

# Fingerprinting Channel Dynamics in Indoor Low-Power Wireless Networks

Marcel Bosling, Matteo Ceriotti, Torsten Zimmermann,  
Jó Ágila Bitsch Link, Klaus Wehrle  
Communication and Distributed Systems Group  
RWTH Aachen University, Germany  
<http://comsys.rwth-aachen.de>

## ABSTRACT

Wireless low-power embedded devices are populating indoor environments, where everyday activities drastically impact communication. We explore a statistical approach to identify changes to the communication links state during system operation. The long-term behavior of the link RSSI is modeled with a normal distribution and compared against the model of the most recent measurements. A Welch's t-Test is then employed to identify whether the short-term and long-term link evolutions stem from the same distribution. Upon significant divergence, the long-term model is updated and a significant change in the underlying communication state is inferred. We investigate this technique to efficiently store a compressed fingerprint of the evolution of communication. Considering the memory constraints of low-power embedded systems, this approach allows to gather extensive information on the behavior of communication directly from the deployed network. This fingerprint could then be used to replay the network dynamics in simulation. We implemented the introduced techniques to prove their feasibility. In controlled experiments, we evaluate the reactivity and sensitivity of the approach to changes in the environment, as well as the accuracy of the resulting channel fingerprint.

## Categories and Subject Descriptors

C.2.1 [Computer-Communication Networks]: Network Architecture and Design—*Wireless Communication*

## Keywords

sensor networks; IEEE 802.15.4; real-world measurements; indoor testbeds; communication models

## 1. INTRODUCTION

In low-power embedded networks, the environment affects the communication quality between devices. This holds in particular in indoor scenarios, i.e., rapidly changing settings where network quality suffers from events such as people

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*WiNTECH'13*, September 30 2013, Miami, Florida, USA  
Copyright 2013 ACM 978-1-4503-2364-2/13/09 ...\$15.00.  
<http://dx.doi.org/10.1145/2505469.2505477>.

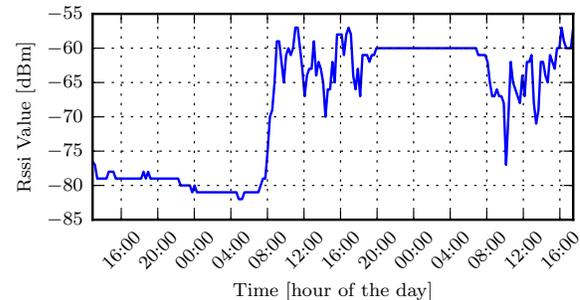


Figure 1: Evolution of a single 802.15.4 link between Sunday 22 and Tuesday 24, in April 2012.

crossing a wireless link and doors being opened or closed. Figure 1 depicts the RSSI trace of a representative IEEE 802.15.4 [1] wireless link, as measured in an indoor testbed at our department. During working hours, people activities manifest in a variable and short-term impact on the observed RSSI; during the weekend or nights, the link exhibits instead a stable behavior. Interestingly, there is a clear difference in the average RSSI between two consecutive nights, which must have been caused by more structural changes in the scenario, e.g., doors or windows.

Adaptive network algorithms try to cope with these effects by controlling parameters such as transmission power or routing paths. Their effectiveness is dependent upon the ability to accurately and rapidly characterize the link behavior. All the approaches from the literature (Section 2) focus on specific, short-term dynamics of the communication links. Moreover, these studies are typically done in controlled environment; once deployed, continuously monitoring the details of the communication behavior is considered impractical. As a result, several dynamics, caused by the environment in which the technology is immersed once operational, are disregarded. These become visible only from a long-term perspective, as shown in Figure 1 in the link dynamics over two consecutive nights.

In this paper, we introduce a strategy to estimate the current communication state and detect significant changes in its development over time (Section 3). The work is based on the long- and short-term evolution of the RSSI of a wireless link. We exploit its normal distribution in stable links and its high sensitivity to environmental changes to employ traditional statistical analysis techniques, which can be effectively implemented on resource-scarce devices. By observing online the RSSI of incoming links, a Welch's t-Test reveals if the short-term evolution statistics deviate from the long-

term ones. In this case, the recent model replaces the old one, and this information is exposed to other network layers.

We executed several experiments to evaluate the approach in a real indoor low-power wireless testbed (Section 4). The results shows that environment dynamics can be followed with a delay of 5 to 10 seconds, depending on the configuration of the stability test and the sample size (Section 5). Furthermore, we investigate the possibility to use the recorded sequence of identified models as a compressed fingerprint of communication over time, e.g., to replay wireless links in simulation a posteriori (Section 6). Such an effective compression of the incoming RSSI measurements, with limited accuracy loss, will enable us to observe long-term behavior of communication in deployed networks (Section 7).

## 2. RELATED WORK

This work relates to modeling wireless communication, online computation of model states, and estimation of link state in adaptive protocols.

