Practical Evaluation of Cooperative Communication for Ultra-Reliability and Low-Latency

Martin Serror*, Sebastian Vaaßen*, Klaus Wehrle*, James Gross†
*Chair of Communication and Distributed Systems, RWTH Aachen University, Germany
†School of Electrical Engineering & Computer Science, KTH Royal Institute of Technology, Sweden
{serror, wehrle}@comsys.rwth-aachen.de, sebastian.vaassen@rwth-aachen.de, james.gross@ee.kth.se

***Abstract*—Existing wireless communication systems are not able to meet the stringent requirements for critical machine-to-machine communications regarding ultra-reliability and low-latency. Since increasing the communication reliability often comes at the price of increasing the latency as well, new mechanisms must be proposed that consider both challenges together. A promising approach, according to analytical work, is to increase the reliability by using cooperative diversity, where all stations within range help each other in the transmission process. Theoretical analyses, however, only provide a limited insight regarding the actual performance due to the strong assumptions they make to model such complex systems. In this paper, we thus evaluate the practical feasibility of ultra-reliable low-latency communication through cooperation by designing a data link protocol that incorporates a best relay selection mechanism. We implement our protocol in a real-world testbed, consisting of software-defined radios, to gain a better understanding of how future ultra-reliable low-latency systems should be designed and implemented. Our measurement campaigns show that at a given low target latency of $1\,\mathrm{ms}$, we achieve a packet error rate between 10^{-5} and 10^{-7} with a standard 802.11a physical layer.**

I. Introduction

Wireless communication is well established in business and home environments offering mobility and high data rates at low installation and maintenance costs. In other domains, however, wired communication is still prevalent since, in contrast to wireless, it ensures a high reliability and a low-latency. In critical Machine-to-Machine Communications (M2M) as can be found in industrial automation, for example, time-critical messages are exchanged between sensors, actuators, and controllers requiring a communication latency of a few milliseconds or even in the sub-millisecond range and a Packet Error Rate (PER) down to 10^{-9} [1]. In the context of 5G, this type of communication requirements is often referred to as Ultra-Reliable Low-Latency Communication (URLLC) [2]. Existing wireless communication systems for industrial automation, such as WirelessHART and ISA100.11a [3], enable a more reliable and periodic data delivery, but do not reach the aforementioned stringent communication guarantees.

A well-known technique to increase the communication reliability is diversity—either in time, frequency, or space. While time diversity spreads information over time instances, frequency diversity builds upon different communication channels, and spatial diversity leverages uncorrelated transmission paths through the network. Each technique, however, has its drawbacks: Time diversity increases the communication latency, frequency diversity depends on complex coordination schemes for the different transmission channels, and spatial diversity requires additional hardware (i. e., multiple antennas), which might not be available due to costs and size constraints. A promising approach to tackle this challenge is cooperative diversity, a special form of spatial diversity. There, instead of using additional antennas, messages are relayed via cooperating stations in the network, which thus form a virtual antenna array leveraging the broadcast nature of the wireless channel. The benefits of relaying on the communication reliability are well-researched and analyzed [4], [5]. Likewise, for URLLC, the potential of relaying in multi-user scenarios with a stringent communication deadline has been shown analytically [6], [7]. It is known that full diversity order can be achieved when selecting the best relay out of all available relays [8]. In general, each additional station increases the diversity degree, which leads to a higher reliability of the system. However, this implies that instantaneous Channel State Information (CSI) of the links is available at the senders such that, for each connection, the best available relay can be selected.

Such analytical approaches, nevertheless, often abstract from the significant challenges that arise when implementing these systems in real-world deployments. For instance, the synchronization of the stations, the collection of CSI, and the best relay selection must be ensured within a (sub-)millisecond latency bound, while also guaranteeing a high reliability. Furthermore, the dynamics and unpredictable nature of the propagation environment in which URLLC systems operate, can hardly be modeled adequately in theoretical analyses. These challenges must be addressed in the design of such protocols since strong abstractions may lead to false conclusions regarding their performance evaluation. Therefore, we focus on the *practical* feasibility of URLLC by means of cooperative transmission schemes based on instantaneous CSI, to verify empirically to what extend already existing analytical findings can be applied to the real world.

In this paper, we thus design and implement an experimental, cooperative URLLC protocol on Software-Defined Radio (SDR) boards to evaluate in different scenarios how to achieve URLLC in practice with a low protocol overhead. In the proposed data link protocol, an Access Point (AP) centrally schedules transmission slots for stations according to Time Division Multiple Access (TDMA). To ensure

978-1-5386-4725-7/18/$31.00 ©2018 IEEE

computational feasibility within the ultra-low latency bound, each station locally decides how to transmit its data packet within its given time slot. That is, according to instantaneous CSI, it decides whether to transmit the packet once with a stronger Modulation and Coding Scheme (MCS), twice via the direct link (weaker MCS), or via the best available relay (also weaker MCS). To enable cooperative diversity through relaying, each station including the AP overhears ongoing transmissions and may act as a relay on demand. To the best of our knowledge, this is the first real-world implementation of a URLLC protocol. According to our results, cooperative diversity significantly contributes to a high reliablity within the fixed low-latency bound, compared to solely relying on time diversity. In this, the best relay selection based on instantaneous CSI can be efficiently integrated into the data link protocol. The experimental results also indicate, however, that further reliability techniques, e. g., on the physical layer, are needed to fully achieve the anticipated guarantees.

The remaining structure of this paper is as follows: First, an overview of related work in the domain of URLLC and cooperative diversity is given (cf. Sec. II). Then, we provide a detailed design description of our relaying decision approach (cf. Sec. III). This design is empirically evaluated on SDRs regarding achieved reliability with a given low-latency in different scenarios (cf. Sec. IV). Finally, the main results of this paper are concluded (cf. Sec. V).

II. Related Work

In this section, we shortly present the related work in the realm of URLLC. All presented approaches have in common that they leverage cooperative diversity to increase the reliability while also specifying a fixed latency bound. We begin with presenting promising analytical approaches, then we continue with an overview of prototypical implementations.

