Time and Frequency Domain Description of Gilbert-Elliott Data Loss Models

András Palkó
Department of Measurement and Information Systems
Budapest University of Technology and Economics
Budapest, Hungary
palkoandras96@gmail.com

László Sujbert
Department of Measurement and Information Systems
Budapest University of Technology and Economics
Budapest, Hungary
sujbert@mit.bme.hu

***Abstract*—Nowadays, the usage radio channel or Internet connection grows for real-time transmission of measurement data or signals. For these applications, not reliable communication needs to be analyzed from a signal processing point of view. Our paper contributes to this topic with the analysis of data loss as a signal distorting phenomenon. The paper contains a possible mathematical description of the data loss via the indicator function. Moreover, the Gilbert-Elliott data loss model class is presented, which gives a framework to analyse numerous data loss phenomena in general. In order to analyse the data loss in the frequency domain, the spectrum of the indicator function is required. This paper presents the derivation of the power spectral density of the Gilbert-Elliott model. The paper reviews a possible identification method for the simpler models derived from the Gilbert-Elliott model. Theoretical analysis and simulations are supported by measurements. The measurement results are used to show how the data loss models function in practice.**

Keywords—data loss, data loss model, spectrum, identification, hidden Markov model, Gilbert-Elliott model

I. Introduction

Traditional communication and measurement systems offer high-precision, reliable and fast data transmission. Recently, due to the technological development, sensor networks with less reliable communication appeared. E.g., this means that some measurement data gets damaged during the transmission. The idea of the Internet of Things takes this a step further: connect every electronic equipment to the Internet to create an enormous sensor network, measure and control the physical world from a distance. These systems give motivation to examine data loss phenomena from a signal processing point of view.

Data loss can be viewed as a measurement error caused by the not reliable equipment and procedures. A radio signal may be received incorrectly due to interference, packets can be lost over Internet connections or external circumstances can hinder the measurements. Data loss can mean invalid samples, for example an overdrive AD-converter. Lost packages over an Internet connection can cause missing samples. External circumstances (e.g., weather for an astronomical measurement) may also lead to missing or invalid samples. In a system with multiple clock domains, synchronization issues are able to cause for a sample to be doubled or omitted in a transition over the clock domain.

In the prevoius works [1] [2], data loss models chosen by the physical models have been examined. These were the random independent, Markov, block-based models and in general, models with some memory. In this paper, the Gilbert-Elliot model [3] is investigated as it comprises all the previously analyzed models. After the investigation, the results can be simplified with deductive reasoning for the simpler models: random independent, two-state Markov, Gilbert [4] and complementary Gilbert. The benefit of this methodology is the unified description of these models in time and frequency domain.

If an application transmits data via a communication channel and the useful information is the spectrum of the signal, even one lost sample is crucial. Here the spectra of the signal and the indicator function are convolved. In order to analyze the data loss in the frequency domain, the spectrum of the indicator function is required. Data loss appears as a noise in the spectrum which can hinder the detection and accurate measurement of weaker spectral components. On the other hand, this noise can be used to identify the data loss model.

When data are lost, a natural idea is to resend, remeasure, but in some cases it cannot be done. When a signal is measured in real-time, the lost samples cannot be replaced. For reliable communication, a complex state machine needs to be implemented. Moreover, reliable communication in a finite time is impossible, thus in real-time systems data loss needs to be taken into account.

There are methods which require a whole block of samples, e.q. spectral analysis, where even one lost sample is crucial. When processing in blocks, it is a possibility to discard blocks with lost samples. However, even with weak data loss, the time required to acquire a whole block can result in an unacceptable measurement time.

In section II we give the mathematical description of data loss and some basic definitions are introduced. Section III contains the description of the examined data loss models. Time domain and statistical properties are given in section IV while frequency domain analysis results are presented in section V. Section VI deals with the identification problem

The project was funded by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013).

978-1-7281-0702-8/19/$31.00 ©2019 IEEE

and section VII concludes with measurement results.

