Adaptive Hidden Markov Modelling for Digital Wireless Channels

Huang Ruoheng (Harley)
harleyhwan@yahoo.com

***Abstract*—Modelling of digital wireless channels is a popular application in the usage of testing and analyzing error-control schemes, as well as high-layer wireless communication protocols. For instance, modelling of digital wireless channels gives the benefit of reducing the simulation timing cost significantly. This is absolutely crucial to the industry since the product cycle can significantly be reduced and a leading market can be secured. Similarly, this model is greatly helpful for communication standards design and evaluation.**

The employment of Markov or hidden Markov technologies, which are very popular in pattern recognition area, enables to greatly improve the generating of error models and this could increase accuracy and performance of the overall communication system. However, the limitation of the samples training process in Markov model results in the highly amount of the error sequences to be needed for this purpose. For generating the different generative models, the different descriptive models are needed for this purpose, and the problem is that, it is very time consuming to get a sample error sequence which also known as the descriptive model from the computer simulation which covers the whole transmission links. In order to reduce the numbers of sample error sequences used for training process in the generation process, a research called the adaptive-modelling have been carried out, and a method named adaptive-HMM has been suggested during the research semester. This proposed technique is able to generate an error sequence in any carries-to-interference ratio, which also known as CIR or radio transmission channel condition by using limited numbers of descriptive models. In this project, the statistical properties of digital wireless channels in 12 different in door EGPRS protocol radio transmission channel environments will be investigated in detail based on the error models, as well as a generative model based on hidden Markov and the proposed adaptive-generative model.

Keywords*：*Adaptive, Hidden Markov model, Wireless Channels

I. INTRODUCTION

A. Aims and objectives

A time-discrete digital channel comprises a complete communication chain, which include the transmitter, the physical transmission channel, and the receiver in the complex baseband. However, due to the complexities of this telecommunication process, errors have always been encountered in the digital wireless channels. And in order to study those telecommunication processes, channel models are introduced, for digital channels, those models are called the error models or the error sequences models, which are the Exclusive disjunction operation results of the input signals with the output signals in the format of ‘0’s and ‘1’s.

The error models have wide applications on the designing and performance evaluation of error-control schemes, as well as high-layer wireless communication protocols. One important application of error models is in wireless communication system performance simulations: by using generated error sequences model to replace to the previous model which contains different channel conditions and different physical layer techniques, the simulation timing cost reduced greatly. This is absolutely crucial to the industry since the product cycle can significantly be reduced and a leading market can be secured, so, this model is greatly helpful for communication standards design and evaluation.

However, those previous error models are gain from the simulation process which also includes the whole transmission process in different communication environment (different CIR), and this process is also very time-consuming. In order to get more error models which has the same statistics distributions in some aspects, the (hidden) Markov generative models has been introduced by several researchers, those kinds of generative models could produce the error models with very similar statistics features with the ideal descriptive error models in the limit CIR condition. So, in this project defines the problem of the study and formulates has been studied, and several hidden Markov generative models suggested by other researchers in generating error sequence in the fixed CIR condition are also been learned. Then the studies of the feature of descriptive error models is also given, which is very helpful on developing the adaptive models. As the result a methodology based on an existing hidden Markov generative model has been suggested, and this model is able to generate error sequence in any CIR condition, which called the adaptive Hidden Markov generative model. The following chapter has presents the results from the suggested adaptive hidden Markov generative model, as well as conclusion of the whole thesis which stating the significance of thesis work and possible improvements.

B. Background information

Communication channel can be modeled mathematically or physically by calculating the physical communication processes which could have modified the transmitted signal. For example, in the wireless communications, the channel can be modeled by calculating the reflection of every object in the environment, and a sequence of random signals, which represent the simulation of noise, could also be added in the channels. By doing these kinds of research about characters of wireless channels allowing us to use a common and simple model instead of the real-model, which gives a great help for the simulation works on knowing how the channel affects transmitted signals. And in this thesis, kinds of the efforts were carried out for optimizing the parameters of the channel as well as getting an accurate simulation result.

Wireless communication channels can be classified in two major categories. One category is the analog channels also known as the physical channel, this kind of the channel is mainly interesting in receiving the signals like signal strength, the noise and/or interference power, etc. So, the channel models for analog channels are emphasis on describing the fading characteristics of the received signal.

The other category is called the digital channels, in which we are interested in the number and distribution of error events in a sequence of bits or packets. A digital or a time-discrete channel comprises the complete transmission chain, including the transmitter, such as modulator, which convert the signals in a format which easier to transmit, the physical channel, and the receivers, such as demodulator, which convert those signals back to original information signals in the complex baseband. And in this digital wireless channels, errors always encountered due to the noise's effects in the transmit channels, and those errors are not independent but occur in bursts or clusters. So, the channel models for digital channels are also called the error models, which aim at describing the statistical properties of the underlying burst error sequences. Error models have wide applications to the design and performance evaluation of error-control schemes, as well as high-layer wireless communication protocols.

