Analytical Estimation of Packet Delays in Relay-based IMT-advanced Networks

Daniel C. Schultz, Ralf Pabst, Bernhard Walke
RWTH Aachen University, Faculty 6, Chair of Communication Networks (ComNets), Aachen, Germany
e-mail: {dcs|pab|walke}@comnets.rwth-aachen.de

Abstract—The paper shows that relays as inherent technology of IMT-Advanced system can provide low delays as required for future multimedia services. The results shown prove that the proposed relaying scheme is able to keep relay requirements on the air interface and to support delay sensitive high interactive services. The analytical results are validated by system level simulations, which show that the relay might become even more important for real scenarios in order to connect the User Terminal (UT) on the uplink.

I. INTRODUCTION

Fourth-Generation (4G) radio system such as investigated in the EU funded IST WINNER project aim at the provision of ubiquitous broadband access for nomadic users.

It is well known that the range of a 4G broadband radio interface will be limited by the high attenuation at the envisaged carrier frequencies around 2.4 GHz and beyond 3.4 GHz as identified by the World Radiocommunication Conference (WRC) 2007. Further the rather limited amount of spectrum identified for broadband radio communication will force mobile network operator to deploy an increased number of site with limited transmission power owing to in order to achieve the necessary capacities per area element especially in densely populated urban areas. Fixed layer-2 Relay Node (RN) appear to be the most promising technology to overcome the unfavorable radio propagation conditions and bring the capacity available at the Base Station (BS) into the area cost efficiently [1][2].

Relays will be cost efficient due to their deployment flexibility and the missing fixed, either by cable, fibre wire or microwave, backbone connection. Thus they provide the optimal means for initial network roll out of broadband wireless systems.

On the other hand RNs need extra radio resource to connect to the backbone network. That these resources are justified in terms of capacity has been proven in several papers [3]. Next to the achievable throughput, delay is an important Quality of Service (QoS) measure for future broadband internet services. To achieve a good insight in the delay characteristics the Complementary Distribution Function (CDF) of delay in a relay based OFDMA system will be determined.

The basis for the analytical delay calculation will be the WWI New Radio (WINNER) Medium Access Control (MAC) super frame as briefly introduced in the following section. In Section III for the delay calculation in MAC frame based Orthogonal Frequency Division Multiple Access (OFDMA) systems will be derived and calculated analytically. In Section IV the calculated results are validated by means of system level simulations.

II. WINNER MAC SUPER-FRAME STRUCTURE

In the EU funded WINNER project an OFDMA based MAC frame has been developed (for further details see [4] and [5]). The WINNER MAC frame concepts has large similarities with the currently developed 3rd Generation Partnership Program (3GPP) Long Term Evolution (LTE) OFDMA protocol as well as with the approach followed be mobile WiMax.

Figure 1 for the Time Division Duplex (TDD) mode the super frame consists of a set of eight sub-frames. Each sub-frame is further divided into an Downlink (DL) and an Uplink (UL) slot. In the configuration assumed for this paper each slot consists of 15 OFDM symbols and is 0.3372 ms plus 8.4 µs duplex guard time long, resulting in a MAC frame duration of 0,6912 ms. The frames in the Frequency Division Duplex (FDD) mode are of similar length, but subdivided in the frequency domain for UL and DL operation.

III. DELAY CHARACTERISTICS OF THE TDD MAC SUPER-FRAME

To achieve a good insight in the delay characteristics the CDF of delay will be determined. To better classify the results for the multi-hop scenario the results for the 1-hop scenario will be presented firstly as reference.

A. One hop Delay TDD

Figure 2 shows an extract of the investigated MAC SuperFrame (SF), which is composed of a number of MAC Frames each again subdivided into DL and UL phase.

For the investigation of delay characteristics the traffic flows are not fully loaded in order to avoid an impact of scheduling strategies and buffer waiting times.

