

MICROSCOPIC AND MACROSCOPIC APPROACHES IN SECTOR FAILURE RATE ESTIMATION

Alexander KUZNETSOV¹ and Raman VENKATARAMANI²

1) Seagate Research, 1251 Waterfront Place, Pittsburgh, PA, Alexander.V.Kuznetsov@Seagate.com

2) Seagate Research, 1251 Waterfront Place, Pittsburgh, PA, Raman.C.Venkataramani@Seagate.com

I. INTRODUCTION

The sector failure rate (SFR) is extremely small at normal operating conditions of hard disc drives. In practice it cannot be obtained by counting as that would require numerous simulations. Therefore, appropriate statistical models are applied for the distribution of error symbols in a sector to estimate the SFR. In this paper we look at the underlying philosophy of existing estimation methods and classify them into macroscopic and microscopic types. We show how to estimate the model parameters optimally from a library of sample error indicator vectors. We observe that the microscopic models are well suited for certain iterative channels.

II. MACROSCOPIC AND MICROSCOPIC MODELS

Fundamentally, all models for error vectors may be classified as either *microscopic* or *macroscopic* models. A microscopic model captures the precise correlation between symbol errors locations while a macroscopic model captures the weight of the error indicator vectors alone ignoring the actual bit or symbol error locations. In general, precise microscopic models are difficult to find for iterative coded channels with interleavers due to the complex long range correlations present in the symbol errors. In contrast, macroscopic models are also simpler to analyze and are generally *universal* in the sense that it can work well for a variety of codes but they require larger training data sets

An example of the macroscopic model is the *multinomial model* in which (a) an error indicator vector consists of N independent error events, and (b) a single error event is of a string of up to L symbols in error. The resulting error vector weight has the multinomial distribution. This is evidently a macroscopic model. Sometimes, the symbol errors are modeled as a Markov random process. An example is the Gilbert model [1, 2] or the generalized Gilbert model [3]. These models are microscopic because they explicitly model the error locations, not the error weight. An example of a slightly richer model than the Gilbert model is shown in Figure 1. According to this model, the probability of seeing a symbol error is q_{m+1} where m is the run-length of preceding symbol errors leading in that error event. For reduced model complexity we let $q_m = q_M$ for all $m \geq M$ where M is a fixed parameter. When the chain returns to the zero state from the nonzero state m , the generation of an error event $E(m)$ is completed by attaching the final "0" to the run of m consecutive ones. This is a microscopic model because it characterizes the symbol error locations.

Having chosen any model (microscopic or macroscopic) its parameters can be estimated from the library of actual error vectors as follows. Let the data set consists of K error indicator vectors. Let \hat{r}_k and r_k denote the histogram of the error vector weights and the modeled weight distribution respectively. The optimal estimate of the model parameters is the so called *minimum relative entropy (MRE) solution* where we minimize the *Kullback-Liebler distance* between the two weight distributions [4].

$$\min D(\hat{r} \parallel r(p)) = \sum_n \hat{r}_n \log(r_n(p) / \hat{r}_n)$$

III. MICROSCOPIC MODEL WITH TWO COMPONENT TYPES

We shall now extend multinomial model to SFR estimation of iterative channels as well. Let $E(m)$ denote an error event (EE) of length m : it consists of $m \geq 1$ ones followed by a zero, e.g., $E(3)=1110$. A sequence of $n \geq 1$

Alexander KUZNETSOV
Seagate Research
E-mail: Alexander.V.Kuznetsov@Seagate.com
fax: +1-412-918-7111
tel: +1-412-918-7157
1251 Waterfront Place, Pittsburgh, PA
Preferred category: #8

consecutive zeros $I(n)$ is called an *error-free run of length n*. Thus, we can always parse an error indicator vector into a unique sequence of interlacing error events and error free runs as follows: $\xi = (E_1, I_1, E_2, I_2, \dots, E_L, I_L)$.

Main Assumptions: The events are independent identically distributed random objects and the number of such error events in a sector is an independent random variable. Assume that E_i are described by the microscopic model described in Section III-B. Estimation of the parameters of the Markov model described above of can be done using the minimum relative entropy method described in Section III-C. We do not assume that PDFs of error events and error free runs are identical.

The word failure rate is $\rho_w(T) = \sum_k Q(k) \rho_w(T, k)$ where $Q(k)$ and $\rho_w(T, k)$ are the distribution of number of error events per word and the conditional word failure rate given k error events in the word. The latter quantity is computed using the MRE method described in Section II. The PDF $Q(k)$ is estimated from statistics obtained from real or simulated data however its tail is extrapolated using a pre-specified distribution. Numerically, iterative channels using TPC and LDPC codes have exponential tails ($Q(k) = e^{-\lambda k}$) and iterative channels using SPC have tails with a Poisson distribution ($Q(k) = \lambda^k e^{-\lambda} / k!$). The parameter λ is estimated to fit the observed tails.

IV. SIMULATIONS

In the following example we evaluate the performance of the *descriptive* and *predictive* capabilities of various SFR estimation methods for a TPC system operating at 24dB. Each sector consists of 500 symbols including the ECC parity symbols. The full data set contains 79111 sectors and the partial data set contains 4% of this data. Figure 5 shows the SFR estimates using several methods from the partial data and counted values from full data set. These methods appear to be consistent with each other and match the extrapolated counted values well.

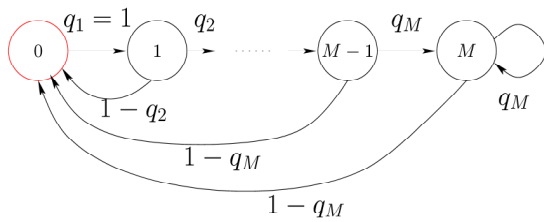


Figure 1: Model for error event

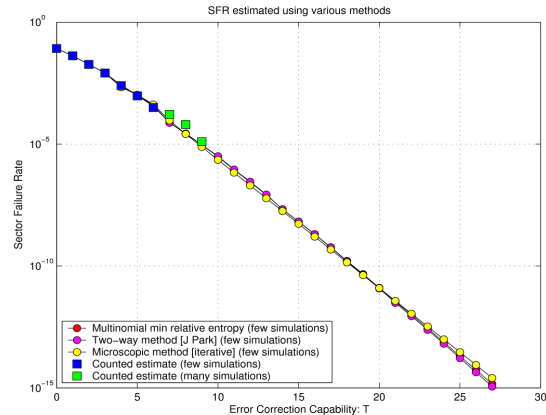


Figure 2: SFR estimates for the TPC system

REFERENCES

- 1) J. R. Yee and E. J. Weldon, "Evaluation of the performance of error-correcting codes on a Gilbert channel," *IEEE Trans. Comm.*, vol. 43, pp. 2316–23, Aug. 1995.
- 2) E. N. Gilbert, "Capacity of burst noise channel," *Bell Systems Tech. J.*, vol. 39, pp. 1253–66, Sept. 1960.
- 3) J. K. Wolf, "ECC performance of interleaved RS codes with burst errors," *IEEE Trans. Magnetics*, vol. 34, pp. 75–79, Jan. 1998.
- 4) T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.