Experimental Evaluation of Techniques to Lower Spectrum Consumption in Wi-Red

Gianluca Cena 1, Stefano Scanzio 2, and Adriano Valenzano 1

1 Affiliation not available
2 CNR-IEIIT

October 30, 2023

Abstract

Seamless redundancy layered atop Wi-Fi has been shown able to tangibly increase communication quality, hence offering industry-grade reliability. However, it also implies much higher network traffic, which is often unbearable as the wireless spectrum is a shared and scarce resource. To deal with this drawback the Wi-Red proposal includes suitable duplication avoidance mechanisms, which reduce spectrum consumption by preventing transmission on air of inessential frame duplicates. In this paper, the ability of such mechanisms to save wireless bandwidth is experimentally evaluated. To this purpose, specific post-analysis techniques have been defined, which permit to carry out such an assessment on a simple testbed that relies on plain redundancy and do not require any changes to the adapters’ firmware. As results show, spectrum consumption decreases noticeably without communication quality is impaired. Further saving can be obtained if a slight worsening is tolerated for latencies.
Experimental Evaluation of Techniques to Lower Spectrum Consumption in Wi-Red

Gianluca Cena, Senior Member, IEEE, Stefano Scanzio, Member, IEEE, and Adriano Valenzano, Senior Member, IEEE

Abstract—Seamless redundancy layered atop Wi-Fi has been shown able to tangibly increase communication quality, hence offering industry-grade reliability. However, it also implies much higher network traffic, which is often unbearable as the wireless spectrum is a shared and scarce resource. To deal with this drawback the Wi-Red proposal includes suitable duplication avoidance mechanisms, which reduce spectrum consumption by preventing transmission on air of inessential frame duplicates.

In this paper, the ability of such mechanisms to save wireless bandwidth is experimentally evaluated. To this purpose, specific post-analysis techniques have been defined, which permit to carry out such an assessment on a simple testbed that relies on plain redundancy and do not require any changes to the adapters’ firmware. As results show, spectrum consumption decreases noticeably without communication quality is impaired. Further saving can be obtained if a slight worsening is tolerated for latencies.

Index Terms—IEEE 802.11, Wi-Fi, seamless redundancy, PRP, experimental evaluation, communication efficiency.

I. INTRODUCTION

Wireless communications are currently employed in many different contexts. Besides personal connectivity and office/home environments, they are becoming increasingly popular also in other scenarios, like intelligent transportation systems [1], Internet of Things (IoT) [2], environmental monitoring [3], and precision agriculture [4], to cite a few. Although wireless networks have the potential to bring tangible benefits also to time-sensitive applications, they are quite unreliable and scarcely deterministic, and so are often deemed unsuitablein contexts like factory automation. As a matter of fact, with the exception of WirelessHART and ISA100.11a [5], real-time distributed control systems based on wireless technologies are currently quite unusual in industrial plants.

Unlike multimedia applications, where data can typically be buffered on both sides of a network connection, control applications (relying on, e.g., the request-response or producer-consumer data exchange paradigms [6]) demand that every single piece of information is delivered timely and reliably. This can be hardly achieved with conventional wireless solutions, since they are typically based on random access mechanisms, which unavoidably lead to collisions. Behavior can be improved by using deterministic access schemes, which prevent intra-network interference, e.g., Time Slotted Channel Hopping (TSCH) for IEEE 802.15.4 [7] and Hybrid-coordination-function Controlled Channel Access (HCCA) for IEEE 802.11 [8], not counting the many proposals found in literature, for instance [9][10][11][12]. It is worth noting that most of these mechanisms are unable to face electromagnetic disturbance (including multipath fading effects) and, in several cases, not even interference of nearby non-compliant wireless devices. These phenomena are unpredictable, and can be effectively counteracted by exploiting diversity, in time, frequency, space, and so on. For instance, automatic retransmission mechanisms are customarily included in any wireless protocol, frequency hopping is adopted by TSCH and Bluetooth, while multiple spatial streams are exploited in recent IEEE 802.11 versions.

In the following, we consider the use of IEEE 802.11, also known as Wi-Fi, for connecting devices in time-sensitive control applications. The reasons of this choice are threefold: 1) Wi-Fi is extremely popular, and consequently not expensive, 2) it features very high throughput, and 3) it achieves complete interoperability with Ethernet. We specifically focused on mechanisms based on time and frequency diversity (i.e., retransmissions and channel redundancy), both of which bring improvements on communication quality at the cost of higher spectrum consumption. Techniques based on message scheduling, like HCCA or the like, were not taken into account as, to the best of our knowledge, there is currently little availability of commercial equipment offering proper support.

Seamless redundancy, as defined by the Parallel Redundancy Protocol (PRP) [13], is meant to increase availability of real-time Ethernet networks. Nonetheless, it can be also used to improve communication quality of wireless links. In [14], an arrangement was proposed, we denote PRP over Wi-Fi (PoW), which layers end-to-end seamless redundancy as per PRP atop IEEE 802.11, by using commercial Wi-Fi equipment, like access points (AP) and wireless adapters (WA), as well as specialized PRP devices (RedBoxes) [15]. Behavior, in terms of the fraction of frames delivered to destination correctly and timely, was experimentally shown to improve substantially. Similar approaches were proposed for streaming phasor measurements in smart grids [16].

While relying on the same principles as PoW, the Wi-Fi Redundancy (Wi-Red) approach [17] is a link-level solution where PRP and the IEEE 802.11 Medium Access Control (MAC) mechanism are intertwined in order to improve performance further. In particular, specific duplication avoidance (DA) mechanisms are introduced to prevent transmission on air of identical copies of the same packet, when doing so
is useless, so as to reduce network load. Basically, Wi-Red can be seen as an holistic approach aimed at optimizing time and frequency diversity in Wi-Fi. In theory, the same approach could be also used with wireless technologies other than Wi-Fi, like those employed in wireless sensor networks (WSN). However, this is typically pointless, as WSNs are mainly aimed at ensuring low power consumption, and not high performance. For example, solutions like TSCH offer time-frequency diversity through channel hopping, without the need to have a dual radio block, at the expense of larger transmission latencies.

This paper, which grounds on the preliminary work presented in [18], experimentally evaluates the effects of DA mechanisms on the overall spectrum consumption (combined traffic on all physical channels) of a redundant link between two wireless stations, and analyzes the trade-off between bandwidth saving and communication quality for a basic proactive heuristic. With respect to [18]: 1) communication efficiency is evaluated considering the actual number of transmission attempts on air, inclusive of retries; 2) runs include more than twice samples and employ more realistic interfering traffic patterns; 3) assumptions on virtual analysis of proactive approaches are validated; and 4) both frame losses and more-than-duplex redundancy are considered. The paper is organized as follows: Section II summarizes Wi-Fi seamless redundancy basics, while Section III describes our experimental testbed. Quantitative evaluation of the ability of DA mechanisms to reduce spectrum consumption, based on measurements carried out on the testbed, is reported in Sections IV and V, which concern two types of approaches, that are reactive and proactive ones, respectively. Finally, some conclusions are drawn in Section VI.

II. LINK-LEVEL WI-FI SEAMLESS REDUNDANCY

In its simplest embodiment, Wi-Red corresponds to a thin layer, located just above Wi-Fi adapters and implemented by bringing software modifications to their device drivers. However, to unleash all its potential, adapters’ firmware has likely to be rewritten in part.

A. Redundant Network Architecture

An IEEE 802.11 Basic Service Set (BSS) is a group of wireless stations (STA) located in the same place, operating on the same channel, and enabled to exchange data among each other. Similarly, a redundant BSS (RBSS) is defined as a set of co-located redundant STAs (RSTA) communicating on multiple physical channels according to seamless redundancy principles. An RSTA comprises two or more sub-STAs, operating on distinct channels, and one link redundancy entity (LRE), which coordinates their operations so as to achieve seamless redundancy. Each sub-STA includes both a MAC and a radio block, and behaves mostly the same as a conventional STA. More details can be found in [17].