A. Analytical Approaches

Occupy CoW [9] aims at low-latency and high-reliability through simultaneous relaying, i. e., multiple stations relay a packet simultaneously. The protocol is organized in communication cycles consisting of seven phases, which are either for uplink (from stations to controller), for downlink (from controller to stations), or for scheduling. Every message gets relayed at least once simultaneously by all stations that were able to decode it. The analytic performance evaluation reveals that even with a low cycle time of $2\,\mathrm{ms}$, a high reliability with a PER below 10^{-9} can be achieved. Moreover, the authors extended the protocol with network coding, which leads to further improvements in reliability [10]. Nevertheless, the protocol makes strong assumptions regarding the time synchronization of the stations and the analytic evaluation only offers an upper bound. It thus remains open how this protocol performs on real hardware.

Similar to our approach, the authors of [11] propose a TDMA-based approach with relaying to address reliable wireless industrial networks. To further increase the reliability with a low impact on the delay, the authors use Luby coded packets in the relaying process, where k original packets are encoded into $k+m$ packets and any subset of k correctly received packets suffices to retrieve the k original packets. The simulation results show that Luby coded packets reduce the number of needed transmissions while achieving the same reliability, as long as the transmission channels do not suffer from a high PER. These results thus indicate that the scalability of URLLC systems can be further improved by applying network coding techniques, which we, however, leave for future work.

The authors of [12] propose a wireless extension of the IO-Link standard, which is based on Bluetooth Low-Energy. Reliability is achieved through frequency hopping and up to two possible retransmissions. The authors analyze that, given a PER of 10^{-3} for one subcycle, a PER of 10^{-9} can be achieved within a communication latency of $5\,\mathrm{ms}$. Nevertheless, this result is not validated experimentally.

Such analytical approaches reveal interesting new techniques towards reliable and low-latency wireless communication. However, existing prototypical implementations do not yet target a (sub-)millisecond communication bound, as discussed in the next section.

B. Prototypical Implementations

EchoRing [13] is a distributed wireless token passing protocol for mission-critical communication. The token-passing in combination with failure detection and recovery enables, for each station, a deterministic access to the wireless medium. Furthermore, it uses cooperative diversity, by implementing a best relay selection scheme, to increase the reliability. Experimental evaluation results show that the prototypical implementation of EchoRing on SDRs achieves a PER below 10^{-6} with a latency bound of $10\,\mathrm{ms}$. Through its distributed organization, EchoRing does not suffer from a single-point-of-failure. A centralized approach, however, offers a lower coordination overhead and might enable tighter latency bounds.

The authors of [14] propose Real-time Network Protocol (RNP), a hybrid Medium Access Control (MAC) that uses cooperative diversity to achieve reliable communication. Communication is organized into superframes, which contain a TDMA and a Carrier Sense Multiple Access (CSMA) phase. The superframe thus guarantees a deterministic latency for the communication. Retransmissions of unsuccessful packet deliveries occur in the CSMA phase, which is initiated by the central gateway. Evaluation results show that within a latency of $100\,\mathrm{ms}$ a PER below 10^{-4} can be reached. Safety- and mission-critical applications in industrial automation, however, have more stringent latency and reliability bounds [1].

Marchenko et al. [15] evaluate different best relay selection schemes for URLLC. They differentiate between periodic, adaptive, and reactive relay selection. In the first scheme, relay selections are updated strictly after a certain time. Adaptive relay selection, in turn, takes the success ratio of the current relay into account and subsequently updates the relay selection if the performance deteriorates. Finally, reactive relay selection determines the best relay only after a failed direct transmission. In the evaluation, the reactive scheme

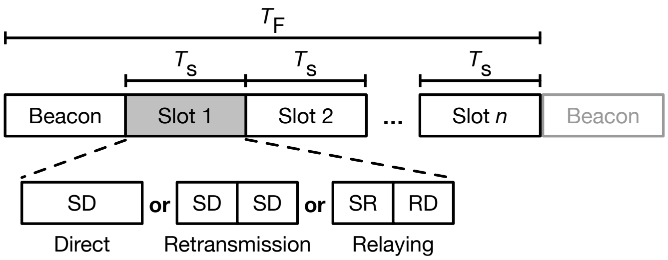


Fig. 1. TDMA superframe of the proposed MAC protocol. After the beacon slot each station is assigned a data slot where it can transmit one packet using Direct, Retransmission, or Relaying.

Prot ID	Type	Dest	Relay	Source	Length	CSI	Realloc	Alloc
1	1	1	1	1	1	*N*-1	1	*N*-1

(a) Structure of beacon frame.

Prot ID	Type	Dest	Relay	Source	Length	CSI	ACK	Payload
1	1	1	1	1	1	*N*-1	1	1...255

(b) Structure of data packet.

Fig. 2. Structure of beacon frame and data packet including the sizes of the different fields in bytes, where N denotes the number of participating stations.

shows the best performance in terms of reliability since the selection is based upon fresh and accurate CSI. However, the time overhead is considerable, introducing additional delays to gather and process the current CSI. In our approach, we therefore opt for a periodic scheme with a low periodicity, where CSI collection and relay selection are integrated into the anticipated latency bounds.

III. Design

In this section, we present the design of the proposed URLLC protocol. This design is the basis for our implementation on SDRs and the performance evaluation in Sec. IV. Firstly, we give a general overview of the main characteristics of the protocol (cf. Sec. III-A). Then, we present the different transmission options in more detail (cf. Sec. III-B). Finally, we describe how a transmission option is selected and discuss the resulting overhead (cf. Sec. III-C).

A. Overview

The proposed URLLC protocol is a MAC layer protocol enabling *reliable* and *time-bounded* communication for the participating stations. Time-critical communication is achieved through a deterministic TDMA scheme, where a centralized AP coordinates at which point in time a station is allowed to transmit. Reliability for the latency-bounded communication is achieved through locally selecting the best available transmission option based on instantaneous CSI. To continuously measure CSI, all stations and the AP overhear the transmissions from all other stations. This also enables cooperative communication, i. e., best relay selection, as explained in the further course of this section.

Regarding the TDMA protocol, we introduce time-bounded superframes, where every station is assigned one slot to transmit a data packet. The structure of such a superframe is depicted in Fig. 1. The superframe consists of a beacon slot and one data slot for each station. We assume the length of the superframe, denoted by T_{F}, to be short, e. g., $T_{\mathrm{F}} = 1\,\mathrm{ms}$. This length thus defines the communication latency of the system. In the beacon slot, the AP broadcasts a control message to all participating stations. This beacon serves as synchronization reference to the stations and includes a transmission schedule, which is assumed to be valid for at least several superframes, such that every station is notified in advance about schedule changes, even when it misses some beacons. We choose to announce schedule changes 50 superframes in advance, i. e., if a station misses 50 superframes in a row, it assumes that it lost connectivity to the AP and refrains from sending data packets. Note that this time bound, which corresponds to $50\,\mathrm{ms}$, can be increased or decreased depending on the deployment.

