An efficient learning technique to predict link quality in WSN

Dana Marinca, Pascale Minet, Nesrine Ben Hassine

HAL Id: hal-01094472

https://inria.hal.science/hal-01094472

Submitted on 23 Feb 2015

An Efficient Learning Technique
to Predict Link Quality in WSN

D. Marinca[1], P. Minet[1] and N. Ben Hassine[1,2]

(1)Inria, Rocquencourt, 78153 Le Chesnay, France

(2)PRiSM, UVSQ, 78000 Versailles, France

dana.marinca@inria.fr, pascale.minet@inria.fr, nesrine.ben-hassine@inria.fr

Abstract— **In this paper, we apply learning techniques to predict link quality evolution in a WSN and take advantage of wireless links with the best possible quality to improve the packet delivery rate. We model this problem as a forecaster prediction game based on the advice of several experts. The forecaster learns on-line how to adjust its prediction to better fit the environment metric values. Simulations using traces collected in a real WSN show the improvement of the prediction when the experts use the SES prediction strategy, whereas the forecaster uses the EWA learning strategy.**

Index Terms— **Machine learning, prediction, forecaster, expert, cumulated loss, wireless sensor networks, link quality.**

I. INTRODUCTION

With the tremendous progress in wireless technologies and the increasing miniaturization, we are witnessing a large deployment of wireless sensor networks (WSNs) in various domains: environmental, home automation, human or structural health monitoring, industrial process control, assisted maintenance, etc. The large majority of these applications gather data originated from sensors in charge of sampling their environment.

To ensure the accuracy of the monitoring, the data gathering process must meet delay, throughput and robustness requirements. Since multichannel communications increase parallelism of transmissions and offer a better protection against interferences, they are an attractive solution for data gathering. However, wireless channels with external interferences as well as links of poor quality must be avoided to meet these requirements. The challenge comes from the high versatility of the wireless medium, the signaling cost of switching between channels and the limited resources of sensor nodes.

Observing the behavior of a wireless link it is possible to infer predictions about the future link quality and dynamically adjust radio parameters to improve the communication in an energy-efficient manner. Hence, each instantaneous variation of link quality must not be followed by an immediate channel switching.

For that purpose, in this paper we are interested in predicting the quality of wireless links in WSNs. This allows the WSN to anticipate link degradation and adequately react by switching on an alternative link / channel of better quality. More precisely, we focus on prediction with expert advice [9], where several agents, called experts, try to predict values close to the outcomes of the environment. A special expert, called forecaster, takes into account the experts predictions to compute its own prediction. Its goal is to compute values as close as possible to the real environment outcomes values. The

simulations show that the forecaster prediction is accurate because it adapts with agility to better fit the environment evolution. Due to significant quality fluctuations of wireless links, the importance of a metric value decreases with its age and we investigate in this paper the impact of this aspect on the accuracy of the prediction technique.

II. RELATED WORK

Learning techniques have been adopted to improve the adaptivity of systems to dynamic changes of their environment or modifications in application requirements. In the literature, several metrics [1] have been used to estimate the quality of a WSN link (Signal over Noise Ratio (SNR), Received Signal Strength Indication (RSSI), Link Quality Indicator (LQI), Packet Success Rate (PSR), etc.). Different quality prediction techniques have been investigated such as pattern matching, regression, filtering and smoothing techniques, as we can see further.

Farkas et al. [2] proposed an approach based on pattern matching to predict the variation of the link quality based on SNR metric. Thus, the forecasters try to detect patterns similar to the current situation in the history of SNR values. In [3], the authors use the regression technique that can be applied to any measurable WSN metric. This technique consists in modeling the relationship between an input variable and its response, such that the future response variable can be estimated by a regression model applied to the corresponding input variable. In order to estimate the PSR values in [4], BLITZ (Preamble-Based Link Quality Estimation) operates at two time-scales. At the preamble level, chip errors of received symbols are first averaged and fitted to a polynomial model to obtain an estimation of the instantaneous PSR. At the packet level, chip error statistics from multiple transmitted packets are filtered according to a Weighted Moving Average (WMA) function to smooth short-term fluctuations of the estimation method. The authors of [6] have defined F-LQE to estimate link quality by combining four different metrics (PSR, link asymmetry, link stability and SNR) into a single quality indicator. The quality of the link is specified as a fuzzy rule whose evaluation returns the membership of the link in the fuzzy subset of good links. Liu et al. [5] combine physical parameters (RSSI, SNR, and LQI) and logical parameter (PSR). They use three prediction models namely Naive Bayes, Logistic Regression Classifier and Artificial Neural Networks to predict the link quality. Based on PSR, Holistic Packet Statistics [7] estimates the short and long-term quality of link, the variation and the trend of the link quality. PSR is filtered using a WMA function to smooth fluctuations of the estimated method. Boano et al.

propose in [10] a geometrical combination of PSR, LQI, and SNR into a robust estimator of link quality, called triangle metric. They showed that this metric provides fast assessment and reliable estimation based on a reduced number of data packets.