Generally speaking, effectiveness of seamless redundancy in networks including multiple uncoordinated RSTAs is typically not as good as on point-to-point links. In fact, intra-network interference they generate on physical channels is clearly not independent. The easiest way to face this situation is to exploit for RSTAs a deterministic protocol overlay (e.g., HCCA or, possibly, solutions specifically engineered for seamless redundancy), which prevents them from transmitting on air at the same time. This is not unfeasible, as these devices are usually part of the same control system. Actual advantages of redundancy in presence of many nodes operating with random access is still an open issue, and is left as future work. For such reason, the following analysis takes into account a unidirectional redundant link, denoted $L_1$, between two RSTAs (originating RSTA and recipient RSTA), and $C_L = \{A, B, \ldots\}$ is the set of the related physical channels. To keep the testbed simple, only duplex redundancy has been considered in the experiments, in which case $L$ is denoted $A+B$. We assume that fragmentation, aggregation, and block acknowledgment mechanisms are not exploited, since they are not meant to meet the requirements of time-sensitive control applications.

B. Seamless Redundancy Basics

Although seamless redundancy also applies to unconfirmed transmissions, this paper only focuses on acknowledged exchanges. Besides being practically mandatory to support reliable data transfers in Wi-Fi, the latter also enable effective optimizations when paired with seamless redundancy. For our purposes, the term packet denotes a unit of data to be exchanged over the network, as seen above the redundancy and data-link layers, and approximately corresponds to a MAC Service Data Unit (MSDU). When a request is made in the originating RSTA to transmit a packet over a redundant link $L$, the related LRE issues concurrent requests on the adapters of its sub-STAs for sending identical copies of the packet on every channel $C \in C_L$. Sub-STAs’ operations are mostly disjoint [19], [20], and the transmission of each packet copy on the related adapter obeys MAC rules. This involves sending specific frames on air, corresponding to MAC Protocol Data Units (MPDU), as well as dealing with carrier sensing, interframe spaces, and random backoff, issuing retransmissions upon missing acknowledgment, and so on.

Transmission of each packet copy on the related channel consists of one or more attempts: exactly one initial try plus, upon unsuccessful delivery, a variable number of retries. To avoid starvation, a retry limit is defined to bound the number of transmission attempts for any packet copy. Each attempt is made up of a DATA frame (carrying user data), sent by the originator, immediately followed by an ACK frame (acknowledgment), returned by the unique recipient after a Short Interframe Space (SIFS). Failed attempts are detected by the MAC in the originating sub-STA by setting a specific ACKTimeout whenever a DATA frame is sent. If the ACK frame is not received before the timeout expires, the attempt is considered as failed.

On the recipient RSTA, the LRE takes care of removing duplicates, reordering packets, and delivering them to the upper layers. To do so, packet copies received on the physical channels of the redundant link are paired using specific sequence numbers added on the originating side. For any packet, the DATA frame (either the initial try or a retry) that is
correctly received first among all its copies is retained, while the following ones are discarded.

C. Duplication Avoidance Mechanisms

In PoW, every packet is always sent on all channels. Since channels can be assumed to be reasonably independent in a well-configured RBSS [19], overall spectrum consumption in PoW coincides with the sum of the traffics that would be induced if non-redundant, conventional Wi-Fi were used to send packets separately on each physical channel. The main improvement Wi-Red [17] brings over PoW [14] are duplication avoidance mechanisms. Basically, they are aimed at reducing the amount of frames sent on air. This is a strict requirement when smooth coexistence has to be ensured with nearby networks, as the wireless spectrum is a scarce resource not to be wasted, especially in the 2.4 GHz band and, to a lesser (but increasing) extent, 5 GHz band.

Two kinds of DA mechanisms exist, namely reactive and proactive, both of which operate on the originating side of a redundant link. The former simply prevents duplicate transmissions on air that are recognized to be useless, since an acknowledgment has already been received from the recipient, on any physical channel of the redundant link, for the information they convey. Instead, the latter also includes heuristics aimed at complementing the reactive approach, so as to increase its effectiveness.

III. EXPERIMENTAL EVALUATION

The experimental setup used for assessing Wi-Red performance resembles the one in [20]. However, the analysis carried out in this paper does not focus on communication quality, but rather on the ability of DA mechanisms to reduce spectrum consumption (which is the main performance metric here). As shown in Fig. 1, the testbed is made up of a PC running the Linux OS (kernel v. 3.16.0) and provided with two identical Wi-Fi adapters (dual-band TP-Link TL-WDN4800), which emulates the behavior of the originating RSTA. Two identical APs (NETGEAR WAC120), set to operate on completely disjoint frequencies to avoid interference (channel 1 in the 2.4 GHz band and 44 in the 5 GHz band), partially mimic the behavior of the recipient RSTA. Each Wi-Fi adapter on the PC is associated to a distinct AP, which, on correct reception of a DATA frame, returns an ACK frame to the related adapter and relays the packet to the PC Ethernet port through a Gigabit switch.

On the PC, a measurement task produces a cyclic pattern of packets. Generation period $T_M$ is set to 100 ms. To obtain reliable statistics, each run lasted 24 hours and included $N = 864000$ packets, which are sent on the redundant link according to PoW rules (each packet is concurrently fed to both Wi-Fi adapters). Timestamps are taken on every packet transmission and reception. Since they refer to the same time base, i.e., the CPU time stamp counter (TSC) register of the PC, transmission latencies can be computed. Following the approach in [12], receive times were not taken on packet arrival to the Ethernet port. Instead, we modified Wi-Fi adapters’ drivers so as to acquire a precise timestamp in kernel space upon ACK frame reception or ACKtimeout expiry. Doing so improves both accuracy and precision of measured transmission latencies. In fact, although APs were, on average, quite fast in forwarding packets, they also introduced non-negligible jitters and delays. This is no surprise, since they were not conceived for time-sensitive applications. Results were checked against timestamps and losses obtained on the Ethernet port.

A. Modeling Experimental Results

Each experimental run in the testbed is modeled as a packet sequence $M = \{m_1, ..., m_N\}$, where $m_i$ is the $i$-th packet generated by the measurement task. For every packet, a transmission request is separately issued on all adapters. The copy of $m_i$ sent on channel $C \in C_i$ (either $A$ or $B$, in our testbed) is denoted $m_{C,i}^C$, and its transmission outcome is described as a pair $(t_{C,i}^C, l_{C,i}^C)$. Boolean value $l_{C,i}^C$ is the loss indication, either true if $m_i$ went lost on $C$ or false when it was successfully delivered. In the following, Booleans are treated as integers, i.e., true $\equiv 1$ and false $\equiv 0$. The non-

### TABLE I

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>generic physical channel (either $A$ or $B$, in our testbed)</td>
</tr>
<tr>
<td>$L$</td>
<td>redundant link ($A+B$, in our testbed)</td>
</tr>
<tr>
<td>$C_i$</td>
<td>set of physical channels of $L$ (${A, B}$, in our testbed)</td>
</tr>
<tr>
<td>$T_{ACKtimeout}$</td>
<td>LRE delay (from XACK to early termination)</td>
</tr>
<tr>
<td>$T_{SF}$</td>
<td>SIFS duration on channel $C$</td>
</tr>
<tr>
<td>$T_{ACKtimeout}$</td>
<td>ACKtimeout duration on channel $C$</td>
</tr>
</tbody>
</table>

- $m_i$: $i$-th packet generated in the experiment
- $x_i$: first received ACK (XACK) for packet $m_i$
- $Q_i$: quickest channel, on which $x_i$ is received (either $A$ or $B$)
- $T_x,i$: cross acknowledgment (XACK) time for packet $m_i$
- $m_{C,i}^C$: copy of packet $m_i$ sent on channel $C$
- $T_w,i$: acknowledgment for packet copy $m_{C,i}^C$
- $z_{w,i}^C$: number of transmission attempts of packet copy $m_{C,i}^C$
- $T_{t_{C,i}^C}^C$: transmission time of packet copy $m_{C,i}^C$
- $T_{s_{C,i}^C}^C$: starting time on air of the final attempt of packet copy $m_{C,i}^C$
- $T_{R_{C,i}^C}^C$: receive time of packet copy $m_{C,i}^C$
- $T_{s_{C,i}^C}^C$: acknowledgment time of packet copy $m_{C,i}^C$
- $T_{D_{C,i}^C}^C$: duration of the DATA frame in the final attempt of packet copy $m_{C,i}^C$
- $T_{A_{C,i}^C}^C$: duration of the ACK frame in the final attempt of packet copy $m_{C,i}^C$
negative value $d_C^i$ is the transmission latency of $m_i$ on $C$, measured from the time $t_{R,i}^F$ when the transmission request is issued on the related adapter to the time $t_{R,i}^L$ when the packet is correctly received by the recipient, that is, $d_C^i \triangleq t_{R,i}^L - t_{R,i}^F$. Latency is only defined for packets for which $t_{R,i}^L = 0$, since on the contrary $t_{R,i}^L$ does not exist.