Error sequence is the comparison result which using exclusive disjunction between an input sequence and the output sequence of the digital channel, and this is based on either bit levels. An example of the error sequence is shown as Figure 1 below:

input sequence :	**000111010101100**
output sequence:	**001001101101101**
XOR output:	**001110111000001**

Figure 1 An example of getting the error sequence

Let's assume X_{in} is the input sequence while X_{out} stands for the corresponding output sequence, the binary error sequence *Eh* could be gain from the equation 1.1 below:

$$Eh(i) = X_{in}(i) \oplus X_{out}(i) \qquad (1)$$

($\oplus$: Exclusive Or, XOR. Where $1\oplus 1=0$, $1\oplus 0=1$, $0\oplus 0=1$, $0\oplus 1=1$)

There are two categories of the error models which are discussed in this thesis, namely the descriptive models and the generative models: descriptive statistics models or the descriptive models is a model which analyzes the statistical behavior or features of a selected error sequences which obtained directly from a real digital channel or a computer simulation of the overall communication link. A generative model is where a model specifies a mechanism or an algorithm which generating the error sequences with similar statistically behavior to the target sample error sequences. Generative models are parameterized mathematical models capable of generating error sequences that are statistically similar to those produced by descriptive models. Accurate and efficient generative models are significant helpful for the designing and performance evaluating of wireless communication protocols as well as error-control schemes. Comparing with a descriptive model, the main advantage of the generative model is that it can greatly reduce the computational effort or the time-cost for generating long error sequences, and therefore speed up simulations. [19]

In this paper, 12 different descriptive models which have been obtained from the computer simulation of the overall communication link in different CIR situations are considered as the reference target models, so the proposed generative models in this thesis should have the similar statically performance with the previous descriptive models.

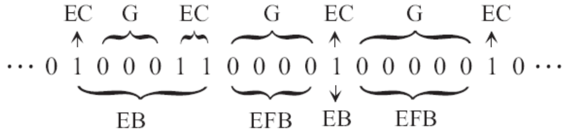

Figure 2 An extract from a hard error sequence.

Figure 2 shows an extract from a hard error sequence. In which, some statistics are been investigated during the whole thesis research process. A hard error sequence are represented by the binary sequences of '1' and '0', with a "1" denoting an error bit, and a "0" a correctly received bit. And in this hard error sequence, some statistics are noticed:

A gap (G) is defined as a string of consecutive zeros between two ones, having a length equal to the number of zeros.

An error cluster （EC） is a region where the errors occur consecutively, and has a length equal to the number of ones.

Threshold value (ε): the minimum length of EFB

An error-free burst (EFB) is defined as an all-zero sequence with a length of at least ε bits, where ε is a positive integer. Compared with a gap, an error-free burst has a ε minimum length of and is not necessarily located between two errors.

An error burst (EB) is a sequence of zeros and ones starting and ending with a "1," and separated from neighboring error bursts by error-free bursts. It should be observed that the minimum length of an error burst is 1 and the number of consecutive error-free bits within an error burst is less than ε. The following burst error statistics will be relative to the research:

1) G(m/g): the gap distribution (GD), which is defined as the cumulative distribution function (CDF) of gap lengths mg.

2) $P(0_{m0}/1)$: the error-free run distribution (EFRD), which is the probability that an error is followed by at least m0 error-free bits. The EFRD can be calculated from the GD. Obviously, $P(0_{m0}/1)$ is a monotonically decreasing function of m0 such that $P(0_{m0}/1)$ = 1 and $P(0_{m0}/1)$ $\to$ 0 as $m0 \to \infty$.

3) $P(1_{Ne}/0)$: the error cluster distribution (ECD), which is the probability that a correct bit is followed by mc or more error bits. Similar to $P(0_{m0}/1)$, $P(1_{Ne}/0)$ is also a monotonically decreasing function of Ne.

4) PEB(n): the error burst distribution (EBD), which is the CDF of error burst lengths n.

5) PEFB(m): the error-free burst distribution (EFBD), which is the CDF of error-free burst lengths m.

6) P(I; n): the block error probability run distribution (BEPRD), which is defined as the probability that a block of n bits will contain exactly I errors. This quantity is important for determining the performance of error-correcting schemes.

7) P(m; n): the block error probability distribution (BEPD), which is defined as the probability that a block of n bits will contain at least m errors. This could be considering as the cumulative distribution function of BEPRD which is also important for determining the performance of error-correcting schemes.

