Classification of Packet Transmission Outcomes in Wireless Sensor Networks

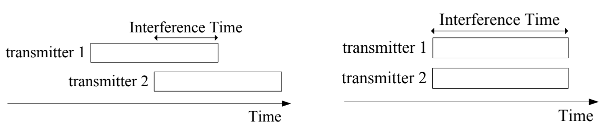
Zhi Ang Eu
NUS Graduate School for Integrative Sciences and Engineering
National University of Singapore
Email: euzhiang@nus.edu.sg

Pius Lee, Hwee-Pink Tan
Institute for Infocomm Research, Singapore
Agency for Science, Technology and Research (A*STAR)
Email: {wqplee,hptan}@i2r.a-star.edu.sg

***Abstract*—In wireless networks, it is important to determine the outcome of packet transmissions for networking protocols. In this paper, we design a transmission outcome classifier for IEEE 802.15.4 wireless networks based on received signal strength indicator and link quality indicator values. Our classifier performs loss differentiation by analyzing statistical differences between weak signal and collision losses. We implement our proposed classifier using the CC2500 RF transceiver and evaluate it experimentally. The results show that our classifier can accurately detect packet transmissions as well as distinguish wireless losses due to weak signals and multiple access collisions, with a maximum error rate of 15%. We apply the classifier to probabilistic polling, which is a MAC protocol designed for energy harvesting wireless sensor networks, and show experimentally that it is able to achieve close to or even exceed the theoretical throughput due to packet capture effect.**

I. Introduction

In wireless sensor networks (WSNs), IEEE 802.15.4 is the most commonly used physical layer protocol with many RF transceivers (e.g., CC2420, CC2500) based on this standard. Many contention-based MAC protocols have been proposed to coordinate access to the medium to achieve the performance required by the application. It is important to determine the outcome of packet transmissions, i.e., whether (i) no packet is received (**E1**); (ii) a packet is correctly received (**E2**); (iii) a packet is lost due to weak received signal (**E3**); or (iv) a packet is lost due to multiple access collision (**E4**) for the protocol to take appropriate action, e.g., by adapting the protocol parameters or adjusting transmission power. There are typically two types of packet collisions: (a) partially overlapping, arising from MAC protocols such as RI-MAC [1] and (b) fully overlapping, arising from MAC protocols such as probabilistic polling [2], as illustrated in Fig. 1.



(a) Partially overlapping collision (b) Fully overlapping collision

Fig. 1. Different types of collisons

There are several prior work that address loss differentiation in wireless networks. In [3], Received Signal Strength Indicator (RSSI) values and error patterns within a physical-layer symbol in IEEE 802.11 were proposed to diagnose wireless losses. However, these techniques are not directly applicable to IEEE 802.15.4 which uses different modulation schemes and hardware. In [4], a RTS/CTS and packet fragmentation mechanism is used in IEEE 802.11 to isolate the physical packet error rate. However, this requires additional transmission overheads which our classifier does not incur. In [5], it is shown that collision can be detected for partially overlapping collisions if the power levels of two concurrent transmissions differ significantly. In comparison, our classifier works for fully overlapping collisions without any restriction on the power levels of concurrent transmissions. In [6], RSSI values were used to determine if a collision occurred by observing changes in these values during packet transmissions; however, their method only works where the collisions do not fully overlap (see Fig. 1a). RSSI values and Link Quality Indicator (LQI) values have also been used to estimate link quality (e.g., [7]) but they are not used to differentiate wireless losses.

In this paper, based on extensive empirical measurements, we exploit both the LQI and RSSI values to design a novel classifier that can distinguish between different types of packet transmission outcomes in IEEE 802.15.4 networks. To the best of our knowledge, there is no prior work that can classify packet transmission events for fully overlapping collisions. We experimentally evaluate the efficacy of our proposed classifier, and further quantify its benefits by incorporating it into a probabilistic polling scheme [2], which was designed to achieve good performance in energy harvesting WSNs.

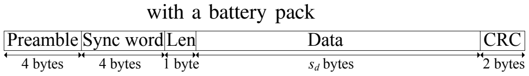
II. Design of Transmission Outcome Classifier

In this section, we describe our design of a transmission outcome classifier that can distinguish between the four types of events (**E1-4**) as defined in Section I. We consider the CC2500 low-cost 2.4 GHz RF transceiver [8] (Fig. 2a) which is highly suitable for low-power applications such as wireless sensor networks. The corresponding packet format comprises the preamble, sync word, data payload length, data payload and CRC, as illustrated in Fig. 2b.