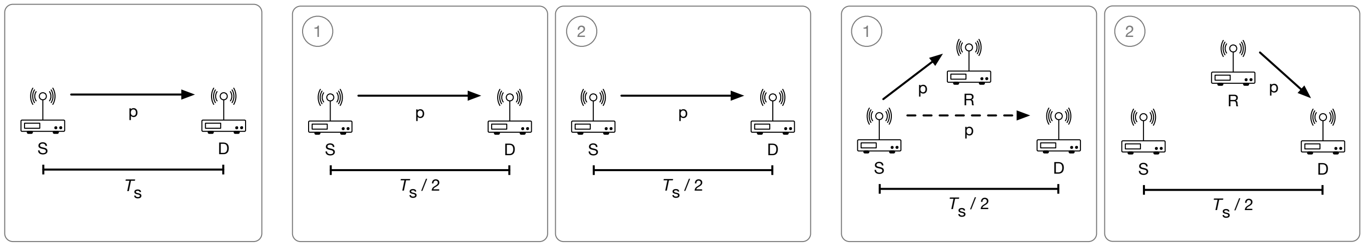
After the beacon slot, each participating station possesses one data slot according to the transmission schedule. In each data slot, the assigned station (S) sends a data packet (p) to a destination (D) using the best transmission option according to instantaneous CSI. Possible options are *Direct*, *Retransmission*, or *Relaying*, which are further explained in Sec. III-B. Each option must be performed within the time-bounded data slot, i. e., the MCS of transmissions and retransmissions must be adapted such that they conclude before the data slot expires.

To reduce the communication and time overhead, data packets are acknowledged by the receiver in the subsequent superframe. Therefore, the receiver piggybacks the Acknowledgement (ACK) in the header of its own data packet transmission. After sending a data packet, a sender thus receives, in the worst case, after less than $2\,\mathrm{ms}$ the corresponding ACK.

Fig. 2 shows the structures of the beacon frame and the data packet. Both start with a protocol ID to distinguish our packets from other wireless protocols, followed by a packet type field. The next three fields specify the destination, a potential relay, and the source of the packet. Then follows the length of the payload. The CSI is encoded, link by link, in the order of the station IDs. For the beacon, we introduce a reallocation counter to indicate the remaining number of superframes until a new schedule will become effective. Afterwards follows the transmission schedule. For the data packet, we reserve one byte to acknowledge the previously received packet. Then follows the payload with a maximum length of $255\,\mathrm{bytes}$.

B. Transmission Options

For a sending station (S), we define three distinct transmission options to convey its data packet (p) to the destination (D). Either of these transmission options must complete within the data slot of S, which has a fixed duration of T_{s}. The length of T_{s} can be individually set for each station, e. g., depending on the load of a station. Fig. 3 depicts the different transmission options, *Direct*, *Retransmission*, and *Relaying*, which we shortly present in the following.



(a) Direct. (b) Retransmission. (c) Relaying.

Fig. 3. Transmissions options for a sending station S. In (a), S uses the whole transmission time T_s for a direct transmission of the packet p with a strong MCS to the destination D. In (b), S transmits two copies of p directly to D within T_s. In (c), S sends p via the best available relay R to D, where p might also be overheard by D when S initially transmits p to R.

1) Direct: As shown in Fig. 3(a), S uses the entire data slot to transmit p directly to D. That is, S applies the strongest available MCS for p that fits into the slot duration of T_s, e. g., Binary Phase-shift Keying (BPSK) with coding rate $1/2$, to minimize transmission errors on the link between S and D.

2) Retransmission: Instead of transmitting p only once, S uses its data slot to transmit two copies of p, one after the other, via the direct link to D (cf. Fig. 3(b)). Consequently, the time for each transmission of p corresponds to roughly half of the data slot, which, in turn, means that a weaker MCS must be chosen to transmit the same data in half the time, e. g., to Quadrature Phase-shift Keying (QPSK) with coding rate $1/2$.

3) Relaying: For this transmission option, S leverages cooperative diversity (cf. Fig. 3(c)). In a previous step, S determined the best available relay R to transmit p to D. Then, S transmits p to R. Upon reception, R immediately sends p to D. Once again, an MCS for p must be selected, such that both transmissions (S to R and R to D) fit into the data slot of S, e. g., QPSK with coding rate 1/2. Note that during the first transmission p might also be overheard by D, additionally increasing the reliability of this transmission option. Moreover, one could apply maximum-ratio combining at D to benefit even further from both transmission paths, which we, however, did not yet implement in our prototype.

C. Selecting a Transmission Option

Based on instantaneous CSI, each station determines the best transmission option, i. e., the one with the highest success probability, for transmitting its data packet p to destination D. This implies that accurate CSI of the entire network is available at each station to take an informed decision on how to transmit p to D. In general, this selection process is divided into three parts: the *Measurement Phase*, the *Reporting Phase*, and the *Transmission Phase*. In the following, we describe how these three phases are implemented in our design.

1) Measurement Phase: All stations and the AP measure the quality of a link implicitly when receiving a packet via this link using the Received Signal Strength Indicator (RSSI), which is provided by the hardware upon reception of a packet. This occurs either when the respective node is the intended receiver of the packet, or simply through overhearing transmissions from others. With a known noise floor, the RSSI can be used to determine the instantaneous Signal to Noise Ratio (SNR), denoted by γ. Since every station and the AP transmit at least once per TDMA superframe, we can assume that at the end of such a superframe, the current quality of each link has been measured at least once. If a station currently does not have any data to send, it would have to transmit a *dummy packet* containing its recent measurements. After the Measurement Phase, the current direct link qualities are thus located at the receiving stations. With asymmetric links, these link qualities have to be reported to the transmitting stations in a timely manner to enable an accurate best relay selection. In the following, we describe how this Reporting Phase is organized.