II. Preliminaries

A. Mathematical Description of Data Loss

1) Indicator Function: Data loss can be modeled in discrete time with an availability indicator function:

$$K_n = \begin{cases} 1 & \text{if the sample is available at } n \\ 0 & \text{if the sample is lost at } n \end{cases} \quad (1)$$

2) Statistical Properties: In this paper, consecutive samples are called a sequence, a sequence with a fix length is a block. Using the indicator function we can define the μ data availability and γ data loss rates:

$$\mu = \Pr(K_n = 1) \quad \gamma = \Pr(K_n = 0) \quad (2)$$

where $\Pr(\cdot)$ is the probability operator. The $R(L)$ reliability and $R'(L)$ complementary reliability functions give information about the time distribution of the lost samples. They give the probability of L consecutively available and lost samples:

$$R(L) = \Pr(\forall i \in \{1, \ldots, L\} : K_i = 1) \quad (3)$$

$$R'(L) = \Pr(\forall i \in \{1, \ldots, L\} : K_i = 0) \quad (4)$$

The average length of available and lost sequences will be marked with E_{N1} and E_{N0}, respectively.

B. Modeling of a Data Loss Process

With the indicator function we can easily model a data loss process. If an x_{0n} discrete signal and the K_n indicator function are given, the x_n signal with lost samples is the product of them:

$$x_n = x_{0n} K_n \quad (5)$$

Thus, the spectrum of a signal with lost samples is the convolution of the spectra of the original signal and the indicator function.

III. Gilbert-Elliott Data Loss Model Family

A. Gilbert–Elliott Model

The Gilbert-Elliott data loss model is a two-state two-output hidden Markov-model. In this paper the two states will be marked with A and B. The state transition probabilities are p and q. The probabilities of the available samples in the two states are a and b, respectively. These probabilities will be called as the output probabilities.

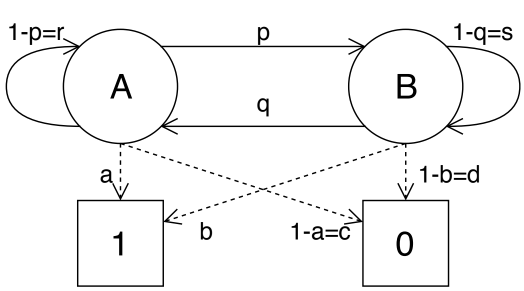

Fig. 1: Gilbert-Elliott data loss model

The probabilities of staying in the same state will be marked with $r = 1 - p$ and $s = 1 - q$. Similarly, the probabilities of lost samples will be marked with $c = 1 - a$ and $d = 1 - b$.

This model can be used for such systems which have two states, an 'Up', where the system mostly works and a 'Down', where the system is mostly faulty.

B. Simpler Models

The simpler models are different from the Gilbert-Elliott in their behavior in the 'Up' and 'Down' states. In either or both states can be the data availability deterministic, or the two states can generate the same output.

1) Gilbert Model: Gilbert model is the simplified version of the Gilbert-Elliott model when $a = 1$. This means that in one state, the sample is always available. This model is applicable for a system which works perfectly in state 'Up' and in a mostly faulty way in state 'Down'.

2) Complementary Gilbert Model: We get the complementary Gilbert model from the Gilbert-Elliott when $b = 0$. In other words, the sample in state B is always lost. The model can describe a system which is mostly working in the 'Up' state and totally faulty in the 'Down' state.

3) Two-state Markov Model: The two-state Markov-model is a special case of Gilbert-Elliott model with $a = 1$ and $b = 0$. It is also the special case of the Gilbert and complementary Gilbert models. Here the output is a deterministic function of the state. The two-state Markov-model is useful to describe a system which works perfectly in state 'Up' and not at all in state 'Down'.

4) Random Independent Model: Random independent data loss model is the simplest data loss model. Each sample is available with a fixed probability. Unlike the previous models, here the availability of different samples is independent from each other. This also means that the random independent model is memoryless. In the 'Up' and 'Down' terminology: the behavior of the system is the same in the two states.

We can get a random independent model from the Gilbert-Elliott in two ways: the simpler one is when $a = b$. This means that the probability of an available sample is equal in the two states, thus from outside the two states are indistinguishable. The other possibility is when $p + q = 1$, which means that the next state is independent from the current one. It can be shown using the stationary distribution of the underlying Markov-chain.