The sync word plays a key role in our proposed classifier. Upon detection of the sync word, the measured LQI value, L_m, [8] gives an estimate of how easily a received signal can be demodulated by accumulating the magnitude of the error between ideal constellations and the received signal over the 64 symbols immediately following the sync word. It is a metric used to measure the quality of the received signal and ranges from 0 (high quality link) to 127 (low quality link), and is sometimes correlated with the RSSI: a strong signal (i.e., high RSSI value) is less likely to be affected by noise and therefore indicative of a high quality link (low LQI value). Since the sync word is typically much smaller than the whole packet, the likelihood of detecting it remains much higher under event **E3** (weak signal from a single transmitter) than **E4** (multiple access collision). Hence, detection of the sync word is a necessary condition for event **E2** and **E3** (single transmitter). Following this, we can infer that event **E2** occurred if the CRC is correct; otherwise, we propose to distinguish between packet losses due to weak signal and multiple access collisions by using both the RSSI and LQI values - this is described in Section II-B.

On the other hand, if the sync word cannot be detected, then the measured RSSI value, R_m, which indicates the strength of the received signal, can be used to distinguish between **E1** (no transmission) and **E4** (multiple transmissions). Assuming R_b to be the background noise, then it is reasonable to expect that $R_m \leq R_b$ for **E1** and $R_m > R_b$ for **E4**.

(a) CC2500 RF Transceiver with a battery pack



(b) Packet format

Fig. 2. CC2500 transceiver and packet format

A. LQI vs. RSSI for Weak Signals and Multiple Access Collisions

We begin by conducting experiments to obtain the RSSI-LQI characteristics under the conditions of (i) weak signal losses and (ii) multiple access collisions. Our experimental setup comprises one receiver and n_s transmitters as shown in Fig. 3. A control node is placed close to the transmitters to synchronize the nodes while the receiver is placed at a distance of d_{tr} from the transmitters. The control node will broadcast control packets at fixed intervals to the receiver and transmitters. After receiving the broadcast packet, every transmitter will send out a data packet of size s_d while the receiver will start detection of data packets.

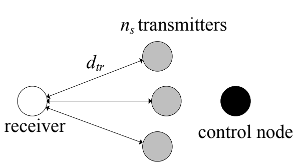
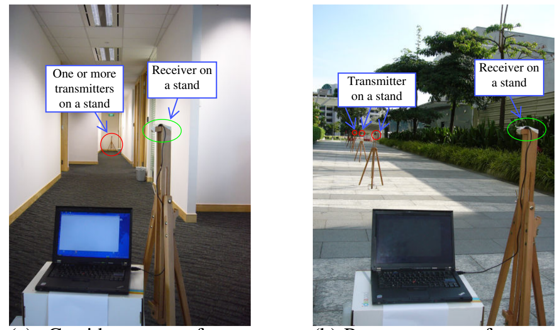


Fig. 3. Experiment setup to classify collision and weak signal losses

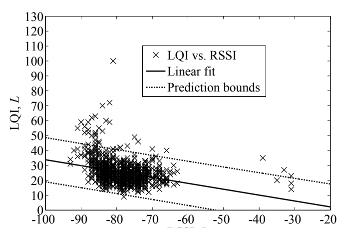
1) Weak signals (n_s = 1): In the first experiment, we have a single transmitter (i.e., all packet losses will be due to weak signal) and vary d_{tr} from 0m to 30m in steps of 5m. For each transmitter-receiver distance, the control node will send out control packets until 1,000 control packets are received by the receiver. The size of the control and data packet is 23 and 51 bytes respectively which are inclusive of the physical layer overheads. We carry out our experiments at a corridor (indoor) and a pavement (outdoor) as shown in Fig. 4. We collect all the incorrect packets as indicated by a wrong CRC value with a total of 1,012 and 1,182 packets for the indoor and outdoor environment respectively. The corresponding RSSI and LQI values are plotted in Figs. 5a and 5c.