2) Reporting Phase: Once the link qualities have been measured, they must be conveyed to the transmitting stations for the scheduling decisions. This implies that every station must transmit its measurements periodically to every other station. Since each station transmits at least once in a superframe, it simply piggybacks its measurements in such a regular data transmission, which is overheard by all other stations. The AP, which may also act as relay, piggybacks its measurements in the beacon frame. The size of a data or beacon message thus linearly increases with the number of participating stations in the network, as shown in the beacon and packet structures in Fig. 2.

3) Transmission Phase: Based on the reported link quality measurements, a station calculates for its current receiver which transmission option should be applied. Therefore, it first determines the best available relay in the network. The expected error probability of the relaying process for a given relay R can be expressed as follows

$$\mathbb{P}_{\text{SD}}(\text{R}) = \mathbb{P}_{\text{SD}} \cdot (\mathbb{P}_{\text{SR}} + (1 - \mathbb{P}_{\text{SR}}) \cdot \mathbb{P}_{\text{RD}}) \quad , \tag{1}$$

where $\mathbb{P}_{\text{SD}}$ denotes the expected error probability for a transmission from S to D, and so forth. Consequently, the best available relay can be determined by solving the optimization problem

$$\min_{\text{R} \in \mathcal{R}} \mathbb{P}_{\text{SD}}(\text{R}) \quad , \tag{2}$$

where $\mathcal{R}$ denotes the set of available relays, i. e., overhearing stations including the AP. According to [13], this optimization

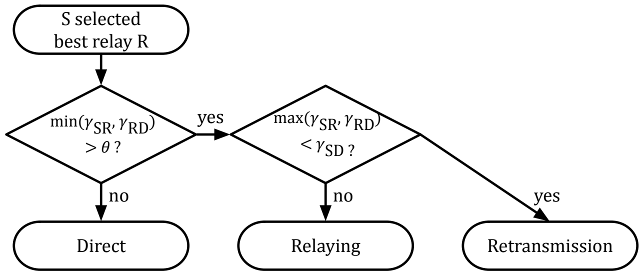


Fig. 4. Local decision tree for each station to select a transmission option based on instantaneous CSI.

problem can be simplified to

$$\min_{\mathrm{R} \in \mathcal{R}} \frac{1}{\gamma_{\mathrm{SR}}} + \frac{1}{\gamma_{\mathrm{RD}}} \quad , \tag{3}$$

where γ denotes the instantaneous SNR in the linear domain. Assuming that $\mathcal{R}$ is relatively small, e. g., $5-20$ stations, this optimization problem can be solved by trying out all possible solutions without significant time overhead. For larger systems, however, a suitable heuristic should be applied.

Once the station has selected a best relay, it locally determines the best transmission option for the given time slot. We opt for a comparably simple heuristic in the form of a decision tree, which can be applied with low computational effort. The implemented decision tree, which is depicted in Fig. 4, is based on observations when one transmission option should be favored over another.

In the first step, the decision tree checks if both links on the relay path are stronger than a *decoding threshold (θ)*. This threshold is the minimal required SNR to successfully decode a transmission with the selected MCS of Relaying/Retransmission. If not, the packet must be transmitted with a stronger MCS and therefore only Direct is possible. Otherwise, it checks whether the direct link is stronger than both links on the relaying path to decide between Relaying and Retransmission. A strong direct link thus indicates that a retransmission should be performed via the same link, while a strong relaying path indicates that the relay should be used for the retransmission. Finally, the station writes its decision into the packet header to inform other (overhearing) stations during the transmission process. This is necessary since another station might be selected as a relay and therefore needs to know how to handle the overheard packet.

4) Complexity: A crucial part of best relay selection and the subsequent transmission option selection is the needed time from the measurement of link qualities until the corresponding scheduling decision. It has been shown that relay selection based on outdated CSI suffers from performance losses [16]. In order to rely on instantaneous CSI at the transmitter, this delay must be kept below the coherence time of the wireless channel. The coherence time depends, among other external influences, mainly on the mobility of sender and receiver, i. e., a higher mobility typically leads to a shorter coherence time.

In the following, we consider the time delay and the message overhead for collecting CSI in a wireless network of N stations including the AP. Furthermore, we assume that all stations are in transmission range and that the wireless links between the stations are asymmetric, i. e., the instantaneous SNR from station i to station j, denoted by γ_{ij}, is not necessarily equal to the instantaneous SNR from j to i, i. e., $\gamma_{ij} \neq \gamma_{ji}$. Therefore, the link qualities must be measured for both transmission directions while for symmetric links, i. e., $\gamma_{ij} = \gamma_{ji}$, it suffices to measure the link quality of one direction.

a) Time Delay: Since we defined that every station (including the AP) transmits at least once during a superframe, we know that at the latest after T_{F} each station i measured $\gamma_{ji}, \forall j \in \{1, \ldots, n\} \backslash \{i\}$. To report the measured link qualities to all other stations such that every station has the link qualities of all links in the network, another T_{F} is required. To summarize, the best relay and the transmission option selection are based on CSI, which is delayed by (at most) $2 \cdot T_{\mathrm{F}}$, where our design envisions a low T_{F} of a millisecond and below.

b) Message Overhead: Regarding the message overhead, link measurements are piggybacked in the header of regular data transmissions. That is, when assuming $1\,\mathrm{byte}$ per link measurement, the MAC header of every data packet and the beacon increases by $N-1\,\mathrm{bytes}$. When assuming symmetric links, the message overhead can be roughly reduced by a factor of 2, which, however, may lead to performance losses as discussed in Sec. IV-C. Furthermore, the message overhead may be reduced by reporting less frequently the measured CSI. The performance loss when trading accurate CSI for low message overhead depends on the considered deployment, which we further investigate in Sec. IV-E.

IV. Performance Evaluation

In this section, we experimentally validate the performance of cooperation in URLLC, based on the proposed protocol. As a main metric to assess the performance, we consider the achieved reliability for a low latency, i. e., the observed Packet Error Rate (PER) for a given deadline. We begin with a detailed description of our real-world testbed (Sec. IV-A). Then, we quantify the achieved reliability of the distinct protocol variants in different evaluation environments (Sec. IV-B). Subsequently, we provide details on an efficient parametrization of the target system by first addressing the question of whether symmetric link qualities can be assumed in a practical deployment (Sec. IV-C). We continue by identifying the cases in which a direct transmission should be preferred over a relay transmission (Sec. IV-D). Finally, we quantify the impact of outdated CSI on the reliability (Sec. IV-E).