8) Δk: the bit error correlation function (BECF), which is defined as the probability of two error bits occurring at a distance of Δk bits apart. This quantity is of great usefulness for the design of bit interleavers.

C. *Adaptive model*

The idea of this master's project is to explore a totally new research area in wireless communication error modelling, which is called the 'adaptive' error modelling.[21] [22] [23] Because of the designing and performance evaluation of error-control schemes or wireless communication protocols needs thousands of the error sequence/models in different transmission quality (CIR) for the evaluation process. However, the currently methods is to obtain those error sequence by the simulation process includes the whole transmission process in different communication environment, this process could cost more than eight hours to get a simple error sequence on a high performance computer, and another drawback is that, this simulation process could only generate the certain error sequence under the given communication environment condition. So, a new solution is need, which could reduce the time significantly of generating those error sequences as well as more error sequence in more communication environment conditions could be produced, which this is the aims of the whole project.

There are 12 different descriptive models also known as error sequences used in the project, which are obtained from the simulation of the whole transmission process with the carrier to noise ratio (CIR) of 5,7,8,9,11,13,15,17,19,21,23 and 25dB respectively. The previous suggested generated model by other researchers could only generate the error sequences within the 12 CIRs which mentioned before, however, the proposed 'adaptive' process is a generative model which could generate an error sequence within any CIR condition between 5 to 25 dB. In order to suggest a solution for adaptive error model, several pre-works should be done, such as investigating the statistics for those 12 descriptive models, knowing the relations between CIR and different statistics, finally, a proper method could be suggested for the 'adaptive' requirement.

II. HIDDEN MARKOV GENERATIVE MODEL

The mother model has been selected, this model is based on J. G. Frias and P. M. Crespo's [4] proposed model, and in the following thesis, this model will be called HMM while the adaptive model with developed from HMM is call adaptive-HMM. Before deep into the adaptive method, the HMM should be introduced in details, the main idea of the HMM is generating a new error sequence with the same carrier-to-interference ratio as the sample error sequence or descriptive model, therefore, the generative model should have almost the same statistics feature as the descriptive model.

A. *Define the error bursts and error-free bursts*

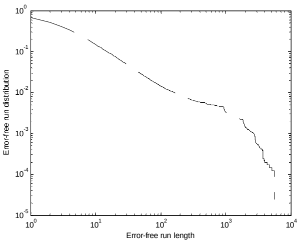

Figure 3 A typical error-free run distribution

Figure 3 shows a distribution of error-free intervals $P(0_{m0}/1)$ or the probability of numbers of '0's occurring between two '1's. The typical distribution of numbers of '0' is decreasing gradually, which is also similar to the real radio communications: errors occupies quite often, so the probability of a very long transmission without making any mistake is extremely small, to be noticed that, there is a 'plateau' in the distribution curve, and from the plateau in the shape of this graphs, it is concluded that the probability of encountering error-free intervals of lengths ranging from η i=300 to η f=1000 ,where η i, η f are the beginning and end of the plateau, respectively. From this fact that, this channel model could be qualitatively described into the well-differentiated modes of operation: good and bad. In the bad situation, the channel generates bursts with a density of errors inside the burst always quite large while in the good situation of operation the channel generates fewer errors with a density always small. Therefore, a very reasonable method of defining

---

Identify applicable sponsor/s here. If no sponsors, delete this text box. (sponsors) a burst has been gain from this fact: first of all, select a value of ε which η i< ε < η f; so, a error-free burst (EFB) is the continuous ‘0’s with the length of ‘0’s larger than ε; therefore, a error burst (EB) is the sequence of ‘1’s and ‘0’s which begins and ends with ‘1’s. After having defining the EFB and EB, EB is then partitioned in to blocks which will be introduced later.

B Partition of EBs

The basic idea behind the proposed HMM method is to partite all bursts encountered in the original error profile into classes, which have been previously defined and whose elements share some common characteristics.

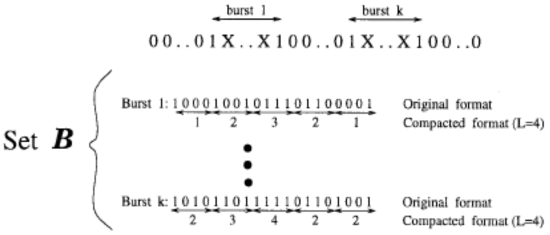

Figure 4 Example of bursts in a particular error profile. The original format and the compacted format (L=4) representation are shown

For the sake of clarity, let Figure 4.2 above be a particular error profile obtained from the radio channel under study. A ‘1’ or ‘0’ in the error profile means an error bit or a correct bit has been detected by the receiver. From this Figure4.2, one thing may be noticed that an error bursts could be defined as a sequence of ‘1’s and ‘0’s which begins and ends with ‘1’s, and the error bursts are located between two long error-free intervals which also known as the error-free bursts.

