, there are many unclear issues as yet, and the WiMAX Forum will have much work to do to determine appropriate profiles for different scenarios. The following, then, are some of the issues that warrant further study within this context.

**Centralized/distributed mode:** Clearly, the choice of a centralized or distributed system has profound ramifications throughout the entire system. For larger systems in which there are more than two hops to the BSs, the benefits of distributed systems are most probably greater. However, many deployments are likely to be smaller topologies in which most communications will be via one or two hops; in these cases the centralized mode of operation seems preferable as it will result in lower-complexity RSs, ultimately leading to lower overall costs.

**Relay modes:** The essential difference between the relay modes is quite clear, and the choice of relay modes is largely based on whether the objective is to provide increased coverage or capacity. However, there is a need for systems that can support both relay modes, and the standard is only recently considering this.

**Forwarding scheme:** In 802.16e-2005 the MAC overheads increase as the number of connections increases, and this is compounded in 802.16j systems. A key advantage of the tunnel mode over the CID-based mode is the aggregation of data through packet concatenation, packing, and fragmentation mechanisms. However, tunnel creation and management leads to further complexity at the RSs and overhead associated with tunnel management in the system. The CID-based approach must be used in transpar-
Table 3. Design options of 802.16j systems.

<table>
<thead>
<tr>
<th>Transparent mode relay</th>
<th>Non-transparent mode relay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling</td>
<td>Centralized/distributed</td>
</tr>
<tr>
<td>Hops</td>
<td>2 or more</td>
</tr>
<tr>
<td>Forwarding scheme</td>
<td>CID-based</td>
</tr>
<tr>
<td>Channels</td>
<td>1 or 2</td>
</tr>
<tr>
<td>Path management</td>
<td>Embedded/explicit mode</td>
</tr>
<tr>
<td>Routing</td>
<td>Left to vendor design</td>
</tr>
<tr>
<td>RRM</td>
<td>Less complex</td>
</tr>
<tr>
<td></td>
<td>More complex</td>
</tr>
</tbody>
</table>

ent relay mode as specified in the standard, whereas in non-transparent mode both schemes can be used depending on the system configuration. The difference in terms of MAC efficiency provided by these schemes is accentuated as the number of hops in the system increases. Thus, the tunnel mode eventually leads to higher system efficiency as the number of hops in the system increases.

Validity of channel quality information: Channel quality information is quite widely used throughout the system. However, this information may potentially vary quite rapidly, especially in the context of a large amount of mobile users. While the standard does make some optimizations of signaling traffic in the system, there is still a potentially large amount of signaling traffic that needs to be communicated within the system to ensure that appropriate routing and handover decisions are made. It is necessary to find the right balance between minimizing signaling traffic and ensuring that good decisions can be made at the BS in a timely fashion.

Path management: The key difference of the two path management mechanisms defined is the signaling message required. The embedded scheme simplifies operation at the RS with no routing table, allows fast data forwarding, and generates less overhead and delay in route update than the explicit scheme. Thus, the embedded scheme provides sufficient information in order to manage the path in centralized scheduling. The explicit scheme should be used in distributed scheduling as it allows the distribution of the routing table and service flow parameters to the RSs along the path.

Radio resource management: Resource allocation in multihop networks requires design of novel scheduling policies under certain QoS constraints such as reliability, fairness, and latency. There are many aspects to this issue, all of which require further investigation; these include the approaches to realize distributed systems, ways to maximize spatial reuse, and dynamic mechanisms to control the amount of resources allocated to each of the zones in both the transparent and non-transparent modes.

Overall, the substantial amount of choice coupled with a general lack of understanding of the impact of the different design decisions make system design difficult. While some initial work on the performance of 802.16j systems is beginning to appear [7, 8], much research remains to be carried out in order to understand how 802.16j systems perform under different configurations and at what cost compared to 802.16e systems.

**Summary**

IEEE 802.16j offers a potential solution to some of the classic problems in deployment of radio access networks. Comprising low-cost relays associated with BSs, they can be used both to realize larger coverage areas for BSs and to increase capacity for congested areas. Development of the standard is in the final stages, and it should be ratified in late 2008 or early 2009. Here, some of the essential aspects of the standard have been presented, including the different modes of operation, the frame structures used, the network entry procedures etc. This was followed by a discussion of the different aspects that must be considered when thinking about how to design such systems. The relay network architecture, however, is a relatively new design and introduces many complexities within the already challenging environment of radio access networks with mobility support. Many of the issues remain unanswered, and substantial work is necessary to really understand the cost/benefit trade-offs that arise in IEEE 802.16j systems.

**Acknowledgments**

V. Genc gratefully acknowledges the support of IBM and the Irish Research Council for Science, Engineering and Technology.

S. Murphy gratefully acknowledges the support of the EU FP7 ICT CARMEN project.

**References**


**Biographies**

VASKEN GENC (vasken.genc@ucdconnect.ie) received a B.Sc. in computer science in 2003 and an M.Sc. degree in networks and systems engineering in 2005 from the University of Versailles Saint-Quentin, France. He is a Ph.D. student at
the School of Computer Science and Informatics, University College Dublin, Ireland. His research interests are in broadband wireless access systems, multihop wireless networks, and quality of service.

SEAN MURPHY (sean.murphy@iname.com) obtained his primary degree from Dublin College University (DCU) in 1989. He proceeded to work on his postgraduate studies with Teltec Ireland, a research centre within DCU. While involved with Teltec, he spent some time working on EU research projects. He obtained his Ph.D. from DCU in 2001. He spent a year working with Ericsson Expertise Ireland and left to join the Performance Engineering Lab at University College Dublin. His interests are in mobile and wireless communications systems, with an emphasis on multihop systems and video distribution in peer-to-peer systems.

YANG YU (yang.yu@ucd.ie) received a 1st Class Honors B.Sc. in computer science at Dublin Institute of Technology, Ireland, in a joint program with Harbin Institute of Technology, China, in 2006. He is currently a Ph.D. student with the Performance Engineering Laboratory at University College Dublin. His research topic is WiMAX network planning and optimization.

JOHN MURPHY [SM] (j.murphy@ucd.ie) received his B.E. in electronic engineering from University College Dublin, an M.Sc. in electrical engineering from the California Institute of Technology, and a Ph.D. in electronic engineering from Dublin City University. He is a Fellow and a Chartered Engineer with Engineers Ireland, and holds an IBM Faculty Fellowship. He is a member of the Editorial Boards of IEEE Communications Letters, IET Communications, and Telecommunications System Journal. His research is in the area of performance engineering of networks and distributed systems, with almost 100 peer-reviewed journal articles or international conference full papers published, and is funded by 17 competitive research grants with funding in excess of $4 million. He has supervised 12 Ph.D. students to completion.