In a previous paper [8], we proposed a prediction model based on several experts, each expert following its own strategy to predict the future values of a given metric. The forecaster, a special expert, follows an Exponential Weighted Average (EWA) strategy to compute the predicted values based on experts prediction. The model has the advantage of being very flexible, able to support various prediction methods and can be extended to another link quality metric as well as a combination of several metrics. Simulation results from [8] shows that this model also presents a high reactivity to dynamic changes of the environment. In this paper, we adopt this model and propose new prediction methods for both the experts and the forecaster. We have selected the LQI metric because of its reliability in link quality estimation [1] and its availability: the value can be obtained at physical layer from each received packet. The forecaster and experts performance is evaluated by a loss function, measuring the discrepancy between the prediction and the real value.

This paper is organized as follows. Section III presents the problem model with new prediction methods (SES and AMW) and forecasting strategies (EWA and BE). Theoretical results are proved in Section IV for different experts and forecasters. In Section V we provide a comparative evaluation of EWA and BE forecasters, as well as of different expert prediction methods. We finally conclude in Section V.

III. PROBLEM MODELING

A. Theoretical Model

We model the link quality prediction problem in the theoretical framework of a prediction game [9]. A prediction game is played in successive trials and the goal at the beginning of each trial t is to predict the outcome value y_t which will be given by the monitored environment at the end of the trial. We will define further the prediction and the learning strategies.

1) Prediction game framework

Denote $\mathcal{E} = \{E_i, i = 1..N\}$ the set of N experts The prediction game is defined as a 4-tuple (Y, D, T_{set}, l), where $Y = \{y_t\}$ is the set of possible outcomes, $D = \{\hat{p}_t\}$ the set of possible predictions, T_{set} the set of trials and $l(\hat{p}_t, y_t)$ is the loss function that evaluates the accuracy of the prediction $\hat{p}_t$ with regard to the real outcome y_t. At the beginning of the trial t, using its own prediction strategy PS_i, each expert E_i computes and predicts a value $p_{i,t}$ to anticipate the outcome y_t. The cumulative loss $L_{i,t}$ for each E_i evaluates the prediction accuracy from the beginning up to trial t as follows: $L_{i,t} = \sum_{j=1}^{t} l(p_{i,j}, y_j)$. Knowing the values predicted by the experts, the forecaster F computes its predicted value $\hat{p}_t$ applying its learning strategy LS_F. The accuracy of the forecaster predictions is evaluated by the cumulative loss $L_{F,t} = \sum_{j=1}^{t} l(\hat{p}_j, y_j)$. Table I summarizes the prediction game model,

where the smallest possible outcome is denoted l_{min} and the largest l_{max}.

TABLE I
FRAMEWORK FOR PREDICTION GAME

Prediction Game	
$T_{set} = \{1,2,3,\dots,t,\dots\}$	set of trials
$Y = \{y_1, y_2, \dots, y_t, \dots\} \subseteq [l_{min}, l_{max}]$	set of environment outcomes
$D = \{\hat{p}_1, \hat{p}_2, \dots, \hat{p}_t, \dots\} \subseteq [l_{min}, l_{max}]$	set of possible predictions
$l: D \times Y \rightarrow R^+$	loss function
Players – The experts	
$\mathcal{E} = \{E_i, i = 1..N\}$	set of experts
$\{p_{i,t}\} \subseteq [l_{min}, l_{max}], for\ t \in T_{set}$	predictions set of E_i
$PS_i \in \{SES, AMW\}$	prediction strategy of E_i
$w_{i,t} \in [0,1]$	weight of E_i during the trial t
$L_{i,t} = \sum_{j=1}^{t} l(p_{i,j}, y_j)$	cumulated loss of E_i
Player – The forecaster	
$F = (\mathcal{E}, D, l, LS_F)$	forecaster definition
$LS_F \in \{EWA, BE\}$	forecaster learning strategy
$L_{F,t} = \sum_{j=1}^{t} l(\hat{p}_j, y_j)$	cumulated loss of F

