Leveraging MAC Preambles for an Efficient Link Estimation

Camilo Rojas and Jean-Dominique Decotignie
Swiss Center for Electronics and Microtechnology (CSEM)
Neuchâtel, Switzerland
camilo.rojas@epfl.ch, jean-dominique.decotignie@csem.ch

***Abstract*—Estimating the quality of a link is a key primitive in WSNs, as upper layers use this piece of information in making performance-critical decisions. State-of-the-art estimators extract a single sample of the link state per packet. Thus, improving the accuracy through multiple samples requires multiple packets, resulting in a significant energy and traffic overhead and delay.**

To address this issue, this paper proposes Rep, a novel sampling scheme able to extract the link quality from the packet repetitions of low-power preamble sampling MACs. The experiments show that Rep reduces the energy and traffic used for link estimation by one order of magnitude, and increases the speed of the process by one order of magnitude, while maintaining state-of-the-art accuracy.

I. Introduction

Collection protocols' ability to identify and make use of the most reliable routes available relies on the computation of a Link Quality Estimation (LQE). The LQE thus has a fundamental impact on the network performance, affecting the design of higher-layers of the protocol stack [1], and constituting a central piece of state-of-the-art collection protocols, such as RPL [2], Contiki-Collect [3] and CTP [4].

It is widely accepted that making use of physical metrics, such as RSSI (Received Signal Strength Indication) and LQI (Link Quality Indication), on top of the traditional PRR, in the computation of the LQE results in significant accuracy, and thus Wireless Sensor Network (WSN) reliability increases [1]. However, the accuracy improvement potential of RSSI and LQI, and thus composite metrics relies on their access to multiple samples.

State-of-the-art LQE algorithms, still operate by extracting a single sample of the link state with each individual packet. Thus, increasing the number of samples in order to mitigate the interference and obtain higher accuracies of the LQE would necessitate a proportional increase in the number of packets. This translates in an increase of the energy overhead, as well as the traffic and the duration of the LQE operation [1] [5] [6]. These factors thus become a limitation to the number of samples obtainable, hindering the use of promising information sources. Moreover, this accuracy-energy consumption trade-off creates additional challenges in mobile WSNs, since the high variability of the medium imposes more pressing constraints on the duration of link estimation [6].

In this paper, we propose a mechanism to overcome this trade-off, which we have named Rep. Rep manages to obtain the high number of samples necessary for a reliable computation of the RSSI, while lowering the number of packets transmitted, and thus the energy consumption. The novelty of this sampling mechanism lies in its ability to leverage the already existing packet repetitions (hence the name Rep) in preamble sampling medium access control (MAC) protocols in order to extract the link quality information. The experimental evidence provided in this paper shows that Rep can reduce the traffic, energy consumption and duration of a link estimation by one order of magnitude, without compromising on the metric's accuracy.

II. Related Work and State-of-the-art

The problem of accurately and efficiently estimating the link quality in WSNs is considered an essential [1], yet open research challenge [7]. For example, several studies suggest that LQI is a promising LQE [8] [9] [10]. Nevertheless, LQI needs to be averaged over a high number of samples (40 up to 120) in order to provide a result accurate enough to be used by routing protocols [8] [6]. More concretely, each node in a WSN operating with this metric would need to transmit 40 to 120 packets for a single estimation, and to repeat this process periodically, generating a prohibitive energy and traffic overhead. Additionally, the delay caused by obtaining this information hinders the taking of critical decisions in the upper layers, such as bootstrapping and repairing the topology, potentially resulting in packet losses.

Common approaches to improving the LQE attempt at either increasing the accuracy of metrics that can be obtained from the radio hardware or from the reception statistics with a low overhead, or at combining multiple such metrics into composite metrics ([1] presents a comprehensive review of these efforts). Nevertheless, the information required by many of these metrics is often not available in commonly used WSN protocols. The most cited reasons include the prohibitive number of samples necessary for an accurate use of LQI and RSSI inputs [8], the fact that some metrics' inputs are only available upon reception, thus of no use to the sender [1], or the fact that the different inputs necessary for the computation of the composite metric are not simultaneously available [7].

A previous approach at solving these limitations is F-LQE [7], a link estimator using fuzzy-logic to combine multiple link quality metrics. Its authors also proposed several design changes that must be implemented in Collection Tree Protocol in order to extract the required input information. F-LQE necessitates many samples of RSSI and LQI, which the authors solve with frequent beaconing (every 1s). The resulting high energy consumption limits the practical applications of F-LQE.

Additionally, [5] proposes LQR, a novel paradigm for coping with unreliable links. Instead of performing a link quality estimation which would require multiple packets, LQR only performs a ranking (i.e., identifying the best link available at that given moment), thus needing less packets. LQR operates by probing all the neighbors of a mote with multiple broadcasts, receiving their replies with the LQE, and comparing these inputs in order to create a ranking. Therefore, it significantly improves the efficiency in terms of the number of sample packets, at the expense of not quantifying the link quality. The authors also address the problem of designing a tailored mechanism for extracting the required input information. However, by only performing the ranking and not the estimate, LQR cannot support collection tree protocols, perhaps the most widespread application of WSNs. The reason is that these protocols require quantifying and comparing routes, which is an operation that necessitates the estimation of the links along each route.

The shared concern for the energy and traffic cost of link quality estimation, as evidenced by multiple efforts to minimize the transmission overhead [5] [6], has also stimulated the research on fine-granularity link estimators. An example of such an effort is presented in [8], which studies the estimation accuracy of LQI and RSSI with sampling periods of 10ms (the lowest period experiments found in the literature). To the best of our knowledge, the studies that analyze the accuracy of link estimation with a sampling period inferior to 100ms, do not develop a technique to integrate the fast sampling in MAC protocols. The proposed mechanism, Rep will build upon these efforts and propose a concrete mechanism to integrate these link estimators in low power MAC protocols and use them in practical applications.

Additionally, state-of-the art protocols tend to isolate the experimental system from external disruptions. In the frame of this study, we have chosen to keep the system subject to stress from wireless interference in order to analyze its effect. This approach is not commonly reported in link estimation literature.

III. Proposed Protocol - Rep

A. Key Features

Rep extracts useful link quality statistics, such as the average LQI and RSSI, from a single packet in low-power preamble sampling MAC (LP-MAC) protocols (e.g. BMAC [11], WiseMAC [12] and ContikiMAC [13]) (Fig. 3). This is achieved by leveraging a resource that is commonly available in state-of-the-art protocol stacks, yet so far under-used, namely the fact that LP-MACs transmit the exact same packet several times (hereafter referred to as packet repetitions) in order to wake the intended receiver(s) up.

Estimating the link quality from each of these repetitions enables a reduction in the number of packets required for link estimation, and consequently the scaling down of the energy consumption. Moreover, it significantly reduces the duration of the link estimation. This study analyses the impact of sampling with Rep over the accuracy of the link estimation, as compared to the state-of-the-art coarse-grain approach.