**Modeling wireless communication.** Wireless communication can be described by different flavors of propagation models [4], and through experimental studies [11, 13]. With more complex environments, obstacles and materials need also to be taken into account, adding complexity to the resulting model [5]. While these models provide static description of communication, its actual behavior manifests a time dependent factor, mostly caused by the movement of obstacles in the environment. In [2], the authors matches different statistic distributions, e.g., Normal, Laplace, to wireless measurements collected in indoor scenarios. The authors then only envision the possibility of detecting online the best fitting distribution. While these studies offer great insights on the actual behavior of wireless communication, they provide off-line techniques that describe a posteriori the evolution of the network. We aim, instead, at a general channel fingerprinting method capable of identifying changes in the communication state at runtime, without prior knowledge. Moreover, we aim at a computationally efficient solution that can be executed on embedded devices with limited memory and computational requirements.

**Resource-efficient online computation of models.** The computation of models online, based on current measurements, is a promising technique that can be exploited in a wide range of application scenarios. In [10], the authors use derivative-based prediction to match light measurements against a linear model and reduce the amount of data communicated to the base station. The work clearly demonstrates the feasibility and effectiveness of employing modeling techniques on constrained devices. However, while a linear approximation is sufficient to fingerprint the evolution of light in road tunnels, it cannot be employed in our scenario to model communication dynamics.

**Online estimation of link quality state.** The authors of TALENT [8] try to combine long- and short-term quality metrics in order to exploit bursty links, i.e., intermediate quality links that exhibit times of high reliability. Bursty links were initially explored in [3], where such links were exploited as shortcuts in multi-hop networks. Similarly, in [8], the authors employ an online machine learning algorithm for predicting short time intervals with high reliability. The information about the predicted availability of such links can then be exploited by the routing algorithm. These ap-

proaches exploit short-term behavior of the links and focus on specific link characteristics.

## 3. MODELING COMMUNICATION STATE

In typical indoor settings, heterogeneous factors, e.g., people walking and modifying the environment, produce very different and unpredictable impact on individual wireless link characteristics. In this section, we propose an approach to model link states based on the observation of RSSI over time. The simplicity of the approach allows an effective implementation for resource-constrained devices, which can then directly observe online the dynamics of communication.

### 3.1 Link State Model

As analyzed in [2], a stable link RSSI can be approximated with a normal distribution. By observing a specific incoming link, we can take a sequence of measurements,  $\mathbf{X} = \{X_{t_1}, \dots, X_{t_n}\}$ , at the time messages are received. It is then possible to compute the sample mean,  $\bar{X}_n$ , and the sample variance,  $S_n^2$ , over  $n$  observations and use them as point estimators for the underlying normal distribution of the measured link,  $\mathcal{N}(\mu, \sigma^2)$ . When an event in the environment happens, the wireless propagation characteristics may change. The incoming link RSSI can then be newly sampled,  $\mathbf{X}' = \{X_{t_{(n+1)}}, \dots, X_{t_{2n}}\}$ , and the point estimators of the current underlying distribution,  $\mathcal{N}(\mu', \sigma'^2)$ , recomputed.

To analyze whether the actual properties of the link have changed, the two distributions need to be compared and significant differences need to be identified. The problem is redolent of a hypothesis test. It could be interpreted as testing two samples  $\mathbf{X}$  and  $\mathbf{X}'$  as belonging to the same underlying distribution. In the case they originate both from the same period of stability, no difference should be found. If one of the samples originates from a period in which the RSSI distribution differs from the sample of the other period, the difference should be detected.

### 3.2 Sampling RSSI

We define the *history* of a link RSSI as the sequence  $X_{t_1}, \dots, X_{t_n}$  of measurements for different points in time  $t_1 < \dots < t_n \leq t$  with  $t$  the current time and  $t_{i+1} - t_i = c$ . Most environmental events affect the link RSSI quickly; we then use a fixed size sliding window over the immediate past history, denoted as  $h_{\text{head}}^s$  with size  $s$ . Each time a measurement is taken, it is added to the history, and the oldest measurement is removed to maintain a fixed size  $s$ . The removed entry is added to another history,  $h_{\text{tail}}$ , of variable size, which captures the long-term evolution of the RSSI.

After updating the two histories their respective statistics are compared through an affinity test. If the affinity test succeeds, the short-term evolution ( $h_{\text{head}}^s$ ) supports the stable long-term evolution ( $h_{\text{tail}}$ ) and the next measurement is taken. Otherwise, the two distributions  $\mathcal{N}(\mu_{\text{head}}, \sigma_{\text{head}}^2)$  and  $\mathcal{N}(\mu_{\text{tail}}, \sigma_{\text{tail}}^2)$  are different such that the stable long-term evolution is not supported by the immediate past anymore, forcing a *model change*. In this case  $h_{\text{tail}}$  is replaced by  $h_{\text{head}}^s$  and the  $h_{\text{head}}^s$  has to be rebuilt from new measurements.