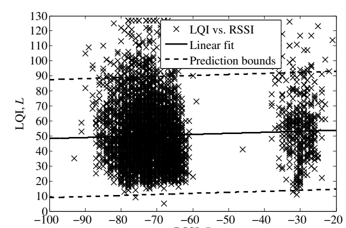
(a) Corridor setup for n_s transmitters with fixed d_{tr} (b) Pavement setup for transmitters with different d_{tr}

Fig. 4. Experimental setups in indoor and outdoor environments

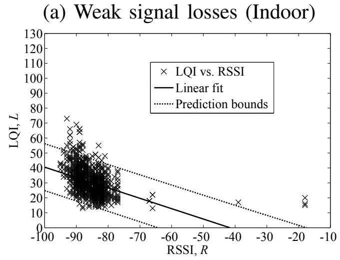
2) Multiple access collisions (n_s >1): In the next experiment, we have multiple transmitters where $n_s \in (2, 3, 4, 5)$, therefore all packet losses will be due to multiple access collisions. We vary both n_s and d_{tr}. A total of 28,000 control packets (7,000 for each n_s value with 1,000 for each d_{tr} value) are collected by the receiver. We collect all the incorrect packets with a total of 5377 and 3685 packets for the indoor and outdoor environment respectively. The corresponding RSSI and LQI values are shown in Figs. 5b and 5d.



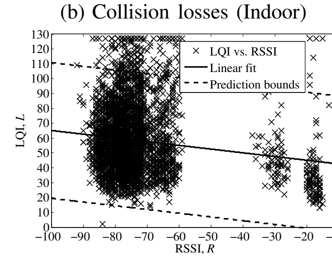
(a) Weak signal losses (Indoor)



(b) Collision losses (Indoor)



(c) Weak signal losses (Outdoor)



(d) Collision losses (Outdoor)

Fig. 5. RSSI versus LQI values for weak signal and collision losses in different environments

From Fig. 5, we observe that at a given RSSI value, the LQI value for collision losses are higher than that for weak signal losses, and use this as a basis to devise an algorithm to differentiate between these wireless losses.

B. Joint RSSI-LQI based packet loss classifier

For a given environment and known packet loss type y, where $y \in \{w, c\}$ corresponds to $w$eak signal (**E3**) and $c$ollision (**E4**) losses respectively, we first try to fit the data given in Fig. 5. Although various types of data fitting functions can be applied to (R_m,L_m), we adopted a linear fit, $\hat{L_0} = aR_0 + b$, as it is simple and only marginally less accurate than an exponential fit (e.g., the summed square of residuals, a goodness-of-fit statistic, are 57,121 and 55,336 respectively for the dataset in Fig. 5a). Accordingly, we compute the 95% prediction interval [9], (L_{yl},L_{yu}) that reflects the range of LQI values possible for a future RSSI reading as follows:

$$
\begin{aligned}
L_{yl} &= \hat{L_{y0}} - t_{0.025} s \sqrt{1 + \frac{1}{n_r} + \frac{(R_0 - \bar{R})^2}{\sum_{i=1}^{n_r} (R_i - \bar{R})^2}} \\
L_{yu} &= \hat{L_{y0}} + t_{0.025} s \sqrt{1 + \frac{1}{n_r} + \frac{(R_0 - \bar{R})^2}{\sum_{i=1}^{n_r} (R_i - \bar{R})^2}}, \quad (1)
\end{aligned}
$$

where $\bar{R}$ is the average RSSI reading, $\hat{L_{y0}}$ is the predicted LQI value for loss type y, n_r is the number of readings, $t_{0.025}$ is a value of the t-distribution with n_r degrees of freedom, and s is an unbiased estimate of the variance.

If (R_m,L_m) is the measured RSSI and LQI value of the data packet for which a sync word is received and the CRC is incorrect, our proposed packet loss classifier works as follows: (i) if $L_m \in (L_{cl}, L_{cu})$ and $L_m \notin (L_{wl}, L_{wu})$, then the packet loss is due to collision; (ii) if $L_m \in (L_{wl}, L_{wu})$ and $L_m \notin (L_{cl}, L_{cu})$, then the packet loss is due to weak signal; (iii) else, the event is classified as $y_p = \arg \min_{y=\{w,c\}} |L_m - \hat{L_{y0}}|$. Our proposed algorithm to distinguish between the events **E1** to **E4** is illustrated in Fig. 6.

