Optimal Estimation of Training Interval for Channel Equalizations

Dongyan Chen, Yiguang Hong, and Kishor S. Trivedi
Center for Advanced Computing and Communications (CACC)
Electrical and Computer Engineering Department, Duke University
Durham, NC 27708-0294

Abstract – In this paper, an optimal training equalization for wireless communication is proposed and analyzed. By our scheme, the training of the equalizer is carried out periodically, with the training interval optimized for a maximal channel utilization. A closed-form expression for the optimal training interval is derived via a semi-Markov process (SMP) which requires the knowledge of the channel equalization failure time distribution. A statistical estimation algorithm is presented and applied to adaptively estimate and track the optimal interval when the failure time distribution is not available. Numerical results show that by choosing the optimal training interval, the channel utilization can be improved, and the statistical estimation algorithm can effectively approach the optimal solution with a reasonable number of failure time data points.

I. INTRODUCTION

Broadband wireless networks tend to be an integral part of the global communication infrastructure with the rapid growth in popularity of wireless data services. With limited spectrum resources, there is an urgent need for developing new techniques for better bandwidth utilization.

In wide-band digital communication applications, modulation pulses will spread and result in inter-symbol interference (ISI) when modulation bandwidth exceeds the coherence bandwidth of the radio channel. Equalization algorithms are usually built into the receiver to compensate for the channel amplitude and delay variations and combat the ISI, in order to reduce the bit error rate (BER) [1]. Generally speaking, the equalization algorithms can be categorized into training based equalization and blind equalization [1], in the sense of whether the training sequence is needed for the initial adjustment of the coefficients. For the training based schemes, training is carried out periodically due to the time-varying nature of the wireless channel, and hence it is called periodic training equalizer. Different adaptive equalization methods are given in references such as [2] [3]. Since no training is needed for blind equalization algorithms, higher bandwidth utilization may be achieved by the blind equalization as the channel can be fully devoted to data packet transmission. However, the blind equalization is more complicated than the periodic training equalization, and its performance suffers from the slower convergence rate [4]. On the other hand, the bandwidth utilization of the periodic training equalizer is lower due to the requirement of training sequence. What is worse, a careless selection of training interval might bring about either redundant training sequences when the channel varies relatively slowly, or excessive packet retransmissions when the channel varies relatively fast. In [5], an adaptive training equalization algorithm was proposed to determine a training decision. The abrupt change detection algorithm [6] was adopted to inspect channel changes, and once an abrupt change in the channel parameters is detected, a training sequence is sent. This scheme can be called condition-based training, because the training decision is directly based on the channel conditions. However, it is constrained by the complexity of implementation of the abrupt change detection algorithm, and may be prone to performance degradation due to false-alarms and missed-alarms.

In this paper we propose an optimal time-based training scheme. By this scheme, the equalizer is still trained periodically. However, the optimal training interval is derived from the equalization failure time distribution, so that the channel utilization is maximized. Since channel estimation plays an important role in the performance of equalizers [7], and may cause the equalization failure when the estimation significantly deviates from real channel parameters, here it is used as the index for the incapability of the equalizer to decode received symbols.

The proposed equalization policy is formulated as a semi-Markov process (SMP), which is a generalization of Markov process [8]. In fact, SMP has been widely used in hardware preventive maintenance, whose optimal maintenance policies have been studied extensively in the references such as [9] and [10]. The explicit information on the parameters and distribution function of estimation failure time is needed for the analytical optimal solution to an equalization training. In many situations, the failure distribution data related to channel estimation under study can not be obtained beforehand, due to the time-varying nature of wireless channels. To tackle this difficult problem, we propose a non-parametric statistical estimation method for the unknown channel estimation failure distribution. Similar methods were introduced by Dohi et al in the
study of software rejuvenation [11], where the strictly increasing failure rate (IFR) failure time distribution was treated. Our approach extends theirs by including some important cases of more general failure time distributions. It is also worthwhile to point out that such estimation algorithms further illustrate a certain level of adaptivity in the case of time varying failure distribution, which is also explored in this paper.