IV. Time Domain and Statistical Properties

In this section the time domain properties defined in section II are given for the Gilbert-Elliott model family.

A. Properties of the Gilbert-Elliott Model

1) Data Availability and Data Loss Rates: To give the data availability rates, we need the π stationary distribution of the underlying Markov-chain, which is

$$\pi = \begin{bmatrix} \pi_{\mathsf{A}} & \pi_{\mathsf{B}} \end{bmatrix} = \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix} \quad (6)$$

Combining with the probabilities of available samples in both states, we can get the data availability rates:

$$\mu = a\pi_{\mathsf{A}} + b\pi_{\mathsf{B}} = \frac{aq+bp}{p+q} \quad \gamma = c\pi_{\mathsf{A}} + d\pi_{\mathsf{B}} = \frac{cq+dp}{p+q} \quad (7)$$

2) Reliability Functions: The reliability function takes the following form:

$$R(L) = \frac{aq\left[\lambda_2\lambda_1^{L-1} - \lambda_1\lambda_2^{L-1} + \left(\lambda_2^{L-1} - \lambda_1^{L-1}\right)(ra+pb)\right]}{(p+q)(\lambda_2-\lambda_1)} + \frac{bp\left[\lambda_2\lambda_1^{L-1} - \lambda_1\lambda_2^{L-1} + \left(\lambda_2^{L-1} - \lambda_1^{L-1}\right)(qa+sb)\right]}{(p+q)(\lambda_2-\lambda_1)} \quad (8)$$

$$\lambda_{1,2} = \frac{ra+sb \pm \sqrt{[ra+sb]^2 - 4aby}}{2} \quad (9)$$

where $y = 1-p-q$. For the derivation of the reliability function see appendix A. Due to the symmetry of the model, the complementary reliability function can be expressed from the reliability function, with the $a \mapsto c$ and $b \mapsto d$ substitutions.

3) Average Length of Avavilable and Lost Sequences: The average length of available sequences is

$$E_{N1} = \frac{aq+bp-ab\,(p+q)\,y}{(aq+bp)\,(1-(ar+bs)+aby)} \quad (10)$$

For the derivation of the average length of available sequences see appendix B. The average length of lost sequences can be expressed from the previous equation using the $a \mapsto c$ and $b \mapsto d$ substitutions.

B. Properties of Simpler Models

The properties of the simpler models generally can be derived from the properties of the Gilbert-Elliott model using appropriate substitutions. For example, the reliability function of the Gilbert model can be expressed from (8) with $a = 1$.

Assuming geometric probability distribution of the data loss, much simpler equations can be derived. In the following expressions, indices G, CG, M and RI indicate the Gilbert, complementary Gilbert, Markov and the random independent model's parameters, respectively.

1) Gilbert Model:

$$R'_G(L) = \gamma_G\,(sd)^{L-1} \quad E_{N0,G} = 1/(1-sd) \quad (11)$$

2) Complementary Gilbert Model:

$$R_{CG}(L) = \mu_{CG}\,(ra)^{L-1} \quad E_{N1,CG} = 1/(1-ra) \quad (12)$$

3) Two-state Markov Model:

$$R_M(L) = \mu_M r^{L-1} \quad R'_M(L) = \gamma_M s^{L-1} \quad (13)$$

$$E_{N0,M} = 1/q \quad E_{N1,M} = 1/p \quad (14)$$

4) Random Independent Model:

$$R_{RI}(L) = \mu_{RI}^L \quad R'_{RI}(L) = \gamma_{RI}^L \quad (15)$$

$$E_{N0,RI} = 1/\mu_{RI} \quad E_{N1,RI} = 1/\gamma_{RI} \quad (16)$$

V. Description in Frequency Domain

This section presents the derivation of the power spectral density of the Gilbert-Elliott model.