Several loss functions were analyzed in the literature [11][9]. We select the square loss function, l_{sq}, because of its simplicity[9]. It is defined by the equation:

$$l_{sq}(a, b) = (a - b)^2 \qquad (1)$$

2) Expert prediction strategies

An expert E_i uses a prediction strategy PS_i to compute the predicted value $p_{i,t}$. We propose hereafter two prediction strategies. We first consider the Average on Moving Window (AMW) strategy. AMW predicts the average of equally weighted outcomes observed in a given past time observation window. More precisely, at the beginning of the trial t, the expert E_i predicts a value from the outcomes observed in $[t - \Delta t_i, t)$, where Δt_i is the size of the *observation window* analyzed by E_i.

AMW: $\qquad p_{i,t} = \sum_{j \in [t - \Delta t_i, t[} \frac{y_j}{Card_{i,t}} \qquad (2)$

where $Card_{i,t}$ denotes the number of outcomes collected in the interval $[t - \Delta t_i, t)$. This prediction is valid until the beginning of the next trial, $t + 1$. AMW has the advantage of supporting a varying number of outcomes in each *observation window*. This property is required for WSNs where the number of messages exchanged over a wireless link may vary over the time. In addition, AMW has no initialization problem and requires keeping only the outcomes of the last window. Practically, this means that each outcome must be time stamped to keep only the relevant ones in the *observation window* moving with time.

The second prediction strategy is the Single Exponential Smoothing (SES). SES predicts a value computed from a weighting of the past outcomes over a given *observation window*. The weight of an outcome decreases exponentially with its age. The SES expert E_i predicts at the beginning of the trial t a value computed according to Eq. (3), for outcome values collected in the window $[t - \Delta t_i, t)$:

SES: $\qquad p_{i,t-\Delta t_i} = y_{t-\Delta t_i}$

$\qquad$ for $k = t - \Delta t_i + 1\ \ to\ \ t$

$\qquad p_{i,k} = \alpha * y_{k-1} + (1 - \alpha) * p_{i,k-1} \qquad (3)$

where $0 < \alpha < 1$ is the smoothing factor. Values of α close to 1 give a greater weight to recent outcomes, whereas values close to 0 have a greater smoothing effect and are less reactive to incoming changes. We will see in Section V how to determine the appropriate values of α and the most appropriate sizes of observation windows. Notice that the size of the observation window determines the storage requirement for the expert.

3) Forecaster learning strategies

The forecaster F will compute its predicted values based on the advice of N experts E_i. We will consider two learning strategies for the forecaster F: Exponentially Weighted Average (EWA) and Best Expert (BE).

The EWA forecaster [9] predicts the value $\hat{p}_t$, computed as a weighted average of the experts predictions $p_{i,t}$, as expressed in Eq. (4).

$$\hat{p}_t = \sum_{i=1}^{N} w_{i,t-1} p_{i,t} \tag{4}$$

Where $w_{i,t-1}$ represents the weight of the expert E_i and reflects the accuracy of its prediction. Initially, all the N experts have the same weight $w_{i,0} = 1/N$. At the end of each trial t, the weight $w_{i,t}$ of each expert is updated according to Eq. (5), where the constant $\eta > 0$ impacts the learning speed.

$$w_{i,t} = \frac{w_{i,t-1} e^{-\eta l(p_{i,t}, y_t)}}{\sum_{j=1}^{N} w_{j,t-1} e^{-\eta l(p_{j,t}, y_t)}} \tag{5}$$

The prediction of the EWA forecaster is equivalent to Eq. (6) (see [9]):

EWA: $$\hat{p}_t = \frac{\sum_{i=1}^{N} e^{-\eta L_{i,t-1}} p_{i,t}}{\sum_{i=1}^{N} e^{-\eta L_{i,t-1}}} \tag{6}$$

We now focus on the alternative forecaster strategy called the Best Expert (BE). The best expert is defined as the expert minimizing the cumulated loss until the current trial. The BE forecaster predicts the value given by the best expert. Let E_i be the best expert at trial $t-1$. Thus, the BE forecaster predicts at trial t the same value as E_i:

BE: $$\hat{p}_t = p_{i,t} \tag{7}$$

with $L_{i,t} = min_{E_j \in \varepsilon} L_{j,t-1}$ and $\hat{p}_1 = min_i (p_{i,1})$

Notice that unlike the EWA forecaster, the BE forecaster is unable to predict a value that has not been predicted by one of its experts.