The contribution of our work may become more significant for such applications as the emerging wireless networking and wireless data services. In these applications, the users tend to be more “static”, and occupy the channel for a longer time. In this case, the channel is less fluctuating than the traditional mobile voice service. Therefore, the selected training interval, once fixed, need not be revised frequently. The optimal training interval is given, with the conditions for the existence and uniqueness of the optimal result shown in Section III. Based on the theoretical results, an online estimation algorithm for optimal training interval is presented in Section IV. Following that, Section V is devoted to the numerical illustrations of the aforementioned techniques with a Weibull channel estimation failure time distribution. Finally, Section VI concludes the paper.

II. MODEL DESCRIPTION

When the discrete-time white noise channel can be represented as the tap delay line model [1], the received symbol is then

\[ r_k = \sum_{i=0}^{q-1} f_k(i)w_{k-i} + v_k = f_k^T w_k + v_k. \]  

(1)

where \( w_k = (w_k, w_{k-1}, \ldots, w_{k-q+1})^T \), with \( w_k \) being the \( k \)th transmitted symbol, and \( f_k = (f_k(0), f_k(1), \ldots, f_k(q-1))^T \), with \( f_k(i) \) being the \( i \)th channel tap coefficient when receiving the \( k \)th symbol, and \( v_k \) is the additive white Gaussian noise.

Practically, the training-based equalizers operate either in training mode or in receiving mode [1]. In training mode, the received symbols are compared with the known transmitted training symbols to determine the channel tap coefficients. In the receiving mode, the operation may be decision-directed, where the decisions on the information symbols are assumed to be correct and are used to track the changes in the channel model.

Because the condition of the considered channel is time varying, the discrepancy between the estimated channel by the equalizer and the real channel also evolves with time. As a consequence, if the inaccuracy in estimation is not detected and corrected promptly, the erroneously decoded symbols will prevail and lead to lost data packets [7] and channel outages [12]. To avoid such losses, the equalizer is trained periodically to rectify its deviated estimations.

In this paper, we use the mean square error (MSE) of channel estimation as a metric for its failure to estimate channel tap coefficients, which is defined as

\[ J_k = (\hat{f}_k - f_k)^T (\hat{f}_k - f_k) \]  

(2)

where \( \hat{f}_k \) is the estimated channel tap coefficient vector when receiving the \( k \)th symbol, and the superscript \( T \) denotes the transpose. With \( J_{th} \) as a threshold given in advance, \( J_k > J_{th} \) defines the failure of the channel estimator to keep track of the channel variations.

This equalization problem can be depicted by the transition diagram shown in Figure 1. The system can be in one of three states:

- **State 0**: In this state, the equalizer works in the receiving mode. Since the information can be conveyed, we call the state of the equalizer available.
- **State 1**: In this state, the equalizer is in the training mode, and it is unavailable to receive user data.
- **State 2**: In this state, the cumulative channel estimation error causes an equalization failure, that is, the BER exceeds the assigned threshold and valid transmission cannot proceed. Like state 1, this state is also an unavailable state for the equalizer to receive user data, since the erroneously received data packet has to be dropped.

When the equalizer in question stays in state 0, its only available state, it will either enter state 1 with a general distribution function \( F_0(t) \) (for transmission of a training sequence), or fail with a distribution function \( F_2(t) \) (due to excessive channel estimation errors). The distribution function for the duration of training is \( F_1(t) \). When the equalizer is in state 2, there will be a certain delay taken by the sender to notice the transmission failure by either a negative acknowledgment or a timeout event. We assume this delay has a distribution function \( F_2(t) \). After that, the sender will have to transmit a training sequence to re-establish the communication. Therefore, as shown in Figure 1, the system enters state 1 after it leaves state 2.

We assume that periodic training is performed in our proposed scheme. Suppose that the training interval is \( t_0 \) (that is, \( t_0 \)
is the time elapsed between two consecutive training triggers. In other words, \( F_0(t) = U(t - t_0) \), where \( U(\cdot) \) is the unit step function. The duration for each training process is generally distributed with mean \( t_1 \), and the time required to recover from the equalization failure is assumed to be generally distributed with mean \( t_2 \). Since each state is a regenerative state, the underlying stochastic process is a semi-Markov process (SMP), which has been scrutinized in the references such as [8].