A. Derivation of the Spectrum of the Gilbert-Elliott Model

The derivation is based on the spectrum of Markov chain driven signals [5]. Let X_n be a discrete time Markov chain with transitions occurring every T seconds. A $z_i(t)$ waveform is assigned to every state of the Markov chain. The $z(t)$ Markov chain driven signal is the following:

$$z(t) = \sum_{n=0}^{\infty} z_{X_n}(t-nT) \quad (17)$$

Firstly, we need to convert the Gilbert-Elliott model to an equivalent Markov chain. The states will be A1, A0, B1, B0 with e.g., A0 meaning getting a zero output from state A. The $\mathbf{P}_{4S}$ transition matrix of this Markov chain is

$$\mathbf{P}_{4S} = \begin{bmatrix} ar & cr & bp & dp \\ ar & cr & bp & dp \\ aq & cq & bs & ds \\ aq & cq & bs & ds \end{bmatrix} \quad (18)$$

A rectangular impulse with amplitude 1 and length T is assigned to the states A1 and B1. The waveform for the other two states is the constant zero. According to [5] the $\phi(f)$ power spectral density of the Gilbert-Elliott model is

$$\phi(f) = T\mathrm{sinc}^2(fT) \cdot \left[\mu\gamma + 2\frac{pq\,(a-b)^2}{(p+q)^2}\frac{y(\cos x - y)}{|1-y\mathrm{e}^{-jx}|^2} + \mu^2\sum_{n=-\infty}^{\infty}\delta(fT-n)\right] \quad (19)$$

where $x = 2\pi fT$. According to [1], the $S(f_k)$ spectrum in discrete time can be expressed as

$$S(f_k) = GH(f_k) + \mu^2\delta(f_k) \quad (20)$$

where $f_k = \frac{k}{N}$ is the relative frequency, $k = 0 \ldots N-1$, N is the DFT size, G is a scale factor, and $H(f_k)$ gives the shape of the spectrum. Equation (19) gives the spectrum of the indicator function restored with a zero order hold. $T\mathrm{sinc}^2(fT)$ comes from the ZOH, and in the brackets we can find the shape of the spectrum:

$$H(f_k) = \mu\gamma + 2\frac{pq\,(a-b)^2}{(p+q)^2}\frac{y\,(\cos x - y)}{|1-y\mathrm{e}^{-jx}|^2} \quad (21)$$

The scale factor G is [1]:

$$G = \frac{\mu\,(1-\mu)}{\sum_{k=0}^{N-1} H(f_k)} = \frac{\mu\gamma}{N\left(\mu\gamma + 2\frac{pq(a-b)^2}{(p+q)^2}\frac{y^N}{1-y^N}\right)} \quad (22)$$

Putting together, the spectrum of the Gilbert-Elliott model is

$$S(f_k) = \frac{\mu\gamma}{N}\frac{\mu\gamma + 2\frac{pq(a-b)^2}{(p+q)^2}\frac{y(\cos x - y)}{|1-y\mathrm{e}^{-jx}|^2}}{\mu\gamma + 2\frac{pq(a-b)^2}{(p+q)^2}\frac{y^N}{1-y^N}} + \mu^2\delta(f_k) \quad (23)$$

B. Discussion

The shape of the spectrum is determined by y. When $y \approx 1$, the spectrum is lowpass. $y \approx 0$ means a nearly uniform spectrum, $y \approx -1$ means a highpass shape.

Furthermore, $a = b$ also results in a uniform spectrum. The smaller the absolute difference between a and b, the flatter the spectrum is. Figure 2 illustrates the spectrum shapes.

The blue continuous line shows the spectrum when the transition probabilities are small ($y \approx 1$) and the output probabilities strongly differ. This results in a lowpass spectrum. Green continuous line is also a lowpass shape, however this is almost entirely flat. This flatness is caused by the only slightly different output probabilities. Blue dashed line is a flat spectrum. Although the output probabilities are different, $y = 0$ thus a random independent data loss occurs, which means a flat spectrum. Green dashed line shows the spectrum when the transition probabilities are high ($y \approx -1$). This with the different output probabilities result in a highpass spectrum. It is worth noting that the DC component contains a unit impulse proportional to the squared data availability rate.