4) Evaluation of forecaster learning strategies

In the literature, several definitions are possible to identify the best forecaster. One definition focuses on the minimization of the external regret, such as $L_{F,t} - min_{E_i \in \varepsilon} L_{i,t}$. The external regret of the EWA forecaster has been proved to be bounded and different bounds can be found in [9][11]. Intuitively, this means that the forecaster predictions are not far from those of the best expert. Other definitions are possible such as: the best forecaster is this minimizing the cumulated loss. That is the definition we adopt in this paper.

B. Problem Statement

In the framework defined in Section III.A we define experts and forecasters able to predict on-line the link quality based on the history of Link Quality Indicator (LQI) metric captured in a real WSN. The prediction of LQI is needed in order to be able to replace a link whose quality is strongly decreasing before it becomes unacceptable. This will improve the packet delivery rate that is perceived by the application. We will use experts predicting according to AMW and SES. The forecasters learning strategies are EWA and BE.

IV. THEORETICAL RESULTS

We now study the prediction accuracy of a forecaster depending on its experts. We distinguish several cases.

A. A Perfect Expert

A perfect expert is an expert E_k that for any trial t, always predicts the exact value. Notice that this expert has a null cumulated loss $L_{k,t} = 0$ for any trial t. It is obvious that the best forecaster is this following the advice of expert E_k.

Property P_1: Assuming the existence of a perfect expert, E_k, there is no difference, after a finite number of trials, between the predictions given by any forecaster F that minimizes the external regret and the prediction given by the BE forecaster. In addition, these predictions are exact.

Proof: The perfect expert E_k meets $L_{k,t} = min_{E_i \in \varepsilon} L_{i,t} = 0$. To minimize the external regret of F means to minimize $L_{F,t} - min_{E_i \in \varepsilon} L_{i,t} = L_{F,t}$. Both forecasters aim at minimizing their cumulated loss. They achieve it by following the perfect expert. Hence the property. ∎

Property P_2: Assuming the existence of a perfect expert E_k, the number of trials after which the BE forecaster exactly knows the perfect expert is upper bounded by s, where s is the number of contiguous trials in which a non-perfect expert gives at least one erroneous advice. This BE forecaster will predict for any trial $t>s$, the exact value and will meet $L_{F,t} = L_{F,s}$.

Proof: After each trial t, the BE forecaster definitely eliminates all the experts that have given an erroneous advice. Since after s trials, each non-perfect expert has given at least one erroneous advice, it has been eliminated. The only remaining expert is the perfect one. Hence, the BE forecaster predicts for any trial $t > s$ the exact value and its cumulated loss remains equal to $L_{F,s}$. Hence the property. ∎

B. A Dominating Expert

Let us assume now that there is no perfect expert, but there is an expert E_k that dominates any other expert. This expert meets for any trial t, $L_{k,t} = min_{E_i \in \varepsilon} L_{i,t}$.

Property P_3: Assuming the existence of a dominating expert E_k, the number of trials after which the BE forecaster exactly knows the dominating expert is upper bounded by s, where s is the maximum number of contiguous trials in which a dominated expert may have a cumulated loss equal to $L_{k,t} = min_{E_i \in \varepsilon} L_{i,t}$, for any trial $t \le s$. For any trial $t > s$, the BE forecaster provides the same prediction as the dominating expert E_k. In addition, the cumulated loss of F is equal to that of the dominating expert plus a constant: $L_{F,t} = L_{k,t} + (L_{F,s} -$

$L_{k,s}$).

Proof: After s trials, each dominated expert has a cumulated loss strictly higher than $min_{E_i \in \varepsilon} L_{i,s}$, it is definitely eliminated by the forecaster. At any trial $t > s$, the only remaining expert is the dominating one. Since after trial s, the BE forecaster behaves exactly as the dominating expert, its cumulated loss is equal to this of the dominating expert plus the difference between the two cumulated losses at trial s. This BE forecaster has a cumulated loss that meets for any trial $t \geq s$, $L_{F,t} = L_{k,t} + (L_{F,s} - L_{k,s})$. Hence the property. ∎

C. A Varying Dominating Expert

In the real world, experts can frequently provide bad predictions. However, it is reasonable to assume that at any trial t, the expert that minimizes its cumulated loss up to trial $t - 1$ predicts a value not far from the outcome. Notice that this expert can change at each trial. As a consequence, the BE forecaster will follow at each trial t the expert minimizing its cumulated loss until the trial $t - 1$. Let $E_{f(j-1)}$ be the expert minimizing its cumulated loss at trial $j-1$, with $j > 2$.