The identified bursts have been extracted from that sample error sequence, which has been demonstrated in set B, which contains all the error bursts. The next step is to partition into disjoint classes or called the burst classes, in such a way that the bursts in each class or training bursts share a common property, as is explained later. In the Figure 4.2 original format is the original error burst appears in the form of ‘1’ and ‘0’s, to get the compacted format, the length L is defined, in the sample, L is equal to 4, so the original error burst are divided into the sequence of small continuous blocks, those blocks could also be represented by the numbers of ‘1’s in this L-length block, such as by 12321 or 23422in Figure 4.2 was shown, which shows on the compacted format. Here e_i is the length of an EB which has been defined previously, and L is the length which to partition an EB and u_i means the numbers of blocks in this i th EB after partition. So U is the vector numbers of blocks in an EB of the whole error sequence.

$$e_i = u_i \times L \tag{2}$$

$$U = \begin{bmatrix} u_1 & u_2 & u_3 & \dots & u_i & \dots & u_n \end{bmatrix} \tag{3}$$

After an EB has been divided by the length L in to u blocks, Ω is the vector used to describe the numbers of error bits in each blocks of an EB.

$$\Omega = \begin{bmatrix} d_1 & d_2 & d_3 & \dots & d_i & \dots & d_u \end{bmatrix} \tag{4}$$

then, the peak number of errors (PNE) of a burst has been defined, which stands for the maximum value of ‘1’s in the block which the error burst has been partitioned in blocks already, in the Figure 4.2, the PNE for burst 1 and k are 3and 4 respectively. Let’s define $P_{1\times\mathrm{n}}$ for the vectors of PNEs, so p_i indicates the PNE of the i th EB after the EBs are partitioned by length L, so it also could be consider as the maximum elements in the i th EB’s Ω vector.

$$p_i = \max(\Omega_i) \tag{5}$$

$$P_{1\times\mathrm{n}} = \begin{bmatrix} p_1 & p_2 & p_3 & \cdots & p_i & p_{i+1} & \cdots & p_n \end{bmatrix} \tag{6}$$

C Baum-Welch training algorithm for HMM sub-models

The Baum-Welch algorithm is a generalised expectation-maximization (GEM) algorithm. It can compute maximum likelihood estimates and posterior mode estimates for the parameters, transition and emission probabilities of a HMM, when given only emissions as training data. The algorithm has two steps: First, calculating the forward probability and the backward probability for each HMM state; On the basis of this, determining the frequency of the transition-emission pair values and dividing it by the probability of the entire string. This amounts to calculating the expected count of the particular transition-emission pair. Each time a particular transition is found, the value of the quotient of the transition divided by the probability of the entire string goes up, and this value can then be made the new value of the transition.[2]

An overview of the Baum–Welch algorithm utilized for the parameterisation of the HMM is presented. The purpose is to make the algorithm ready for a quick use and if needed an excellent explanation and deduction of the algorithm can be found in [1].

D Concatenation of generated HMM sub-models

Finally, the HMM submodels have to be concatenated to form the HMM. Figure 4.3(a) shows a simple and reasonable method of doing the concatenation. It consists of an error-free state, which generates variable length error-free intervals with a probability given by the histogram of error-free interval lengths between training bursts. The zero state is then connected to super-states, each one of them representing an HMM. The transition probabilities give the a priori probabilities of generating a burst of class.

III. A PROPOSED ADAPTIVE-HMM

In order to reduce the cost for training process, the adaptive-HMM has been suggested. The basic idea of adaptive-HMM is using two sample error sequence in A dB and B dB to generate an error sequence at M dB (A<M<B), the flow chart is shown as the Figure 5 below