### III. SMP AND OPTIMIZATION

We define the channel utilization as the goodput divided by the channel capacity, where the goodput means the amount of valid user data retrieved by the receiver in a unit time. From the preceding discussions, it is obvious that only in state 0 could user data be received with negligible errors. Assume the channel capacity is a constant \( C \), the steady-state channel utilization can then be written as

\[
A(t_0) = \frac{c_0}{C} = \pi_0
\]

where \( \pi_0 \) is the steady-state probability that the system is in state 0. Following the standard procedure of SMP analysis and employing the approach in [10], the steady-state channel utilization can be obtained as

\[
A(t_0) = \pi_0 = \frac{\int_{t_0}^{t_3} \overline{F}_2(t)dt}{\int_{t_0}^{t_3} \overline{F}_2(t)dt + t_1 + F_2(t_0)t_3} = \frac{S(t_0)}{T(t_0)}, \tag{3}
\]

where \( \overline{F}_2(t) = 1 - F_2(t) \) is the complementary distribution function.

Define the following function:

\[
q(t_0) = T(t_0) - [1 + t_3r_f(t_0)]S(t_0), \tag{4}
\]

where \( r_f(t_0) = \frac{dF_2(t_0)/dt}{\overline{F}_2(t_0)} \geq 0 \) is the failure rate. By solving \( dA(t_0)/dt_0 = 0 \), we obtain the following theorem on the uniqueness and existence of the optimal training interval (the proof is omitted due to space limitations).

**Theorem 1:**

i) If \( q(\infty) < 0 \) (or \( r_f(\infty) > t_3/t_2\pi_0 \)), then there is a finite optimal training interval \( t_0^* \) satisfying \( q(t_0^*) = 0 \), and its local maximal channel utilization can be taken as

\[
A(t_0^*) = \frac{1}{1 + t_3r_f(t_0^*)}. \tag{5}
\]

Moreover, if the channel failure time distribution is IFR (increasing failure rate), then \( \frac{dr_f(t_0^*)}{dt_0^*} \geq 0 \), then there is a finite and unique optimal training interval \( t_0^* \) such that \( t_0^* = \sup\{t_0\mid q(t_0) = 0\} \), and the maximal channel utilization can be taken as

\[
A(t_0^*) = \frac{1}{1 + t_3r_f(t_0^*)}. \tag{6}
\]

\(^{1}\)For the equalization problem, resulting from the error propagation in the equalization algorithm and the time varying nature of the channel, it is reasonable to assume that the channel estimation errors accrue with time, which leads to the IFR failure time distribution.

In addition, if \( \frac{dr_f(t_0)}{dt_0} \neq 0 \) at \( t_0 = t_0^* \), then \( t_0^* \) is the only element in the set \( \{t_0\mid q(t_0) = 0\} \).

ii) If \( q(\infty) \geq 0 \) and the channel failure time distribution is IFR, then the optimal interval is \( t_0^* \rightarrow \infty \), and

\[
A(\infty) = \frac{t_2}{t_1 + t_2 + t_3}.
\]

### IV. STATISTICAL ESTIMATION ALGORITHM

Deriving the optimal training schedule by Theorem 1 requires the knowledge of the channel estimation failure time to determine \( F_2(t) \). Since the channel estimation failure time depends on the value of estimation failure threshold, estimation algorithm design, and channel variation characteristics, in most cases this information is not available \textit{a priori}. In this section, we follow the ideas in [13] and [11], and present a statistical optimization algorithm to estimate the optimal training schedule online.

To collect the channel estimation failure data in practice, two identical channel estimators are required. One of the estimator (we call it primary estimator) works in normal mode with the equalizer, and it is updated periodically by the training process. The other one (we call it secondary estimator), however, is not trained. We take the output of the primary estimator as the real channel parameters \( f_k \), and the output of the secondary estimator as \( \hat{f}_k \). When the MSE of the output of the secondary channel estimator is larger than \( J_{th} \) as in (2), an estimation failure is indicated.