Behavior of redundant link $L$ can be analyzed starting from its physical channels. Let $t_{T,i}^L$ be the time when packet $m_i$ is generated. With the exception of proactive mechanisms [17], described in Section V, we can assume that, with very good approximation, transmission times for all the copies of any given packet coincide, that is $t_{T,i}^C \approx t_{T,i}^L$, $\forall C \in C_L$. Because of delays introduced at the MAC layer by the shared medium (due to, e.g., the channel sensed busy or retransmissions), the same does not necessarily hold for receive times $t_{R,i}^L$.

Links exploiting seamless redundancy satisfy two basic properties: 1) a packet $m_i$ is lost on $L$ (i.e., it never arrives to the recipient, not even after all allowed retries have been carried out by all adapters) only if its copies are definitely lost on every channel $C \in C_L$; 2) the receive time of any packet $m_i$ correctly delivered on $L$ corresponds to the minimum between the receive times of its copies on all physical channels, i.e., $t_{R,i}^L = \min_{C \in C_L} t_{R,i}^C$. This means that communication quality on a redundant link is never worse than on any of its physical channels.

B. Measurement Technique

From a practical viewpoint, two timestamps were acquired for every packet $m_i$ on each physical channel $C$: $t_{T,i}^C$ is taken immediately before the transmission request is invoked, while $t_{T,i}^L$ is taken in the device driver when transmission ends and passed to the measurement task through a character device. In order to simplify the following analysis, we will assume that errors always affect a transmission attempt as a whole, and not the ACK frame alone. In the real world, this assumption is typically verified because: 1) ACK frames are very short and typically raise interrupts only at the end of packet transmission (ient) of packet $m_i$ on $C$; 2) receipt of the ACK frame is usually dealt with by recovery mechanisms in the upper protocol layers.

In order to simplify the following analysis, we will assume that errors always affect a transmission attempt as a whole, and not the ACK frame alone. In the real world, this assumption is typically verified because: 1) ACK frames are very short and rely on robust modulation and coding schemes, and 2) under the reasonable hypothesis that the BSS does not suffer from the hidden node problem, they cannot incur in collisions not involving also the related DATA frame. This means that, when a DATA frame is correctly transferred, very likely the related ACK frame also does. We checked this assumption posteriori on experimental samples, by comparing the set of packets arrived on the Ethernet port against those for which an ACK was received, and found that it actually holds in almost all the cases. However, it is worth pointing out that the correct behavior of both seamless redundancy and DA mechanisms does not depend in any way on this assumption, which only leads to slightly pessimistic results for communication quality.

Under the above hypothesis, the events raised upon acknowledgment (on the originator) and receipt (on the recipient) of packet $m_i$ on channel $C$ refer to the same (successful) transmission attempt, which corresponds to $m_{i,w_1}^C$.

C. Assumptions on Channel Errors

Both DATA and ACK frames may be corrupted during transmission on air. If an error affects the ACK frame, the MAC on the recipient notifies the correct packet reception, but additional retries are attempted by the originator, until it either receives an ACK or exceeds the retry limit. Typically, this causes the packet acknowledgment time on the originator to be delayed to one of the following retries. If none of these succeed, the packet is considered lost by the originator, hence making its delivery state inconsistent with the recipient. This issue is usually dealt with by recovery mechanisms in the upper protocol layers.

D. Limitations of Commercial Equipment

Real Wi-Fi adapters, including those used in our testbed, typically raise interrupts only at the end of packet transmission

\[
\begin{align*}
&\text{packet } m_1, m_2, \ldots, m_i, \ldots, m_n, m_{i+1}, \ldots, m_m, \ldots, m_{n+p-C}, \ldots
\\&\begin{pmatrix}
&\text{initial} & \text{1-st retry} & \text{(ℓ-th) retry} & \text{final attempt} \\
&m_{1,1}^C, \ldots, m_{1,w_1}^C, \ldots, m_{1,2}^C, \ldots, m_{1,\ell}^C, \ldots, m_{1,w_1}^C, \ldots & m_{i,1}^C, \ldots, m_{i,2}^C, \ldots, m_{i,\ell}^C, \ldots, m_{i,w_1}^C, \ldots & m_{n,1}^C, \ldots, m_{n,2}^C, \ldots, m_{n,\ell}^C, \ldots, m_{n,w_1}^C, \ldots & m_{n+p,1}^C, \ldots, m_{n+p,2}^C, \ldots, m_{n+p,\ell}^C, \ldots, m_{n+p,w_1}^C, \ldots
\end{pmatrix}
\end{align*}
\]
and reception. Thus, it is not possible to obtain timestamps $t^l_{W,i,t} \in T^C_{W,i,t}$ for inner attempts, but only for the final ones, in which case either an ACK frame has been correctly received (success) or the ACKtimeout has expired and the retry limit has been exceeded (failure). These conditions conclude the transmission process for packet copy $m^C_i$. To keep notation simple, the starting time on air of the final attempt on $C$ is denoted $t^C_{W,i,t}$, while the end of transmission coincides with $t^C_{X,i}$. The durations of the related DATA and ACK frames (indicated as $T^C_{D,i}$ and $T^C_{A,i}$, respectively) can be obtained by inspecting the driver.

In case of success ($t^l_{C} = 0$) the starting time on air of the final transmission attempt of $m^C_i$ can be calculated as

$$t^C_{W,i} = t^C_{X,i} - (T^C_{D,i} + T^C_{SIFS} + T^C_{A,i}) \tag{2}$$

and a reliable estimate for the receive time can be obtained as

$$t^l_{R,i} = t^l_{X,i} - (T^C_{SIFS} + T^C_{A,i}) \tag{3}$$

while case of failure ($t^l_{C} = 1$)

$$t^C_{W,i} = (t^C_{X,i} - (T^C_{D,i} + T^C_{ACKto})) \tag{4}$$

where $T^C_{ACKto}$ is the ACKtimeout duration on $C$.

For the purpose of the following analysis, transmission of $m_i$ on physical channel $C$ of the testbed is completely described by a tuple $\tau_i^C = \{ t^C_{i}, t^C_{X,i}, t^C_{D,i}, t^C_{A,i}, T^C_{D,i}, T^C_{A,i} \}$. Overall, every experimental run is modeled as $\mathcal{T} = \{ (\tau_i^C, \tau_j^C) \}_{i=1\ldots N}$. For duplex redundancy (as in the following) can be prevented on any channel $C \in \mathcal{Q}_j$ when

$$t^l_{X,i} + T^C_{LRE} < t^C_{W,i,t}.$$

A. Performance Comparison between RDA and PoW

1) Communication Quality: When the RBSS is correctly dimensioned, so that generated packets seldom suffer from appreciable queuing delays in local buffers, RDA offers the same performance as PoW. In fact, RDA acts after ACK reception, i.e., after the related packet has been delivered to the recipient. This means that, in absence of queuing, RDA operations can not affect in any way neither latencies $d^C_{l}$ nor packet losses $l^C_{l}$ of PoW. This is mostly the case of our experiments, as packets produced by the measurement task were intentionally spaced wide enough. When traffic increases and the RBSS operates in saturated conditions, RDA sensibly improves communication quality over PoW [17], because of its ability to lower network load, which reduces collisions and, hence, packet losses and mean latencies. Latencies also shrink because transmission queues in the originating STA are drained more quickly. Definitely, RDA adoption can never worst communication quality.

2) Spectrum Consumption: Unless interfering traffic and disturbance in the RBSS are very low (in which case redundancy is useless), RDA typically offers tangible advantages with respect to PoW, as our experiments indicate. The ability to reduce wasted bandwidth improves further when network traffic increases and packets are buffered, since in this case they can also be removed from transmission queues. Actual
improvements in real scenarios heavily depend on the ability of the LRE to quickly terminate ongoing transmissions upon ACK reception. As will be shown, delays introduced by, e.g., software LRE implementations, impair RDA performance. In order to reliably assess spectrum consumption on a real RDA testbed, modifications to the firmware of Wi-Fi adapters are likely required, which is unfeasible with most commercial devices (including ours). For this reason, we evaluated bandwidth saving in RDA using a different approach.