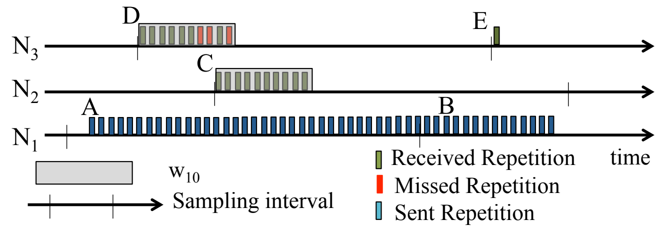

Fig. 1: Rep's working principle: N_2 and N_3 overhear several MAC repetitions from the packet broadcasted by N_1, this samples can be used for fine-grain link estimation.

Throughout this paper, the term *packet* is used exclusively in the sense of the OSI model, namely the data transmission unit at the Network layer, while the transmission unit at the MAC layer is named *frame*. In the case of LP-MAC protocols, each packet is encapsulated in a frame and transmitted through multiple copies, that we refer to as repetitions.

B. Implementation

The repetitions from unicast, anycast and broadcast packets can be used for link estimation. This study explores the use of broadcasts, mostly given the fact that they rely on a large number of repetitions, and the fact that they are ubiquitous in routing protocols for discovering new nodes and routes.

The integration of Rep in LP-MACs is demonstrated through ContikiMAC-R, a mechanism suggested in our previous study [14]. ContikiMAC-R is based in ContikiMAC, modified as follows:

- The sender uses a MAC flag (F_{LE} = true, LE comes from Link Estimation) to identify a Rep broadcast, i.e., it will provide the link quality to the receiver. With Rep, the broadcast transmission is prolonged by a time w_m (see Fig. 1). The extension corresponds to the time needed to send m repetitions, and it is necessary to prevent a mote from receiving less than m repetitions as a consequence of waking up at the end of the series. Receiving less repetitions can hinder the accuracy of the resulting LE; thus the extension has the role of decoupling the resulting estimation from the wake-up offset.
- Upon reception of a broadcast repetition, the radio keeps listening during a time window w_m. If the link has a high quality, it will receive m repetitions.
- When F_{LE} = false, ContikiMAC-R operates as ContikiMAC. Hence, a broadcast transmission will consist of enough repetitions to fill a sampling period (T_w) (plus one extra repetition), and only one repetition will be received (if successful).

The following events illustrate its operation with $m = 10$ (Fig. 1):

(A) Node N_1 broadcasts a beacon by sending repetitions to fill a sampling period (F_{LE} = true).

(B) Since F_{LE} = true, N_1 extends the broadcast time by w_{10}.
(C) N_2 wakes up and receives one of the repetitions. Since it is a broadcast and F_{LE} = true, it extends its listening interval to w_{10}. During the window, it receives ten out of ten repetitions (PRR=100% for this link, PRR is the Packet Reception Ratio).
(D) N_3 receives seven out of ten repetitions (PRR=70%).
(E) N_3 wakes up and receives a repetition from the packet already received in (D) for link estimation, hence N_3 does not extend the listening interval, as there is no guarantee of receiving at most m packets. Otherwise, reusing the packet might provide an erroneous link estimation if there are less than 10 repetitions left for reception.

In case the receiver misses the repetition "in the air" during the channel sampling, it will miss the entire packet and not be able to extract the link quality. The probability of this event happening is the same when Rep is enabled or disabled, as the number of channel samplings of each node able to receive a particular broadcast packet stays constant.

C. Advantages

By leveraging the packet repetitions to extract the link quality, Rep uses a sampling period in the order of milliseconds, resulting in a fine-grain (FG) sampling (Fig. 3). This is faster than traditional link estimation schemes that use a sampling period in the order of seconds (Fig. 2), hereafter referred to as coarse-grain (CG) sampling. This feature provides, by design, the following advantages:

- **Reduction of the duration of link estimation by one order of magnitude**: Rep avoids the delay for extracting multiple link quality samples. This delay is due to contention for accessing the channel, and depends on the network density, thus ranging from 100ms to seconds. The result of using Rep is a significant increase in the speed of the link estimation (section IV-B4 further quantifies this improvement). Moreover, this speed increase holds for different routing protocols, irrespective of the transmission primitive they use for link estimation (i.e., broadcast, unicast or anycast). This happens as the core saving in terms of speed comes from circumventing the delay in accessing the channel, which is determined by the sampling period (T_w), and the subsequent backoff delay.
  Moreover, the presented fine-grain approach can be used in collection protocols that take advantage of the sporadic high quality intervals of unstable links. These schemes necessitate a mechanism to quickly identify a good link in order to use it before it degrades (e.g., [5]). Rep can improve the speed and accuracy of the required link characterization.
- **Reduction of the overhead (traffic) of link estimation by one order of magnitude**: Rep requires a single packet to extract multiple link quality samples, compared to CG schemes which only extract one sample per packet (even though a packet will anyway be composed of multiple repetitions). Therefore, Rep results in a significant overhead reduction in terms of traffic.
  Another relevant factor are the possible changes in the distribution of packet collisions due to the reduction of the sampling timescale. Even though this demands a specific analysis, which is out of the focus of this paper, we highlight that the reduction of the traffic with respect to the CG case should decrease the probability of collisions.
- **Increase in the number of samples**: The fast probing enables the collection of a volume of samples otherwise prohibitive for LQEs based on coarse-grain sampling. For example, Rep can collect 120 samples in less than a second, as compared to the duration in the order of minutes required by coarse-grain schemes (consider a MAC sampling frequency of 2 Hz, details explained in Sec. IV-B4). The difference is even more significant if we consider the channel contention due to multiple nodes within range, all required to transmit for link estimation purposes.
- **Combined used of Rep and CG**: We note that both FG and CG sampling approaches can be used simultaneously by extracting the link quality from multiple repetitions of multiple packets. Therefore, in case that the fine-grain sampling hinders the accuracy of a particular LQE, the information obtained can still be used to complement the coarse-grain scheme. This approach provides a more complete description of the link state, and we expect it to improve the accuracy compared to the sole use of coarse-grain information.

The benefits that several domains can perceive from the reduction in energy, duration and traffic of the channel sampling enabled by Rep, include: increasing the lifetime of collection networks, improving the reliability of multi-hop WSNs, enhancing the dataset for machine learning applications in WSNs and improving the recovery and boostrapping of WSNs. The fast and energy-efficient link estimation provided by Rep can be used in applications beyond data collection, for example: providing new tools in flooding based protocols (such as glossy [15]), or improving the accuracy of RSSI-based wireless localization and Radio Tomographic Imaging [16].