A. Setup

For the evaluation, we implement our experimental protocol on the Wireless Open Access Research Platform (WARP) v3 [17]. WARP boards are SDRs, consisting mainly of a Field Programmable Gate Array (FPGA), two radio interfaces, and several I/O ports. These boards are used for research in wireless communications, to prototype new protocols and test their performance in real-world testbeds. Our

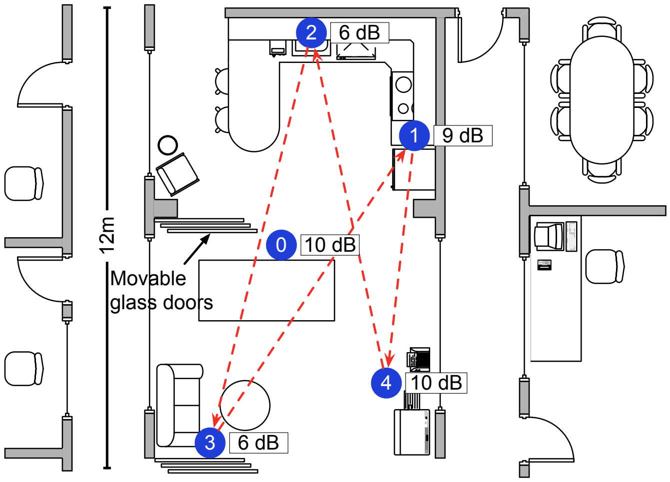


Fig. 5. Evaluation setup. The circled numbers represent the positions of the five WARP boards in the social room, where ID 0 is the AP and the remaining IDs are stations. Next to each board, we indicate the selected antenna attenuator value. The dashed arrows represent the data flow.

implementation is based on the 802.11 Reference Design [18], which realizes IEEE 802.11a/n [19] in a custom FPGA core and on two MicroBlaze Central Processing Units (CPUs), CPU High and CPU Low, where most parts of the Physical Layer (PHY) are realized in the FPGA core and the MAC layer is implemented mainly on the CPUs. Because the focus of this work is not to optimize the PHY and our approach is compatible with any packet-based PHY, we leave the PHY of the 802.11 Reference Design unchanged. Instead, we modify the code running in CPUs High and Low according to the design described in Sec. III.

As evaluation setup, we choose the social room of our research institute, where we place five WARP boards at different locations in the room. The topology of the boards including their IDs is shown in Fig. 5. The board with ID 0 assumes the role of the AP, while boards 1-4 are the stations. Note that the antennas are mounted at the ceiling and all antennas are in line-of-sight. However, the room has a movable glass door which might block the line-of-sight between 1 and 2 to 0, 3, and 4. The stations get assigned a fixed destination for their data packets, shown by the dashed arrows in Fig. 5. To artificially decrease the link qualities, we additionally connect attenuators to the antenna port of each board. The respective attenuator values are also shown in Fig. 5. Regarding the traffic load, we assume that each station generates one data packet every millisecond with a deadline of $1\,\mathrm{ms}$, i. e., packets that arrive after $1\,\mathrm{ms}$ are considered *lost*. Since all stations transmit data packets of the same size, i. e., $64\,\mathrm{bytes}$ of payload, we set for all stations the data slot duration to the same length, i. e., $T_\mathrm{s} = 210\,\mu\mathrm{s}$. Based on this topology, we specify three distinct evaluation environments to asses the performance of our protocol, where the evaluation parameters for the different environments are listed in Table I.

Static Environment. We assume that the stations are in a static environment, i. e., we aim at minimizing external influences on the transmission quality. Therefore, measurements are only performed during the night and on weekends when, most of the time, no one is in the social room. This environment thus serves the purpose to quantify the performance difference of the distinct protocol variants in a controlled setup. To still observe packet errors in a reasonable amount of time, we reduce the transmission power of AP and stations to the lowest available level, i. e., $-9\,\mathrm{dBm}$. The average SNRs of the different links in this environment are listed in Table II, where stars label the considered direct links.

Dynamic Environment. To observe the impact of variations in the wireless channel on our protocol, we perform a continuous performance evaluation of our protocol variants during one week. The measurements are thus performed during busy hours of the social room, e. g., on week days between $12\,\mathrm{pm}$ and $2\,\mathrm{pm}$, as well as on unoccupied hours of the social room, e. g., on Sunday mornings. For this environment, we increase the transmission power to $10\,\mathrm{dBm}$ for all boards.

Mobile Environment. To capture the effects of mobility on the transmission reliability, we define a third environment where the stations vary their transmit power. This allows us to simulate mobility following a predefined pattern, without having to physically move the stations. Therefore, the transmit power of each station (except the AP) oscillates between $-9\,\mathrm{dBm}$ and $0\,\mathrm{dBm}$, where the power is either increased or decreased by $1\,\mathrm{dBm}$ every superframe. We use this scenario mainly to investigate the effects of outdated CSI on the transmission option selection.

TABLE I
EVALUATION PARAMETERS

Parameter	Value
#APs	1
#stations	4
Duration of superframe (T_F)	$1\,\mathrm{ms}$
Duration of transmission slot (T_s)	$210\,\mu\mathrm{s}$
Size of payload (D_pl)	$64\,\mathrm{bytes}$
Size of (MAC) header (D_h)	$11\,\mathrm{bytes}$
Transmission bandwidth (B)	$20\,\mathrm{MHz}$
Center frequency (f_c)	$5600\,\mathrm{MHz}$
Transmission power (P_Tx)	$-9\,\mathrm{dBm}$ (static) $10\,\mathrm{dBm}$ (dynamic) $-9\,\mathrm{dBm}$ - $0\,\mathrm{dBm}$ (mobile)
Noise floor (P_noise)	$-94\,\mathrm{dBm}$
MCS for Direct	BPSK $1/2$
MCS for Retransmission / Relaying	QPSK $1/2$
Decoding threshold (θ)	$4\,\mathrm{dB}$

TABLE II
AVERAGE LINK SNRS [dB] IN THE STATIC ENVIRONMENT

From \ To	AP	STA 1	STA 2	STA 3	STA 4
AP	–	3.4	1.4	11.9	5.4
STA 1	4.6	–	4.4	2.3	2.8*
STA 2	1.4	4	–	2.1*	3.2
STA 3	11.9	1.7*	1.7	–	22
STA 4	5	1.7	1.9*	21.2	–