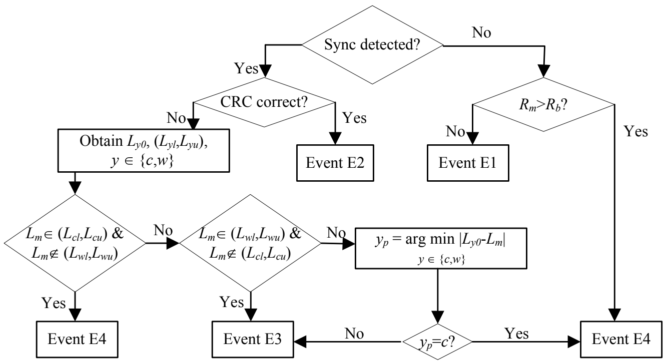


Fig. 6. Transmission Outcome Classifier

To reduce computation overheads, the prediction intervals (L_{yl},L_{yu}) are computed and preprogrammed into the nodes for different environments using a lookup table. The lookup table is small as the measured RSSI values are integers and range from -100 to -10.

III. Experimental evaluation of Transmission Outcome Classifier

We let p_r and p_w be the probabilities of right and wrong event classification respectively. We let n_t be the total number of packet events to be classified, and n_1 to n_4 be the number of events classified under events 1 to 4 (E1-4) respectively. Therefore, we have $p_r = n_r/n_t$ and $p_w = n_w/n_t$. Table I shows the values of n_r and n_w for different number of transmitters. Note that it is possible to receive correct data packets even if there are multiple transmitters due to the packet capture effect when the signal strength of one transmission is significantly higher than other transmissions.

TABLE I
Calculation of n_r and n_w

Number of transmitters (n_s)	n_r	n_w
0	n_1	$n_3 + n_4$
1	$n_2 + n_3$	$n_1 + n_4$
> 1	$n_2 + n_4$	$n_1 + n_3$

The 95% confidence intervals for the background noise are (-105.63,-96.13) and (-105.71,-96.2) for the indoor and outdoor environment respectively for 5,000 samples, therefore R_b is set to -96 for both environments. Two different datasets are used in the evaluation: the first dataset is the training dataset that is used to obtain the linear fit (i.e., values of a and b); the second dataset comprises readings from a new location (e.g., a different corridor or pavement) to verify that the classifier works at different locations. The prediction accuracy for the indoor environment is shown in Fig. 7 for different n_s values. The maximum error rate is 15.1%. As the number of transmitters increases, the quality of the packet received is decreased (i.e., higher LQI values), and hence it is easier to identify collision losses, leading to improved accuracy. For the single transmitter scenario, most misclassifications are a result of a loss of the sync word, therefore the classifier wrongly classifies packet losses as collision losses instead of weak signal losses. However, the misclassification probability is low.

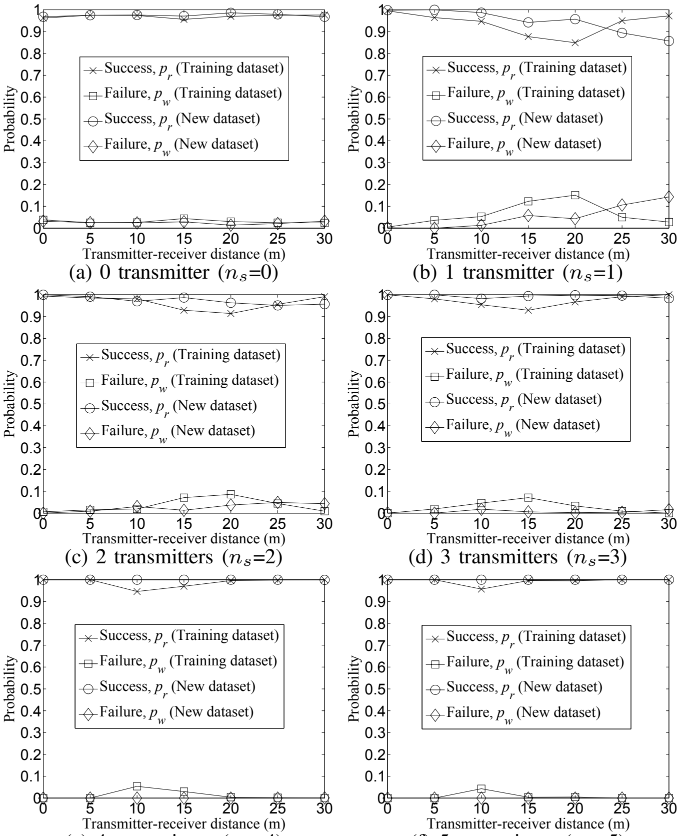
The prediction accuracy for the outdoor environment is shown in Fig. 8. The maximum error rate is 12.6% which is slightly lower than that for the indoor environment. For both environments, we observe that the accuracy obtained with the non-training dataset is comparable to that for the training dataset, therefore this validates that our proposed classifier works well for different locations.