B. Virtual RDA Analysis on PoW Logs

PoW behavior closely resembles basic Wi-Red, i.e., seamless redundancy with no DA mechanisms. Once started, transmission of each packet copy is carried out independently on the related channel as per IEEE 802.11 MAC rules. Interestingly, experimental data acquired in a run from the PoW testbed can be used to analyze RDA performance as well. In particular, the amount of transmission attempts that can be saved by early termination in the very same operating conditions (including environmental ones, which are almost impossible for us to replicate) can be inferred. This provides a reliable estimate of the improvements RDA achieves over PoW in terms of spectrum consumption, without the need to actually implement RDA.

A prerequisite for the following analysis is that, terminating a packet transmission in advance does not affect the following ones. This is mostly true if, on every physical channel $C$, adjacent packet transmissions do not overlap, i.e., if transmission latency $d_{i}^{\text{C}}$ never exceeds the generation period $T_{M}$. Under such hypothesis, if a given PoW experiment was repeated, in exactly the same operating and environmental conditions, using a real RDA testbed, the sequence of transmission attempts on air would simply consist of a subset of the PoW run. In fact, some PoW attempts are possibly saved because of early termination. In other words, for any attempt $m_{i}^{\text{C}}$ carried out when RDA is enabled there would be an equivalent attempt (same starting time on air and duration) carried out by PoW, but not vice-versa, because all attempts for which (5) holds are prevented by RDA. For instance, in the example shown in Fig. 3, RDA cancels the last two attempts carried out by PoW on $B$ in $Q_{i}$, that is, $m_{4,4}^{B}$ and $m_{4,5}^{B}$.

In the above analysis there is no need to define new quantities that explicitly refer to RDA, as they coincide with PoW. The only difference lies in the number of transmission attempts, which for RDA is possibly smaller than PoW. The number of attempts for packet $m_{i}$ on channel $C$ in the RDA case is denoted $w_{i}^{C}$.

C. Experimental Evaluation of RDA

Let $L$ be a redundant link like the one implemented by our PoW testbed, for which the hypotheses that permit to apply virtual RDA analysis hold. In theory, early termination could be completely analyzed from the related experimental logs. However, since starting times $t_{W,i}^{\ell}$ are only available for final attempts, exact $w_{i}^{C}$ values cannot be determined. Nevertheless, an indication can be still obtained on whether, for any given packet, some attempts can be saved. This provides an upper bound on RDA spectrum consumption. As shown in Fig. 3, if (5) holds for the final attempt $m_{i}^{\text{C}}$ on channel $C \in Q_{i}$, then at least one attempt is prevented for sure on $C$. In the sample packet transmission depicted in the figure, where $w_{i}^{B} = 4$, this approach permits to infer that $w_{i}^{B} \leq 3$ (although, actually, $w_{i}^{B} = 2$).

Let $e_{i}^{C}$ (early termination condition) be a Boolean value that is true if and only if, owing to XACK $x_{i}$, RDA is able to terminate the transmission of packet $m_{i}$ on channel $C \in Q_{i}$ in advance with respect to what is foreseen by conventional Wi-Fi. As said before, this is only possible for packets correctly delivered (and acknowledged) on $L$. From (5),

$$
\begin{align*}
   e_{i}^{C} & \triangleq \left[ t_{i}^{L} = 0 \land t_{X,i} + T_{\text{LRE}} < t_{W,i}^{\ell} \right] \\
   (6) & 
\end{align*}
$$

where $t_{W,i}^{\ell}$ is given by (2) and (4), for successful and failed packet copy transmissions, respectively. For the quickest channel, (6) correctly provides $e_{i}^{Q_{i}} = 0$, so it holds for every $C \in C_{L}$. Above considerations identically apply when RDA is coupled with some heuristics (PDA).

D. Metrics about Spectrum Consumption

Reliability through time and frequency diversity comes to the price of a higher spectrum consumption. The holistic approach provided by DA mechanisms offers non-negligible advantages in this respect. The related benefits in terms of bandwidth saving can be evaluated by analyzing the number of transmission attempts actually performed on air during any given experiment.

A number of simple metrics have been introduced to this extent. In the following, when referring to redundant links, the kind of DA mechanism will be specified, if relevant, as superscript, in the place of the link. Term DA generically refers...
to both RDA and PDA, while PoW denotes the lack of any DA. In these cases, the link can be inferred from the context (either \( L \), in general reasoning, or \( A+B \), when referring to our experimental testbed).

1) Saved Bandwidth: The fraction of packets in an experimental run whose transmission is terminated early on physical channel \( C \in C_L \) is

\[
\pi^C \triangleq \frac{1}{N} \sum_{i=1}^{N} e^C_i. \tag{7}
\]

For any packet sent on the redundant link \( L \), the average number of copies that experience early termination, all channels considered, is

\[
\pi^L \triangleq \frac{1}{N} \sum_{i=1}^{N} \sum_{C \in C_L} e^C_i = \sum_{C \in C_L} \pi^C. \tag{8}
\]

When \( L = A+B \), in our duplex testbed, there is only one channel other than the quickest, and \( \pi^A+B \) coincides with the fraction of packets terminated early on either \( A \) or \( B \).

Upon early termination of a packet copy, one or more transmission attempts are canceled for sure on the related channel. This implies that \( \pi^C \) constitutes a lower bound on the average number of useless attempts that, overall, can be prevented for each packet by RDA with respect to PoW, and provides an estimate of the bandwidth saving achieved over basic seamless redundancy. Packets that are dropped on all physical channels but one are those for which the most savings are expected. In fact, in the absence of early termination, all attempts up to the retry limit are (unsuccessfully) carried out on faulty channels.

Our testbed does not make \( T^C_{DL} \), available when transmission on adapter \( C \) fails, nor it can be easily inferred (because of rate adaptation mechanisms), and so \( e^C_i \) cannot be evaluated for the related packet copies. This is not a severe drawback since, due to the effectiveness of MAC retransmissions, the fraction of packets that were dropped on either of the physical channels in our experiments was always very low (less than 0.0033%). Hence, their contribution to \( \pi^C \) can be quietly neglected, by assuming for them \( e^C_i = 0 \). It is worth pointing out that the case when the loss rate is negligible only leads to a pessimistic, smaller value for \( \pi^C \), which is acceptable as this quantity represents a lower bound.

2) Network load: The average number of transmission attempts per packet (including retries) carried out on any channel \( C \) during an experiment on the PoW testbed is

\[
\pi^C \triangleq \frac{1}{N} \sum_{i=1}^{N} w^C_i. \tag{9}
\]

As in Wi-Fi every packet is sent on air at least once, \( \pi^C \geq 1 \).

The value of \( w^C \) for successfully delivered packet copies can be obtained by inspecting the device driver of the related adapter, whereas it is not available for failed copies. In the latter case, using for \( w^C \) the default retry limit as per Wi-Fi specifications [8] is not the best choice, because of the rate adaptation mechanism (Minstrel) [22], which changes transmission parameters dynamically according to some channel statistics. Hence, for dropped packets (i.e., if \( \ell^C_i = 1 \) we assumed \( w^C_i \) to be equal to the maximum among all the values we measured in the experiments, that is \( w^C_i = \max_{V \in \{L,N,C\} \in C_L} w^C_{i,V} \) = 21. Again, since the loss ratio in our experiments was quite low, this pessimistic approximation is perfectly acceptable.

Let \( \overline{\pi^L} \) be the average number of transmission attempts on air per packet, considering all channels. It depends on the overhead caused by both retransmissions and seamless redundancy. In PoW, every packet is always transmitted on all channels, so

\[
\overline{\pi^{\text{PoW}}} \triangleq \frac{1}{N} \sum_{i=1}^{N} \sum_{C \in C_L} w^C_i = \sum_{C \in C_L} \overline{\pi^C} = |C_L| \cdot \overline{n}, \tag{10}
\]

where \( \overline{n} = \frac{1}{|C_L|} \sum_{C \in C_L} \overline{\pi^C} \) is the mean number of attempts per packet and per channel, and describes, on average, the behavior of a single conventional Wi-Fi channel. Due to the different spectrum conditions of physical channels \( C \in C_L \), \( w^C \) values usually do not coincide and the same holds for \( \overline{\pi^C} \). Since \( \overline{\pi^C} \geq 1, \overline{\pi^{\text{PoW}}} \geq |C_L| \), where \( |C_L| = 2 \) for duplex links. Roughly speaking, spectrum consumption in our PoW testbed is, on average, twice as much as Wi-Fi.