### 3.3 Affinity Test

For the affinity test, the Welch's t-Test [12] can be chosen as the statistics on the histories of stable links are normally distributed. Let  $X_1^1, \dots, X_{n_1}^1$  and  $X_1^2, \dots, X_{n_2}^2$  be two samples of different size  $n_1 \neq n_2$ , with unequal variances  $\sigma_1^2 \neq \sigma_2^2$

and  $X_i^k \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mu_k, \sigma_k^2)$ , with  $1 \leq i \leq n_k, k \in \{1, 2\}$ . The test in consideration is given by  $H_0 : \mu_1 = \mu_2$ , and it can be interpreted as testing if the samples being drawn originate from the same population. Both the means and the variances have to be estimated from the samples; therefore,  $\mu_k := \bar{X}_k$  and  $\sigma_k^2 := S_k^2$ . The test statistic, then, becomes

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

which is a normalized distance in terms of the combined variance. It follows a Student's t-distribution with parameter  $\nu$ :

$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{S_1^4}{n_1^2(n_1-1)} + \frac{S_2^4}{n_2^2(n_2-1)}}$$

The hypothesis  $H_0$  on significance level  $\alpha$  is rejected when  $|T| > t_{\nu, 1-\alpha/2}$ , where  $t_{\nu, 1-\alpha/2}$  denotes the quantile function of the t-distribution for quantile  $1 - \alpha/2$  and degrees of freedom  $\nu$  and it is commonly looked up in a table.

Therefore, for each of the two histories employed in our specific scenario, the corresponding statistics can be computed, and the Welch's t-Test can be performed with hypotheses  $H_0 : \bar{X}_{\text{tail}} = \bar{X}_{\text{head}}$  and significance level  $\alpha$ . The test yields two outcomes of acceptance: the two histories either stem from the same normal distribution ( $H_0$ ) or arise from different normal distributions ( $H_1$ ). In the case  $H_0$  is accepted, no model change occurs, and the distribution  $\mathcal{N}(\bar{X}_{\text{tail}}, S_{\text{tail}}^2)$  is updated incorporating the oldest measurement from  $h_{\text{head}}^s$ . In the case  $H_1$  is accepted, a model change is detected, changing the long-term link distribution parameters from  $\mathcal{N}(\bar{X}_{\text{tail}}, S_{\text{tail}}^2)$  to  $\mathcal{N}(\bar{X}_{\text{head}}, S_{\text{head}}^2)$ . The initial time of validity of the new model is then assigned to be the timestamp of the oldest entry in  $h_{\text{head}}^s$ . The higher networking layers can then be notified that a model change occurred with the information about the old and new model. The information consists of the statistics of the old model that is outdated and the statistics of the new model, representing the current communication state.

### 3.4 Implementation for Low-Power Devices

More complex solutions than the one introduced are possible, e.g., employing a different set of distribution models to increase the precision of the acquired link description. As envisioned in [2], we could try to identify at runtime the distribution model, e.g., Normal or Laplace, that best describes the current observations. However, we target a method that could be easily implemented on embedded devices, lacking memory and processing capabilities; the realization of more complex techniques would be then arguably infeasible.

To prove the effectiveness of our approach, we implemented it in TinyOS [7], so to be executed on embedded platforms, e.g., TelosB [9] featuring a CC2420 radio chip [6] as in our own testbed. The history  $h_{\text{head}}^s$  is implemented as a fixed sized ring buffer. Once full, it is updated by removing the oldest entry and making room for a new observation. The old measurement is then added to the history  $h_{\text{tail}}$ . For the long-term history  $h_{\text{tail}}$ , only the statistics are necessary, which are updated each time a measurement is added; this is sufficient as the approach requires measurements only to be added or the complete history to be replaced. For the Welch's t-Test, the estimators  $\bar{X}$  and  $S^2$  for both histories

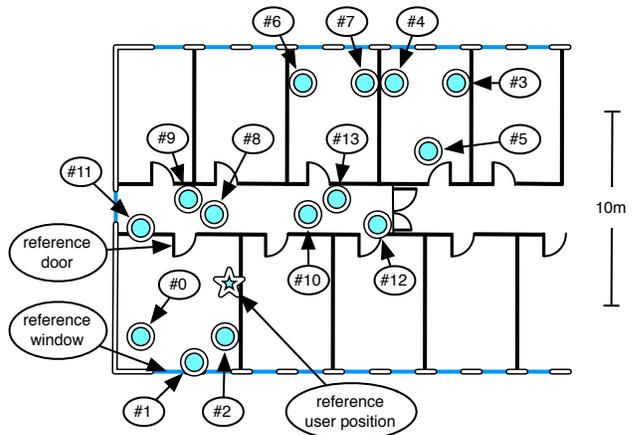


Figure 2: Indoor testbed of 14 TelosB devices.

are calculated once a new measurement is recorded, the test statistics are computed, and the value for a predefined  $\alpha$ , set at compile time, is looked up in a table. The look-up table is built at compile time for a specific  $\alpha$ . The required computation is, therefore, restricted to compute few mathematical operations, i.e., average and standard deviation, and to access simple data structures, i.e., ring buffers and precomputed look-up tables.