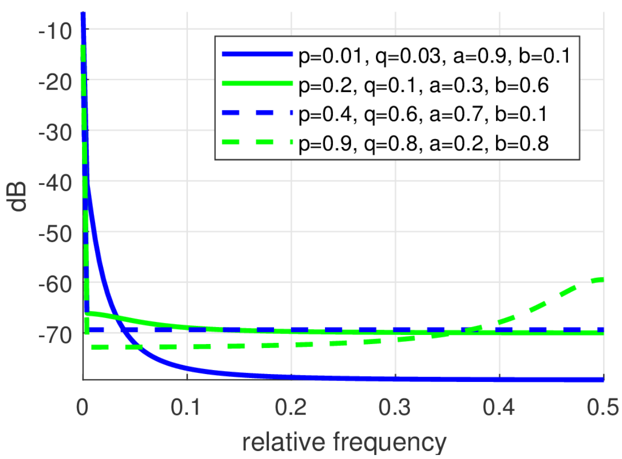

Fig. 2: Examples for the spectrum of the Gilbert-Elliott model

C. Spectra of the Simpler Models

The spectra of the simpler models can be derived from the spectrum of the Gilbert-Elliott model using the conditions to derive said model. For example, we get the spectrum of the Gilbert model from (23) when $a = 1$.

1) Gilbert Model:

$$S_G(f_k) = S(f_k)|_{a=1} = \\ = \frac{\mu_G\gamma_G}{N}\frac{\mu_G\gamma_G + 2\frac{pq(1-b)^2}{(p+q)^2}\frac{y(\cos x - y)}{|1-y\mathrm{e}^{-jx}|^2}}{\mu_G\gamma_G + 2\frac{pq(1-b)^2}{(p+q)^2}\frac{y^N}{1-y^N}} + \mu_G^2\delta(f_k) \quad (24)$$

2) Complementary Gilbert Model:

$$S_{CG}(f_k) = S(f_k)|_{b=0} = \\ = \frac{\mu_{CG}\gamma_{CG}}{N}\frac{\mu_{CG}\gamma_{CG} + 2\frac{pqa^2}{(p+q)^2}\frac{y(\cos x - y)}{|1-y\mathrm{e}^{-jx}|^2}}{\mu_{CG}\gamma_{CG} + 2\frac{pqa^2}{(p+q)^2}\frac{y^N}{1-y^N}} + \mu_{CG}^2\delta(f_k) \quad (25)$$

3) Two-state Markov Model:

$$S_M(f_k) = S(f_k)|_{a=1,b=0} = \\ = \frac{\mu_M\gamma_M}{N}\frac{1-y^N}{1+y^N}\frac{1-y^2}{|1-y\mathrm{e}^{-jx}|^2} + \mu_M^2\delta(f_k) \quad (26)$$

4) Random Independent Model:

$$S_{RI}(f_k) = S(f_k)|_{a=b} = \mu_{RI}\gamma_{RI}/N + \mu_{RI}^2\delta(f_k) \quad (27)$$

VI. Identification

Identification of a data loss process is finding the best fit data loss model and its parameters. As the Gilbert-Elliott model is a hidden Markov model, the Baum-Welch algorithm [6], Viterbi path counting [7] or other procedures can be used for identification.

These methods are applicable for the simpler models too, however, they require higher computational capacity or lack in accuracy. We designed a method for identifying the simpler models of the Gilbert-Elliott model family. This method will be introduced in this section.

A. Spectral Parameters

We can define three spectral parameters using the following substitutions:

$$p + q \mapsto X \quad pq(a-b)^2 \mapsto Y \quad aq + bp \mapsto Z \quad (28)$$

Applying them we can write (23) as

$$S(f_k) = \frac{\frac{Z}{X}\left(1-\frac{Z}{X}\right)}{N}\frac{\frac{Z}{X}\left(1-\frac{Z}{X}\right)+2\frac{Y}{X^2}\frac{(1-X)[\cos 2\pi f_k - (1-X)]}{|1-(1-X)\mathrm{e}^{-j2\pi f_k}|^2}}{\frac{Z}{X}\left(1-\frac{Z}{X}\right)+2\frac{Y}{X^2}\frac{(1-X)^N}{1-(1-X)^N}} + \\ +(Z/X)^2\delta(f_k) \quad (29)$$