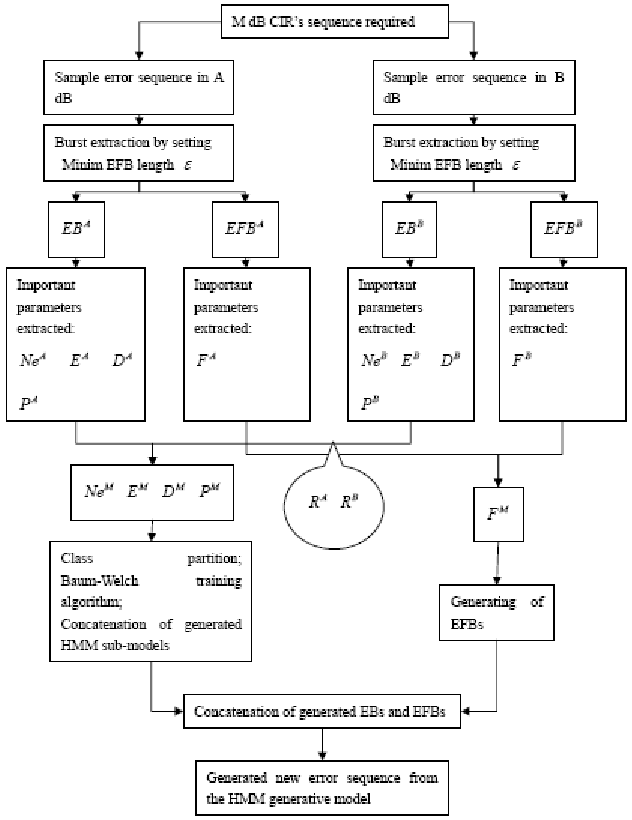

Figure 5 The flow chart of adaptive-HMM generative model

IV. RESULT

In order to scrutinize the Adaptive-HMM operation, according to the methodologies of the adaptive modelling, two reference error sequences are needed to be obtained for its parameterisation. Because only 12 descriptive models existed for evaluation, so, this adaptive model needs to generate an error sequence in the CIR condition which already existed and to achieve this, the underlying channel is tailored to a typical urban environment with carrier to interference ratio of 8 dB (M) which also means the Adaptive-HMM is going to generate an 8dB's error sequence with the reference error sequences 7 (A) and 9 (B) dB, and the mobile speed is 3km per hour, the data are transmitted using the protocol of Enhanced Data Rate for GSM Evolution (EGPRS) with blocks of 116 bits and a transmission rate of Fs=270.8 kb/s. This adaptive generating result is comparing with the 8dB descriptive model as well as the generating result of 8dB from hers mother model, the HMM.

The reference error sequence of 8dB has 15 million bits. It exhibits long error bursts interweaved with long error-free bursts. It has 4150 error bursts and 4149 error-free bursts with maximum lengths of 8235 and 6251 bits, respectively when the threshold value ε =900. We find n from Figure 5.1 which displays the EFRD. From its shape plateau, which could be found is that η_i = 800 and η_f = 1000 hold. So the chosen value of ε is 900.

For the sake of comparison, a HMM, and adaptive-HMM have been implemented. The number of classes used for the HMM is 12. In fact, no better improvement of the HMM statistics could be attained by increasing its classes number to more than 12, and the total number of states is 400. The number bits which should represent each block in the error bursts is chosen to be 103 bits. And the same parameterisation is selected for adaptive-HMM. However, the only difference is that adaptive-HMM is using 7dB and 9dB's error sequences as the examples while the HMM is using 8dB's as the example error sequence.