Property P_4: The cumulated loss of the BE forecaster selecting at each trial $j = 2..t$, the expert $E_{f(j-1)}$, is equal to $L_{F,t} = \sum_{j=2}^{t} (p_{f(j-1),j} - y_j)^2 + (\hat{p}_1 - y_1)^2$.

Proof: By definition, we have $L_{F,t} = L_{F,t-1} + (p_{f(t-1),t} - y_t)^2$. By writing this formula at each trial $j = 2..t$, we get the property. ∎

The following property P_5 proves that the BE forecaster can predict with the same accuracy as the best expert at time t, provided that some initial condition is met by all experts.

Property P_5: If for the first trial, all experts predict the same value that is either l_{min} or l_{max}, the BE forecaster selecting at each trial t the expert E_k with $L_{k,t-1} = min_i L_{i,t-1}$, has a cumulated loss equal to that of E_k: $L_{F,t} = L_{k,t}$.

Proof: By definition, we have

$$L_{F,t} = L_{F,t-1} + (p_{k,t} - y_t)^2$$
$$L_{k,t} = L_{k,t-1} + (p_{k,t} - y_t)^2,$$

We then get:

$$L_{F,t} - L_{k,t} = L_{F,t-1} - L_{k,t-1}.$$

By writing this formula at each trial $j = t..2$, we get:

$$L_{F,t} - L_{k,t} = L_{F,1} - L_{k,1},$$
$$L_{F,1} - L_{k,1} = (\hat{p}_1 - y_1)^2 - (p_{k,1} - y_1)^2$$
$$= (\hat{p}_1 - p_{k,1})(\hat{p}_1 - 2y_1 + p_{k,1})$$

To make this quantity negative or null, we must have:
- First case: $\hat{p}_1 - p_{k,1} \leq 0$ and $\hat{p}_1 - 2y_1 + p_{k,1} \geq 0$.

To meet this condition, we can take $\hat{p}_1 = p_{i,1} = l_{max}$ for any $i=1..N$, since y_1 can take any value in $[l_{min}, l_{max}]$.
- Second case: $\hat{p}_1 - p_{k,1} \geq 0$ and $\hat{p}_1 - 2y_1 + p_{k,1} \leq 0$.

To meet this condition, we can take $\hat{p}_1 = p_{i,1} = l_{min}$ for any $i=1..N$, since y_1 can take any value in $[l_{min}, l_{max}]$.

In both cases, we have $L_{F,1} - L_{k,1} = 0$.

Finally, we get $L_{F,t} = L_{k,t}$. ∎

Since in real wireless sensor networks, it is difficult to propose a perfect expert or even a dominating one, the EWA

and the BE forecaster will select one expert minimizing the cumulated loss computed on the whole past.

V. PERFORMANCE EVALUATION

We now evaluate the performances of the strategies proposed for the experts and the forecasters. For this purpose, we developed a simulator in Python 2.7 and used real traces generated by a real WSN.

A. Simulation parameters

All simulation runs are based on real LQI values collected from a real WSN based on the IEEE 802.15.4 standard. This standard provides 15 channels in the 2.4GHz spectrum. We focus on channel 11 and on the link between two sensor nodes identified by id=110 and id=139. Many simulations were conducted on large collections of LQI values, all showing a very good reactivity and accuracy of the forecaster prediction.

Without loss of generality, we present hereafter simulations results on a short time LQI collection, highlighting important changes in link quality. The LQI values presented hereafter have been collected during data transmissions for 1500 trials and a total duration of 1mn53s. As depicted in Figure 1, we can distinguish three phases. Phases P_1 and P_3 provide good values of LQI, whereas phase P_2 provides bad and unstable values. In [8] we showed that important changes in the profile of LQI values (quality's increase or decrease) are detected very quickly by the forecaster. The beginning of a new phase can be identified by a significant change in the slope of the cumulated loss curve, as illustrated in Figure 9. In the following we study the three different phases separately.
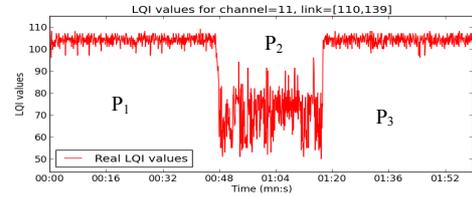

Figure 1. The LQI values provided by real traces

B. Prediction strategies of experts

1) Optimal parameters for SES experts

Before using an expert strategy, it is required to compute the value of its parameters. For a SES expert, there are two parameters to define: α the smoothing factor and Δt_i, the observation window size. We propose to select the parameters values minimizing the cumulated loss of the considered experts during each phase.