Notice that while there are 4 model parameters, there are only 3 spectral parameters. This makes it impossible to identify the Gilbert-Elliott model from its spectrum. However, as the simpler models have at most 3 parameters, they can be identified using the spectral parameters. We can calculate the spectral parameters from the spectrum as

$$X = \left(2 - 2\sqrt{H-1}\right)/(2-H) \quad Z = \mu X \quad (30)$$

$$Y = \frac{(S'(0) - S'(1/2))\,Z(X-Z)(2-X)X}{2(1-X)[X(S'(0)-S'(1/2)) + 2S'(1/2)]} \quad (31)$$

$$H = \frac{S'(0) - S'(1/2)}{S'(1/4) - S'(1/2)} \quad S'(f_k) = S(f_k) - \mu^2\delta(f_k) \quad (32)$$

B. Calculation of Simpler Model Parameters

The parameters of the Gilbert, complementary Gilbert and Markov models can be expressed from the spectral parameters using the definition of the spectral parameters and the conditions used to derive the model.

1) Gilbert Model: We get the X_G, Y_G and Z_G spectral parameters from (28) with $a = 1$. The model parameters:

$$p = \frac{X_G (X_G - Z_G)^2}{(X_G - Z_G)^2 + Y_G} \quad q = \frac{X_G Y_G}{(X_G - Z_G)^2 + Y_G} \quad b = \frac{Z_G (X_G - Z_G) - Y_G}{X_G (X_G - Z_G)} \tag{33}$$

2) Complementary Gilbert Model: We get the X_{CG}, Y_{CG} and Z_{CG} spectral parameters from (28) with $b = 0$. The model parameters:

$$p = \frac{X_{CG} Y_{CG}}{Z_{CG}^2 + Y_{CG}} \quad q = \frac{Z_{CG}^2 X_{CG}}{Z_{CG}^2 + Y_{CG}} \quad a = \frac{Z_{CG}^2 + Y_{CG}}{Z_{CG} Y_{CG}} \tag{34}$$

3) Markov Model: We get the X_M, Y_M and Z_M spectral parameters from (28) with $a = 1$ and $b = 0$. The model parameters:

$$p = X_M - Z_M \quad q = Z_M \tag{35}$$

C. The Proposed Identification Method

The proposed method is the extension of the procedure described in [2]. The method is the following:

1) Handling of the block-based data loss, obtaining of block indicator function.
2) Executing the procedure in [2]: spectral estimation, IDFT, auto-regressive model fitting.
3) If the third and further coefficients are sufficiently small, two-state Markov model or random independent model describes the process. This can be decided using the second coefficient.
4) Attempt of identification of the Gilbert model using the spectral parameters.
5) If the estimated parameters are outside of the range of parameters, attempt of identification of the complementary Gilbert model using the spectral parameters.
6) If the estimated parameters are outside of the range of parameters, other model describes the data loss.

Block-based data loss means that the samples are bundled into fixed size blocks and these blocks are either fully available or fully lost. In this case, it is sufficient to examine the availability of a single sample from each block.

Figure 3 illustrates the identification method. The boxes filled by yellow handle the block-based data loss, red boxes contain the procedure of [2] and the blue boxes show the addition to identify Gilbert and complementary Gilbert models. The method requires relatively low computational capacity and memory.

VII. Measurements

The theoretical investigations are completed by measurements, e.g., by measurement of packet loss over UDP connection and sound transmission over Skype. The UDP measurements will be introduced in this section.

UDP is a well-known, not reliable communication protocol between computers. In some real-time applications there is not enough time for the retransmission of lost packets. In other applications, the resource cost of reliable communication is too high. Thus, UDP protocol is worth examining.

In the measurement setup, the sender computer was sending packets to the receiving one. The receiving computer determined which packets were lost. In some measurements, the computers were in a LAN, in others, they were physically about 250 km apart. The computers were connected via Ethernet, VPN (Hamachi) or the shared internet connection of a mobile phone. In some measurements, as parallel traffic a large file was uploaded to a cloud storage. The size of the packets, the number of simultaneously sent packets and the time between the packets were also varied.

The measurements lasted 10 minutes long each. After the measurements, the indicator function were constructed as the availability of the received packets. The indicator function have been processed by the method described in section VI. The identification results are displayed in table I.