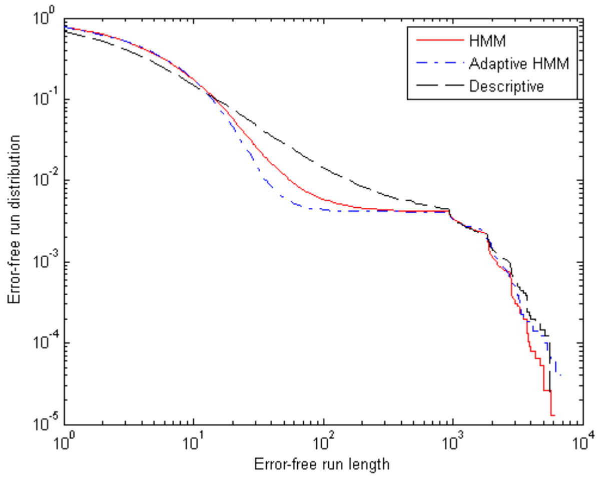

Figure6. EFRDs of the adaptive model, HMM and descriptive model

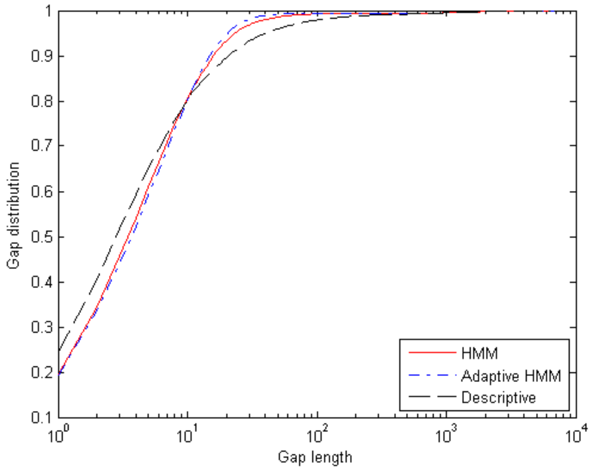

Figure7. GDs of adaptive model, HMM and descriptive model

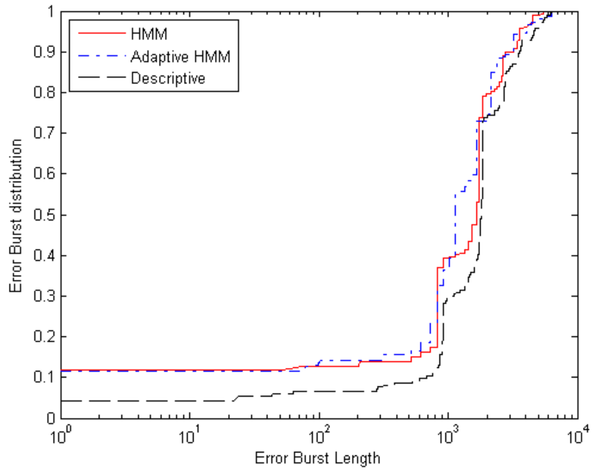

Figure8 EBDs of the adaptive model, HMM Mode and descriptive model

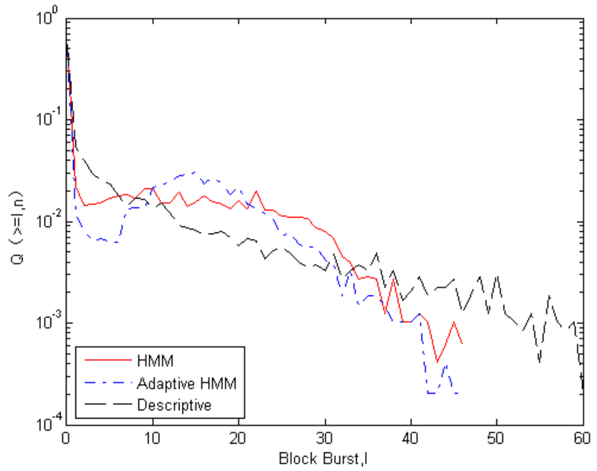

Figure11 BEPRDs of the adaptive model, HMM and descriptive model (n=116)

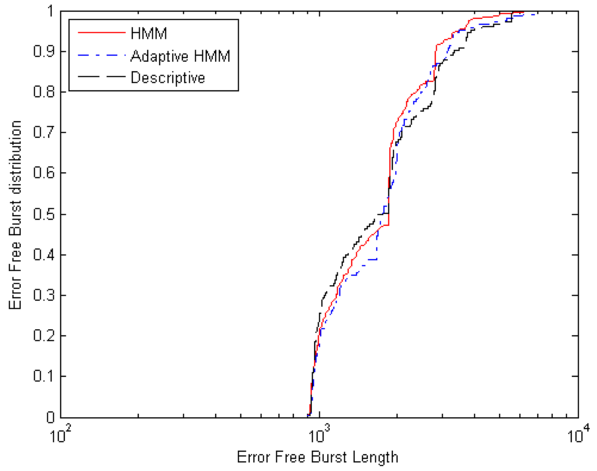

Figure9 EFBDs of adaptive model, and HMM model and descriptive model

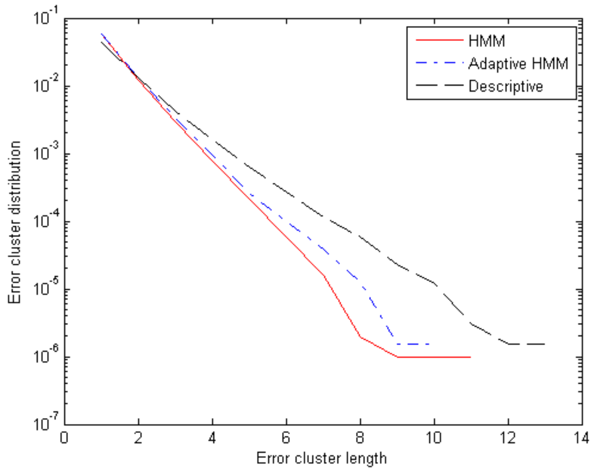

Figure 12 ECDs of adaptive model, HMM and descriptive model

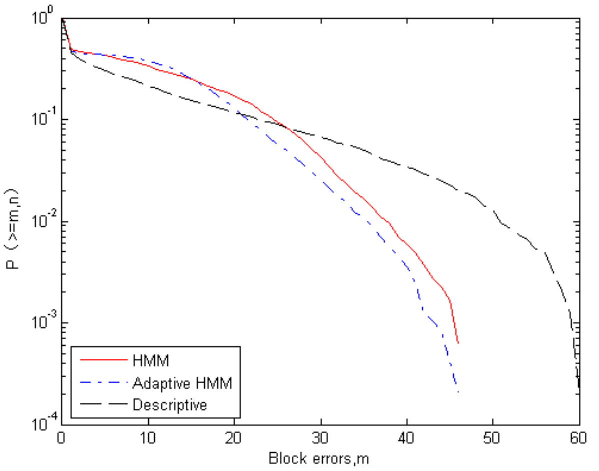

Figure10 BEPDs of the adaptive model, HMM and descriptive model (n=116)

