Measuring and Monitoring Reliability of Wireless Networks

Matthias Herlich, Christian Maier
<firstname.lastname>@salzburgresearch.at

Abstract—**In the future, connected vehicles and automated factories will increasingly use wireless communication. Because the cost of communication failures in these applications is high, the wireless networks must be reliable and continuously monitored.**

To address this need, we propose a general approach to determine the reliability of wireless networks and describe implementations of it. Our approach focuses on checking necessary assumptions, instead of assuming they are fulfilled. We successfully tested our methods in an operating industrial environment. Our tests show that both the initial measurement of the reliability of the wireless network and the continued monitoring are a helpful tool for factory operators.

In the future, methods such as those described in this paper should be used to ensure the reliable operation of wireless networks in critical scenarios.

Index Terms—**Wireless communication, Telecommunication network reliability.**

I. INTRODUCTION

Many use cases of wireless communication benefit from high reliability. In industrial and automotive use cases additionally the impact of insufficient reliability is high [1]. Such use cases often require the wireless communication to be $99.999\,\%$ reliable (≈ 5 minutes unavailability per year). Here reliability is defined as the probability of packet delivery between two end points with a given maximum transmission duration [2]. Because of the high impact of insufficient reliability, it is necessary to determine the reliability with a high degree of certainty. Additionally, the reliability needs to be high continuously and not only initially. Hence, it is necessary not only to initially measure the reliability, but also to continuously monitor it.

The motivation for our approach is threefold: First, it should be statistically well analyzed. This includes checking if the assumptions are fulfilled and providing false-positive and false-negative rates, that is, understanding how often and which mistakes our approach makes. Second, the approach should be light weight. That means the measurements themselves should not disturb the environment, not generate additional data traffic and not require additional hardware. Third, it should determine if the predictions are valid over time, because in general it is unclear when measurements of reliability lose their predictive value as a result of changesk in the environment that influence the propagation of electromagnetic waves. To our knowledge the combination of these features is a unique novelty.

To better understand what influences the reliability we created the causality graph shown in Figure 1. It shows effects and parameters as nodes in the graph and connects them with a directed edge if the source influences the destination.

Measuring only the relevant parameters at the physical (PHY) layer does not provide enough information to determine the reliability. The PHY layer provides an important contribution to the end-to-end reliability of a wireless communication system, but ignoring all other layers leads to false certainty. As an example, we found a printer in an operating production environment that irregularly transmitted data packets so often that communication on that channel was impossible. The only reason that operation did not halt was that the printer used a different channel than most other devices. This example illustrates that a strong signal is only a necessary, but not a sufficient criterion for reliability and that monitoring the reliability is necessary.

Following the related work (Section II), we give an overview of measuring and monitoring (III). This

The published version is available as 10.1109/MCOM.001.2000250.