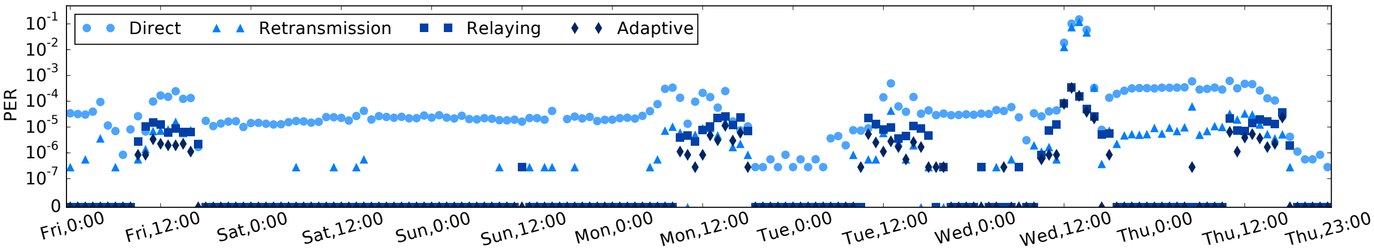


Fig. 6. Hourly PER of the different transmission options in the Dynamic Environment for one week. The first measurement starts on Friday 0:00, while the last one starts on Thursday 23:00.

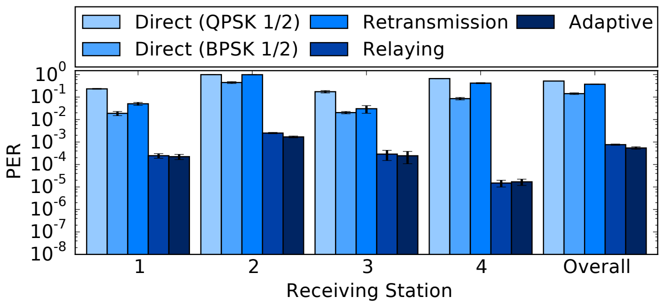


Fig. 7. Avg. PERs of different transmission options in the Static Environment. *Adaptive* represents the dynamic selection of the currently best option.

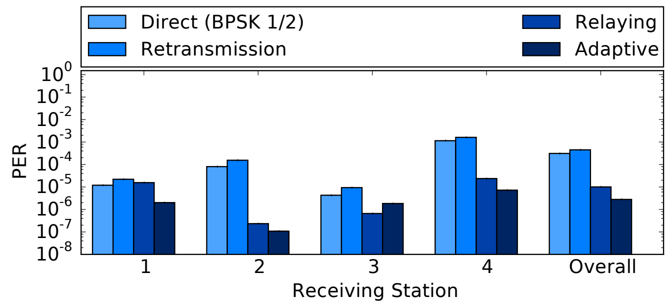


Fig. 8. PER of the different transmission options in the Dynamic Environment for a continuous evaluation over 168 hours.

B. Achieved Reliability

We are interested in the performance of the different transmission options, especially when we let each station dynamically select the best option for each packet transmission (denoted by *Adaptive*). Therefore, we measure the PERs in two distinct environments, namely *Static* and *Dynamic*.

In the Static Environment, each participating station transmits $2 \cdot 10^6$ packets per protocol variant, where each measurement is repeated 11 times. The results are depicted in Fig. 7. Note that Direct with QPSK 1/2 only serves as a reference since the more robust BPSK 1/2 is available for the given time slot when using Direct. We first observe that Retransmission, in general, does not improve the reliability compared to Direct. On the contrary, the reliability even decreases by half an order of magnitude, because a weaker MCS is applied due to timing constraints. When comparing Relaying to Direct, we see a major improvement in the reliability by at least two orders of magnitude due to the additional cooperative diversity. The improvement factor thereby depends on the connectivity to the potentials relays. Finally, we find that Adaptive compared to Relaying only marginally decreases the PER indicating that the vast majority of transmissions in Adaptive is performed using Relaying. Indeed, our measurements show that, in Adaptive, on average only $0.1\,\%$ of the transmissions are performed using Direct or Retransmission, while for the remaining transmissions Relaying is selected.

In the Dynamic Environment, we start our measurements on a Friday at midnight and let them run for exactly a week. We change the protocol variant every superframe in a round-robin fashion such that every variant is affected by variations in the wireless channel in a similar manner. Fig. 6 shows the PER of each variant for every hour, i. e., the first measurement starts on Friday 0:00, while the last one starts on Thursday 23:00.

We see that every protocol variant is affected by changes in the environment, i. e., the PERs at night and on the weekend are smaller than on typical working hours when the social room is busy. In almost every case, nevertheless, the PER of Adaptive is either below or not worse than the other protocol variants. Note that on Wednesday afternoon, there was a social gathering where the room was very crowded and the glass doors were moved from the left side of the room to the right side. Therefore, the PERs of Direct and Retransmission increased by almost three orders of magnitude compared to Adaptive. These results thus show that although none of the protocol variants is completely immune against harsh changes in the environment, Adaptive achieves, in general, a lower PER than the other variants.

To further investigate this observation, we repeat the measurements after removing the antenna attenuators of all boards. We are thus interested in the achievable PER of this scenario, when we are not artificially deteriorating the link qualities. Same as before, the measurements were performed in the Dynamic Environment continuously over 168 hours, sending a total of $1.755 \cdot 10^8$ packets per link and protocol variant. The resulting PERs are shown in Fig. 8. This plot reveals that, although operating in a dynamic environment, we are able to achieve PERs between 10^{-5} and 10^{-7}. Nevertheless, these results also indicate that further techniques increasing the reliability, e. g., on the PHY, are required to achieve URLLC.

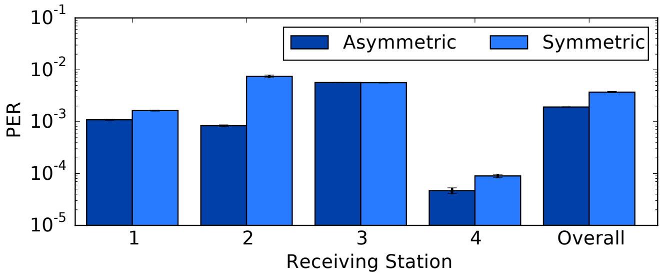


Fig. 9. PER of the relaying process when collecting asymmetric CSI compared to symmetric CSI.