When a DA mechanism is in effect, equation (10) can be rewritten as

\[
\overline{\pi^{\text{DA}}} = \frac{1}{N} \sum_{i=1}^{N} \left( w^{Q_i} + \sum_{C \in Q_i} w^C_i \right). \tag{11}
\]

Unlike PoW, \( w^C_i \) can be equal to 0 for \( C \in Q_i \), as the only constraint is that every packet has to be sent on air at least once on at least one channel. Therefore, like plain Wi-Fi, \( \overline{\pi^{\text{DA}}} \) could be as low as 1. This happens when, for every packet \( m_i \), \( w^{Q_i} = 1 \) and \( w^{C_i} = 0 \). In theory, for some specific runs, spectrum consumption for \( L \) may be even lower than any of its physical channels when used as simplex links. In other words, condition \( \overline{\pi^{\text{DA}}} < \min_{C \in C_L} (\overline{\pi^C}) \) might possibly hold.

For instance, when \( L = A+B \), consider the case \( w^{Q_i} < w^{C_i} \) and \( w^{C_i} = 0 \) (i.e., the number of attempts that would be performed on \( Q_i \), in the absence of DA is higher than on \( Q_i \), but all of them are prevented by DA), where \( Q_i = A \) if \( i \) is even and \( Q_i = B \) otherwise.

In general, \( e^C_i = 0 \rightarrow w^C_i = w^C_i \), while \( e^C_i = 1 \rightarrow w^C_i < w^C_i \). This implies that \( w^C_i \leq w^C_i - e^C_i \). For a given DA mechanism, from (8), (10), and (11), we have

\[
\overline{\pi^{\text{DA}}} \leq \overline{\pi^{\text{PoW}}} - \overline{\pi^{\text{DA}}}. \tag{12}
\]

Hence, DA mechanisms never worsens spectrum consumption of Wi-Red.

3) Communication efficiency: from the originator viewpoint, it is defined as the reciprocal of the average number of transmission attempts per packet. For physical channel \( C \) it corresponds to \( \eta^C = 1/\pi^C \). While being related to the goodput, they do not coincide, as \( \eta^C \) does not take into account protocol overheads (frame preambles, MAC headers and trailers, acknowledgments, inter-frame spaces, random backoff, etc.), but only retransmissions. Definition for redundant link \( L \) is similar, \( \eta^L = 1/\pi^L \). If DA is used, a lower bound can be found for \( \eta^{\text{DA}} \), defined as

\[
\eta^{\text{DA}} \triangleq \frac{1}{\overline{\pi^{\text{PoW}}} - \overline{\pi^{\text{DA}}}} \leq \eta^{\text{DA}}. \tag{13}
\]
4) Relative network load: The load on the link for a given DA mechanism with respect to PoW is defined as

\[ \vartheta_{DA} \triangleq \frac{w_{DA}}{w_{PoW}} = \frac{|C_L| \cdot \vartheta}{\vartheta} \]  

(14)

where \(1/\vartheta_{PoW} \leq \vartheta_{DA} \leq 1\) (the lower, the better). Similarly, the relative load implied, on average, over all, simple Wi-Fi links, is expressed as \(\Theta_{DA} \triangleq \frac{\vartheta_{DA}}{\vartheta} = |C_L| \cdot \vartheta_{DA}\).

Equation (14) can be rewritten as \(\vartheta_{DA} = \frac{\vartheta_{DA}}{\vartheta} \cdot |C_L| \cdot \vartheta\), which shows that the average number of transmission attempts on air per packet, including all channels, is given by the product of three factors: the relative load of the specific DA technique over PoW, the number of channels, and the mean per-packet load in Wi-Fi due to retransmissions, obtained by averaging all channels of the redundant link. Only the first factor depends on DA effectiveness (which is affected, e.g., by \(T_{LRE}\)). Thus, \(\vartheta_{DA}\) is a valuable metric to assess the ability of a certain DA mechanism to save bandwidth.

Due to the limitations of our testbed, what we can measure is an upper bound on \(\vartheta_{DA}\), i.e.,

\[ \vartheta_{DA} \triangleq \left(1 - \frac{\vartheta_{DA}}{\vartheta} \right) \geq \vartheta_{DA}. \]  

(15)

Therefore, DA performance is never worse than what we found using our approximate analysis.

5) Simplex transmissions on the redundant link: The last quantity we consider is the fraction of packets which, as a consequence of early termination, are only sent on \(Q_i\), that is, for those with \(w_{C}^i = 0, \forall C \in Q_i\). From the point of view of nearby wireless STAs, transmission of these packets appears to take place as on a conventional Wi-Fi (simplex) link, which is not fixed but dynamically selected among channels \(C \in C_L\).

Counting these packets exactly is not possible in our testbed. However, condition

\[ z_i^C \triangleq \left[ e_i^C \land (w_i^C = 1) \right] \]  

(16)

implies that no transmission attempts at all are performed for \(m_i\) on \(C\). In fact, \(w_i^C = 1\) means that only the initial attempt is performed by PoW, but it is canceled by DA since \(e_i^C = 1\). For the quickest channel (16) correctly yields \(z_i^{Q_i} = 0\). Hence, \(\varpi^C = \frac{1}{N} \sum_{i=1-N} z_i^C\) is a lower bound on the fraction of packets whose transmission is fully prevented on \(C \in C_L\) by DA.

For the redundant link \(L\), Boolean condition \(z_i^C \triangleq \prod_{C \supseteq L} z_i^C\) implies that \(m_i\) transmission is completely prevented on every channel other than the quickest. We are interested in the metric

\[ z_i^L = \frac{1}{N} \sum_{i=1-N} z_i^L, \]  

(17)

which represents a lower bound on the fraction of packets that were sent on air on one channel only (the quickest), as on a simplex link. Since in our duplex testbed there is only one channel other than \(Q_i\), \(z_i^A\) and \(z_i^B\) can not be both true, which means that \(z_i^{A+B} = z_i^A + z_i^B\). In this specific case, \(\varpi^{A+B} = \varpi^A + \varpi^B\). Given that \(z_i^{A+B} \rightarrow (e_i^{A+B} = 1)\), where \(e_i^{A+B} = e_i^A + e_i^B\), it also follows that \(\varpi^{A+B} \leq \varpi^A\).

E. Experimental Results for RDA

To assess the improvements RDA can bring in the real world, three experiments were run on the PoW testbed in different operating conditions. Channel \(A\), set in the 2.4 GHz band, was quite cramped, due to the presence of many active wireless devices in nearby premises, which were not under our control. To try ensuring some repeatability, every experiment lasted for a whole day (24 hours). Even so, results concerning channel \(A\) in the different runs can not be reliably compared one another. Channel \(B\), set in the 5 GHz band, had very light load. Thus, random interfering traffic was purposely injected on that channel during experiments. Interference consisted in the repeated generation of packet bursts. Payload size of interfering packets was set to 1500 bytes. The number of packets in each burst was not fixed, but followed an exponential random distribution with mean value 300 and truncated to 1500. Within any burst, transmission requests for packets were spaced by 400μs. The duration of the gap between the end of a burst and the beginning of the following one was randomly selected, according to an exponential distribution with mean value 200 ms and truncated to 20 s.

Runs were carried out with either 1, 2, or 4 interfering STAs on \(B\) (1 Int., 2 Int., and 4 Int., respectively). Experimental results, including bounds calculated using virtual RDA analysis, are reported in Table II, under the assumption that \(T_{LRE} = 0\) (i.e., when early termination is assumed to be implemented in hardware in the RSTA). For each experiment, three rows are included, for physical channels (\(A\), \(B\)) and the redundant link (\(A + B\)). The first set of columns includes statistics about the transmission latency \(d\): mean value, standard deviation, median, 99.99 percentile, and maximum. In the second set of columns, the deadline miss ratio \(\epsilon_{C,B}^{\text{RDA}}\) (percentage of packets whose latency exceeded \(H\), where \(H\) is set to either 10 or 100 ms) and the packet loss ratio \(\xi_{C,B}^{\text{RDA}}\) are shown. Although this paper does not focus on communication quality, these quantities provide a glimpse of the improvements seamless redundancy achieves.