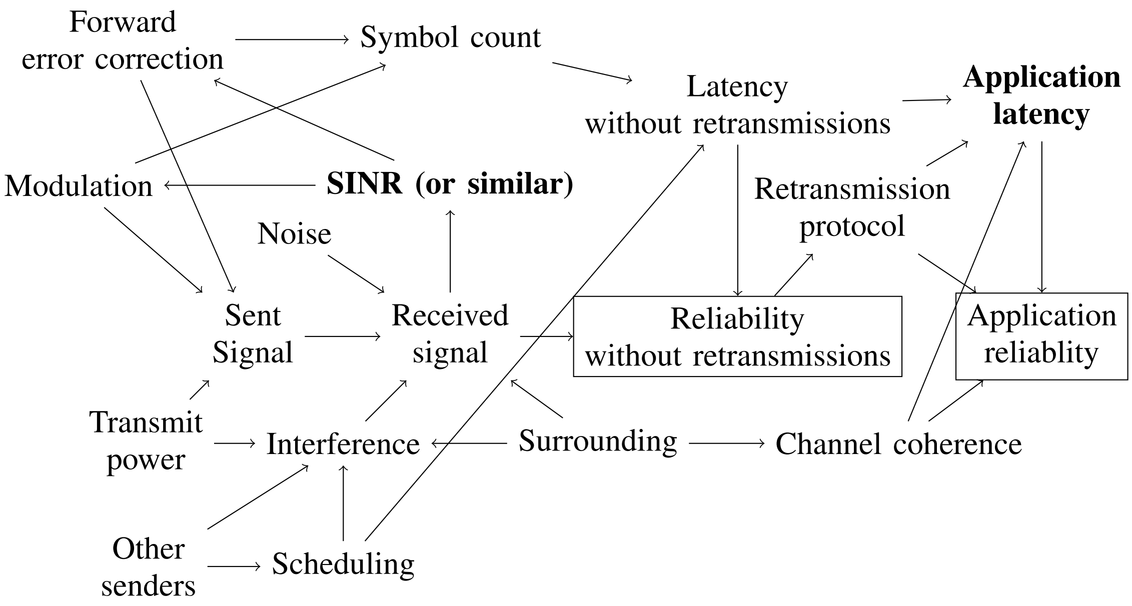

Figure 1. The causality graph of influences on reliability shows "components" as nodes and connects them with an arrow if the source of the arrow influences the destination. This is not a comprehensive graph, but a guide to the effects we consider. Effects we propose to monitor closely are bold (latency and a signal strength indicator).

includes a description of simple approaches and their limits. A description of our methods for measuring (IV) and monitoring (V) follow. We lastly describe a proof of concept to illustrate the method in practice (VI). This paper builds on and reuses parts from our earlier papers [1], [2], [3].

II. RELATED WORK

Work related to this paper comes from wireless channel models, wireless measurements and a statistical background. Avicienis et al. provide a comprehensive overview of definitions of dependability, reliability and availability [4] and Ji et al. discuss the physical layer aspects of ultra reliable communication [5].

Many channel models describe averages and quantiles in the order of up to $99\,\%$ reliability of a wireless channel. However, channel models cannot be used to determine if a network provides $99.999\,\%$ reliability, because the behaviour of the models at the edge of their covered range is unclear. Moreover, the models consider only the physical channel. In particular, effects of higher layers such as medium access control (MAC), schedulers and retransmission schemes are not taken into account.

Direct measurement of wireless reliability is an area of active research. Because currently few networks provide a higher reliability than $99\,\%$, reliability measurements are not designed for highly reliable networks (e.g. [6], [7]). Additionally, the methods usually do not analyze false positives and false negatives in detail and do not test the assumptions the measurement methods make. For example, Pocovi et al. [8] describe a measurement framework to quantify end-to-end latency and reliability of wired and wireless communication systems. Similar to our work, their setup consists of low-cost hardware. While their general setup can also be used for our measurements, two additions of our deeper statistical analysis are: (1) They obtain the reliability from the empirical cumulative distribution function. That is, they do not obtain significance levels, which are a key novelty of our approach. (2) Additionally, our work includes an ongoing check of the validity for future transmissions.

Bai and Krishnan [6] analyze reliability of wireless networks in automotive applications. Most of their measurements indicate a reliability (between $80\,\%$ and $99\,\%$) that is lower than required by critical applications ($\geq 99\,\%$). Woo and Kim [7] describe that signal strength indicators (RSSI, SINR) do not provide good estimates of reliability. Salyers et al. [9] describe that the specific hardware is an important factor for the reliability of wireless communication.

In contrast to *measuring*, the work on *monitoring* the reliability of wireless networks consists mostly of patents instead of scientific publications. These patents describe devices and architectures, but do not provide enough technical detail to construct and validate the monitoring systems.

Extensive literature exists on calculation of binomial proportion confidence intervals and their comparison [10], [11]. In this paper we apply their results to the problem of measuring reliability.