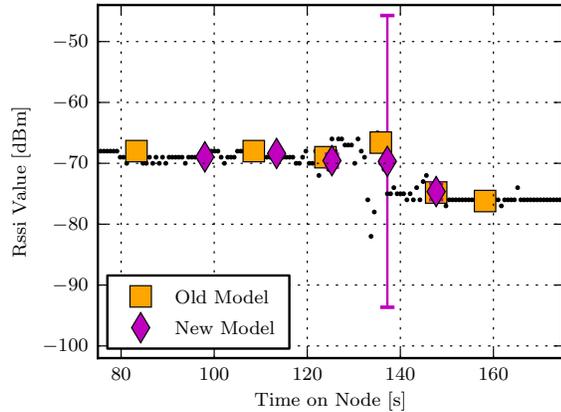
## 4. EXPERIMENTAL SETUP

In this section, we introduce our testbed setup and the set of experiments that we performed.

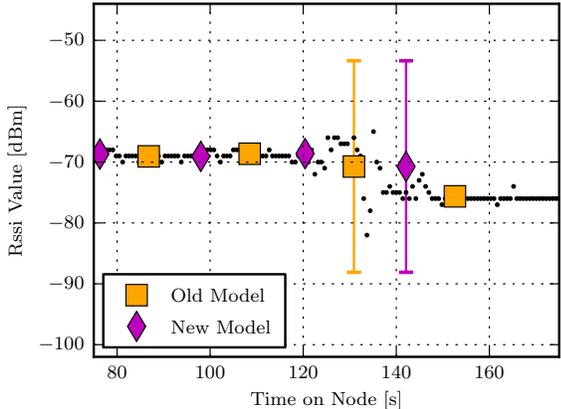
### 4.1 Experiments

The experiments were conducted in an indoor testbed at our department, composed of 14 TelosB wireless sensor devices distributed in a corridor and several offices as depicted in Figure 2. Node #1 lies on a desk; all the other devices are installed on the ceiling. The nodes are running the firmware described in Section 3.4; they are also connected with a serial back-channel to a machine logging the measurements. The same machine acts as an orchestrator for a collision-free communication schedule, selecting the node which is next to perform a send of a packet over the radio. This allows to generate traffic and measure link RSSI without the impact of interference as caused by concurrent transmissions. This is implemented through commands sent via serial connection to the nodes in a round robin fashion with a fixed interval of 50 ms, resulting in each link being checked every 700 ms. We used a configuration for the firmware in which the significance of the Welch's t-Test is  $\alpha = 0.005$ , and the size of  $h_{\text{head}}^s$  is  $s = 15$ . The 802.15.4 radio channel was set to 11, a channel free from concurrent 802.11 networks at the time of the experiments; no radio duty cycling was employed.

To be able to trace single environmental effects, the experiments were conducted at times without people in the offices. The experiments conducted were closing and opening a door (denoted as “reference door” in Figure 2) and opening and closing a window (denoted as “reference window” in Figure 2). We refer to each experiment scenario respectively as DOOR and WINDOW. These events were chosen in order to examine the impact of environmental effects introduced by changes at the perimeter (WINDOW) and inside (DOOR)



(a) Window size = 5 and  $\alpha = 0.005$



(b) Window size = 30 and  $\alpha = 0.1$

**Figure 3: RSSI trace and derived models (average and variance) for different configurations. A new model describes the short-term history at the time of a model change detection. An old model represents the long-term history as computed at the time of a model change; its timestamp corresponds to the one of the most recent sample in the long-term history. Increasing window size and  $\alpha$  results in a decreased number of generated models and a worse adherence to the link dynamics.**

the network. The person running the experiments used a single reference position during all the measurements.

In the DOOR scenario, the person running the experiments (1) timestamped the beginning of the experiment, (2) walked slowly to the already opened door, (3) traced the reach of the door, and (4) closed the door slowly (in approximately 2 s); finally, the user (5) walked back to the initial position, and (6) recorded the end of the experiment. The same procedure was repeated for opening the door few minutes afterwards. A similar procedure was used in the WINDOW scenario. To trace the steps, a dedicated device, held by the person in the experiments, was programmed so that, at the press of a button, a message was sent to the other nodes, timestamped, and recorded at the data server via the serial back-channel. The transmission power levels (in dBm) were chosen to be  $p_d \in \{-25, -15, -10, -7, -5, -3, -1, 0\}$  in the DOOR scenario and  $p_w \in \{-15, 0\}$  for WINDOW. Each combination of experiment and power level was executed three times.

## 4.2 Replaying Measurements in Simulation

The introduced technique presents several parameters: the significance level  $\alpha$ , the windows size  $s$  of  $h_{\text{head}}^s$ , and the transmission power level  $p$ . A reasonable exploration of the parameter space is infeasible through in-field experiments. Furthermore, as the environment is not fully under control, the experiments might not be comparable due to unknown effects that are introduced in between experiment runs. Therefore, we implemented a simulator in charge of replaying offline raw RSSI traces, collected as previously described. In this way, while the transmission power  $p$  clearly affects the measured RSSI signal (and therefore cannot be simulated reliably), different  $\alpha$  and  $s$  values can be explored offline. This approach may introduce deviations due to the different floating point precision between the simulation and the implementation for embedded devices. We measured the differences by comparing the simulated results for the same parameter values used by our firmware during the experiments, observing negligible deviations. In Section 5, we then use simulation to explore the parameter space.