The most interesting results, reported in the sets of columns on the right, concern RDA ability to prevent useless attempts. Relevant measured quantities for physical channels \(A\) and \(B\) and the redundant link \(A + B\) include the fractions of early terminated packets (\(\varpi^C\)) and simplex packets (\(\varpi^C\)), as well as the mean number of transmission attempts performed per packet (\(\varpi^C\)). Communication efficiency when no DA is in use, both for physical channels (\(\eta^A\) and \(\eta^B\)) and the redundant link (\(\eta^\text{PoW}\), is given in column \(\eta^C\). Finally, the rightmost columns report metrics about RDA computed using virtual analysis, i.e., the lower bound on communication efficiency (\(\eta^\text{RDA}\)) and the upper bounds on the relative network load with respect to either PoW (\(\vartheta^\text{RDA}\)) or plain Wi-Fi (\(\vartheta^\text{RDA}\)).

Experimental results are quite valuable and show that, in real operating conditions, RDA cuts redundancy overhead tangibly. The lower bound \(\eta^\text{RDA}\) on communication efficiency (and hence, \(\eta^\text{RDA}\) as well) was always noticeably higher than \(\eta^\text{PoW}\), irrespective of the interfering traffic, and the upper bound \(\vartheta^\text{RDA}\) on the network load relative to PoW ranged from 64.54% to 79.19%. This means that all benefits on commu-
nication quality brought by seamless redundancy are achieved at a fraction of the bandwidth. Since spectrum conditions may vary suddenly and unexpectedly, fair comparison between redundant and simplex Wi-Fi links requires that the worst of the physical channels is taken into account. As can be seen in the table, in the $4$ Int. case communication efficiency of RDA is marginally worse than channel $B$ alone (62.67% vs. 69.33%), but communication quality on $A$ + $B$ is much better than both $A$ and $B$.

An additional benefit of RDA is that, the higher the overall network traffic, the better it works. See, e.g., when interfering traffic on $B$ is varied. The measured fraction $\tau_{\text{RDA}}$ of early terminated packet transmissions with 1, 2, and 4 interferers was equal to 42.98% ($1$ Int.), 65.03% ($2$ Int.), and 87.66% ($4$ Int.), respectively. Concerning the fraction $\tau_{\text{RDA}}$ of simplex packet transmissions, when traffic is not negligible more than half of the packets (54.78% and 58.63%, for $2$ Int. and $4$ Int., respectively) were actually sent only on one channel, hence causing no overhead on the other. This is no surprise, as RDA effectiveness improves when variability of transmission latencies on replicated channels gets higher. Such behavior is highly desirable since, when overall network traffic increases, the overhead due to redundancy consequently shrinks, hence lowering the impact of transmitting RSTAs on the network load. This counteracts congestion phenomena affecting random access techniques, which worsen the quality of communication (see, e.g., the latency on channel $B$ when interfering traffic is increased). As a consequence, network stability (overall traffic on air vs. total offered load) improves.

The effect of using sub-optimal (e.g., software) RDA solutions, modeled by considering non-negligible LRE latencies, is shown in Fig. 4. Three plots are reported, corresponding to the above experimental runs, where $T_{\text{LRE}}$ in (6) is virtually varied in the range $[0$ – $1000] \mu$s. As can be seen, delays due to XACK processing always worsen spectrum consumption as, on average, less transmission attempts can be terminated in advance. For example, in the $1$ Int. case, $\tau_{\text{RDA}}$ decreases by a factor 4 (from 42.98% to 10.18%) when $T_{\text{LRE}}$ is increased by 1 ms. For this reason, implementation of RDA in the hardware (or firmware) of Wi-Fi adapters is highly suggested. This also confirms the advantages brought by our approximate virtual approach to RDA evaluation (based on the analysis of PoW experimental logs) with respect to measurements carried out directly on a non-properly optimized software RDA implementation. Moreover, it is important to remark the ability of the virtual approach to infer results for different techniques and parameters starting from the same set of data, allowing for a fair comparison between them.

V. PROACTIVE DUPLICATION AVOIDANCE

Proactive Duplication Avoidance (PDA) is meant to improve RDA behavior by further reducing the number of useless duplicates sent on air. Unlike RDA, which has no counterindications, PDA may set some trade-off among performance indices (e.g., the approach described below trades responsiveness for communication efficiency). Therefore, when employing these mechanisms, a compromise has to be necessarily found, typically by tuning suitable configuration parameters.

PDA often relies on heuristics, which could be arbitrarily complex and possibly intertwined with the IEEE 802.11 MAC. The class of PDA mechanisms we considered in this paper only concerns duplex redundant links, and operates by intentionally disabling transmission requests on physical channels. In this way, the likelihood that the LRE succeeds in canceling the transmission of the deferred copy may increase tangibly, hence reducing spectrum consumption.

A. Timed Duplicate Deferral

The PDA technique we analyze here is called Timed Duplicate Deferral (TDD). A similar solution, not coupled with early termination, was proposed in [23]. TDD is much simpler than the Dynamic Duplicate Deferral (DDD) technique presented in [17], as its operation does not rely on any indication provided at runtime by the MAC layer. Thus, it can be easily implemented using commercial Wi-Fi adapters.

TDD operates as follows: upon generation of packet $m_i$, the LRE immediately invokes its transmission on $A$ (referred to in the following as primary channel). At the same time, it starts a countdown timer initialized to the deferral time $T_{D,i}$. When $T_{D,i}$ expires, and provided that the XACK from $A$ has not been raised in the meanwhile, transmission is also invoked on $B$ (secondary channel). In formulas

$$T_{D,i}^B = T_{D,i}^A + T_{D,i},$$

A distinct $T_{D,i}$ value can be chosen for each packet $m_i$ depending, e.g., on its payload size and the bit rate chosen at runtime by the rate adaptation mechanism.

TABLE II

<table>
<thead>
<tr>
<th>Exp. Chan.</th>
<th>$\pi^C$</th>
<th>$\sigma^C_d$</th>
<th>$d^C_{p50}$</th>
<th>$d^C_{p99}$</th>
<th>$d^C_{\text{max}}$</th>
<th>$\gamma^C_{\text{p2.5ms}}$</th>
<th>$\gamma^C_{\text{p10ms}}$</th>
<th>$\gamma^C_{\text{p100ms}}$</th>
<th>$\gamma^C_L$</th>
<th>$\rho^C$</th>
<th>$\eta^C_{\text{RDA}}$</th>
<th>$\vartheta^C_{\text{RDA}}$</th>
<th>$\Theta^C_{\text{RDA}}$</th>
<th>$\tilde{\Theta}^C_{\text{RDA}}$</th>
<th>$\tilde{\eta}^C_{\text{RDA}}$</th>
<th>$\delta^C_{\text{RDA}}$</th>
<th>$\tilde{\delta}^C_{\text{RDA}}$</th>
<th>$\hat{\delta}^C_{\text{RDA}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+ B</td>
<td>0.281</td>
<td>0.500</td>
<td>0.076</td>
<td>12.516</td>
<td>40.726</td>
<td>0.2014</td>
<td>0.000</td>
<td>0.000</td>
<td>0.8766</td>
<td>0.5863</td>
<td>2.4723</td>
<td>40.45</td>
<td>62.67</td>
<td>64.54</td>
<td>129.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.576</td>
<td>1.187</td>
<td>0.076</td>
<td>40.721</td>
<td>201.39</td>
<td>4.951</td>
<td>0.0014</td>
<td>0.000</td>
<td>0.0503</td>
<td>0.5478</td>
<td>2.1438</td>
<td>46.65</td>
<td>66.96</td>
<td>69.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>2.542</td>
<td>6.753</td>
<td>1.021</td>
<td>126.92</td>
<td>219.09</td>
<td>41.683</td>
<td>0.7268</td>
<td>0.000</td>
<td>0.7427</td>
<td>0.4626</td>
<td>1.4424</td>
<td>69.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A+ B</td>
<td>0.660</td>
<td>1.384</td>
<td>0.084</td>
<td>48.306</td>
<td>209.41</td>
<td>2.7142</td>
<td>0.0359</td>
<td>0.0244</td>
<td>0.221</td>
<td>0.2382</td>
<td>1.0422</td>
<td>96.14</td>
<td>61.14</td>
<td>79.19</td>
<td>158.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.488</td>
<td>0.754</td>
<td>0.229</td>
<td>19.677</td>
<td>91.892</td>
<td>0.5046</td>
<td>0.000</td>
<td>0.000</td>
<td>0.4298</td>
<td>0.3804</td>
<td>2.0653</td>
<td>48.42</td>
<td>64.82</td>
<td>78.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.108</td>
<td>0.135</td>
<td>0.081</td>
<td>2.193</td>
<td>52.242</td>
<td>0.0012</td>
<td>0.000</td>
<td>0.000</td>
<td>0.4298</td>
<td>0.3904</td>
<td>2.0653</td>
<td>48.42</td>
<td>64.82</td>
<td>78.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.407</td>
<td>1.113</td>
<td>0.083</td>
<td>26.138</td>
<td>135.78</td>
<td>1.7385</td>
<td>0.0313</td>
<td>0.0028</td>
<td>0.1917</td>
<td>0.1760</td>
<td>1.0231</td>
<td>97.75</td>
<td>61.14</td>
<td>79.19</td>
<td>158.38</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Copyright (c) 2018 IEEE. Personal use is permitted. For any other purposes, permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.
B. Experimental Evaluation of TDD