In summary, a statistical analysis of the methods to measure and monitor end-to-end reliability in wireless networks is missing. We will provide an overview in this paper.

III. Overview

Our *measurement* method consists of a black box test, which needs two assumptions, and white box tests of these assumptions. Our *monitoring* method observes characteristic parameters of the wireless communication, while the initial measurement of the reliability is done. Later the monitoring passively observes the same characteristic parameters and notifies the operator when the parameters change, because a change in the channel parameters could signal a change in reliability. This way the monitoring estimates how long the initial measurements of the reliability are valid.

A. Measuring and monitoring

A reliability engineering guide states, that "the field environment is the ultimate test for product performance; therefore the reliability should be evaluated based on field failures whenever possible" [12]. To determine the rate of failures of wireless communication in the field, a monitoring system is needed in the field that is operating during production, because the effects of an actively producing environment might influence the reliability of the wireless communication. Thus, we propose an approach that combines initial measurement of reliability with later monitoring to ensure that it correctly represents the reliability of devices in an operating environment.

According to the reliability engineering guide, a "Failure Reporting Analysis and Corrective Action System [...] must provide for:

(1) Reporting of all production test and inspection failures with sufficient detail to enable investigation and corrective action to be taken.
(2) Reporting the results of investigation and action.
(3) Analysis of failure patterns and trends, and reporting on these.
(4) Continuous improvement (kaizen) by removal of causes of failures." [12]

Achieving these criteria for lost packets in a wireless network is more complicated than for a produced item, because it is impossible to preserve the environment at the moment of a packet loss. Nevertheless, we think it is necessary to build monitoring systems that approximate these criteria. The contribution of this paper is a description of such a combined measurement and monitoring system. The general system architecture is independent of the method used to measure and monitor, but we describe one method for measurement and one for monitoring.

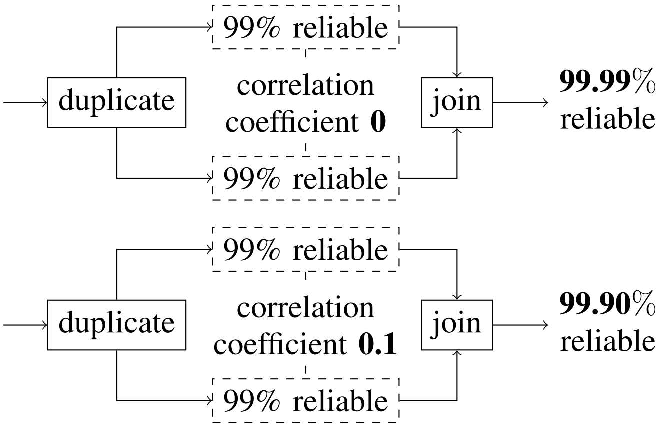

Figure 2. The effect of correlation on resulting reliability can easily be underestimated. In this case a correlation coefficient of 0.1 reduces the resulting reliability by an order of magnitude.

B. Other approaches and limits

To explain the complexity of measuring and monitoring, we next describe approaches that are unable to determine the reliability.

1) Ignoring low correlation: Treating low correlation as zero correlation leads to large errors in the estimated reliability. The following example illustrates this behaviour with a simple calculation.

Assume data is transmitted over two 99%-reliable channels (Figure 2). If the channels are independent, this results in the transmission arriving with probability 99.99%. If the state of the transmission has a correlation coefficient of 0.1, the resulting probability of successful transmission is $\approx 99.90\%$. Thus, ignoring a low correlation coefficient can lead to concluding the channel provides four nines reliability when it actually provides only three nines.

2) Not considered effects: Quantities like S(I)NR, RSSI and latency do not fully determine the result of a packet transmission (success or failure). Hence, they provide only an estimate for the reliability. That is, they do not represent all possible effects for reduced reliability, but only part of them (see Figure 1).