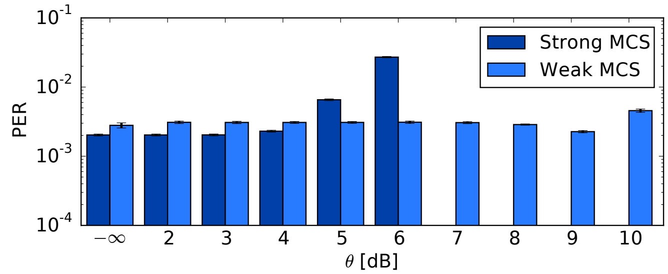


Fig. 10. PER when changing the decoding threshold θ. In Strong MCS, Direct uses BPSK $1/2$, while Relaying uses QPSK $1/2$. In Weak MCS, Direct uses QPSK $1/2$, while Relaying uses 16-QAM $1/2$.

C. Symmetric versus Asymmetric Links

When collecting CSI for the best relay selection, an important question is whether it can be assumed that the link qualities are symmetric, i. e., the instantaneous link quality from station i to station j is the same as the one from j to i. Real-world measurements, e. g., [20], show that, in general, wireless links are asymmetric. In practice, however, collecting CSI for assumed asymmetric links increases the overhead significantly compared to symmetric links, as already outlined in Sec. III-C4. Therefore, we are interested if, performance-wise, it is worth to collect asymmetric link qualities compared to symmetric ones for the best relay selection process.

Fig. 9 shows the PER in the relaying process for each station assuming asymmetric links compared to symmetric links. Our measurements are performed in the Static Environment using the configuration shown in Fig. 5. For both asymmetric and symmetric, we let each station transmit a total of 10^6 packets using Relaying, where the best relay is determined according to the respective link quality assumption. The measurements are repeated 16 times.

The collection of asymmetric link qualities for the relaying process outperforms, as expected, the relay process based on symmetric link qualities, since more accurate CSI leads to less errors when selecting the "best" relay for a connection. The performance gap between asymmetric and symmetric depends on the respective link, but overall contributes to a lower PER. This leads to the conclusion that both transmission directions of a link should be considered separately, especially when the system is operating at a high reliability, where a sub-optimal relay selection has a higher impact on the PER. Furthermore, the link qualities can only be estimated symmetrically at low performance losses, when the stations also use the same transmission power and hardware characteristics.

D. Stronger Coding versus Relaying

According to our decision tree, cf. Sec. III-C3, once the best relay for a given receiver is selected, the station decides whether to transmit the packet via this relay or using the direct link between sender and receiver. Remember that for both transmission options the same deadline applies, i. e., for Direct only one transmission must be performed (from sender to receiver), while for Relaying two transmissions must be performed (from sender to relay and from relay to destination). Therefore, a weaker MCS is applied for Relaying to fit both transmissions into one slot. To determine how the decoding threshold θ of the local decision process should be set, we vary this parameter in two distinct scenarios. The first scenario, denoted by *Strong MCS*, corresponds to the evaluation parameters shown in Table I where Direct uses BPSK $1/2$ and Relaying uses QPSK $1/2$. In the second scenario, denoted by *Weak MCS*, we increase the packet payload to $128\,\mathrm{bytes}$ while keeping the deadline unchanged, i. e., $T_{\mathrm{F}} = 1\,\mathrm{ms}$. Therefore, we must adapt the MCS of Direct to QPSK $1/2$ and Relaying to 16-Quadrature Amplitude Modulation (QAM) $1/2$. Furthermore, we set $P_{\mathrm{Tx}} = -3\,\mathrm{dBm}$ to have a similar PER on the direct links as in Strong MCS. All measurements are conducted in the static environment.

The PERs for different θs for the two scenarios are shown in Fig. 10. In each measurement, we selected a fixed value of θ and transmitted a total of $2 \cdot 10^6$ packets in the network. The measurements were repeated 7 times. Note that in the plot $-\infty$ represents an arbitrarily low value of θ, thus leading to a deactivation of the direct transmission option. Furthermore, we varied θ in the Strong MCS scenario up to $6\,\mathrm{dB}$, while for Weak MCS we extend the measurements up to $\theta = 10\,\mathrm{dB}$ to better capture the trade-off between Direct and Relaying.

For Strong MCS, already at a very small θ, i. e., at $5\,\mathrm{dB}$, the PER increases, confirming that in most cases relaying is preferred over a direct transmission with a stronger MCS, since relaying already offers a relatively robust MCS. For Weak MCS, however, we see a local optimum at $\theta = 9\,\mathrm{dB}$ showing that when the connection to the best relay is poor, the direct link should be used for a more robust transmission. In the vast majority of cases, however, the transmission path over the best relay offers a higher reliability than the stronger coded direct path. Therefore, θ should be set to a low threshold depending on the respective MCS for Direct and Relaying.

E. Outdated Channel State Information

Finally, we are interested in how the freshness of CSI influences the PER of Adaptive. Remember that CSI is used to determine the best relay and the best transmission option. Since the reporting of CSI is costly in terms of transmission resources, one could increase the time interval with which local CSI measurements are reported at the cost of a less accurate

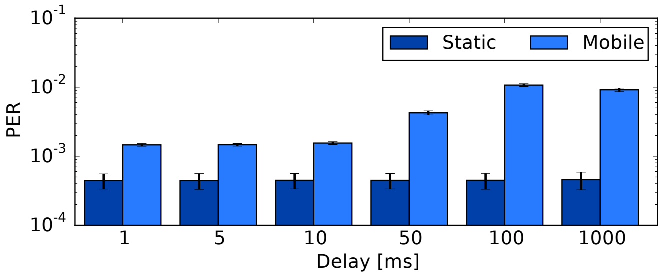


Fig. 11. Impact of outdated CSI on the PER for the Static Environment and the Mobile Environment.

relay and transmission option selection. To assess the impact of outdated CSI on the PER, we conduct measurements in two distinct environments: Static and Mobile, where we increase the delay for reporting CSI between $1\,\mathrm{ms}$ and $1000\,\mathrm{ms}$. In each measurement, a total of $8 \cdot 10^6$ packets are transmitted and the measurements are repeated 11 times.

The results are depicted in Fig. 11. For the Static Environment, an increasing delay does not deteriorate the PER, since for a given connection the best relay is seldomly changed. Indeed, we measured that, on average, relays are only switched approximately every $3\,\mathrm{s}$. In the Mobile Environment, in turn, we observe an increase in the PER of one order of magnitude with an increasing delay. This also matches our measurements that with instantaneous CSI, on average, relays are switched approximately every $4.3\,\mathrm{ms}$. The interval for reporting CSI can thus be adapted depending on the mobility of the target deployment, without significant performance losses.