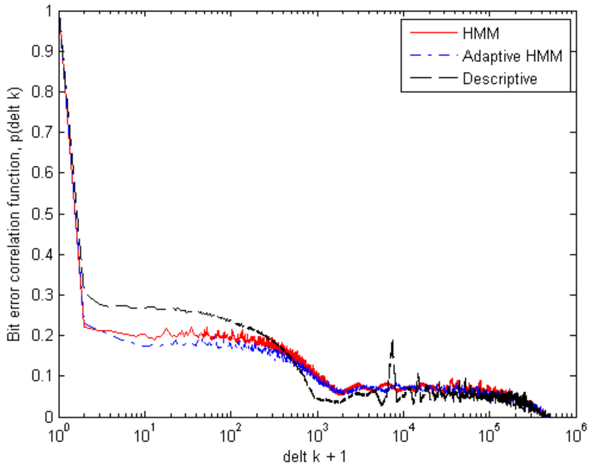

Figure13 BECFs of adaptive model, HMM and descriptive model

Figures list above (Figures 6-13) illustrate the behavior of the burst error statistics, which have been mentioned in Chapter 1, the introduction part, for the descriptive model with carrier to interference ratio of 8 dB which are the curves in black, the HMM generative, curves represents in red and the adaptive-HMMs curves which shows in blue addressed before.It is apparent that all the statistics of adaptive-HMM are very similar as HMM, however, both of them are having some difference in the statistics of EFRD, EBD and ECD compared with descriptive model. The highly similarity of adaptive-HMM from 7dB and 9dB's sample error sequences with the HMM from 8 dB's sample error sequence shows that a very good result of adaptive generative model in generating an 8dB's error sequence has been achieved by the proposed adaptive methods. Although the idea adaptive generative model should have the very high similarity in all the statistics with the descriptive model, due to the drawback of its mother model: HMM, the adaptive-HMM couldn't have even better results that HMM. In general, this proposed adaptive modelling has successfully full up the suggested requirement.

V. CONCLUSIONS AND FUTURE WORK

In this study, a procedure for the parameterisation of adaptive-generative models based on the interconnection of hidden Markov submodels (HMM) which from sample error sequences has been introduced. The hidden Markov method is the well-studied issue for speech recognition of isolated words. However, in here, instead of dealing with words, this model deals with error bursts and error free bursts, the final goal of this model is to generate bursts rather than to recognize words. Comparing to the previous HMM, this adaptive-HMM is particularly suitable to simulate error profiles for the un-known transmission situations (CIR), rather than the sample error sequences in the fixed transmission situations (CIR) are required for the HMM.

To corroborate the accuracy of this model, three examples of indoor EGPRS (Enhanced Data Rate for GSM Evolution) protocol links have been shown. The resulting HMM for the first example (carrier to interference ratio 8dB) consists of 12 classes. And the adaptive-generative model for the second case (carrier to interference ratio 8dB required, and the samples are the 7dB and 9dB error sequences) also divided into12 classes. Those two generative models are compared with the descriptive model which obtains from the simulation of whole communication links in CIR of 8dB. The error sequences generated by the adaptive-HMM is extremely alike as the result from HMM and it is statistically close to the corresponding underlying error process, showing a great accuracy.

However, due to the lack of time, some ideas have been suggested to prove and enhance this adaptive modelling research, but don't have enough time to work them out, so those ideas are suggested as the future works to do which have been listed below:

1) The curve fitting of relations of some important parameter for HMM in different CIR has been done in chapter 4.2. However, because of the lack of the time, in the implementation of adaptive-HMM, the average fitting has been used. Maybe the adapting of curve fitting of those parameters could be used in the adaptive models to improve its' accuracy.

2) To use other wireless communication systems to get the sample error sequences rather than the EGPRS like WiMax, LTE. Then comparing these results, to see if the results for those protocols are as accuracy as the one for EGPRS.

3) The drawback of this model maybe due to its' mother model. The accuracy could be improved, by using other models to enhance the performance such as DPBGM [13] and DEPHMM [12]. However, the potential difficulty of hierarchical or layered hidden Markov models could be caused by the complexity of the states and substates in their models.