TABLE I: Identification results

Model	**Number of occurrences**
No data loss	499
Negligible data loss	238
Random independent	18
Two-state Markov	46
Gilbert	10
Other model	41
Total	852

Negligible data loss means that there were too few samples or too few lost sequences to perform an accurate identification. The cases that generated the other models were mostly deterministic losses caused by the overloading of the VPN software.

VIII. Conclusion

In this paper, not reliable communication was analyzed from a signal processing point of view. This supports the usage of radio channel or Internet connection for real-time transmission of measurement data or signals. We analyzed the data loss as a signal distorting phenomenon.

The Gilbert-Elliott data loss model family was presented. These models were the Gilbert-Elliott, Gilbert, complementary Gilbert, two-state Markov and random independent data loss. They were examined in time and frequency domain. We derived the data loss and availability rates, the reliability functions and the average sequence lengths for the models. The power spectral density of the models were also derived. We proposed an identification method for the simpler models in the family. Finally, some measurement results were presented. In further research the identification method can be refined or extended to include other models. The time or frequency domain properties of other data loss models can also be examined.

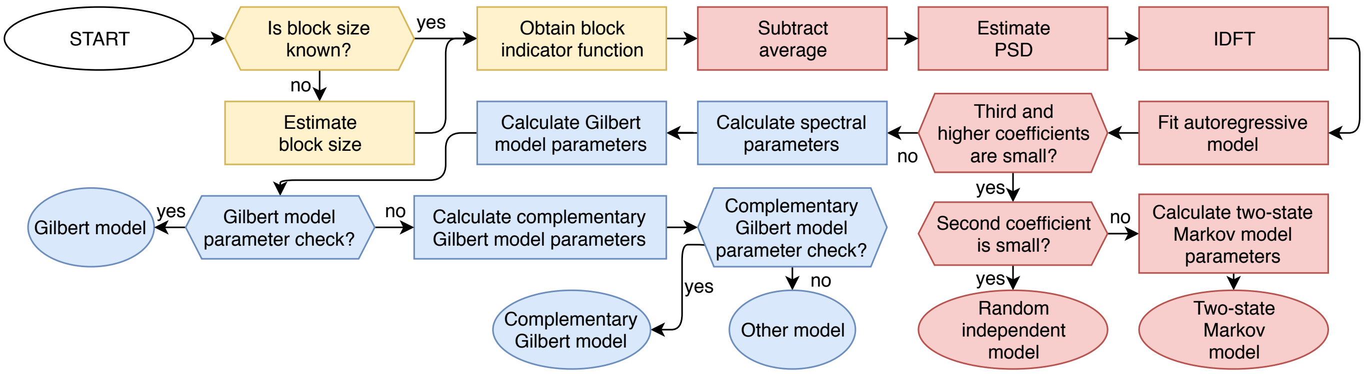

Fig. 3: The proposed identification method

Appendix

A. Reliability Function

For the derivation of the reliability function we introduce the probability of being in the A (B) state after k available samples:

$$A_k = \Pr\left(\forall i \in \{1, \dots, L\} : K_i = 1 \wedge X_k = \mathsf{A}\right) \quad (36)$$

$$B_k = \Pr\left(\forall i \in \{1, \dots, L\} : K_i = 1 \wedge X_k = \mathsf{B}\right) \quad (37)$$

where X_k is the state in the kth time instant. From the stationary distribution, for $k = 1$ these values are

$$A_1 = \frac{aq}{p+q} \quad B_1 = \frac{bp}{p+q} \quad (38)$$

For A_k and B_k the following recursive equation holds:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} ra & qa \\ pb & sb \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix} = \mathbf{P}_1 \begin{bmatrix} A_k \\ B_k \end{bmatrix} \quad (39)$$

with the starting probabilities:

$$\begin{bmatrix} A_k \\ B_k \end{bmatrix} = \mathbf{P}_1^{k-1} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} \quad (40)$$