## 5. MODELING EVALUATION

Before discussing the use of the introduced technique for link fingerprinting, we evaluate the approach in terms of sensitivity and reactivity of the model change detection.

### 5.1 Metrics: Sensitivity and Reactivity

Figure 3 depicts the impact of different  $\alpha$  and window sizes on the generation of the link models for the same RSSI trace. As it can be seen, a large window size and a high  $\alpha$  corresponds to a decreased number of generated model. Instead, a lower  $\alpha$  and a smaller window size are more susceptible to marginal fluctuations in the measurements, with a better adherence to the major deviations caused by real events happening in the environment. To better analyze these effects, we introduce the *sensitivity* and *reactivity* metrics.

**Sensitivity.** The events happening in the environment during the experiments, e.g., the door opening or the beginning of the person movement, have reference timestamps. Ideally, one single model change is produced for each event happening in the environment. In reality, it happens that models are produced without a real event to which they can be mapped (false positives), or, on the contrary, that events do not have a corresponding model (false negative). We then compute the number of false positives  $fp$  as

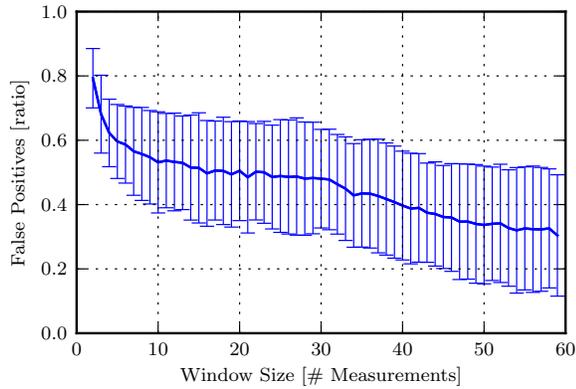
$$fp = \frac{n_{\text{models}} - n_{\text{mappings}}}{n_{\text{models}}}$$

where  $n_{\text{models}}$  is the number of all computed model changes and  $n_{\text{mappings}}$  denotes the number of successfully mapped events. We can similarly compute for the false negatives

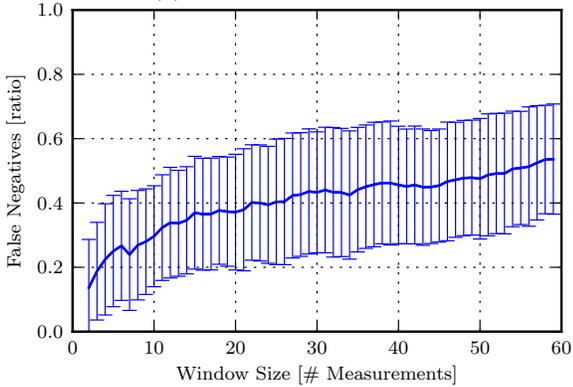
$$fn = \frac{n_{\text{events}} - n_{\text{mappings}}}{n_{\text{events}}}$$

where  $n_{\text{events}}$  is the number of recorded events.

**Reactivity.** In each experiment run, six events are recorded. We further characterize the behavior of the approach through the delay between each event and a corresponding model change. First, the model changes have to be mapped to events with a constrained maximum delay. We therefore match the closest model change  $mc_m$  within a distance threshold of  $d_{\text{max}} = 40$  s to an event  $e_m$ . If the model change  $mc_m$  is situated between two events within the distance threshold,



(a) Relative false positives



(b) Relative false negatives

**Figure 4: Sensitivity results for the DOOR scenario with power level  $-7$  dBm with  $\alpha = 0.005$ .**

it will be mapped to the closer one. If a model change  $mc_m$  is successfully mapped to an event  $e_m$ , it is not available for further mapping to other events. The reactivity measure for a matching is given by

$$r_m = t(e_m) - t(mc_m)$$

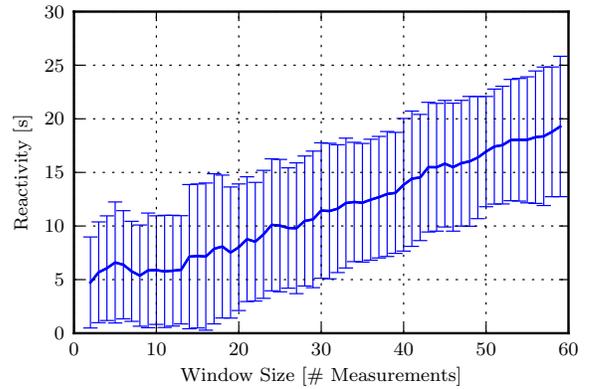
where  $t(e_m)$  denotes the timestamp of the event taking place. The value of  $t(mc_m)$ , instead, denotes the time of the oldest entry in  $h_{\text{head}}^s$  over which the model change was computed.