3) Reliability definitions: There are various (formal and informal) definitions in the literature for the reliability of a wireless channel [13]. They can be grouped as follows: the reliability is…

(1) …the number of received packets divided by the number of sent packets.
(2) …a lower bound for the number of received packets divided by the number of sent packets.
(3) …the probability with which a sent packet arrives at its destination.

The first is a purely empirical one. After transmitting n packets, it is given as the percentage of successfully transmitted packets. Hence, this quantity can be easily determined in a real environment. However, the drawback is that its relevance for future transmissions is unclear.

To get precise definitions for the second and third definitions, a mathematical model for the transmission process is required. In such a model, one may derive formulas or algorithms for confidence intervals and hypothesis tests for the reliability. Then an empirical experiment determines these confidence intervals (resp. the result of the hypothesis test) for the reliability up to a specified significance level. The advantage of the these approaches (they allow statements about future transmissions), need to be carefully considered, because they do not guarantee (or provide a significance level) that the mathematical model actually describes reality sufficiently accurate. In other words: they provide estimates given the correctness of the model, but do not test the correctness of the model.

In summary, the first definition is easy to evaluate, but hard to apply to future transmissions. While the second and third definition are hard to evaluate, but easy to apply to future transmissions. We are not aware of any definition of reliability that is both easy to evaluate and easy to apply to future transmissions.

In contrast to most related work, our method is based on the third type of definition. While the measurements are strictly only valid for the time of the measurement, we combine it with the monitoring which provides confidence that the measurements are still valid later.

IV. MEASURING RELIABILITY

In this section we investigate reliability measurements of a wireless channel for two different transmission schemes: Periodic packet transmissions and Poisson distributed packet transmissions. Periodic transmissions are fully deterministic, while Poisson distributed transmissions provide no information about other transmissions. In this aspect both transmission schemes are on different sides in the predictability spectrum. We provide an overview for periodic transmissions and refer to earlier work for the details. The second is an adaption of the first. We focus especially on the underlying assumptions and propose a method to test if the reliability is above a certain target reliability.

A. Periodic transmissions

In our earlier work [2] we investigated a test method that determines if the reliability r of the wireless transmission channel is below or above a certain target reliability t for periodic packet transmissions. To determine false positive and false negative rates of this test, it is necessary to assume a mathematical model for transmission failures. Whether this model actually holds is checked as part of the test method. This is done by investigating the recorded signal strength indicators of each transmission. We then determined for a given target reliability t the number of required packet transmissions n, such that the validity of the inequality $r \geq t$ can be checked with sufficient accuracy.

1) Simple test: The simple test is positive if all packet transmissions are successful. We define

(*Property* $\mathcal{P}$) If the real reliability r is less than the target reliability t, the test will provide a positive result with a probability less than $1 - t$ [2].

To ensure the bound for the false positive rate holds, when the real reliability is lower than the target reliability, the probability of at least one failure has to be higher than the significance level. With the assumption that repeated packet transmission is a Bernoulli process, this yields an equation for the minimal number of necessary measurements [2]. Table I lists the number of measurements needed

Table I
The number of measurements for target reliability t and significance level α with the simple test; measurements for property $\mathcal{P}$ are bold [2].

		Target reliability t				
		0.9	0.99	0.999	0.9999	0.999 99
α	0.1	**22**	230	2 302	23 025	230 258
	0.01	44	**459**	4 603	46 050	460 515
	0.001	66	688	**6 905**	69 075	690 773
	0.000 1	88	917	9 206	**92 099**	921 030
	0.000 01	110	1 146	11 508	115 124	**1 151 287**

Table II
The number of measurements using the extended test and necessary time based on Wi-Fi Beacon frames [2].

	Target reliability t				
	0.9	0.99	0.999	0.999 9	0.999 99
Measurements	38	1 157	19 620	293 056	3 751 160
Allowed errors	1	4	7	11	14
Necessary time	4 s	2 min	34 min	8 h	4 d

to measure the target reliability t given the significance level α [2]. This describes the mathematical complexity of the measurement.