Simulation results have shown that for $0 < \alpha \leq 0.5$ and large window size $\Delta t_i \geq 550ms$ the forecaster predictions undergo excessive smoothing effect and are far from the real LQI values. The experts cumulated loss is minimized for values of $\alpha \in]0.5, 1[$ and observation window size $\Delta t_i < 550ms$. We consider hereafter several discrete values of α and Δt_i, more exactly $\alpha = 0.5 + k * 0.05$ for $k \in [0..9]$ and $\Delta t_i = i * d$, where $i = 1..6$ and $d = 80ms$. We have defined 60 SES experts, one expert for each couple $(\alpha, \Delta t_i)$.

We analyzed the cumulated loss at the end of each phase P_1, P_2 and P_3 and present theses values in Figures 2, 3 and 4

respectively. We can observe in all cases that the lowest cumulated loss is determined by values of α close to 1, more exactly: $0.75 \leq \alpha \leq 0.95$ for the phases P_1 and P_3, and $0.85 \leq \alpha \leq 0.95$ for the phase P_2. Notice that the cumulated loss is higher during the phase P_2 because of the higher variations of LQI values. During P_1, the experts having the best performances for the specified values of α are defined on windows $\Delta t_1, \Delta t_2, \Delta t_5, \Delta t_6$ as depicted in Figure 5. The most efficient experts during P_3, are defined on $\Delta t_1, \Delta t_2, \Delta t_3, \Delta t_6$, and the performance of $\Delta t_4, \Delta t_5$ is slightly lower (see Figure 4). During P_2 the less efficient experts are defined on $\Delta t_4, \Delta t_5$, as we can observe in Figure 3. Notice that E_1 is the best expert in phases P_2 and P_3, whereas E_2 is the best expert in P_1, as depicted in Figure 5.
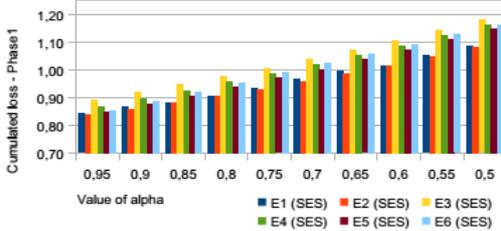

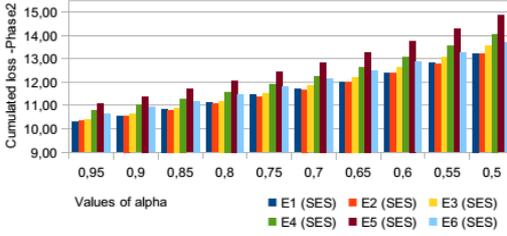Figure 2. Cumulated loss for SES experts during P_1
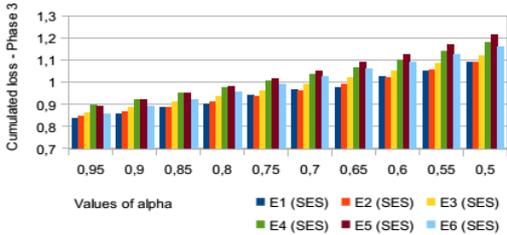

Figure 3. Cumulated loss for SES experts during P_2

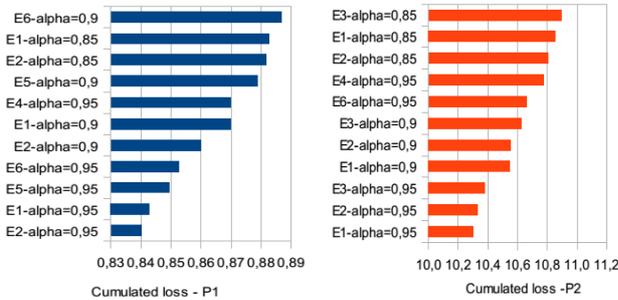
Figure 4. Cumulated loss for SES experts during P_3