## 5.2 Filtering Unaffected Links

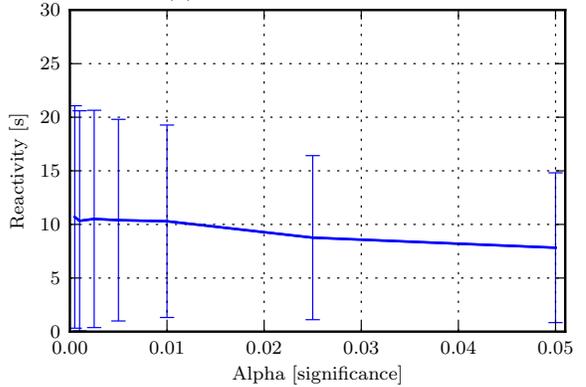
From the gathered measurements, it is evident that some links are not affected by the occurrence of any event. For these links, the model changes are triggered by noise. This is due to the nature of the t-Test, which becomes more responsive to outliers for more stable statistics. The stability can be measured through the variance of a model. In the case of unaffected links, the long-term statistics exhibit a small variance. Noise affects the short-term statistics to diverge from the stable long-term model, generating a model change with almost unobservable magnitude. The difference in the variance between unaffected and affected links is typically significant. Therefore, in the following discussion, we analyze links exhibiting a variance higher than  $t = 4$ . Below this value, noise was indistinguishable from real events.

## 5.3 Results

Figure 4a shows the statistics (average and standard deviation) of false positives for different short-term history sizes. It is possible to notice that for small short-term history sizes



(a) Variable window size



(b) Variable significance level

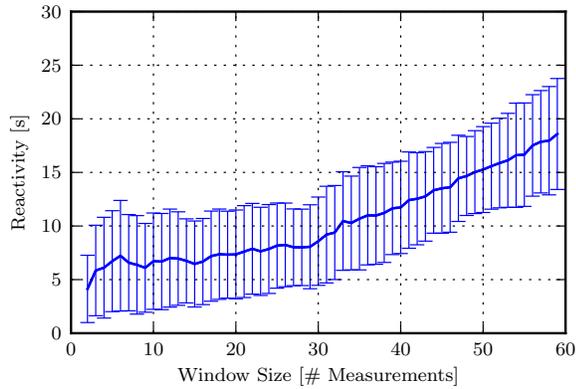
**Figure 5: Reactivity results for the DOOR experiments for power level  $-7$  dBm with  $\alpha = 0.005$ .**

the average of the relative false positives is very high with a low standard deviation. This is motivated by the fact that noise in the RSSI traces affects the statistics for computing model changes significantly more if the short-term history is small. In fact, many more model changes are computed and the mapping to events succeeds only for few of them.

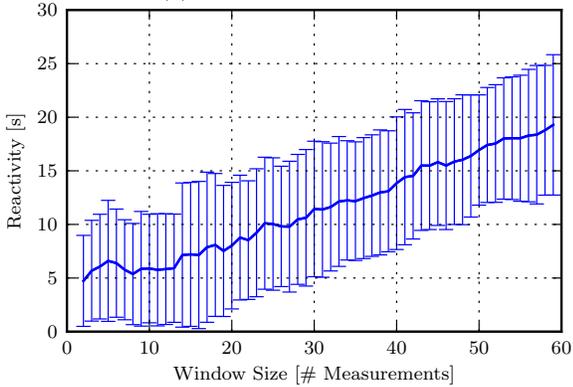
When the short-term history becomes bigger, instead, the average number of false positives decreases but the standard deviation increases. This effect can be explained by the alignment of the model changes. If the remaining few model changes are far away from the events, they become false positives. If they are computed within the provided limits for the mapping, they are matched to an event, decreasing the number of false positives. The overall opposite trend is observable in Figure 4b, which shows the relative number of false negatives for the same experiment.

For reactivity, the average distance of the mapped model changes to the underlying events is shown in Figure 5a. It can be seen that the average increases for increasing short-term history size due to the alignment effect. This motivates the use of smaller short-term history sizes. In Figure 5b, the impact of different  $\alpha$  values is depicted. Due to the clear steps caused by the investigated events on the RSSI signal, the significance of the Welch's t-Test has little impact; the new model distribution clearly differs from the old one.