IV. Application of Transmission Outcome Classifier to Probabilistic Polling

In this section, we apply the transmission outcome classifier to probabilistic polling and evaluate its efficacy for the scenario illustrated in Fig. 9a with one sink and n_s sensor nodes.

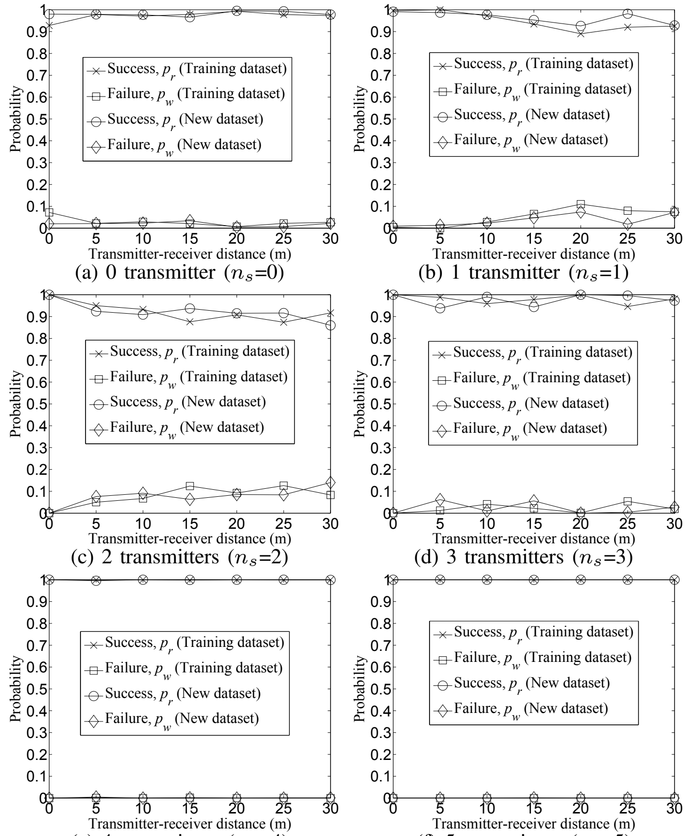
A. Probabilistic Polling

In probabilistic polling, the sink sets a contention probability, p_c in the polling packet to indicate the probability that a node should transmit its data packet. Upon receiving this packet, a node would generate a random number $x \in [0, 1]$, and transmit its data packet if $x < p_c$. Ideally, only *one* out of all the nodes that receive the polling packet should transmit a data packet. Collisions that may occur otherwise are fully overlapping, which reduces the time needed to recover from



(a) 0 transmitter (n_s=0) (b) 1 transmitter (n_s=1) (c) 2 transmitters (n_s=2) (d) 3 transmitters (n_s=3) (e) 4 transmitters (n_s=4) (f) 5 transmitters (n_s=5)

Fig. 7. Classification accuracy for different number of transmitters (indoor)



(a) 0 transmitter (n_s=0) (b) 1 transmitter (n_s=1) (c) 2 transmitters (n_s=2) (d) 3 transmitters (n_s=3) (e) 4 transmitters (n_s=4) (f) 5 transmitters (n_s=5)

Fig. 8. Classification accuracy for different number of transmitters (outdoor)

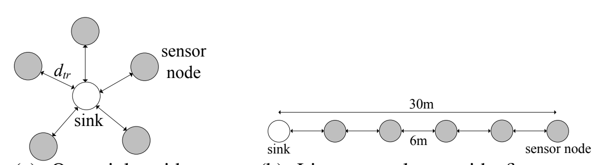
them. The value of p_c is reduced when there are collisions and increased when no transmissions are detected, as shown in Algorithm 1. We apply our proposed classifier to probabilistic polling to allow the sink to distinguish between different packet outcomes in order to achieve high throughput.