This process enables us to collect the channel estimation failure data, and therefore enables the online optimization process, as will be shown in the following.

In this section, we assume \( \frac{dr_f(t_0)}{dt_0} \geq 0 \) and \( r_f(t_0)|_{t_0=0} \neq 0 \). The scaled total time on test (TTT) transform [13] is defined as

\[
\phi(p) = \left(\frac{1}{t_2}\right) \int_0^{P_{2}^{-1}(p)} \overline{F}_2(t)dt, \quad 0 \leq p \leq 1 \tag{7}
\]

Suppose the channel is monitored continuously, and an ordered complete observation of the time when \( J_k > J_{th} \) (as in (2)) is obtained as \( 0 = x_0 \leq \cdots \leq x_n \). Then, the scaled TTT statistics based on this observation is defined by \( \phi_{n_j} = \Phi_j / \Phi_n \), where

\[
\Phi_j = \sum_{k=1}^{j} (n - k + 1)(x_k - x_{k-1}) \tag{8}
\]

We also use the empirical distribution function

\[
\hat{F}_n(x) = \begin{cases} 
\frac{j}{n} & \text{for } x_j \leq x < x_{j+1} \\
1 & \text{for } x_n \leq x 
\end{cases}
\]

to play the same role as \( F_2(x) \) in (7). In fact, define \( \hat{F}_n^{-1}(p) = \inf\{x \mid \hat{F}_n(x) > p\} \), then we have

\[
\lim_{n \to \infty, p \to p} \int_0^{\hat{F}_n^{-1}(j/n)} \overline{F}_n(t)dt = \int_0^{F_2^{-1}(p)} \overline{F}_2(t)dt
\]
uniformly in $p$ with probability one (see [13]).

The following result gives non-parametric statistical estimation algorithm for the optimal training interval.

**Theorem 2**: Suppose that $F_2(t)$ is IFR (i.e., $\frac{dF_2(t)}{dt} > 0$) and $r_f(t) \neq 0$ if $t \neq 0$.

i) With above notation, obtaining the optimal training interval $t_0^*$ is equivalent to obtaining $p^*$ ($0 \leq p^* \leq 1$) such that

$$p^* = \max \{ p \mid \max_{0 \leq p \leq 1} \frac{\phi(p)}{p + t_1/t_3} \}$$

(9)

ii) Suppose that the optimal training interval is to be estimated from ordered complete sample of size $n$, $0 = x_0 \leq x_1 \leq \cdots \leq x_n$, of the failure times from an absolutely continuous distribution $F_2(t)$, which is unknown. Then, a non-parametric estimator of the optimal training interval $t_0^*$ which maximize $A(t_0)$ is given by $x_{j^*}$, where

$$j^* = \max \{ j \mid \max_{0 \leq j \leq n} \frac{\phi_{nj}}{j/n + t_1/t_3} \}$$

(10)

Moreover, the estimator given above is strongly consistent, i.e., $x_{j^*}$ converges to the optimal solution $t_0^*$ uniformly with probability one as $n \to \infty$.

**Proof**: With the given conditions and the discussions in [13], it is not difficult to see that $F_2(t)$ is IFR if and only if if $\phi(p)$ is concave on $p \in [0, 1]$. Let

$$\psi(p) = \frac{\phi(p)}{p + t_1/t_3}$$

After differentiation with respect to $p$ and setting it equal to zero (considering that $t_0 = F_2^{-1}(p)$, and $F_2^{-1}(p)$ is strictly monotonically increasing when $p \neq 0$ because of the given conditions), we have

$$\frac{1}{t_2 \tau_f(t_0)} F_2(t_0) = \int_0^{t_0} \frac{F_2(t)dt}{r_2} + \frac{t_1}{t_3}$$

which can be found to be equivalent to $q(t_0) = 0$ after a few algebraic manipulations. Note that $t_0^*$ is the maximal value satisfying $q(t_0) = 0$ and $F_2$ is strictly monotonically increasing. Therefore, $p^*$ defined in (9) equals to $F_2(t_0^*)$.