TDD performance can be experimentally assessed in our testbed using two different approaches, which will be shown to provide very similar results. In both cases, bandwidth saving is evaluated when varying the deferral time, by analyzing logs in the same way as for RDA.

1) Real transmission deferral: The most straightforward way to analyze TDD behavior is to rely on a slightly modified PoW testbed, where the transmission request issued by the measurement task on $B$ for each packet $m_i$ is actually delayed by $T_{D,i}$ with respect to $A$. To make timing accuracy independent from the implementation of the measurement task and the underlying computing platform (including the operating system), the exact displacement between the copies of every packet $m_i$ (that in theory corresponds to $T_{D,i}$) was evaluated from measured timestamps as $t^B_i = t^A_i - T_{D,i}$.

2) Virtual transmission deferral: Approximate yet satisfactory results can be obtained from the logs acquired on the unmodified PoW testbed (i.e., without any displacement between $t^A_i$ and $t^B_i$). Results in [19] show that, in a well-designed redundant link, channel conditions can be considered statistically independent. This means that deferring transmission requests on $B$ does not affect the behavior of $A$, and vice-versa. Moreover, channel behavior in the short term is mostly stationary, hence, statistics on experimental results do not change upon displacement in time, as long as the offset is short enough (in the order of 1 ms).

According to the above hypotheses, one can reasonably assume that the sequence of samples obtained by: a) adding the deferral time to every timestamp $(t^B_i)$ and $(t^A_i)$ in the tuples $\tau^B_i, i = 1...N$ of a log acquired from channel $B$ in the unmodified PoW testbed, and b) leaving all the other quantities $(l_i, w_i, T_{D,i}, and T^A_i)$ unmodified, could satisfactorily mimic the sequence of tuples acquired on $B$ in a real setup like the one described in Section V-B1, where transmissions of copies are actually displaced. This approach permits to create a bunch of virtual TDD experiments from a single real PoW experimental run, where $T_{D,i}$ values (provided that they are small, so as not to contradict stationary hypothesis) can be decided during post-analysis.

In the following, timestamps with the double prime symbol refer explicitly to virtual TDD experiments. For example, the estimated ACK reception time for packet $m_i$ on the secondary channel $B$ in a virtual TDD experiment (characterized by the sequence of $T_{D,i}$ values) is

$$t^B_{i,X,i} = t^B_{X,i} + T_{D,i} \quad (19)$$

where $t^B_{X,i}$ is the timestamp taken in the unmodified PoW testbed. Timestamps $t^B_{D,i}$ and $t^B_{W,i}$ are defined similarly. For the primary channel $A$, virtual and measured timestamps coincide.

In presence of TDD deferral, the transmission latency on the redundant link for packet $m_i$ is $t^{A,TDD} = \min(d^{A,i}a^{A,i} + T_{D,i})$. This implies that, unlike RDA, $t^{A,TDD} \geq d^{A,TDD}$.

The same approach used for RDA analysis can be applied to estimate the effects of early termination in TDD, by replacing real samples in the log with virtually deferred ones. Since channel usage is no longer symmetric, (6) splits into

$$e^A_i \triangleq \begin{cases} t^A_i & = 0 \land t^A_{X,i} + T_{LRE} < t^B_{W,i} \\ \text{otherwise} \\ \end{cases} \quad (20)$$

and, from (19),

$$e^B_i \triangleq \begin{cases} t^B_i & = 0 \land t^B_{X,i} + T_{LRE} < t^A_{W,i} \\ \text{otherwise} \\ \end{cases} \quad (21)$$

C. Validation of Virtual TDD Analysis

A preliminary experiment was carried out to assess if results obtained through virtual deferral of samples performed offline are comparable to actually displacing transmission requests in the testbed at runtime. Since the payload size of packets sent by the measurement task was fixed, all $T_{D,i}$ offsets in experimental runs have to be set to the same value $T_D$. Jitter induced by the operating system on the actual displacement between the transmission times of the two copies of each packet in the modified PoW testbed described in Section V-B1 was found to be quite small. Therefore, we can practically assume that the deferral time in any run is fixed and coincides with the mean value of the measured offset, that is $T_D \approx T_D = \frac{1}{N} \sum_{i=1}^{N} (t^B_{X,i} - t^A_{X,i}).$

We performed several runs, where $T_D$ was varied in the range $[-250...+250] \mu s$. The nominal displacement used in each run was selected in such a way that the related $T_D$ approximately corresponds to one among 17 predefined values. In the case $T_D = 0$ (which corresponds to pure RDA), transmissions were not displaced, as in the unmodified PoW testbed. Negative $T_D$ values refer to experiments where $B$ was the primary channel, while $A$ was deferred by $|T_D|$. These
cases are explicitly included because behavior of channels is not symmetric.

To mitigate discrepancies due to slow variations of the wireless spectrum conditions, runs for the different deferral times were interleaved in a single experiment, which lasted one day. Such a kind of assessment was lacking in [18]. Two interferers were activated on B during the experiment (as in the 2 Int. case). Results, concerning both the fraction of early terminated packets (\(\pi^{\text{TDD}}\)) and mean transmission latency (\(\bar{T}^{\text{TDD}}\)), are reported in Fig. 5 using “×” symbols.

The very same logs, restricted to the subset of samples for which \(T_D = 0\) (i.e., when the testbed operated according to pure PoW, with no deferrals), were then used to generate virtual TDD runs according to Section V-B2, where \(T_D\) was varied with fine granularity in the range \([-250 \ldots +300]\) µs. Resulting plots are superposed to Fig. 5 using solid lines. As can be seen, there is a very good agreement between the results obtained with the two approaches. This means that virtual deferral carried out during post-analysis provides a reliable approximation of real deferral performed by TDD. For this reason, the relation between spectrum consumption, communication quality, and \(T_D\), will be analyzed in the following by using the virtual approach.

D. Experimental Results for TDD

To provide coherent results, the very same experimental PoW logs used in Section IV to analyze RDA (1 Int., 2 Int., and 4 Int., depending on the amount of interference on B) were employed to carry out virtual TDD analysis. First, they were processed to cope (virtually) with transmission deferral; then, the same analysis as for RDA was performed to estimate the effects of early terminations. In particular, we analyzed the tradeoff between spectrum consumption and communication quality when \(T_D\) is varied between 0 (that corresponds to pure RDA) and 500 µs.

Results for the three runs are reported in Fig. 6. Plots in the upper part of the figure show some metrics about bandwidth consumption, i.e., the fraction of early terminated packets (\(\pi^{\text{TDD}}\)), the fraction of simplex packets (\(\hat{\pi}^{\text{TDD}}\)), and the relative network load with respect to PoW (\(\gamma^{\text{TDD}}\)). Instead, those in the lower part concern the average latency on the redundant link (\(\bar{\tau}^{\text{TDD}}\)).