ACKNOWLEDGMENT

My grateful thanks to my supervisor, Dr Cheng-Xiang Wang, for his support throughout this research project and to MSc Omar whose generous suggestions helped make changes and improvements in this project.

REFERENCES

[1] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications on speech recognition", Proc. IEEE, vol. 77, no. 2, pp. 257–285, 1989.

[2] L. R. Rabiner, B. H. Juang, "An introduction to hidden Markov models", IEEE ASSP Mag., vol. 3, no. 1, pp. 4–16, Jan. 1986.

[3] Y Ephraim, N. Merhav, "Hidden Markov Processes", IEEE trans. Inform. Theory, vol. 48, June 2002.

[4] J. G. Frias and P. M. Crespo, "Hidden Markov models for burst error characterization in indoor radio channels", IEEE Trans. Veh. Technol., vol. 46, no. 6, pp. 1006--1020, Nov. 1997.

[5] W. Zhu, J. G. Frias, "Stochastic context-free grammars and hidden Markov models for modelling of bursty channels", IEEE Trans. Veh. Technol., vol. 53, no. 3, pp. 666--676, May. 2004.

[6] W. Turin, "Digital Transmission Systems: Performance Analysis and Modelling", New York: McGraw-Hill, 1999.

[7] W. Turin, M. M. Sondhi, "Modelling error sources in digital channels," IEEE J. Select. Areas Commun., vol. 11, pp. 340–347, Apr. 1993.

[8] N. Nefedov, "Discrete channel models for wireless data services", Proc. IEEE VTC'98, Ottawa, Canada, pp. 683—687, May. 1998.

[9] N. Nefedov, "Generative Markov models for discrete channel modelling", in Proc. IEEE PIMRC, Helsinki, Finland, pp. 7–11, Sep. 1997.

[10] N. Nefedov, "Realistic discrete channel models for error control strategies evaluation", in Proc. SBT/IEEE ITS, Sao Paulo, Brazil, pp. 37–42, Aug. 1998,.

[11] N. Nefedov, "Discrete channel models for evaluation of error control strategies in CDMA", in Proc. IEEE ISSSTA, Sun City, South Africa, pp. 245–249, Sep. 1998.

[12] O. S. Salih, C.-X. Wang, D. I. Laurenson, "Double embedded processes based hidden Markov models for binary digital wireless channels", Proc. ISWCS'08, Reykjavik, Iceland, Oct. 2008.

[13] H. H. Bui, D. Q. Phung and S. Venkatech, "Hierarchical hidden Markov models with general state hierarchy", Proc. AAAI'04, San Jos, CA, USA, pp. 1–6 , Jul. 2004.

[14] N. Oliver, A. Garg and E. Hovitz, "Layered representations for learning and inferring office activity from multiple sensory channels", Comp. vision and Img. understanding, vol. 96, pp. 163–180, Feb. 2004.

[15] A. Beverly, K.S. Shanmugan, " Hidden Markov models for burst errors in GSM and DECT channels", in Proc IEEE Globcom'98, 8-12, PP:3692 – 3698, Nov. 1998.

[16] B.M. Hochwald, P.R. Jelenkovic, "State learning and mixing in entropy of hidden Markov processes and the Gilbert-Elliott channel", IEEE Trans. Inform. Theory, vol. 45, no. 1, PP. 128 – 138, Jan. 1999.

[17] C.-X. Wang, “Sum-of-sinusoids mobile fading channel models with applications to error models”, PhD thesis, Aalborg University, Aalborg, Denmark, 2003.

[18] E. N. Gilbert, “Capacity of a burst-noise channel”, Bell Syst. Tech. J.,vol. 39, pp. 1253–1265, Sept. 1960.

[19] B. D. Fritchman, “A binary channel characterization using partitioned Markov chains with applications to error correcting codes” , Ph.D. dissertation, Lehigh Univ., Bethlehem, PA, 1966.

[20] http://en.wikipedia.org/wiki/Baum-Welch_algorithm

[21] C. R. Anderson, P. Domingos and D. S. Weld, “Relational Markov Models and their Application to Adaptive Web Navigation”,Dept. of Comp. Sci. & Eng. University of Washington.

[22] F. Gales, Y. Heights, “Cluster adaptive training of hidden Markov models”, IEEE Transactions on Speech and Audio Processing, 2000.

[23] J.T. Chien, “Adaptive hierarchy of hidden Markov models for transformation-based adaptation”, Speech Communication, vol 36, pp. 291 – 304, Mar. 2002.

[24] http://en.wikipedia.org/wiki/Markov_chain