When a system successfully transmitted as many packets as stated in Table I, it provides the stated certainty that communication was reliable in the measured time interval. When a single transmission fails, the simple test fails (to have a low false positive rate). Therefore, the false negative rate is high, when the real reliability is only slightly above the target reliability. For a trustworthy result it is necessary to also test the assumptions (individual measurements are independent and identically distributed) [2].

2) Extended test: To improve over the simple test, the extended test also limits the false negative rate:

(*Property* $\mathcal{Q}$) If the real reliability of the system is higher than or equal to s (for a fixed $s > t$), the test will provide a negative result with a probability less than $1 - t$ [2].

To reduce the false negative rate, the extended test fails only when more than a constant number of failures a occur. To keep the false positive rate low, the number of measurements n has to be increased accordingly. However, instead of selecting the number of allowed failures a it is more intuitive to limit the false negative rate. Thus, the extended test uses the lowest number of measurements n and allowed transmission failures a that fulfill properties $\mathcal{P}$ and $\mathcal{Q}$. The Properties $\mathcal{P}$ and $\mathcal{Q}$ can be directly described as inequalities based on the cumulative distribution function of the binomial distribution [2]. We propose to arbitrarily set $s = 0.1t + 0.9$. Table II shows configurations of our test for often used reliability values. While the extended test increases the complexity, it reduces the false-negative rate. For example, to test for a reliability of 0.999 it is necessary to transmit $19\,620$ packets and if 7 or fewer of those are lost, the test is passed. Based on WiFi Beacon Frames with $102.4\,\mathrm{ms}$ interpacket time this test would take 34 minutes.

In summary the extended test has a high probability to give a correct result, when the real reliability is below t or above s.

B. *Poisson distributed transmissions*

To determine the reliability of a channel with a Poisson distributed transmission scheme, we model the channel by two states. In the good state, packets are transmitted successfully, whereas in the bad state transmissions fail. Our aim is to determine the reliability of the channel in a specified time interval. If transmitting packets does not change the state of the channel, then the probability of a successful arrival of a packet transmitted at a random moment in the considered interval (the reliability of the channel) is the same for each packet. This assumption is approximately fulfilled if the coherence time of the channel is smaller than the inter-packet time. It follows that repeated packet transmission with exponential inter-packet time constitute a Bernoulli process. Hence to check if the reliability r of the channel is above a given target reliability t, one can again use the method we have outlined in Section IV-A. Hence, the complexity (measured in transmissions) is the same.

V. MONITORING RELIABILITY

A. *General approach*

As demonstrated in the previous section, the effort of the proposed reliability measurement method is large. Hence, it is impossible to repeat such a measurement frequently. This should only be performed if there are indications that the reliability of the channel has changed (compared to an initial measurement). Possible quantities whose change might constitute such an indication are: Received signal strength indicator, channel utilization, traffic patterns, and the physical environment. The relevance of the parameters will depend on the scenario.

As an extension of the measurement method in Section IV, we propose the following approach to monitor the reliability during operation: First, the operator tests if the reliability r of the system is above a given target reliability t (using the method explained in Section IV-B). Second, the parameters that are considered relevant are monitored and compared to their reference values from the initial measurement. If one of these quantities changes significantly, it cannot be guaranteed that the required reliability is still provided. Hence, it is necessary to test the reliability again.

Next, we consider the upper and lower bounds of a binomial proportion confidence interval as reliability indicators. The considered process is the success or failure of the transmission of beacon frames from a fixed access point. There is no evidence that these are independent. Hence, the calculated interval may not be a confidence interval for the reliability. Nevertheless, changes of this interval might indicate changes of the reliability.