D. Other Issues

On the other hand, the influence of fine-grain link dynamics over Rep data must be considered. The links of intermediate quality are known to have a bursty behaviour [1]. This means that they switch between extreme qualities (good or bad) in the order of seconds [17]. Therefore, the link metrics behave differently as the interval between samples decreases (the granularity becomes finer). For example, the PRR tends to assume extreme values (closer to 0% or 100%), as the granularity becomes finer. Hence, it is not a good indicator of intermediate links at this time-scale [1]. Conversely, several studies suggest that the LQI is able to accurately characterize

TABLE I: Model Parameters

Parameter	Value	Meaning
n	1-9	Mean nb. of neighbors
T_d	1.54ms	Transmission duration of a repetition
T_D	2.2ms	Radio-on time for transmitting a repetition
T_{gap}	0.5ms	Radio-off time between consecutive repetitions
T_{oh}^{tx}	0.872ms	Overhead time for sending a packet
T_{oh}^{rx}	1.27ms	Mean overhead time for receiving a packet
w_m	1 - 120	Nb. of samples
T_w	125ms	ContikiMAC Sampling period
w_{tw}	50	Nb. of repetitions to fill a T_w

intermediate links even when the channel is probed during time-scales in the order of 10 ms [8].

IV. Experimental Results

A. Energy Considerations

The following analysis is meant to validate the energy reduction provided by the LE operation with Rep, compared to CG. We focus on the time that the motes need to keep the radio active (i.e., transmitting or listening), since it is a widely-used hardware-agnostic metric for energy consumption [18].

This energy analysis concerns solely the LE packets, since the Rep mechanism is bypassed for the rest of the traffic by using a flag (explained in Sec. III-B).

The analysis is based on the low-power preamble-sampling MAC protocol ContikiMAC v3.0, because it is the default MAC of ContikiOS [19], the *de facto* standard operating system in the WSN domain. Moreover, we ignore the energy consumed by the microcontroller, which is a common assumption because the consumption of the transceiver is the dominant driver of the overall energy performance (e.g., [20]).

A more precise quantification of the network energy savings depends on the frequency of the LE operation and the use of the collected LQ information, both determined by the routing protocol. While we will leave a more detailed, scenario specific characterization and quantification of energy savings out of the scope of this paper, it remains a valuable next step in our study.

We compare the energy that a mote requires for a LE operation using CG and Rep, by modeling each of the approaches. The model parameters are summarized in Table I and detailed in the following sections. We focus on the case where beacons are sent using broadcasts, because this is the main type of beacons used by state-of-the-art protocols. In this analysis, we ignore the processing of delays in the microcontroller, which are small compared to the transition times between states of the radio.

Using Rep to extract RSSI and LQI over many samples enables the reduction of the number of broadcast packets, which should diminish the overall energy consumption of the link estimation operation. Equation 1 details the percentage of energy savings provided by Rep (S), where E_{CG} and E_{Rep} represent the energy required for obtaining w_m samples using CG and Rep, respectively.

$$S = 100 * (E_{CG} - E_{Rep})/E_{CG} \tag{1}$$

First, we address the CG scheme (Fig. 2) in Eq. 2, and describe the total energy (E_{CG}) as the addition of the energy consumed in transmissions by a single mote (E_{CG}^{tx}) and the reception of n mote(s) within range (E_{CG}^{rx}). The parameter n measures the density of the network.

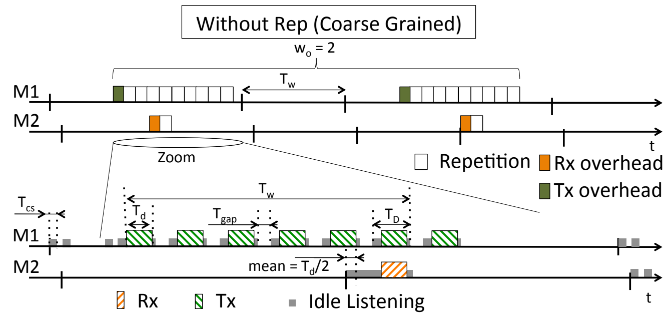

Fig. 2: Timing diagram of the link estimation operation using multiple packets (Coarse-grain sampling).

$$E_{CG} = E_{CG}^{tx} + E_{CG}^{rx} \tag{2}$$

E_{CG}^{tx} (detailed in Eq. 3) represents the energy spent for transmitting w_m packets. Before the transmission, ContikiMAC performs two channel samplings, each with a duration of T_{cs}, thus creating a per-packet transmission overhead of $T_{oh}^{tx} = 2 * T_{cs}$. Assuming the channel is found available, the mote will proceed with the transmission of the w_{tw} repetitions that compose a broadcast, each of them requiring a radio-on time equal to T_D (including the time to turn the radio on and off). P^{tx} is the power consumed by the transceiver while on transmission state.

The values of the previous parameters have been measured using Cooja [21] (the state-of-the-art simulator for WSNs, with cycle-accurate emulators for the microcontroller and radio) for ContikiMAC (version Contiki 3.0 vanilla, transceiver cc2420) and they are shown in Table I.

The energy used for radio transmission decreases significantly, as a result of the reduction in the number of transmitted LE beacons. Even though Rep requires a small energy investment for extending the transmission of packet repetitions by w_m (see Sec. III-B), this is considered negligible in comparison to the reduction in the number of transmitted beacons, since each beacon is composed by tens to hundreds of packet repetitions (w_{tw}, depending on the settings of the MAC protocol).

$$E_{CG}^{tx} = P^{tx} * w_m * (T_{oh}^{tx} + (w_{tw} * T_D)) \tag{3}$$

E_{CG}^{rx} (detailed in Eq. 4) shows the energy spent by n motes within range, to receive one repetition for each of the w_m packets transmitted. Each repetition is composed by the time required for receiving the packet (T_D) and additional radio-on (T_{oh}^{rx}) time for detecting the transmission. T_{oh}^{rx} is described in eq. 5, and it is composed by the mean time required to detect the transmission of a repetition ($T_d/2$, T_d is the transmission duration of a repetition), plus the gap time until the transmission of the first repetition (T_{gap}). T_{oh}^{rx} is also called wake-up probe. P^{rx} is the power consumed by the transceiver

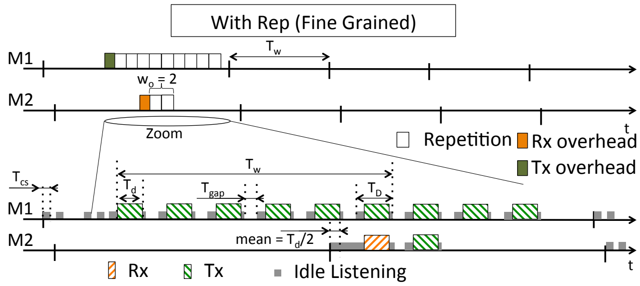

Fig. 3: Timing diagram of the link estimation operation using multiple repetitions of a single packet with Rep (fine grained sampling).

while on reception state. The previous parameters have also been measured using Cooja.

$$E^{rx}_{CG} = P^{rx} * w_m * (T^{rx}_{oh} + T_D) * n \tag{4}$$

$$T^{rx}_{oh} = T_d/2 + T_{gap} \tag{5}$$

Following, we analyze the case of Rep (Fig. 3) by describing its energy in Eq. 6, which follows the same form as E_{CG}.

$$E_{Rep} = E^{tx}_{Rep} + E^{rx}_{Rep} \tag{6}$$

The energy component spent in transmission (E^{tx}_{Rep}, eq. 7) considers a single packet composed by $w_{tw} + w_m$ repetitions, each of them requiring a radio-on time T_D, already explained for the CG case. Nevertheless, in the case of Rep, the transmission overhead T^{tx}_{oh} is factored in only once, due to the transmission of a single packet.