The calculation is based on a given constant packet arrival rate with equidistant packet interarrival times \( t_{DL\text{-}Rate} \) and \( t_{UL\text{-}Rate} \) for the DL and UL direction respectively.

The total delay \( t_{DL\text{-}Delay}(i) \) of packet \( i \) on the DL is calculated as follows:

\[
\begin{align*}
t_{DL\text{-}Delay}(i) &= t_{DL\text{-}Wait}(i) + t_{DL\text{-}trans}(i) + \\
&\quad + pret\text{-}SH \times ret\text{-}SH
\end{align*}
\]  

(1)

with the waiting time

\[
t_{DL\text{-}Wait}(i) = (n - i) \times t_{DL\text{-}Rate}
\]  

(2)
until the packet is scheduled on the DL (n is the number of packet arriving within the duration of one MAC frame). $t_{DL-transmission}$ denotes the time, which is required for the transmission during the DL phase.

The last product denotes the delay, caused by retransmissions, which occur with a probability of $p_{ret-SH}$. In the single hop case one MAC frame is required to retransmit the packet:

$$t_{ret-SH} = t_{MAC-Frame}$$

(3)

On the UL the total delay $t_{UL-Delay(j)}$ for packet $j$ can be written as:

$$t_{UL-Delay(j)} = t_{UL-Wait(j)} + t_{UL-trans(j)} + p_{ret-SH} * t_{ret-SH}$$

(4)

with the latency

$$t_{UL-Wait} = (n - j) * t_{UL-Rate}$$

(5)

until the UL resources have been scheduled and $t_{UL-transmission}$ for the transmission itself.

B. 2-hop Delay TDD

Figure 3 shows the composition of the delays in a 2-hop scenario for UL and DL.

1) Downlink: To calculate the total delay, $t_{2-hop-DL-Delay}$, analytically it has been sub-divided into partial delays. Based on the data rate and the interarrival time $t_{DL-Rate}$ the delay of packet $i$ on the 2-hop DL transmission can be calculated as follows:

$$t_{2-hop-DL-Delay(i)} = t_{2-hop-DL-Wait(i)} + t_{DL-Schedule} + t_{DL-Slot} + t_{UL-Slot} + t_{DL-trans(i)} + p_{ret-2Hop} * t_{ret-1TXMH}$$

(6)

with the delay

$$t_{2-hop-DL-Wait(i)} = (a - i) * t_{DL-Rate}$$

(7)

for the waiting time until the packet is scheduled on the next available DL frame. $t_{DL-Schedule}$ is the time required to schedule the packets and create the resource allocation table. In opposite to the 1-hop case the number of packets $a$ is now collected over time period of two MAC frames, but this should not affect the scheduling time, which is assumed to be fixed in such a way that a full frame can be scheduled.

The time $t_{DL-Slot}$ denotes the time assigned for DL traffic. The DL resources are denoted as $DL-Slot$. In the DL-Slot the packets are transmitted on the first hop, form the BS to the
RN. The RN needs to decode and re-schedule the received packets for the next MAC frame. The time required for this second hop processing is

\[ t_{DL-2nd-hop-Proc.} = t_{decode}(DL-Data) + t_{DL-Sched.} \]  

(8)

t_{decode} is the time required to decode the received data, and t_{DL-Sched.} is the time needed to re-schedule the packets for the next MAC frame. The second hop processing leads to the following design rule for multi-hop systems:

**Design Rule:** When designing a TDD based multi-hop system it has to be assured that the time for the UL-Slot, t_{UL-Slot} is always long enough to allow the RN to decode the previously received packets, code and schedule them for the next MAC frame:

\[ t_{UL-Slot} \geq t_{DL-2nd-hop-Processing} \]  

(9)

Only if Equation 9 is fulfilled the RN is able to forward the packet within one MAC frame.

The time t_{DL-transmission} is required to transmit the packets to its destination User Terminal (UT) (including the time for the resources allocation table).