The proof for the second part of this theorem is quite straightforward, based on the preceding discussion and results in [13].

Q.E.D.

V. NUMERICAL RESULTS

In this section, we evaluate the performance of the proposed optimal training algorithm through numerical examples. We assume that the channel estimation failure time is Weibull distributed, with the distribution function

$$F_2(t) = 1 - e^{-\lambda t^\alpha}$$

(11)

We assume that the data transmission speed of the communication link is 1M bit per second (bps). The length of the equalization sequence is 128 bytes, and the data frame length is 4k bytes. An erroneously received packet would be indicated by the receiver with a negative acknowledgment (NACK) packet sent to the sender, preceded by a training sequence so that the sender can correctly receive the NACK packet. We assume the length of a NACK packet is 256 bytes. When a packet loss is detected, the sender will have to again send out a training sequence, followed by a re-transmission of the data packet. A switching delay of 10ms is introduced for this recovery process.

With the parameters selected as above, the time required for the training, $t_1$, is $128 * 8/2^{20} \approx 0.97ms$, and the average time to recover the system from failure state may include the duration of training for NACK packet transmission, the time taken to transmit a NACK packet, the time taken to re-transmit the previous packet, and the switching delay. Therefore, the time taken to recover from the channel estimation failure is $(4 + 0.25) * 8 * 2^{10}/2^{20} + 0.97 + 10 \approx 44ms$.

The parameters for the channel estimation failure distribution are chosen as $\alpha = 3$, $\lambda = 3$, which will result in a mean time to failure (MTTF) of 522ms.

Applying Theorem 1 and noting that $q(0) = t_1 > 0$ and $q(\infty) \to -\infty$, a finite optimal solution exists. Since the failure time distribution is strictly IFR (that is, $\frac{dF_2(t)}{dt} \mid_{t_0=0} > 0$), the value satisfying $q(t_0) = 0$ is unique (following Theorem 1).

Figure 2 shows the asymptotic behavior for the estimation of $t_0^*$ based on TTT transform, and Figure 3 shows the corresponding channel utilization. For this calculation, we assume that the estimation failure time data can be obtained error-free. From these two figures we see that satisfactory results could be obtained with sample size $n \geq 30$.

Fig. 2. Asymptotic behavior of optimal training interval estimation with sample size $n$.

If a mobile user moves to a new environment, or a new noise source is added, then the considered channel failure model may change correspondingly. However, our design idea can still work well with the assumption that a maximum delay of $n$ observed channel data points suffice from a statistical standpoint, since the optimal training interval can be calculated with $n$ samples. We will study this issue in our second experimental setting, where the channel failure time distribution parameters...
experienced an abrupt change to $\alpha = 2, \lambda = 20$. The number of data samples for the TTT transform, $n$, is chosen to be 40. Figure 4 shows the variation of $t_0^*$ under the abrupt change in channel failure distribution parameters, and Figure 5 shows the corresponding channel utilization. It can be observed from these two figures that the estimation converges to the optimal value when more than 20 new failure time data points are collected on the new distribution parameters. We notice that the optimal training interval decreases from around 90ms to around 20ms. In this case, the training frequency needs to be increased from once for every 3 data packet transmissions to once for every data packet transmission. It is worthwhile to point out that although the estimation error of $t_0^*$ is not negligible, as shown in Figure 4, but since the re-training interval has to be an integer number of data packets, it is actually rounded off in the final decision.

VI. CONCLUSIONS

In this paper we proposed an optimal time-based training scheme. A closed-form solution for the optimal training interval is obtained to maximize the channel utilization. Also a statistical estimation algorithm is proposed and applied to dynamically estimate the optimal training interval with channel estimation data collected online, when the channel estimation failure time distribution is not available. Through numerical evaluations, we found that the proposed training policy effectively increases the bandwidth utilization. The adaptive tracking performance of the estimation algorithm under abrupt changes in the channel estimation failure time distribution is also demonstrated. We focused on the equalization problem under a general class of resource reservation media access control (MAC) schemes. The performance of the optimal training equalization scheme under specific MAC protocols will be addressed in future.

REFERENCES