$$E^{tx}_{Rep} = P^{tx} * (T^{tx}_{oh} + ((w_{tw} + w_m) * T_D)) \tag{7}$$

In the case of the energy spent in idle listening and reception (E^{rx}_{Rep}, eq. 8), we consider the reception of w_m repetitions coming from a single packet, each of them requiring a radio-on time T_D (explained in the CG case). In the case of Rep, the reception overhead T^{rx}_{oh} is only taken into account once, due to the reception of a single packet.

$$E^{rx}_{Rep} = P^{rx} * (T^{rx}_{oh} + (T_D * w_m)) * n \tag{8}$$

The energy used for radio listening in Rep decreases when compared to the CG case, due to the reduction of wake-up probes (each requiring a time T^{rx}_{oh}), which are required for signaling the reception of a packet, and have a duration similar to the one of a packet in ContikiMAC. Therefore, Rep decreases the listening energy by reducing the number of wake-up probes to a single one (from a single packet), instead of w_m (from w_m packets, compare eqs. 4 and 8). Rep does not change the number of repetitions received, remaining w_m in both Rep and CG schemes.

Considering that low power transceivers typically have similar power consumptions during the transmission, reception and listening states, we use the expression $P^{tx} = P^{rx}$ and cancel these factors out of eq. 8.

Figure 4 shows the energy savings (S in eq. 1) obtained by using Rep instead of CG. Note that the bottom of the y-axis is located at 45%. We consider a number of neighbors in the range of 1 to 9, which corresponds to the spectrum between a sparse network and dense one. The limit of 9 is chosen close to the capacity of the neighbor table in CTP (10 neighbors) [22].

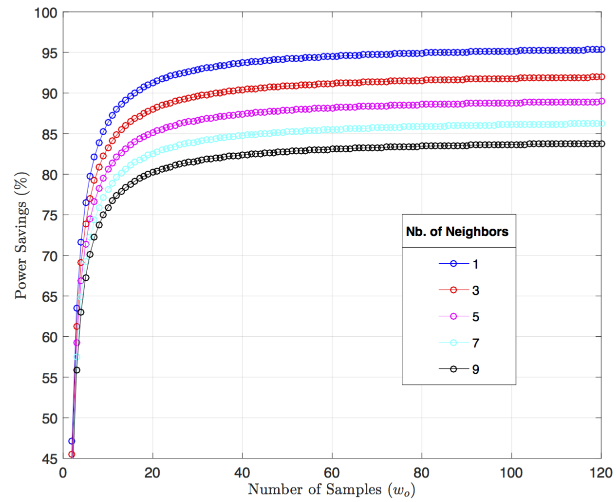

Fig. 4: Power savings of using Rep instead of using CG to obtain a number of samples, as a function of the number of samples and the network density (from the analytic model in Eq. 1). Note that the bottom of the y-axis is located at 45%.

The savings in the figure display an elbow around w_m = 18 samples and they tend to stale beyond this point. The plateau indicates that using Rep instead of CG brings an energy reduction of the link estimation operation of one order of magnitude (83 - 95%).

Figure 4 also shows that the savings decrease as the density of the network increases, thus reaching the value 83% for a $n = 9$ neighbors. We can obtain a deeper insight of this property by analyzing the limits of the Savings equation (eq. 1) with extreme density ($\lim_{n\to\infty} S$) and in the plateau of the curve ($\lim_{w_o\to\infty} S$), simultaneously. The conditions reduce S to $100 * T^{rx}_{oh}/(T^{rx}_{oh} + T_D)$. This represents a theoretical lower bound for the savings, which evaluated according to the values in Table I is 37%, a value that remains significant.

The energy consumption results were not experimentally validated, as a fair comparison would have required the simultaneous execution of the evaluations with Rep and CG, in order to have both schemes testing the same link quality conditions. We leave this further validation for future work.

B. Performance Analysis

1) Overview: By design, Rep improves the speed and reduces the traffic of the link estimation primitive. Thus, we have performed a set of experiments aimed at comparing the accuracy of LQEs when using input information provided by Rep with the one provided by the state-of-the-art sampling scheme (detailed later in this section). The objective is to un-

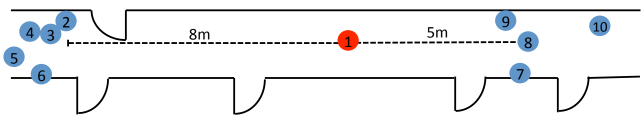

Fig. 5: Location of the experiment in an office hallway. N_1 periodically broadcasts beacons that are used by N_i (i=2..10) to estimate the link quality.

derstand Rep's capacities and limitations under environmental disturbances commonly found in WSN deployments.

2) Setup: We have created a testbed for studying the link estimation accuracy inspired by RadiaLE [23]. RadiaLE is a framework for designing and assessing link quality estimators for WSNs, consisting of a star network, where nodes N_i (i = 2..n) are placed at different distances around a central node N_1, as shown in Fig. 5.

Distance and direction are fundamental factors affecting the link quality, therefore the links between N_1 and the surrounding nodes will have different qualities. We chose a particular direction for each node and determined the appropriate distance empirically, prior to the experiments, aiming to be as close as possible to the transitional region (quantified by means of the broadcast packets PRR).

The transitional region describes the operation of links with intermediate quality and it is generally characterized by a high-variance in reception rates [24]. We aim at setting the links in this range because, in practice, some of them will deviate from this operation point, thus also providing information about the regions below (low quality) and above (high quality) the transitional region. We study the performance of Rep under the most challenging conditions, by focusing on transitional links, as opposed to low- and high-quality links, which are more stable and require significantly less samples for an accurate characterization [8] [25] [5].

Our testbed consists of n = 10 TelosB nodes (Crossbow) and was placed on the ceiling of a hallway in an office building (Fig. 5). Most of the hallway walls are metallic.

We use the following traffic pattern: N_1 is the only node that transmits, the others are listeners during the entire experiment. N_1 sends a broadcast every 2s and the other nodes extract the link quality from the MAC repetitions of each packet ($RSSI$ and LQI).

The focus of this study on $RSSI$ and LQI, over PRR, has a threefold motivation: i) the domain shows a significant and growing interest in exploiting these physical information sources for link estimation [7] [26], ii) there is significant experimental evidence that both metrics preserve a correlation with the link quality at small sampling periods (∼10ms, similar to the reception period of MAC repetitions) [8], and iii) both metrics are widely available in low-power transceivers. We do not elaborate on PRR measurements obtained through REP, because PRR has the tendency to acquire extreme values (e.g., 0-100%) as the period of the sampling becomes shorter. We have observed this in our experiments and has been reported by other authors in [8] [27].

All the nodes are running Contiki OS v3.0 [19] and contikiMAC-R (sampling period 2Hz).