Figure 5. The best SES experts during P_1 et P_2

2) Optimal parameters for AMW experts

We defined 8 AMW experts differing by the size of their observation window. Simulations have shown that large window sizes $\Delta t_i \geq 700\ ms$ are not efficient because the smoothing effect is too important and the predicted values are too far from the real LQI values. For this reason we considered hereafter window sizes defined as: $\Delta t_i = i * d$, where $i = 1..8$, $d = 80ms$.
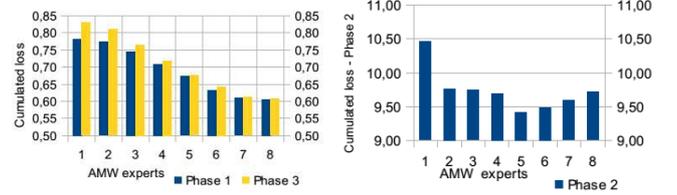

Figure 6. The cumulated loss for AMW experts during P_1-P_3 and P_2

The AMW experts performances are similar for the phases P_1 and P_3 (see Figure 6. left). The cumulated loss decreases while the expert window increases. We notice that windows Δt_7 and Δt_8 have the same cumulated loss. The cumulated loss increases again when the window size increases beyond Δt_8. But we do not present this result for lack of space. The most accurate experts are those defined on observation windows $\Delta t_5, \Delta t_6, \Delta t_7$ and Δt_8 .

During P_2, corresponding to unstable LQI values (see Figure 6. right), the AMW expert using the smallest window Δt_1 provides inaccurate advice. The best expert is the one using Δt_5 that is the optimal window. To compute its prediction, this expert uses a sufficient number of past outcomes, but not too high in order to be sensitive to the most recent outcomes. Experts using Δt_6 and Δt_7 are also acceptable. Notice that E_7 and E_8 are the best experts in phases P_1 and P_3, whereas E_5 is the best expert in P_2.

3) Comparison between SES and AMW experts

We compare AMW and SES experts in terms of cumulated loss, each expert using its optimal parameter values. For the SES experts, in all three phases, the best windows are the smallest ones, Δt_1 and Δt_2, for $0.8 \leq \alpha \leq 0.95$ whereas the AMW experts obtain better results with large windows like $\Delta t_5, \Delta t_6$ and Δt_7.

For both types of experts, the best expert among the experts of the same type depends on the phase considered. In addition, during P_1, P_2 or P_3, the maximum cumulated loss of any AMW expert is slightly lower than of SES experts. Simulation results show that all acceptable AMW experts slightly outperform the SES experts.

With regard to complexity, a SES expert and an AMW expert require to keep at each trial all the outcomes in their window. In addition, a SES expert needs to keep its previous prediction. Both have similar memory requirement, whereas the complexity of a SES expert is a little bit higher.

C. Prediction strategies of forecasters

The forecaster will compute its predictions based on the experts predictions. We evaluate the performances of the EWA and BE forecasters in the purpose of selecting the

forecaster minimizing the cumulated loss. For EWA forecasters, the value of η is set to 15.
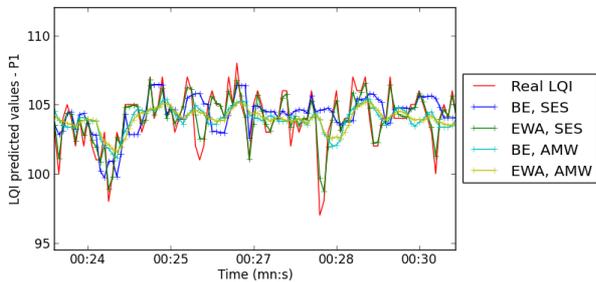

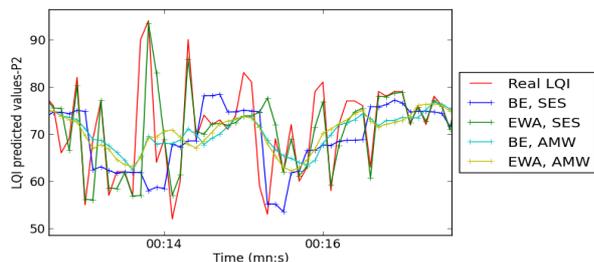Figure 7. The prediction of different forecasters during P$_1$

Figure 8. The prediction of different forecasters during P$_2$

Figures 7 and 8 show that in stable phases as well as in unstable ones, EWA-SES is the combination the closest to reality. However, it can introduce a possible lag in its prediction (see for instance around time 00:15 in Figure 8).

BE-SES can overestimate the link quality (see for instance around time 00:26 and 00:30 in Figure 7), which is not suitable in wireless sensor networks because this can generate packet loss. It can also predict a trough whereas the real value corresponds to a peak, but this is less disturbing.

