PERFORMANCE COMPARISON OF HMM DISCRETE CHANNEL MODELING IN CDMA LINKS

Debjani Mitra
Associate Professor (Electronics & Instt. Deptt.)
Indian School of Mines University,
Dhanbad- 826004, Jharkhand, India

Rakesh Ranjan
J R F (Electronics & Instt. Deptt.)
Indian School of Mines University,
Dhanbad- 826004, Jharkhand, India

ABSTRACT
The performance comparison of two Markov models namely the Baum Welch Algorithm based HMM and Semi Hidden Markov Model has been evaluated for a DS-CDMA link in this work. The simulation includes the effects of AWGN, Multipath and Multiple Access Interference. Validation includes a comparison of the run-length statistic for the original and regenerated error sequence from estimated models. The SHMM approach is seen to be capable of developing a more accurate model as compared to the BWA. The length of number of symbols processed at a time does not affect the accuracy in both the methods.

Categories and Subject Descriptors
I.6.5 [**Simulation and Modeling**]: Modeling Methodologies;
I.6.6 [**Simulation and Modeling**]: Simulation Output Analysis;
H.3.4 [**Information Storage and Retrieval**]: Performance Evaluation (efficiency and effectiveness)

General Terms
Performance, Design, Algorithm.

Keywords
Discrete Channel Model, CDMA, Hidden Markov Model, Baum Welch Algorithm, Run length vector, Log likelihood

1.INTRODUCTION
CDMA technology based on spread spectrum system has strongly linked as a common platform on which 2G and 3G wireless systems have developed, and currently even in the roadmap to 4G wireless, CDMA has emerged as a strong contender to be a significantly rapidly growing technology. The assignment of the same time and frequency allocation in a given RF band with separation in the code domain via unique orthogonal pseudo-random spreading leads to several advantages. It not only enables the reuse of frequencies in every cell, but leads to other benefits like mitigation of multipath fading and interference, soft handoff capability, the ability to exploit voice activity and increase in spectrum usage[1]. CDMA systems are widely deployed and also serve as building blocks for more advanced systems.

The performance evaluation of CDMA system through simulation has been widely popular in the literature [2, 3, 4]. In the simulation of any digital communication system especially in wireless environment, one of the most computationally exhaustive blocks is the modeling and analysis of the multipath fading mobile radio channel. Several performance measure evaluation of the system such as BER, dynamic signal processing blocks, channel correlation properties, network protocol designing, error control coding, handoff algorithms, estimation of packet error rates etc are highly dependent on the channel modeling employed. Broadly we have the waveform channel model (WCM) and Discrete Channel Model (DCM) approaches for this purpose. Basically DCM is a high level abstraction of the physical WCM and this reduces the computational burden a lot reducing the simulation time as well. The discrete channel models are simulated at symbol rate while the waveform level simulation operates on sample-by-sample basis which is 8 to 16 times the symbol rate [5]. In DCM, the main interest is to model the temporal correlations that create burst errors. Such generated error sequences are used to capture the channel as a finite state system. Thus the simulation of the whole physical channel is entirely avoided making the approach computationally efficient and appropriate in the context of all digital transmissions. Gilbert [6] had initially proposed the two-state model for the discrete channel while Fritchman’s state partitioned model [7] also gained wide usage in this area. In this model there is more than one good state and only one state has error with probability of one. Fritchman partitioned the state space into n good states and N-n bad states. The different partitioning schemes are provided in [8].

Hidden Markov Model (HMM) is one of the most popular amongst all DCM approaches being used in wireless channel modeling. Even otherwise, HMM is an established powerful mathematical model with a wide range of other applications including signal processing, speech recognition, pattern recognition, wavelets, queuing theory and others [9]. It has the capability to model a set of ordered observed data generated by any unobservable statistical process. Several types of results based on error patterns are analyzed using HMM technique in different wireless system applications as discussed in [10, 11, 12]. An approach for constructing HMM model for burst errors in GSM and DECT channels is provided in [13]. A HMM technique for a fast and accurate simulation of error patterns in wireless communication systems is presented in [14]. To train HMM for discrete channel modeling a hybrid approach using genetic algorithm (GA) and simulated annealing (SA) is proposed in [15].

In this work we investigate specifically the HMM based discrete channel modeling approach for the analysis of a CDMA system with respect to some of its system parameters like spreading factor, number of interferers, ratio of bit energy to noise PSD, etc. We also explore the validity of two of its important estimation techniques namely, the Baum-Welch Algorithm (BWA) and Semi-Hidden Markov Model (SHMM) and present a comparative analysis of these two schemes with respect to the CDMA system model. Specifically the log likelihood and mean square error (MSE) for the predicted model have been used as the comparison criteria.

The rest of the paper is organized as follows. In section II the two popular methods of parameter estimation for Hidden Markov model are discussed. Section III presents the application of these two approaches in a CDMA system model along with the simulation details used for generating the error sequence for Markov model validation and implementation. Some simulation results are depicted in section IV and finally conclusions are summarized in section V.

2.ESTIMATION OF HMM PARAMETERS

A Markov model representation of a discrete channel is strongly dependent on the Transmitter/Receiver pair as well as the underlying waveform channel that is a part of the communication link. The estimation results will strongly be dependent on the specific system being modeled. Even for the same system, several parameters may play a role. To study this aspect, two different estimation techniques are investigated and for a specific CDMA system the comparison results are shown. The flow graph for BWA and SHMM is shown in Figure 1 and Figure 2.