Algorithm 1 Probabilistic Polling

1: Send a polling packet with contention probability p_c.
2: **if** no node responds to the polling packet (E1) **then**
3: increase p_c
4: **else if** a data packet is successfully received from one of the nodes (E2) **then**
5: maintain p_c at current value
6: **else if** there is a packet loss due to a weak signal received from a single node (E3) **then**
7: maintain p_c at current value
8: **else if** there is a collision between two or more nodes as indicated by a corrupted data packet (E4) **then**
9: decrease p_c
10: **end if**
11: Repeat step 1.

B. Numerical Evaluation

Since maximizing the success probability of a poll (p_s) will also maximize throughput, it is used as the performance metric. The sink will send out 2,000 polling packets in each scenario and the experiment is repeated at five different locations for a total of 10,000 polling packets. After each polling packet is sent, the contention probability is adjusted based on the packet type that it received. In [2], we showed that the throughput of probabilistic polling is maximized by setting the contention probability p_c to $1/n_s$. However, since the sink has no advanced knowledge of the number of transmitters, it has to adjust p_c based on the event type after it sends out a polling packet. Although probabilistic polling is designed for energy harvesting WSNs, the nodes in our experiments are powered using batteries to remove the uncertainty in the energy harvesting process.



(a) One sink with n_s sensor nodes (b) Linear topology with five sensor nodes

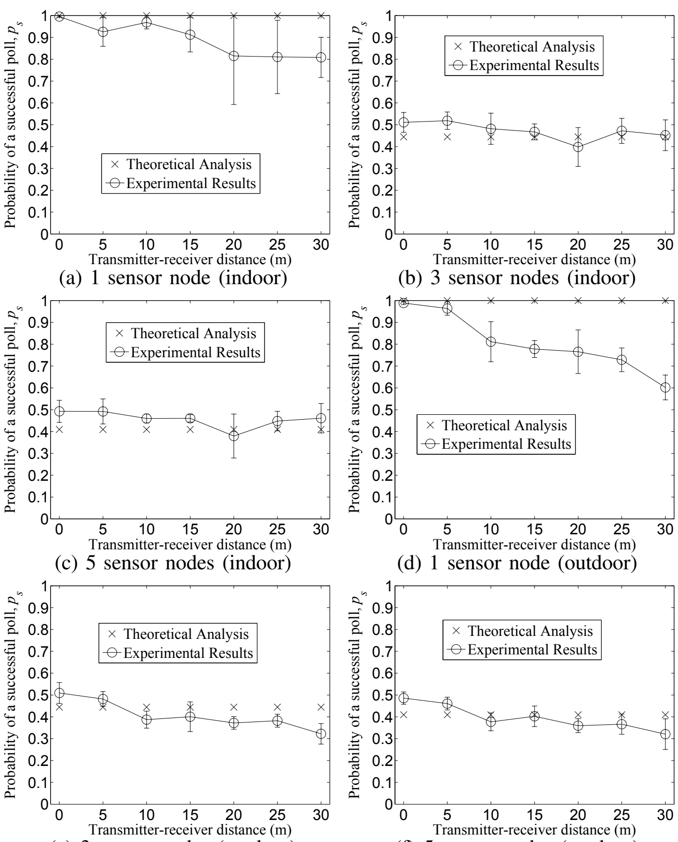
Fig. 9. Experiment setup for probabilistic polling

We let the estimated value of n_s by the sink be n_{est} ($p_c = 1/n_{est}$) with an initial value of 1. The value of n_{est} will be adjusted at the sink based on the event type after a polling packet is sent: (i) If it detects no transmission (**E1**), n_{est} is decreased by 1 subject to a minimum value of 1; (ii) if it receives a correct data packet (**E2**) or detects a weak signal packet loss (**E3**), n_{est} remains unchanged for the next polling packet; (iii) finally, if it detects a collision packet loss (**E4**), n_{est} is increased by 1. The average value of n_{est} will be close to n_s if our classifier is highly accurate. A poll will be successful when the sink receives a correct data packet from one of the nodes, and this occurs with probability

$$p_s = \binom{n_s}{1} p_c(1-p_c)^{(n_s-1)} = (1-\frac{1}{n_s})^{(n_s-1)},$$

since the optimal contention probability is $1/n_s$.