EWA-AMW and BE-AMW have a similar behavior, they give a prediction that is too smoothed in comparison with reality. The drawback of the BE forecaster is that it can have poor performance when in a repetitive manner the best expert chosen at time $t - 1$ gives a bad prediction at time t.

Figure 9 depicts the cumulated loss incurred by different forecasters from the beginning of the simulation. We can draw the following conclusions. Based on the type of forecaster, the experts minimizing the forecaster cumulated loss are different. For EWA forecaster, the EWA-SES combination minimizes the cumulated loss and hence it is the combination we recommend for the forecaster and the experts respectively. If the forecaster is BE, the best experts are AMW.
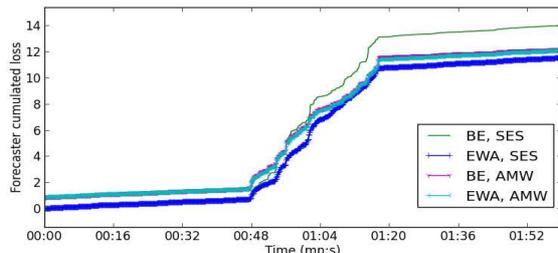

Figure 9. Cumulated loss of different forecasters

With regard to complexity, at each trial the BE and EWA forecasters have to compute the cumulated loss of each expert. In addition, the EWA forecaster needs to compute a weighted average to make its prediction. Hence, the EWA forecaster is a little bit more complex than the BE forecaster.

VI. Conclusion

In this paper, we applied on-line learning strategies to predict the quality of a wireless link in a WSN, based on the LQI metric. A forecaster estimates the LQI value using the advice of experts. Simulation results show that for both types of experts (AMW and SES), the best expert depends on the phase considered. This is the reason why a forecaster is needed. Furthermore, the predictions of the EWA forecaster using SES experts are shown to be reactive and accurate. This combination minimizes the cumulated loss regarding the real LQI values, compared to any other combination such as EWA-AMW, BE-AMW and BE-SES, given by decreasing performance order.

References

[1] N. Baccour, A. Koubaa, L. Mottola, M.A. Zuniga et al. "Radio Link Quality Estimation in Wireless Sensor Networks: A Survey", ACM Transactions on Sensors, Volume 8 Issue 4, September 2012.

[2] K. Farkas, T. Hossmann, F. Legendre, B. Plattner, S. K. Dasc "Link quality prediction in mesh networks", Computer Communications Volume 31, Issue 8, 25 May 2008, Pages 1497–1512, Special Issue: Modeling, Testbeds, and Applications in Wireless Mesh Networks

[3] Y. Xu, W.C. Lee, "Exploring Spatial Correlation for Link Quality Estimation in Wireless Sensor Networks", 4th Annual IEEE Internat. Conf. on Pervasive Computing and Communications, PerCom 2006, Pisa, Italy, March 2006.

[4] M. Spuhler, V. Lenders, D. Giustiniano, "BLITZ: Wireless Link Quality Estimation in the Dark", Proc. of the 10th European Conference, EWSN 2013, Ghent, Belgium, February 2013.

[5] T.Liu, A.E.Cerpa, "Data-Driven Link Quality Prediction Using Link Features", Transactions on Sensor Networks , Vol. 10 Iss. 2, 2014.

[6] N.Baccour, A.Koubaa, H.Youssef, M.Ben Jamaa, et all, "F-LQE: A Fuzzy Link Quality Estimator for Wireless Sensor Networks", Proc. of the 7th European Conf on Wireless Sensor Networks, 2010.

[7] C. Renner, S. Ernst, C. Weyer, V. Turau "Prediction Accuracy of Link-Quality Estimators", EWSN'11 Proceedings of the 8th European conference on Wireless Sensor Networks, 2011.

[8] D. Marinca, P. Minet, "On-line Learning and Prediction of Link Quality in Wireless Sensor Networks ", Inria Research Report, April 2014.

[9] N. Cesa-Bianchi, G. Lugosi, "Prediction, Learning, and Games", Cambridge University Press, 2006.

[10] C.A.Boano, M.A. Zúñiga, T.Voig, A. Willig, "The Triangle Metric: Fast Link Quality Estimation for Mobile Wireless Sensor Networks", Proc. of 19th Int. Conf. on Computer Communications and Networks, 2010.

[11] J.Kivinen, M.K.Warmuth "Averaging Expert Predictions", Comp. Learning Theory: 4th European Conference, EuroCOLT 1999