3) Experiment Variables: WSNs are subject to multiple environmental disruptions with different time and intensity patterns. Therefore, we have designed experiments to study the effects of common disruptions by analysing the following variables:

- Wireless Interference: all the experiments were repeated in two channels: 11 (2.405 GHz) and 26 (2.480 GHz) of the TelosB radio. The former is susceptible to interference from Wi-Fi and Bluetooth networks, while the latter is out of the interference spectrum. This technique for studying the effects of interference has been previously used in RadiaLE [23].
- Number of Repetitions: the maximum number of repetitions that a node can receive from a packet (length of the observation window) determines the accuracy of link estimation. All experiments were evaluated using broadcasts (F_{LE} = true) and repeated with w_m=20 and w_m=80 repetitions. Moreover, the link quality from the first repetition was registered separately in order to evaluate the link estimation accuracy when F_{LE} = false. We select the size of the windows motivated by the promising LE accuracy reported in [6] (for w_m=20) and [8] for (for w_m=80).

The experiments presented in this section have been designed to understand the effect of the previous variables on the accuracy of Rep. We have analysed each combination of conditions during at least 2h, resulting in 3600 broadcasts (sampling opportunities).

We have measured the link estimation accuracy by evaluating the PRR in the near future (PRR_f) and studying its correlation to the link estimation metrics, for each received packet. This methodology has been borrowed from the LE study in [6].

PRR_f was calculated offline over a window of 100 packets that are expected after each reception, as this scheme emulates the application traffic. This methodology is inspired by previous work in the link estimation domain [6].

4) Sampling Mechanisms: The experiments compare the accuracy of link estimation based on the following schemes:

- Coarse-grain (Standard Method): This is the traditional scheme for link estimation. The information is obtained by each receiver (N_i, i=2...10) by considering the link quality from the first repetition received from each packet, and then averaging the values obtained from the packets received during a time interval when m packets are expected (m packets result in $\overline{RSSI}$ and $\overline{LQI}$).
- Fine-grain (Rep): A receiver node considers the link quality from the repetitions received during a time interval when m repetitions are expected, as part of a single packet (a window w_m provides $\overline{RSSI}$ and $\overline{LQI}$).

Given that ContikiMAC uses a standard sampling period (512ms), the transmission of 80 packets for a CG link estimation requires at least 40.96s (= 80 x 0.512s, assuming there is no contention to access the channel). In our experiment, the packets are transmitted with a 2s period, which is closer to values used in practical applications, therefore a single CG link estimation requires 160s (= 80 x 2s). With Rep, a single link estimation takes less than 1.024s (= 2 x 0.512s), as it requires less than two sampling periods: the sampling period required for sending a broadcast, plus the w_m window (e.g., the train of transmitted repetitions from M1 exceeds T_w in Fig. 3).

Upon reception of a packet, ContikiMAC-R provides the link quality data from the first repetition (CG) and the aggregated data from m repetitions (FG). We use the latter for studying the Rep link estimation, and the former in two different ways: for studying the link estimation with a single sample (for example, Fig. 6 row 1) and for studying the CG estimation by averaging it over multiple packets (for example, Fig. 6 row 2).

5) Results: This experiment studies the effect of interference from wireless networks over the estimation accuracy of link quality metrics obtained from Rep.

Figure 6 displays several scatter plots showing the dependency of PRR_f on the $\overline{LQI}$, for $m = 20$ repetitions (left column) and for $m = 80$ repetitions (right column). Each row in the figure displays the results of a different sampling method: single-sample (first row, plots a and b), CG (second row, plots c and d) and Rep (third row, plots e and f). The fourth row displays the LQI variance vs. $\overline{LQI}$ for Rep (plots g and h).

Plots in Fig. 6 obtained from averaging multiple samples (rows 2 and 3) display 4 separate data point islands. It was not possible to obtain data to connect them as there were no links available in this operation region. We overlap a regression curve (displayed in red), which is detailed in Fig. 8 and explained later in this section. The discretization in the x-axis for single-sample and Rep plots is noticeable, the reason is that $\overline{LQI}$ is an integer because it is calculated in the node's microcontroller, while the $\overline{LQI}$ from CG plots (row 2) is a floating point value because the average is calculated offline.

We note that CG (row 2) and Rep (row 3) plots in Fig. 6, tend to have a sigmoid shape as described in [8] [28], and display a clear transition region at 80 < $\overline{LQI}$ <90. These features come from the variance introduced by wireless interference, since they have not been observed in analogous experiments in channels that are less affected by interference (e.g. 26). This effect has been reported in previous studies [1].

The comparison between CG and Rep plots for each w_m value does not show a significant difference in terms of trend of data and its dispersion. A comparison of the regression curves shows that the corresponding CG and Rep plots highlight similar underlying sigmoids. This behavior supports our assertion that the accuracy of Rep is equivalent to the one of the standard method. However, while each point of the former required transmitting 20 and 80 packets, Rep only required a single packet.

The sigmoid trend in the results validates the RadiaLE-like approach to study the quality of the links, since each link only provides information from a particular operation region. Thus, using several links gives insight about multiple portions of the characteristic curve.

The comparison between the plots in Fig. 6 obtained from a single sample and the ones obtained from multiple samples, shows that the averaging of multiple samples significantly reduces the dispersion in the x-axis, thus improving the accuracy of the $\overline{LQI}$ calculation.

Figure 7 is analogous to Fig. 6, but displays the dependency of the future PRR to the RSSI. The regression curves are compared in Fig. 8. The transitional region in Fig. 7 rows 1 and 2 is well defined between -93 and -90 dBm and it is significantly steeper than the transitional region for LQI in Fig. 6. These features agree with previous work reporting that the average $\overline{LQI}$ has a stronger correlation with PRR when compared to $\overline{RSSI}$, in the transitional region [8] [9] [10].

Additionally, Fig. 7 displays a lower dispersion in the y-axis for high RSSI values than its LQI counterparts in Fig. 6. This result is aligned with previous work that reports a smaller variance of RSSI compared to LQI for a given number of samples [8]. This explains the negligible reduction of Var[RSSI] when it is calculated with 80 samples instead of 20 (Fig. 7 g and h).

The data displayed in Figs. 6 and 7 was analyzed through a non-linear regression, in order to better study the link estimation capabilities of the Rep and CG methods. The fitting curves chosen are the sigmoids in Eq. (9) (LQI) and Eq. (10) (RSSI). Eq. (10) (RSSI) has been proposed as a theoretical model for the PRR of a low-power wireless link as a function of RSSI [24]. We extend the use of this model for the LQI, based on the reports of multiple authors that show a sigmoid-like behavior in LQI similar to RSSI's [8] [29]. The position of the transition zone of the sigmoid (x_t in the equations) has been chosen to maximize the r-squared coefficient (i.e., minimize the regression error).