We now look at the results from different transmission powers for the DOOR scenario and focus on power levels  $-10$  and  $-7$  dBm, which present the major differences among all the pairs of adjacent power levels. Figure 6 shows, however, that the trend for different short-term history sizes is the



(a) Power level  $-10$  dBm



(b) Power level  $-7$  dBm

**Figure 6: Reactivity for different power levels with  $\alpha = 0.005$  in the DOOR scenario.**

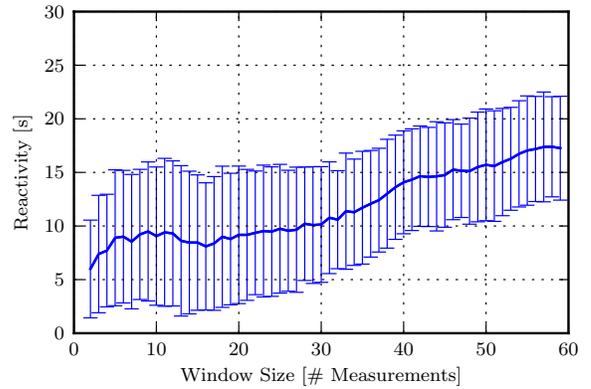
same. The most noticeable differences are only small contrary changes for contiguous short-term history sizes. This holds for all other power levels and results in the power level having marginal effect on the behavior of the approach for different short-term history sizes or significance levels  $\alpha$ .

Lastly, we compare the DOOR and WINDOW scenarios. In Figure 7a, the distributions for the window experiment with varying short-term history size, significance level  $\alpha$  configured to 0.005 and power level set to  $-15$  dBm is shown. For comparison, the plot for the door experiment with the same settings is shown in Figure 7b. Minor differences in the average and in the standard deviation can be identified for small short-term history sizes. Nonetheless, the overall trend is comparable. The same findings also apply to the significance level  $\alpha$ . Overall the behavior of the approach in response to the different events does not differ significantly.

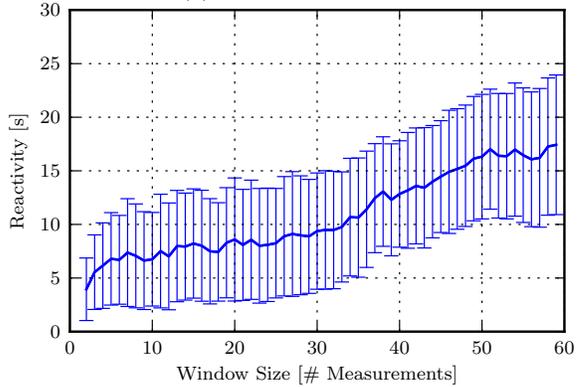
## 6. FINGERPRINTING COMMUNICATION

In the previous section, we analyzed the presented approach in relation to the events happening in the environment. This is the typical direction followed in the design of adaptive network protocols, in charge of exploiting and reacting to changes in the communication characteristics. For these specific mechanisms to be effective, it is mostly important to reduce the delay in detecting significant variations while minimizing the number of ineffective reconfigurations, as caused by false or marginal model change detections.

We, instead, envision a different possible use for the introduced approach. In particular, we consider a deployed network in a typical indoor environment. Such system is



(a) WINDOW scenario



(b) DOOR scenario

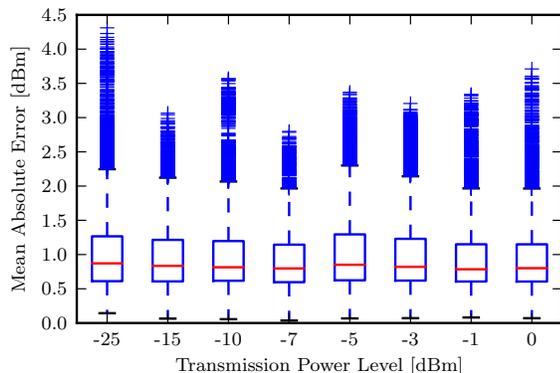
**Figure 7: Reactivity for different scenario with power level  $-15$  dBm and  $\alpha = 0.005$ .**

subject to different combination of the events with which we experienced in our study. Most importantly, these events may cause the communication characteristics to bring the system to a malfunctioning or failure. Little (if any) knowledge is currently available on communication during long periods of time in real scenarios where extensive and controlled measurements can not be taken. This is even more emphasized in the context of resource-constrained devices, which are foreseen to densely populate indoor environments.

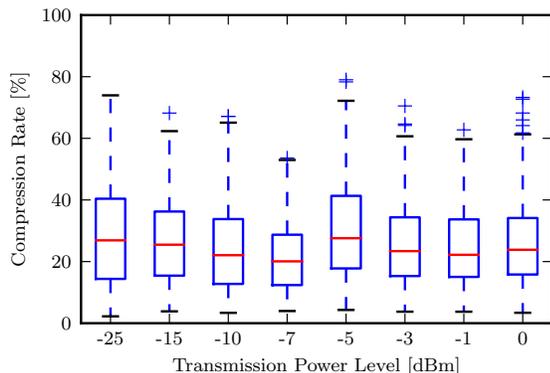
Therefore, we investigate the possibility to integrate the proposed approach in the normal functioning of the system. Each node could observe the incoming links and record the information about the detected communication models, instead of the raw RSSI observations. Upon failure or malfunctioning, or even during normal system operation, these fingerprints could be provided as information of the system evolution and possibly be related to root causes of faults.

To evaluate the effectiveness of this fingerprinting technique, we measure the accuracy of the sequence of models in closely describing the experienced link dynamics. To achieve this, we generate in simulation an artificial trace by drawing observations from the recorded models. Such artificial traces are then compared with the originally measured ones and the error computed. Figure 8 shows that the absolute error is for most links around 1 dBm. Considering that the measurement on real hardware do not have decimal precision, the replayed trace very closely follow the original one.