In the 2-hop scenario the packet is delayed with the probability

\[ p_{ret-2hop} = \sum_{i=1}^{N} p_{ret-hop} \]  

(10)

with \( N = 2 \) hops. In other words the probability for retransmission in a multi-hop system is increased by each hop, whereby it can be assumed that the probability of retransmissions on the relay link between a BS and a RN, or two RNs is rather low.

The retransmission delay in the two scenario is

\[ t_{ret-1TxMH} = 2 * t_{MAC-Frame} \]  

(11)

a) N-hop DL delay: In an N > 2-hop system each additional hop will require one MAC frame compared to the 2-hop system. Thus, for DL transmission each additional RN adds the duration of one MAC frame to the delay:

\[ t_{N-hop-DL-Delay}(i) = t_{2-hop-DL-Wait}(i) + n_{RN} * t_{MAC-Frame} + t_{DL-trans}(i) + p_{ret-Nhop} * t_{ret-1TxMH} \]  

(12)

where \( n_{RN} \) is the number of RNs participating in the multi-hop connection.

With a retransmission probability

\[ p_{ret-Nhop} = \sum_{i=1}^{N} p_{ret-hop} \]  

(13)

for \( N \) hops, whereby \( p_{ret-hop_i} \) is the retransmission probability of hop \( i \).

2) Multi-hop Uplink Delay: To calculate the total delay, t_{2-hop-UL-Delay}, analytically it has been sub-divided into partial delays. Based on the data rate and the interarrival time \( t_{UL-Rate} \) the delay of packet \( j \) on the 2-hop UL transmission
is composed as follows:
\[ t_{2-\text{Hop-UL-Delay}}(j) = t_{2-\text{Hop-UL-Wait}}(j) + t_{UL-\text{Sched.}} + t_{UL-Slot} + t_{DL-Slot} + t_{ULtransm}(i) + p_{ret-2\text{Hop}} \cdot t_{ret-1\text{TxMH}} \] (14)

with the delay
\[ t_{2-\text{Hop-UL-Wait}}(i) = (a - i) \cdot t_{UL-Rate} \] (15)

for the waiting time until the packet has been scheduled for the coming UL Slot. The UL resources have been requested by sending a Reservation Request (RR) message to the RN during the last active frame.

The scheduled packets are transmitted on the first hop during the UL-Slot consuming \( t_{UL-Slot} \). As described for the DL in Section III-B.1 the RN has to decode and re-schedule the received packets for the next hop, further referred to as second hop UL processing.

in case of two transceiver compared to the single transceiver solution. Design Rule: When designing a TDD based multi-hop system it has to be assured that the time for the DL-Slot, \( t_{DL-Slot} \), is always long enough to allow the RN to decode the previously received packets, code and re-schedule them to be transmitted in the following UL-Slot:

\[ t_{DL-Slot} \geq t_{UL-2\text{nd-hop-Processing}} \] (16)

Only if Equation 16 is fulfilled the RN is able to forward the packet within one MAC frame.

During the next UL-Slot the data is transmitted to the BS consuming \( t_{UL-Transmission} \). Finally a decoding delay of \( t_{Decode} \) has to be added.

C. Delay for 2 Transceiver Relay Nodes

In general the delay is corresponding to the delay of the on transceiver solution:

But the second transceiver allows to transmit packets to the on the Relay Link (RL) between the BS and the RN every frame. Thus the waiting times \( t_{2-\text{Hop-DL-Wait}} \) and \( t_{2-\text{Hop-UL-Wait}} \) will decrease. On average these waiting times are only half of the time required in the one transceiver solution.

Further the retransmission delay is reduced to one MAC frame as the RN is able to immediately retransmit or Acknowledgment (ACK) the packets:

\[ t_{ret-2\text{Tx}} = t_{MAC-frame} \] (17)

D. TDD Delay: Results

From Section III-B.1 and Section III-B.2 it can be seen that the delays for DL and UL are the same, if the packet interarrival times for DL and UL are similar. The ratio between DL and DL transmission times has no impact.