The non-linear regression does not intend to prove the goodness of the fit of the measured data points to a theoretical model. Instead, the objective is to show that the two link estimation mechanisms, Rep and CG, result in similar best fit curves, thus stressing the equivalent accuracy of both methods. The technique of fitting PRR vs. LQI or RSSI curves is widely used for predicting the behavior of a link from LQI or RSSI measurements through sigmoid models (e.g. [8] and [24]), or piece-wise linear models (e.g., in [29] and [30]).

$$PRR_f = 100 * (1 - (0.5 * e^{-(\overline{LQI} - x_t)/b}))^a, x_t = 65 \tag{9}$$

$$PRR_f = 100 * (1 - (0.5 * e^{-(\overline{RSSI} - x_t)/b}))^a, x_t = -93 \tag{10}$$

Table II summarizes the coefficients obtained from the regressions, and Figs. 8 and 9 display the regression curves. The Rep and CG curves are very similar for a given w_m, thus supporting the equivalent accuracy of both methods. Moreover, the distinct w_m values bring more significant differences between the curves compared to the use of a given data extraction method (Rep vs. CG). These observations hold for both LQI and RSSI.

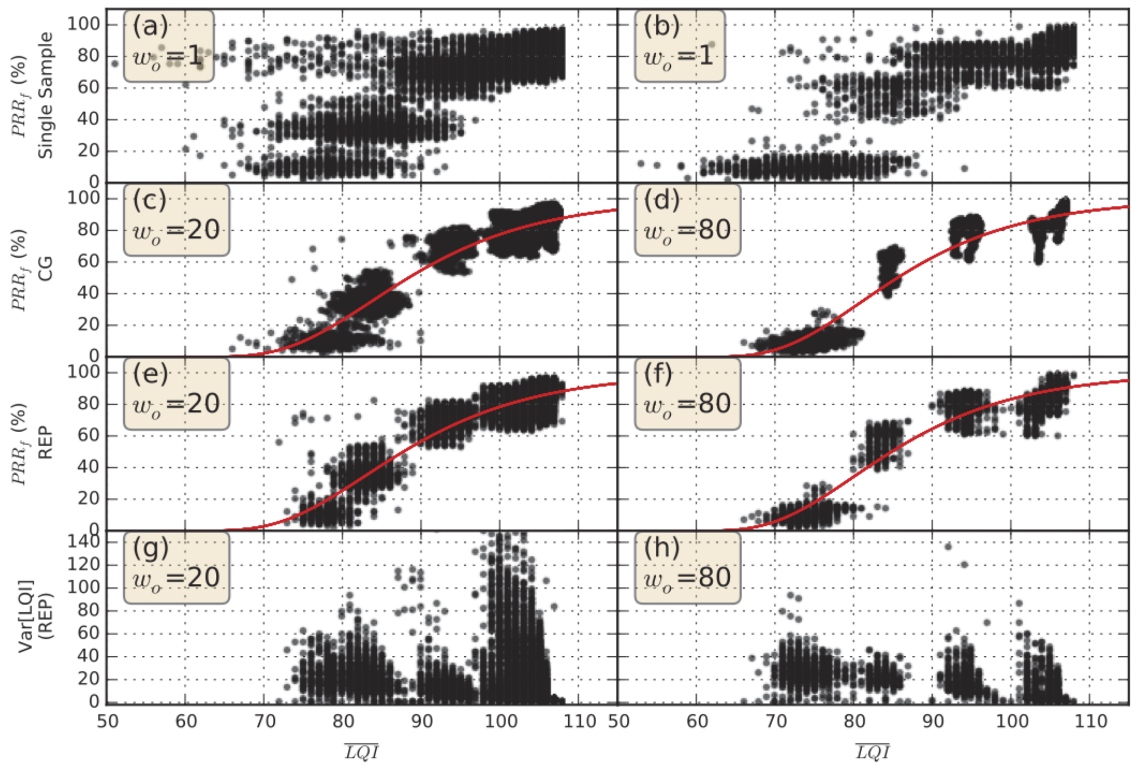

Fig. 6: Future PRR vs. the current value of $\overline{LQI}$ to study the effect of wireless interference. The plots from CG (c and d) and Rep (e and f) show both schemes result in equivalent link estimations (same trend and comparable dispersion). $\overline{LQI}$ is calculated with 20 (left column) and 80 samples (right column). Plots a and b use a single sample. The variance of the $\overline{LQI}$ obtained with Rep is shown in row 4. The black dots are semitransparent, therefore the darkness of the tone is proportional to the number of overlapped data points. The regression curves (in red) are compared in Fig. 8.

TABLE II: Results of the regression using the sigmoid in Eqs. (9) (LQI) and (10) (RSSI).

Experiment	coef. a	coef. b	R-squared
LQI, $m = 20$, Rep	8.832	12.02	0.8682
LQI, $m = 20$, CG	9.536	11.91	0.8715
LQI, $m = 80$, Rep	7.1704	11.73	0.9175
LQI, $m = 80$, CG	7.9544	11.5	0.9167
RSSI, $m = 20$, Rep	1.4832	6.592	0.6491
RSSI, $m = 20$, CG	1.6488	6.3	0.6638
RSSI, $m = 80$, Rep	9.568	1.021	0.7881
RSSI, $m = 80$, CG	6.8832	1.494	0.7884

RSSI curves display more significant differences between w_m values, which are clearly visible in Fig. 9. This can be explained considering the lower influence of external interference during the experiment with w_{20} compared to w_{80}, thus resulting in less datapoints in the lower part of the curve in Experiment 1. This difference is noticeable in Figs. 7 (c) and (d) (w_{80}), which display the lower tail of the curve extending to -93 dBm, not present in (e) and (f) (w_{20}).

We do not include a comparison between the results of the experiments w_o = 20 vs. 80 samples, because the measurements were not done at the same time, under the same perturbations over the network. We leave this analysis to be done in the future in a setting designed for controllable perturbations. Nevertheless, we believe that the results presented are a good starting point for validating the controllable perturbations in a future experiment. These observations are valid for all the results presented in this section (IV-B).

We have repeated this experiment in channel 26, which is largely interference free. The results further support an equivalent accuracy of the link estimation obtained through CG and Rep. Both methods exhibit negligible differences, for each w_m value, thus showing that the link estimation accuracy is equivalent when the channel is stable. We have observed that wireless interference increases the variance of PRR_f, both as a function of LQI and RSSI. Comparing the results obtained with w_o = 20 and 80 samples does not display significant differences, due to the constant conditions of the channel. These results are expected since, in the absence of channel perturbations, the accuracy should be independent of the sampling time-scale and the number of samples. A graphical depiction of the results has been omitted due to space limitations.