Finally, we compute the ability of the fingerprinting to effectively compress the raw measurement traces. Excluding timestamps, a single RSSI sample requires 8 bits; each com-



**Figure 8: Average absolute error between replayed simulated traces and original measurements. The simulation closely follows the raw RSSI traces.**



**Figure 9: Memory requirements for storing the link models compared to storing the raw RSSI values. Most of the signals can be compressed to fingerprints of less than one fourth of the initial size.**

munication change detection provides information about the new and the old model (average and variance) for a total of 128 bits. For the results collected in our experiments and the models generated directly by the devices, the fingerprints take in most of the cases less than 25% of the memory required for the raw observations, as shown in Figure 9. This allows to save 4 times more information about the evolution of communication in the running system.

## 7. CONCLUSION AND FUTURE WORK

We proposed an approach to fingerprint wireless communication online in an operational network for indoor scenarios. During normal human activities, periods of stability and variability of RSSI typically alternate. Our technique detects such state changes relying on statistical methods based on a long-term model of the link RSSI history and a short-term model of the immediate past RSSI measurements. A Welch’s t-Test then states whether the short-term and long-term link evolutions stem from the same distribution or the short-term model deviates from the long-term one. In the latter case, a communication model change is inferred.

We implement the technique for resource-constrained devices, demonstrating the feasibility of the approach. In controlled experiments, we analyze the behavior of the approach under different conditions and configurations. Starting from these results, we investigate the possibility to employ the approach for communication fingerprinting. The final out-

come supports the idea of using the introduced approach to compress the evolution of the incoming links directly on the devices once deployed in their real environment.

We foresee different directions as future work.

**Stability of models.** When a new model change is detected, the link might not be stable yet. By measuring the stability of a new model, the link description could be further refined. Higher layers could then avoid unstable links or tune the radio parameters only once stability is reached.

**Link correlation.** Links in an area of an environmental change may observe the same effect on the wireless channel. The correlation of different link models can then be used as an activation pattern for events. The activation pattern can then fingerprint network states instead of individual links.

**Multi-dimensional link modeling.** Communication can be measured through different metrics, e.g., LQI and SNR for short-term estimation, or PRR for long-term estimation. Those estimators could be combined in a multi-dimensional fingerprint, increasing the precision of the link description.

## Acknowledgments

This work has been co-funded by the German Research Foundation (DFG) in the Collaborative Research Center (SFB) 1053 “MAKI - Multi-Mechanism-Adaptation for the Future Internet” and the Alexander von Humboldt Foundation.

## 8. REFERENCES

- [1] Low-rate wireless personal area networks (lr-wpans), 2011.
- [2] Z. Abbas, J. Nasreddine, J. Riihijärvi, and P. Mähönen. Long-term indoor propagation models for radio resource management. In *WOWMOM*, pages 1–9. IEEE, 2012.
- [3] M. H. Alizai, O. Landsiedel, J. Á. Bitsch Link, S. Götz, and K. Wehrle. Bursty traffic over bursty links. In *SenSys*, pages 71–84. ACM, 2009.
- [4] J. Andersen, T. Rappaport, and S. Yoshida. Propagation measurements and models for wireless communications channels. *Communications Magazine, IEEE*, 33(1):42–49, 1995.
- [5] P. Bahl and V. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *INFOCOM*, pages 775–784. IEEE, 2000.
- [6] Chipcon Tech. CC2420 Datasheet. [focus.ti.com/docs/prod/folders/print/cc2420.html](http://focus.ti.com/docs/prod/folders/print/cc2420.html).
- [7] J. Hill, R. Szcwcyk, A. Woo, S. Hollar, D. Culler, and K. Pister. System architecture directions for networked sensors. In *ASPLOS*, pages 93–104. ACM, 2000.
- [8] T. Liu and A. Cerpa. TALENT: temporal adaptive link estimator with no training. In *SenSys*, pages 253–266. ACM, 2012.
- [9] J. Polastre, R. Szcwcyk, and D. Culler. Telos: enabling ultra-low power wireless research. In *IPSN*. IEEE, 2005.
- [10] U. Raza, A. Camerra, A. L. Murphy, T. Palpanas, and G. P. Picco. What does model-driven data acquisition really achieve in wireless sensor networks? In *PerCom*, pages 85–94. IEEE, 2012.
- [11] K. Srinivasan, P. Dutta, A. Tavakoli, and P. Levis. An empirical study of low-power wireless. *ACM Trans. Sen. Netw.*, 6(2):16:1–16:49, Mar. 2010.
- [12] B. L. Welch. The generalization of student’s problem when several different population variances are involved. *Biometrika*, 34(1/2):28–35, 1947.
- [13] M. Woehrle, M. Bor, and K. Langendoen. 868 MHz: A noiseless environment, but no free lunch for protocol design. In *INSS*, pages 1–8, 2012.