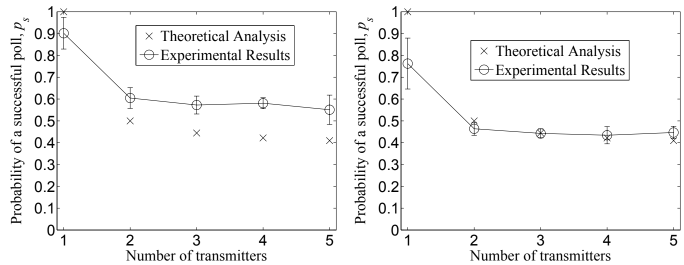
The experimental success probabilities with 95% confidence intervals obtained for indoor and outdoor scenarios are compared with the theoretical analysis in Fig. 10 with 1, 3 and 5 sensor nodes. With a single sensor node, p_s decreases with increasing transmitter-receiver distance even when the accuracy of our classifier is high due to wireless losses. For multiple nodes, the experimental p_s may be higher than the theoretical p_s due to the packet capture effect (not modeled in our theoretical analysis) which allows the sink to receive a correct packet from one node even when there are multiple nodes sending data packets concurrently.



(a) 1 sensor node (indoor) (b) 3 sensor nodes (indoor)
(c) 5 sensor nodes (indoor) (d) 1 sensor node (outdoor)
(e) 3 sensor nodes (outdoor) (f) 5 sensor nodes (outdoor)

Fig. 10. Success probabilities for indoor and outdoor environments

Next, we vary the transmitter-receiver distances using a linear topology with five sensor nodes as illustrated in Fig. 9b (pavement setup shown in Fig. 4b). The results are illustrated in Fig. 11 with varying number of transmitters which are randomly selected from the set of five sensor nodes. Unlike the previous case when all the nodes are equidistant from the sink, they are placed at different distances to the sink, giving rise to a more significant packet capture effect. This results in higher experimental success probability than the theoretical success probability in some scenarios.



(a) Indoor (b) Outdoor

Fig. 11. Success probabilities with different n_s for a linear topology

V. Conclusion

This paper presents a design for a transmission outcome classifier that is able to accurately distinguish among packet transmissions and losses in IEEE 802.15.4 networks. The classifier uses RSSI and LQI (link quality indicator) values to distinguish between collisions and weak signal losses. Unlike other approaches, our classifier works well for fully overlapping collisions and do not require modifications to the hardware. By incorporating the classifier into a probabilistic polling MAC protocol, we are able to achieve close to, or even exceed the theoretical throughput due to packet capture effect. Since our results are validated by actual experiments in both indoor and outdoor scenarios, the insights from our study would be valuable in the design of networking protocols such as MAC, transmit power adaptation and routing protocols.

References

[1] Y. Sun, O. Gurewitz, and D. B. Johnson, "RI-MAC: A Receiver-Initiated Asynchronous Duty Cycle MAC Protocol for Dynamic Traffic Loads in Wireless Sensor Networks," in *SenSys*, 2008, pp. 1–14.

[2] Z. A. Eu, H.-P. Tan, and W. K. G. Seah, "Design and Performance of MAC Schemes for Wireless Sensor Networks Powered by Ambient Energy Harvesting," *to appear in Ad Hoc Networks Journal*, 2010.

[3] S. Rayanchu, A. Mishra, D. Agrawal, S. Saha, and S. Banerjee, "Diagnosing Wireless Packet Losses in 802.11: Separating Collision from Weak Signal," in *IEEE INFOCOM*, 2008.

[4] M. A. Y. Khan and D. Veitch, "Isolating Physical PER for Smart Rate Selection," in *INFOCOM*, 2009, pp. 1080–1088.

[5] K. Whitehouse, A. Woo, F. Jiang, J. Polastre, and D. Culler, "Exploiting The Capture Effect For Collision Detection And Recovery," in *EmNets*, 2005, pp. 45–52.

[6] J.-H. Hauer, A. Willig, and A. Wolisz, "Mitigating the Effects of RF Interference through RSSI-Based Error Recovery," in *EWSN*, 2010, pp. 224–239.

[7] K. Srinivasan, P. Dutta, A. Tavakoli, and P. Levis, "An Empirical Study of Low-Power Wireless," *ACM Transactions on Sensor Networks*, vol. 6, no. 2, Feb. 2010.

[8] CC2500 Datasheet, "Low-Cost Low-Power 2.4 GHz RF Transceiver." [Online]. Available: http://focus.ti.com/docs/prod/folders/print/cc2500.html

[9] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability & Statistics for Engineers & Scientists, Seventh Edition*. Prentice Hall, 2002, pp. 369-371.