C. Consequences of Reducing the Sampling Granularity

The equivalence between the LE obtained from CG or Rep will depend on the temporal dynamics of the perturbations over the links (in the case of non-mobile networks). We have identified three cases where the two methods should yield similar results. First, there are no perturbations, therefore the LE will be independent of the sampling period. Second, the changes of the perturbations occur progressively and over a time-scale that is significantly bigger than the one of CG sampling, so that Rep (being faster) will still capture the same behavior of a link. Third, the changes follow a pattern, with

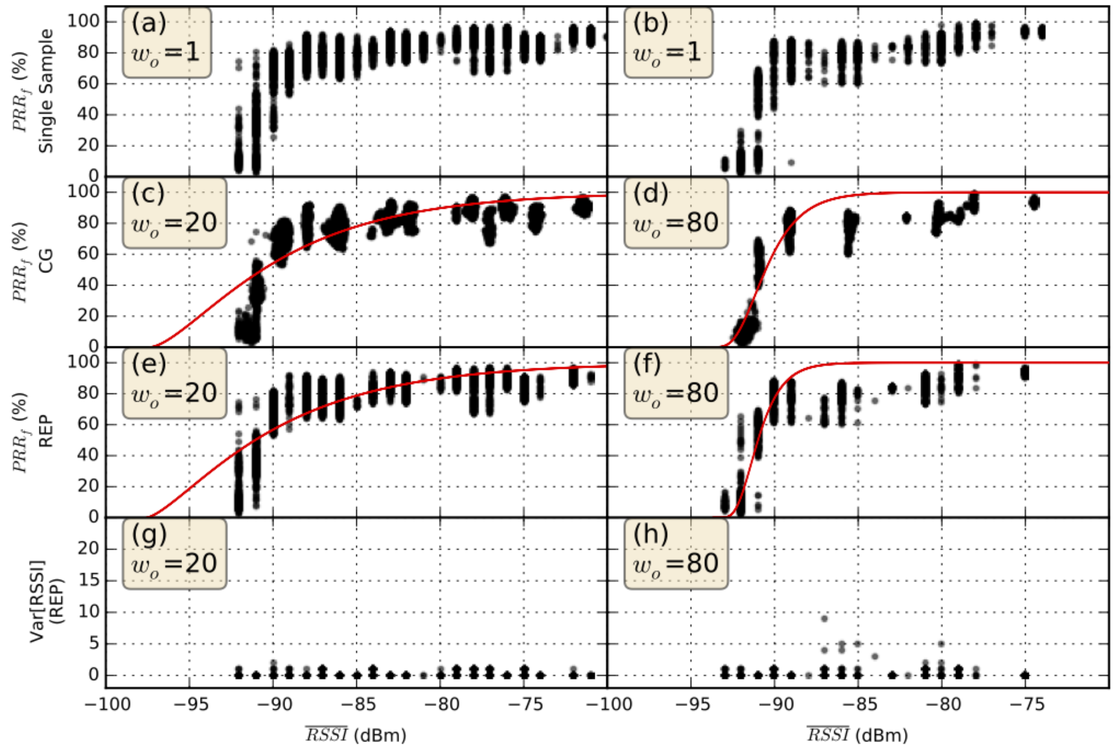

Fig. 7: Future PRR vs. the current value of $\overline{RSSI}$ to study the effect of wireless interference. The plots from CG (c and d) and Rep (e and f) show both schemes result in equivalent link estimations (same trend and comparable dispersion). For further interpretation details of the plot refer to the caption of Fig. 6. The regression curves (in red) are compared in Fig. 9.

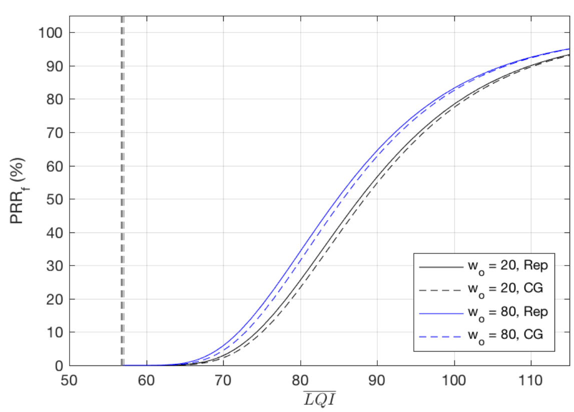

Fig. 8: Regression from Future PRR vs. the current value of $\overline{LQI}$ datasets (Fig. 6) using a sigmoid model (Eq. 9). The similarity of the resulting curves for Rep and CG stress the equivalent accuracy of both data extraction schemes.

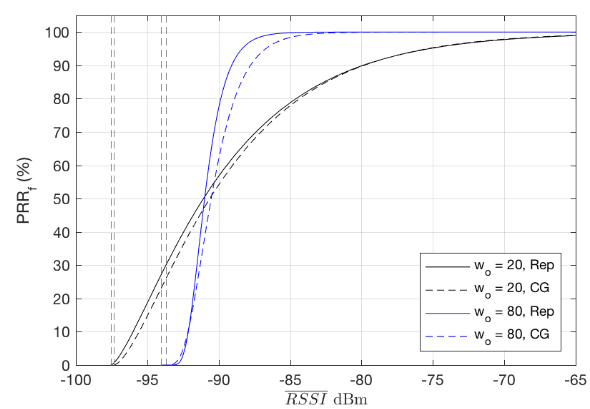

Fig. 9: Regression from Future PRR vs. the current value of $\overline{RSSI}$ datasets (Fig. 7) using a sigmoid model (Eq. 10). The similarity of the resulting curves for Rep and CG stress the equivalent accuracy of both data extraction schemes.

a period small enough to be fully captured with the short duration of Rep.

We will elaborate on the third case to interpret the results obtained in section IV. In an office environment, the dominant source of interference creating the need for multiple samples for a reliable RSSI and LQI measurement is represented by the use of WiFi [31]. Potential other interferences such as Bluetooth and Zigbee are both far less ubiquitous, and microwave ovens tend to have a sporadic use, mostly around lunchtime.

Studies characterizing WiFi variations over time, such as [17] have found that WiFi typically displays a radiation pattern composed of spikes with a period of 24.67ms, each with a width of 8ms (extrapolated from the figures in the reference). Moreover, when constructing empirical models, studies based on JamLab [31], the state-of-the-art mechanism for generating typical interference patterns have also considered that, when downloading a large file, with 97% probability the clear channel period is smaller than 1ms. The overlap of multiple WiFi users, typical for an office environment, creates a continuous pattern during the working hours, while also potentially increasing the frequency of the spikes.

Therefore, the dominant interference sources present a periodic pattern with temporal dynamics that should be fully captured by Rep. This behavior explains our ability to achieve equivalent noise mitigation effects with our fine-grain sampling (of 2.725ms per sample in the Contiki 3.0 vanilla code), as the ones of the traditional coarse-grain sampling in the scale of minutes.

V. Conclusions

This paper has elaborated on Rep, a new sampling mechanism that leverages existing resources of LP-MAC protocols in order to reduce the duration and overhead of link estimation schemes (in terms of traffic and energy) by one order of magnitude. The experimental evidence provided has proven that the link estimation accuracy, when using Rep, remains equivalent to the state-of-the-art methods' in a wide range of challenging scenarios. They have also shown, that, under the perturbation pattern of an indoors office deployment, a fast and energy efficient fine-grain sampling, can result in equivalent measurement accuracies, as a more time and energy expensive coarse-grain one.

For future work, we will use Rep to improve routing protocols by integrating novel composite metrics that, in spite of outperforming the state-of-the-art, are not used in practical applications because they are not supported by the sampling mechanism. We will use these implementations to study the improvements in the energy consumption. We will also further explore the applications of Rep for indoor localization, mobile networks and Radio Tomographic Imaging.

Acknowledgments

This work was developed in the framework of WiseSkin, a Swiss nano-tera project, partly funded by nano-tera and partly by the project partners. The authors would like to thank nano-tera for its support.