V. Conclusion

In this paper, we focus on the practical feasibility of Ultra-Reliable Low-Latency Communication (URLLC) using best relay selection. Therefore, based on existing theoretical findings, we implement a data link protocol for URLLC that adaptively chooses for each connection the best transmission option given a fixed-length time slot. We efficiently integrate the selection of the best transmission option, based on instantaneous channel state information, into the protocol without affecting the low communication latency. To the best of our knowledge, this is the first URLLC implementation on real hardware. We conduct our evaluation in different scenarios showing that, in the vast majority of cases, relaying is the transmission option with the lowest PER compared to a direct transmission with a stronger modulation and coding scheme or a retransmission of the packet by the sender. There are, however, some edge cases where the latter two outperform relaying. With the proposed Adaptive scheme, we are able to reduce the PER even under harsh channel conditions. In other words, our practical results confirm that systems solely relying on time diversity require, by orders of magnitude, more transmission resources than systems using cooperative diversity to achieve the same reliability. For future work, we propose to refine the transmission options, e. g., with network coding, and to consider use cases from industrial automation.

Acknowledgments

We would like to thank the DFG for the support within the Cluster of Excellence “Integrative Production Technology for High-Wage Countries”.

References

[1] A. Frotzscher, U. Wetzker, M. Bauer, M. Rentschler, M. Beyer, S. Elspass, and H. Klessig, “Requirements and current solutions of wireless communication in industrial automation,” in *IEEE Int’l Conf. on Comm. Workshops (ICC’14)*, Jun. 2014, pp. 67–72.

[2] C. Hoymann, D. Astely, M. Stattin, G. Wikstrom, J. F. Cheng, A. Hoglund, M. Frenne, R. Blasco, J. Huschke, and F. Gunnarsson, “LTE Release 14 Outlook,” *IEEE Communications Magazine*, vol. 54, no. 6, pp. 44–49, Jun. 2016.

[3] S. Petersen and S. Carlsen, “WirelessHART Versus ISA100.11a: The Format War Hits the Factory Floor,” *IEEE Industrial Electronics Magazine*, vol. 5, no. 4, pp. 23–34, Dec. 2011.

[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior,” *IEEE Trans. on Inform. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[5] S. S. Ikki and M. H. Ahmed, “Performance Analysis of Adaptive Decode-and-Forward Cooperative Diversity Networks with Best-Relay Selection,” *IEEE Trans. on Comm.*, vol. 58, no. 1, pp. 68–72, Jan. 2010.

[6] M. Serror, C. Dombrowski, K. Wehrle, and J. Gross, “Channel Coding Versus Cooperative ARQ: Reducing Outage Probability in Ultra-Low Latency Wireless Communications,” in *IEEE Global Comm. Conf. (GLOBECOM) Workshops (ULTRA^2)*, San Diego, USA, Dec. 2015.

[7] M. Serror, Y. Hu, C. Dombrowski, K. Wehrle, and J. Gross, “Performance Analysis of Cooperative ARQ Systems for Wireless Industrial Networks,” in *IEEE 17th Int’l Symp. on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Jun. 2016.

[8] A. S. Ibrahim, A. K. Sadek, W. Su, and K. J. R. Liu, “Cooperative Communications with Relay-Selection: When to Cooperate and Whom to Cooperate With?” *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2814–2827, Jul. 2008.

[9] V. N. Swamy, S. Suri, P. Rigge, M. Weiner, G. Ranade, A. Sahai, and B. Nikolić, “Cooperative Communication for High-reliability Low-latency Wireless Control,” in *IEEE Int’l Conf. on Communications (ICC)*, Jun. 2015, pp. 4380–4386.

[10] V. N. Swamy, P. Rigge, G. Ranade, A. Sahai, and B. Nikolić, “Network Coding for High-reliability Low-latency Wireless Control,” in *IEEE Wireless Communications and Networking Conf.*, Apr. 2016.

[11] S. Girs, E. Uhlemann, and M. Björkman, “Increased Reliability or Reduced Delay in Wireless Industrial Networks Using Relaying and Luby Codes,” in *IEEE 18th Conf. on Emerging Technologies Factory Automation (ETFA)*, Sep. 2013.

[12] R. Heynicke, D. Krush, G. Scholl, B. Kaercher, J. Ritter, P. Gaggero, and M. Rentschler, “IO-Link Wireless Enhanced Sensors and Actuators for Industry 4.0 Networks,” in *AMA Conferences*, Jun. 2017, pp. 134–138.

[13] C. Dombrowski and J. Gross, “EchoRing: A Low-Latency, Reliable Token-Passing MAC Protocol for Wireless Industrial Networks,” in *Proc. of 21st European Wireless Conference (EW15)*, May 2015.

[14] J. Silvo, L. M. Eriksson, M. Björkbom, and S. Nethi, “Ultra-reliable and Real-time Communication in Local Wireless Applications,” in *IEEE Conf. of the Ind. Electr. Society (IECON’13)*, Nov. 2013, pp. 5611–5616.

[15] N. Marchenko, T. Andre, G. Brandner, W. Masood, and C. Bettstetter, “An Experimental Study of Selective Cooperative Relaying in Industrial Wireless Sensor Networks,” *IEEE Trans. on Ind. Inform.*, vol. 10, no. 3, pp. 1806–1816, Aug. 2014.

[16] M. Seyfi, S. Muhaidat, J. Liang, and M. Dianati, “Effect of Feedback Delay on the Performance of Cooperative Networks with Relay Selection,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4161–4171, Dec. 2011.

[17] Rice University and Mango Communications, “The WARP Project,” [Online] https://www.warpproject.org/trac.

[18] Mango Communications, “802.11 Reference Design,” [Online] http://mangocomm.com/802.11.

[19] IEEE, “Standard for Information Technology–Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks,” *IEEE Std 802.11-2012*, Mar. 2012.

[20] L. Sang, A. Arora, and H. Zhang, “On Link Asymmetry and One-way Estimation in Wireless Sensor Networks,” *ACM Transactions on Sensor Networks*, vol. 6, no. 2, pp. 12:1–12:25, Mar. 2010.