As can be seen, in each run there are two thresholds for \(T_D\), located around \(±100\) µs, within which spectrum consumption does not vary appreciably, this meaning that TDD adoption does not bring any improvement over pure RDA. However, passed those thresholds, behavior improves suddenly, as witnessed by, e.g., the noticeable increase of \(\pi^{\text{TDD}}\) values. This is because transmission requests on the secondary channel are deferred enough by TDD, so that most of them can be prevented by the LRE. TDD benefits are higher when channels are not heavily loaded, as in the 1 Int. case, mostly because saving due to RDA alone is less tangible. In this respect, TDD adoption improves Wi-Red stability versus network traffic.

In theory, as long as queuing phenomena in Wi-Fi adapters are negligible, parameter \(T_D\) has no effect on the amount of lost packets in TDD, which is the same as RDA and PoW. Instead, it affects transmission latencies, as they are measured from the time the transmission request is issued on the primary channel. As can be seen in the plots of Fig. 6, mean latency \(\bar{\tau}^{\text{TDD}}\) progressively worsens as \(|T_D|\) increases (in either direction). This is quite obvious, since TDD intentionally delays one of the two channels.

Plots about \(\bar{\tau}^{\text{TDD}}\) are not symmetric about the y-axis. Their shape (but not the values) in the different runs is mostly the same when \(T_D > 0\). However, this is no longer true when \(T_D < 0\). This can be explained by remembering that, in the latter case, transmissions are delayed on A, and so a higher number of packets arrive first on B when \(|T_D|\) is increased. As a consequence, communication quality on B matters more. When interference on that channel is higher, as in the 4 Int. case, the related latency is larger. This implies that, for negative \(T_D\) values, the slope of the plots concerning the latency on the redundant link is steeper.

The optimal values for parameter \(T_D\) are located just above said thresholds, and approximately correspond to the overall duration on air of the initial attempt for \(m_i\), on the primary channel (including DATA and ACK frames, SIFS, and \(T_{\text{LRE}}\)), plus a safety margin. With this arrangement, all packet transmissions that find such channel idle and immediately succeed (a non-unusual condition in a well-dimensioned Wi-Fi network, as experimentally shown in [24]) are not replicated on the secondary channel. This particular choice provides significant benefits concerning spectrum consumption, by keeping the latency increase, on average, to the minimum.

For example, if \(T_D\) is set to either \(-100\) µs or \(+100\) µs (near to the optimal values for our experimental conditions), the values in Table III are obtained. With these configurations, \(\pi^{\text{TDD}}\) was always higher than \(\sim 90\%\) for all the experiments reported in the plots (and, in the 4 Int. case, it exceeded 97\% when \(T_D = +100\) µs). This means that less than 10\% of the packets did not incur in early termination. At the same time, \(\bar{\tau}^{\text{TDD}}\) never grown more than 76 µs with respect to pure RDA (see the 4 Int. case, \(T_D = -100\) µs). In the case adaptive TDD algorithms were implemented, able to dynamically select
This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.

The final version of record is available at http://dx.doi.org/10.1109/TWC.2018.2884914

The best primary channel, $d_{\text{TDD}}$, increase would likely remain below 25 $\mu$s (see the 2 Int. case, $T_D = +100$ $\mu$s).

VI. CONCLUSION

Seamless redundancy applied to IEEE 802.11 greatly improves its reliability and makes it suitable for time-sensitive control applications, preserving backward compatibility with Wi-Fi. When PRP is layered directly above wireless adapters, as in the PoW approach, communication quality improves dramatically, but spectrum consumption becomes twice as much as non-redundant Wi-Fi. The Wi-Red proposal exploits synergies between the MAC retransmission mechanism and channel redundancy in order to improve behavior further. In particular, duplication avoidance mechanisms reduce spectrum consumption by preventing unnecessary transmission attempts.

This paper specifically focuses on the ability of such mechanisms to cut down the additional traffic implied by seamless redundancy. Suitable metrics have been defined to this purpose, which were evaluated experimentally on a real testbed. In order to be effective, early termination of packet transmissions has to be implemented in the hardware/firmware of Wi-Fi adapters, which is currently out of our reach. For this reason, an approximate method was devised, where experimental data are acquired on a real testbed that implements PRP on commercial Wi-Fi devices and then processed under the (verified) hypotheses of independence between subsequent transmissions on air and slowly-varying spectrum conditions.

Although in this way we can only obtain a lower bound on the amount of bandwidth saved by duplication avoidance mechanisms (which means that the related results are pessimistic), improvements we measured are nevertheless noteworthy. Reactive approaches, which cancel transmission of duplicate copies upon reception of the related ACK on at least one channel, provide all the benefits of PoW (and often behave even better in terms of improved reliability and timeliness) but using less bandwidth. In our experiments, RDA traffic (or, better, its upper bound as given by $\Theta_{\text{RDA}}$) was 29.08% to 58.36% higher than what generated, on average, by non-redundant Wi-Fi. For the proactive TDD approach, additional spectrum consumption with respect to Wi-Fi, as per $\Theta_{\text{TDD}}$, lay in the range from 9.72% to 25.56%. In the TDD case, the price to pay is a slight increase in the transmission latency (some tens of $\mu$s).

In a world where the wireless spectrum is more and more congested, adoption of seamless redundancy to increase Wi-Fi reliability practically mandates the inclusion of duplication avoidance mechanisms. As future work, we plan to investigate adaptive PDA solutions, aimed at providing optimal performance in spite of variations of the generated traffic pattern and the wireless spectrum conditions, and to analyze networks including multiple RSTAs.

REFERENCES


Gianluca Cena (SM’09) received the Laurea degree in electronic engineering and the Ph.D. degree in information and system engineering from the Politecnico di Torino, Italy, in 1991 and 1996, respectively. Since 2005 he has been a Director of Research with the Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy (CNR–IEIIT), Torino. His research interests include wired and wireless industrial communication systems, real-time protocols, and automotive networks. In these areas he has coauthored about 130 technical papers, three of which awarded as Best Papers of the 2004, 2010, and 2017 editions of the IEEE Workshop on Factory Communication Systems, and one as 2017 Best Paper for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, plus one international patent. Dr. Cena served as a Program Co-Chairman for the 2006 and 2008 editions of the IEEE International Workshop on Factory Communication Systems, and as a Track Co-Chairman in six editions of the IEEE International Conference on Emerging Technologies and Factory Automation. Since 2009 he has been an Associate Editor of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.

Stefano Scanzio (S’06-M’12) received the Laurea and Ph.D. degrees in Computer Science from Politecnico di Torino, Torino, Italy, in 2004 and 2008, respectively. He was with the Department of Computer Engineering, Politecnico di Torino, from 2004 to 2009, where he was involved in research on speech recognition and, in particular, he has been active in classification methods and algorithms. Since 2009, he has been with the National Research Council of Italy (CNR), where he is a tenured Researcher with the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), Torino. Dr. Scanzio served as a Work-in-Progress Co-Chairs in the 2018 edition of the IEEE International Workshop on Factory Communication Systems (WFCS 2018). He teaches several courses on Computer Science at Politecnico di Torino. He has authored and co-authored of more than 50 papers in international journals and conferences, in the area of industrial communication systems, real-time networks, wireless networks and clock synchronization protocols. He received the 2017 Best Paper Award for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and the Best Paper Awards for the papers he presented at the 8th and 13th IEEE Workshops on Factory Communication Systems (WFCS 2010 and WFCS 2017).

Adriano Valenzano (SM’09) received the Laurea degree magna cum laude in electronic engineering from Politecnico di Torino, Torino, Italy, in 1980. He is Director of Research with the National Research Council of Italy (CNR). He is currently with the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), Torino, Italy, where he is responsible for research concerning distributed computer systems, local area networks, and communication protocols. He has coauthored approximately 200 refereed journal and conference papers in the area of computer engineering.

Dr. Valenzano is the recipient of the 2013 IEEE IES and ABB Lifetime Contribution to Factory Automation Award. He was also awarded for the best paper published in the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS during 2016, and received the Best Paper Awards for the papers presented at the 5th, 8th and 13th IEEE Workshops on Factory Communication Systems (WFCS 2004, WFCS 2010 and WFCS 2017).

Adriano Valenzano has served as a technical referee for several international journals and conferences, also taking part in the program committees of international events of primary importance. Since 2007, he has been serving as an Associate Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.