Figure 4 shows the analytical results for UL and DL delays of the MAC SF with \( t_{MACFrame} = 0.693ms \), \( t_{Schedule} = 0.1ms \) and \( t_{decode} = 0.175ms \). The results show that, although the single-hop delays are beneficial the multi-hop delay is, with less than 2.5 ms in the 2-hop and less than 3.1 ms in the 3-hop case on the air interface, still meeting the expectations of future multi-media services.

The retransmission probabilities have been set to \( p_{ret-SH}=0.02 \), \( p_{ret-hop1}=0.01 \) and \( p_{ret-hop2}=0.01 \). The retransmission probabilities for the multi-hop scenario are assumed to be lower due the reduced distance between UT and Radio Access Point (RAP).

IV. SIMULATION RESULTS

Figure 5 plots the Complementary Cumulative Distribution Functions (CCDFs) function of the observed DL end-to-end packet delay. End-to-end in this context means from the arrival at the BS until the successful delivery at the UT. The statistics involve only the successfully transmitted packets, not the ones discarded at the receiver after the Cyclic Redundancy Check (CRC) check. The CCDFs were measured at an average DL cell load of 50MBit/s - a load condition where saturation is not yet reached for any of the compared schemes. Since the system model does not include any load control, measuring the delay in overload conditions would not have produced meaningful results (i.e. infinite delays). The figure shows that under the given load conditions, 99.9% of the successfully transmitted packets reach the UT in less than 1.8ms in the single hop case. Approximately 50% of all packets are even transmitted within 1ms. The introduction of relay nodes raises the 99.9th percentile of the delay to only slightly above 3ms for those cases while 70% of the packets are even received in less than 2ms.

Thereby the following

Although the absolute values differ between the simulations and analysis the tendency is the same. The difference can be explained by the more accurate protocol model and the more
Fig. 5. CCDF of DL Packet delay for 80Mbit/s DL traffic

<table>
<thead>
<tr>
<th>RNs</th>
<th>Scenario</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>“None”</td>
<td>Full reuse, single hop, serves as comparison case.</td>
</tr>
<tr>
<td>3</td>
<td>“No Reuse”</td>
<td>Can be considered as the baseline case; no intra-REC interference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>between RN subcells, all RNs active in the same frame and its</td>
</tr>
<tr>
<td></td>
<td></td>
<td>resources are shared among them</td>
</tr>
<tr>
<td>3</td>
<td>“Full Reuse”</td>
<td>Full reuse i.e. higher intra-REC interference between RN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>subcells, but higher diversity Frame 1: BS 1 serves UTs and RNs Frame 2:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RNs serve UTs, BS silent</td>
</tr>
</tbody>
</table>

Table I
Parameter for delay simulations

detailed interference and retransmission cases. Thus it can be shown the 95%-tiles for the two hop transmission and the one hop transmission in the analytical case are almost identical for the multi-hop case. The pure single-hop scenario should not be considered, as the cell size is so large that many of the UT suffer from very bad channel conditions on the UL feedback channel, which has not been considered in the analysis.

V. CONCLUSION AND OUTLOOK

The paper illustrates that the proposed MAC protocol will easily be able to meet delay requirements of future internet services. Thus it could be shown that relays do not have any considerable QoS drawback compared to single-hop. In fact the simulation results have shown that it is likely to be vice versa, i.e. multi-hop solutions are required to support the uplink channel in order to achieve sufficient channel quality and therewith avoid unnecessary retransmissions, which cause high delays.

ACKNOWLEDGMENT

The presented work has partly been funded by the European Commission in the FP6 IST-Project WINNER II (IST-4-027756). The views expressed are those of the authors and do not necessarily represent the project.

REFERENCES