B. Reliability Indicators

In IEEE 802.11 based WLANs, access points periodically send *beacon frames* to announce their presence and to synchronize the members of the service set. The time interval between two such frames is usually $102.4\,\mathrm{ms}$ [14]. To assess the reliability of the wireless environment during monitoring, we propose to create for each access point a sequence of observations, which are 1 if the beacon frame was received and 0 otherwise. From such a sequence we propose to calculate confidence intervals based on all observations I_n and based only on the previous k observations $\widetilde{I}_n$. We interpret these confidence intervals as indicators for the reliability of the wireless channel between the considered access point and the monitoring station. Here I_n describes the long-term behavior of the channel, whereas $\widetilde{I}_n$ describes the short-term behavior.

Instead of indicators based on the reception of beacon frames, it is possible to construct indicators based on signal strength or latency. However, setting the retest/warning thresholds and determining false positive/negative rates is non-trivial.

VI. PROOF OF CONCEPT

To demonstrate the feasibility of our method we measured and monitored the reliability of wireless communication in an actively operating industrial environment. Next we describe where and how we collected the data and how we analyzed it.

A. Scenario

We analyzed the wireless communication in the order processing department of a food processing factory. In the order processing department the packaged food is compiled according to orders of customers by operators on electric fork lifts. Tablets mounted on the fork lifts show the operator which items have to be collected in the next step. The tablets communicate with the server infrastructure using wireless communication based on WiFi. This scenario provides a good testing environment, because it is an environment in which a set of heterogeneous wireless communication systems are used for practical purposes and thus can be passively monitored. The requirements on latency are low, because of the human in the loop. Higher requirements could be tested, but would not be provided by the underlying WiFi technology.

B. Collecting measurement data

During our measurements the environment was actively operating. To keep the wireless communication intact, we did not create additional traffic, but only passively monitored already existing traffic. We placed three WiFi-Modules in the environment and collected all WiFi traffic over the course of approximately 18 hours. We collected data from stationary points as well as from operating (and hence moving) fork lifts. That is, the physical complexity and overhead were low.

C. Analysis

From the collected data we created traces of: the received signal strength indicator from each transmitting device; the utilization of the channel; the total number of transmitting devices and the reliability indicators for each source of beacon frames. After an initial measurement of the reliability, we continuously monitored the indicators.

Figure 3 shows the results of our monitoring method. In this example, the wireless environment

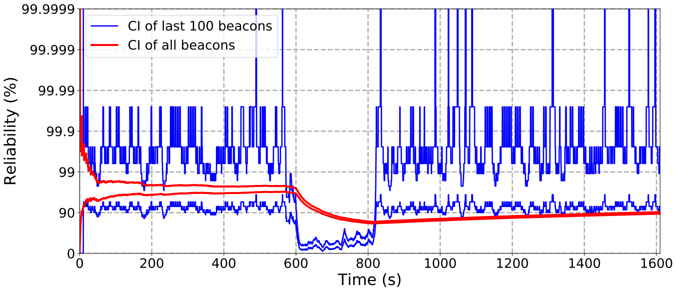

Figure 3. Monitoring of an operating factory shows changes at ≈600 s and ≈800 s.

does not change significantly in the first 600 seconds after the initial reliability measurement. However, after 600 seconds a large change to the reliability indicator is visible. At this time a new measurement of the reliability is necessary.

D. *Caveats, validation and future work*

We tried to guarantee the results of our method, but some statistical assumptions remain [2]. A simple way to validate the methods is to determine the reliability by transmitting many packets. However, when the environment changes faster than the necessary number of packets can be transmitted, we are not aware of any method that is able to validate the measurements. Given these caveats we still consider our measurement method trustworthy.

To fulfil the requirements of the 5G-ACIA [15] the concept needs to be adapted for 5G. This includes selecting appropriate resource blocks which provide the same information (timing, signal strength information) we extract from WiFi Beacon Frames. We expect the necessary information to be available from existing data at base stations (gNodeB). How frequently the information needs to be collected needs to be determined empirically. A full specification of the method for 5G is future work.