References

[1] Baccour, Nouha and Kouba, Anis and Mottola, Luca and Ziga, Marco Antonio and Youssef, Habib and Boano, Carlo Alberto and Alves, Mrio, "Radio Link Quality Estimation in Wireless Sensor Networks: A Survey," *ACM Trans. Sen. Netw.*, vol. 8, 2012.

[2] T. Winter, P. Thubert, and A. Brandt, "Rpl: Ipv6 routing protocol for low power and lossy networks," *IETF Request for Comments 6550*.

[3] N. Tsiftes, J. Eriksson, and A. Dunkels, "Low-power wireless ipv6 routing with contikirpl," in *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2010.

[4] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection Tree Protocol," in *Proc. of the 7th ACM Conference on Embedded Networked Sensor Systems*, ser. SenSys '09, New York, NY, USA, 2009.

[5] M. Zuniga, I. Irzynska, J.-H. Hauer, T. Voigt, C. A. Boano, and K. Roemer, "Link quality ranking: Getting the best out of unreliable links," in *Distributed Computing in Sensor Systems and Workshops (DCOSS), 2011 International Conference on*. IEEE, 2011, pp. 1–8.

[6] C. A. Boano, M. A. Zuniga, T. Voigt, A. Willig, and K. Romer, "The Triangle Metric: Fast Link Quality Estimation for Mobile Wireless Sensor Networks," in *2010 Proc. of 19th International Conference on Computer Communications and Networks (ICCCN)*, Aug. 2010, pp. 1–7.

[7] N. Baccour, A. Kouba, H. Youssef, and M. Alves, "Reliable link quality estimation in low-power wireless networks and its impact on tree-routing," *Ad Hoc Networks*, vol. 27, pp. 1–25, Apr. 2015.

[8] K. Levis, "Rssi is under appreciated," in *Proc, of the 3rd Workshop on Embedded Networked Sensors, Cambridge, MA, USA*, 2006.

[9] J. Polastre, R. Szewczyk, and D. Culler, "Telos: Enabling ultra-low power wireless research," in *Proc. of the 4th International Symposium on Information Processing in Sensor Networks (IPSN)*, NJ, USA, 2005.

[10] L. Tang, K.-C. Wang, Y. Huang, and F. Gu, "Channel characterization and link quality assessment of ieee 802.15. 4-compliant radio for factory environments," *IEEE Trans. on Industrial Informatics*, vol. 3, 2007.

[11] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proceedings of the 2Nd International Conference on Embedded Networked Sensor Systems*, ser. SenSys '04. New York, NY, USA: ACM, 2004, pp. 95–107.

[12] A. El-Hoiydi and J.-D. Decotignie, "WiseMAC: an ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks," in *Proc. of the 9th International Symposium on Computers and Communications (ISCC), 2004.*, vol. 1, Jun. 2004, pp. 244–251 Vol.1.

[13] A. Dunkels, "The contikimac radio duty cycling protocol," 2011.

[14] C. Rojas, D. Piguet, and J.-D. Decotignie, "Poster: Single Packet Link Estimation," in *Proc. of the International Conference on Embedded Wireless Systems and Networks*, USA, 2016.

[15] F. Ferrari, M. Zimmerling, L. Thiele, and O. Saukh, "Efficient network flooding and time synchronization with glossy," in *Proc. of the Int. Conf. on Information Processing in Sensor Networks*, 2011, pp. 73–84.

[16] G. Mao, B. Fidan, and B. D. Anderson, "Wireless sensor network localization techniques," *Computer Networks*, vol. 51, no. 10, 2007.

[17] K. Srinivasan, M. A. Kazandjieva, S. Agarwal, and P. Levis, "The β-factor: Measuring wireless link burstiness," in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '08. New York, NY, USA: ACM, 2008, pp. 29–42.

[18] M. Schuß, C. A. Boano, M. Weber, and K. Römer, "A competition to push the dependability of low-power wireless protocols to the edge," in *Proc. of the International Conference on Embedded Wireless Systems and Networks (EWSN)*, 2017, pp. 54–65.

[19] A. Dunkels, B. Grönvall, and T. Voigt, "Contiki - a lightweight and flexible operating system for tiny networked sensors," in *29th Annual IEEE International Conf. on Local Computer Networks*. IEEE, 2004.

[20] K. Langendoen and A. Meier, "Analyzing mac protocols for low data-rate applications," *ACM Trans. Sen. Netw.*, vol. 7, pp. 10:1–10:34, 2010.

[21] F. Osterlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt, "Cross-level sensor network simulation with cooja," in *Proc. of the 31st IEEE conference on Local computer networks.*, 2006.

[22] U. Colesanti and S. Santini, "The collection tree protocol for the castalia wireless sensor networks simulator," Department of Computer Science, ETH Zurich, Zurich, Switzerland, Technical Report 729, Jun. 2011.

[23] N. Baccour, A. Koubía, M. Ben Jamía, D. do Rosário, H. Youssef, M. Alves, and L. B. Becker, "Radiale: A framework for designing and assessing link quality estimators in wireless sensor networks," *Ad Hoc Netw.*, vol. 9, no. 7, pp. 1165–1185, Sep. 2011.

[24] M. Zuniga and B. Krishnamachari, "Analyzing the transitional region in low power wireless links," in *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*. IEEE, 2004, pp. 517–526.

[25] R. Fonseca, O. Gnawali, K. Jamieson, and P. Levis, "Four-bit wireless link estimation." in *HotNets*, 2007.

[26] T. Liu and A. Cerpa, "Temporal adaptive link quality prediction with online learning," *TOSN*, vol. 10, pp. 46:1–46:41, 2014.

[27] N. Baccour, D. Puccinelli, T. Voigt, A. Koubaa, C. Noda, H. Fotouhi, M. Alves, H. Youssef, M. A. Zuniga, C. A. Boano, and K. Römer, *Overview of Link Quality Estimation*, 2013, pp. 65–86.

[28] D. LaI, A. Manjeshwar, F. Herrmann, E. Uysal-Biyikoglu, and A. Keshavarzian, "Measurement and characterization of link quality metrics in energy constrained wireless sensor networks," in *Global Telecommunications Conference (GLOBECOM), 2003.*, vol. 1, 2003, pp. 446–452.

[29] B.-r. Chen, K.-K. Muniswamy-Reddy, and M. Welsh, "Ad-hoc multicast routing on resource-limited sensor nodes," in *Proc. of the 2nd International Workshop on Multi-hop Ad Hoc Networks: From Theory to Reality*, ser. REALMAN '06, New York, NY, USA, 2006, pp. 87–94.

[30] C. Gomez, A. Boix, and J. Paradells, "Impact of lqi-based routing metrics on the performance of a one-to-one routing protocol for ieee 802.15.4 multihop networks," *EURASIP J. Wirel. Commun. Netw.*, 2010.

[31] C. A. Boano, T. Voigt, C. Noda, K. Römer, and M. Zúñiga, "Jamlab: Augmenting sensornet testbeds with realistic and controlled interference generation," in *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*. IEEE, 2011, pp. 175–186.