For the expression of these probabilities we need the $\lambda_{1,2}$ eigenvalues of $\mathbf{P}_1$, which are

$$\lambda_{1,2} = \frac{ra + sb \pm \sqrt{[ra+sb]^2 - 4aby}}{2} \quad (41)$$

where $y = 1 - p - q$. Because the eigenvalues are different, according to [8] and [9] we can write $\mathbf{P}_1^k$ as

$$\mathbf{P}_1^k = \frac{\left(\lambda_2\lambda_1^k - \lambda_1\lambda_2^k\right)\mathbf{I} + \left(\lambda_2^k - \lambda_1^k\right)\mathbf{P}_1}{\lambda_2 - \lambda_1} \quad (42)$$

Using this result we can write the introduced probabilities as

$$A_L = \frac{aq\left[\lambda_2\lambda_1^{L-1} - \lambda_1\lambda_2^{L-1} + \left(\lambda_2^{L-1} - \lambda_1^{L-1}\right)(ra+pb)\right]}{(p+q)(\lambda_2 - \lambda_1)} \quad (43)$$

$$B_L = \frac{bp\left[\lambda_2\lambda_1^{L-1} - \lambda_1\lambda_2^{L-1} + \left(\lambda_2^{L-1} - \lambda_1^{L-1}\right)(qa+sb)\right]}{(p+q)(\lambda_2 - \lambda_1)} \quad (44)$$

The reliability function is the sum of these two probabilities: $R(L) = A_L + B_L$.

B. Average Available Sequence Length

Let A and B the lengths of the available sequences starting from A and B states. According to the law of total expectation the next equations hold:

$$A = 1 - ar - bp + ar(1 + A) + bp(1 + B) \quad (45)$$

$$B = 1 - bs - aq + bs(1 + B) + aq(1 + A) \quad (46)$$

$1 - ar - bp$ and $1 - bs - aq$ are the probabilities of immediately terminating the available sequences. The other terms are the probabilities of continuing the sequence in the same or in the other state than the current one. From these equations we can express A and B:

$$A = \frac{1 - by}{1 - (ar + bs) + aby} \quad B = \frac{1 - ay}{1 - (ar + bs) + aby} \quad (47)$$

With another application of the law of total expectation, the average length of available sequences is

$$E_{N1} = \left(\frac{aq}{p+q}A + \frac{bp}{p+q}B\right)\frac{1}{\mu} = \frac{aq + bp - ab(p+q)y}{(aq+bp)(1 - (ar+bs) + aby)} \quad (48)$$

References

[1] L. Sujbert and G. Orosz, "FFT-based spectrum analysis in the case of data loss", IEEE Transactions on Inst,rumentation and Measurement, pp. 968–976, Jan. 2016.

[2] L. Sujbert and G. Orosz, "Frequency domain identification of data loss models", Acta IMEKO, vol. 6. nr. 4., pp. 61–67, Dec. 2017

[3] O. Elliott E., "Estimates of error rates for codes on burst-noise channels", The Bell System Technical Journal, vol. 42, Sept. 1963.

[4] Gilbert E. N., "Capacity of a burst-noise channel", Bell System Technical Journal, vol. 39. nr. 5. pp. 1253–1265, 1960.

[5] P. Galko and S. Pasupathy, "The mean power spectral density of Markov chain driven signals", IEEE Transactions on Information Theory, vol. 27. nr. 6., pp. 746–754, Nov. 1981.

[6] Baum, L. E., Petrie T., Soules G. and Weiss N., "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains", Ann. Math. Statist., vol. 41. nr. 1., pp. 164–171, Feb. 1970.

[7] Davis R. I. A. and Lovell B. C., "Comparing and evaluating HMM ensemble training algorithms using train and test and condition number criteria", Pattern Anal. Appl. vol. 6. nr. 4., pp. 327–336, Feb. 2003.

[8] "Formula for matrix raised to power n", https://www.freemathhelp.com/forum/threads/55028-Formula-for-matrix-raised-to-power-n?p=227129\&viewfull=1#post227129, Accessed: 2018-10-18

[9] "Using Cayley-Hamilton theorem to calculate matrix powers", https://www.physicsforums.com/threads/using-cayley-hamilton-theorem-to-calculate-matrix-powers.671687/, Accessed: 2018-10-18