VII. CONCLUSION

In this paper we approached the problem of measuring and monitoring the end-to-end reliability of a wireless network. Our approach provides desirable statistical properties (false-positive/negative rates and testing assumptions), non-intrusive design and checking of predictive value. These properties provide unique benefits for long term measurements and monitoring of reliability in industrial environments. In a proof of concept in an operating environment we showed the general feasibility to monitor the reliability of a wireless network in operation. In the future, methods like these should be used to monitor wireless networks in critical applications.

Acknowledgments: This research is partly funded by the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK) and the Austrian state Salzburg.

REFERENCES

[1] M. Herlich, T. Pfeiffenberger, J. Du, and P. Dorfinger, "Survey of Scenarios for Measurement of Reliable Wireless Communication in 5G," in *Proceedings of the Computer Safety, Reliability, and Security*, Springer, 2018.

[2] M. Herlich, C. Maier, and P. Dorfinger, "A Black Box Measurement Method for Reliability of Wireless Communication," in *Proceedings of the EM-5G: Workshop on Experimentation and Measurements in 5G*, (New York, NY, USA), ACM, 2018.

[3] M. Herlich, T. Pfeiffenberger, J. Du, and P. Dorfinger, "5G-MLab: Developing a Wireless Reliability Measurement Framework," in *Wireless Congress: Systems and Applications*, 2017.

[4] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," *IEEE transactions on dependable and secure computing*, 2004.

[5] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects," *IEEE Wireless Communications*, 2018.

[6] F. Bai and H. Krishnan, "Reliability analysis of DSRC wireless communication for vehicle safety applications," in *Intelligent Transportation Systems Conference*, IEEE, 2006.

[7] S. Woo and H. Kim, "Estimating link reliability in wireless networks: An empirical study and interference modeling," in *Proceedings of the INFOCOM*, IEEE, 2010.

[8] G. Pocovi, T. Kolding, M. Lauridsen, R. Mogensen, C. Markmller, and R. Jess-Williams, "Measurement framework for assessing reliable real-time capabilities of wireless networks," *IEEE Communications Magazine*, vol. 56, no. 12, pp. 156–163, 2018.

[9] D. C. Salyers, A. D. Striegel, and C. Poellabauer, "Wireless reliability: Rethinking 802.11 packet loss," in *World of Wireless, Mobile and Multimedia Networks*, IEEE, 2008.

[10] R. G. Newcombe, "Two-sided confidence intervals for the single proportion: comparison of seven methods," *Statistics in medicine*, 1998.

[11] S. Wallis, "Binomial confidence intervals and contingency tests: mathematical fundamentals and the evaluation of alternative methods," *Journal of Quantitative Linguistics*, 2013.

[12] P. O'Connor and A. Kleyner, *Practical Reliability Engineering (fifth edition)*. Wiley, 2012.

[13] ZDKI Industrial Radio TG1, "Aspects of Dependability Assessment: 'Applications, Requirements and Validation'," 2017.

[14] IEEE, "802.11-2016 Part 11: Wireless LAN Medium Access Control and Physical Layer Specifications," 2016.

[15] 5G-ACIA, "Key 5G Use Cases and Requirements," 2020.

BIOGRAPHIES

MATTHIAS HERLICH received a Ph.D. in Computer Science from the Universität Paderborn. After working at the Palo Alto Research Center and the National Institute of Informatics in Tokyo he went to Salzburg Research. His experience ranges from analytic and simulation-based approaches to evaluation of measurements. His focus is the measurement of reliability in wireless communication.

CHRISTIAN MAIER received a BSc in Mathematics from the Technical University of Munich. He is now a researcher at Salzburg Research. His work focuses on reliability and bandwidth measurements in wireless networks, and on the application of machine learning to communication technologies.