2.1 Baum-Welch Algorithm

In this section we represent the Baum-Welch algorithm (BWA), the most efficient and prevalent method for the estimation of three important parameters, state transition matrix $[A]$, error generation matrix $[B]$ and π_i (expected number of times in state S_i at time t = 1). The error sequence is consisting of the combination of 0's and 1's. The implementation of BWA involves the computation of the "forward variable" and the "backward variable". The calculations of the forward and backward variable are described in the following steps [5].

1. Assume the initial model $\Gamma = [A, B]$
2. With $\Gamma = [A, B]$ as the model, we first define the forward variable as the probability of partial observation sequence in state S_i at time t given the model Γ, i.e.

$$\alpha_t(i) = \Pr\left[O_1, O_2, \ldots, O_t, s_t = i \middle| \Gamma\right] \quad (1)$$

and , then the backward variable

$$\beta_t = \Pr\left[O_{t+1}, O_{t+2}, \ldots, O_T \middle| s_t = i, \Gamma\right] \quad (2)$$

for $t = 1,2,...T$ and $i = 1,2,....N$.

Forward variables:

Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1), \qquad i = 1,2,...N \quad (3)$$

Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \quad (4)$$

where, $1 \le t \le T-1, \quad 1 \le j \le N$

Termination:

$$\Pr[O|\Gamma] = \sum_{i=1}^{N} \alpha_T(i) \beta_T(i) \quad (5)$$

As,

$$\sum_{i=1}^{N} \alpha_T(i) = \sum_{i=1}^{N} \Pr\left[O_1, O_2, \ldots, O_T, s_T = i \middle| \Gamma\right] = \Pr[O|\Gamma] \quad (6)$$

where, $O = [O_1, O_2, \ldots, O_t \ldots O_T]$ is an error sequence obtained through the simulation.

Backward variables:

Initialization:

$$\beta_T(i) = 1, \quad i = 1,2,...N \quad (7)$$

Induction:

$$\beta_t(i) = \sum_{j=1}^{N} \beta_{t+1}(j) b_j(O_{t+1}) a_{ij}, \quad (8)$$

where, $1 \le t \le T-1, 1 \le j \le N$

3. Next compute $\gamma_t(i)$:

$$\gamma_t(i) = \Pr[s_t = i | O, \Gamma] = \frac{\alpha_t(i)\beta_t(i)}{\Pr[O|\Gamma]}, i = 1,2,\ldots\ldots, N \tag{9}$$

4. Next compute $\xi_t(i, j)$:

$$\xi_t(i,j) = \Pr[s_t = i, s_{t+1} = j | O\Gamma] = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1})\beta_{t+1}(j)}{\Pr[O\Gamma]} \tag{10}$$

5. Now the parameters are estimated as follows:

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{11}$$

and

$$b_j(e_k) = \frac{\sum_{t=1 | O_t = e_k}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)} \tag{12}$$

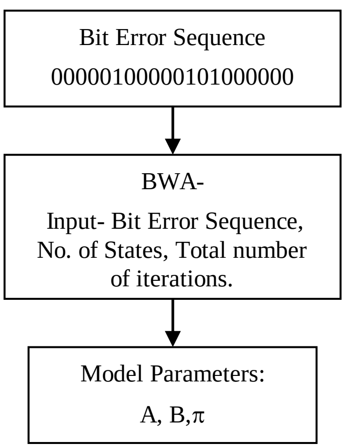

Figure 1: (Flow graph – BWA)

2.2 Semi Hidden Markov Model

The motivation for the Semi Hidden Markov Model is that it leads to a substantial reduction in the time required to derive the model from simulated data. The input data to this model is in the form of a runlength vector and hence the length of input sequence is greatly reduced. Only two parameters out of three as in the case of BWA are required to be estimated, namely state transition matrix $[A]$ and π_i (expected number of times in state S_i at time t = 1). Similar to BWA, it also involves the calculation of forward and backward variables.

3. THE CDMA MODEL

CDMA is an interference limited system and the near-far effect is also a big obstacle in this system. Hence a perfect power control is assumed here and at receiver each signal has the equal average power. The simulation of CDMA system also includes the effect of Additive White Gaussian Noise (AWGN). Users other than the desired user act as interference. As shown in Figure 3, the effects of the simultaneous presence of desired user signal, interfering user signal and AWGN noise has been considered in the channel.

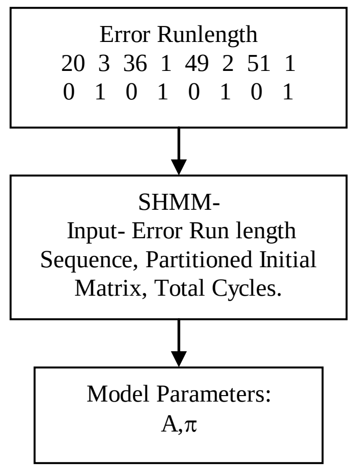

Figure 2: (Flow graph – SHMM)

The desired as well as interfering users use the Direct-sequence spread spectrum (DS-SS) followed by BPSK modulation before the transmission of the signal into the wireless channel. The effect of multipath delay in a Rayleigh fading environment is incorporated in the simulation. The delays of each multipath component are limited to integer multiples of chip duration. For the modulation of user's symbol, Binary Phase Shift Keying (BPSK) scheme is used. The user's information signal is multiplied by the user's spreading sequence to produce a spread spectrum signal, which is then transmitted into the channel. Circular shifted PN- sequences are used for the spreading and dispreading operation. A simple correlation receiver is assumed. At the receiver, it correlates the incoming signal with the user's spreading signal. Then user's information signal pass through the correlation receiver and simultaneously the task of rejecting the interfering users is also done by the receiver.

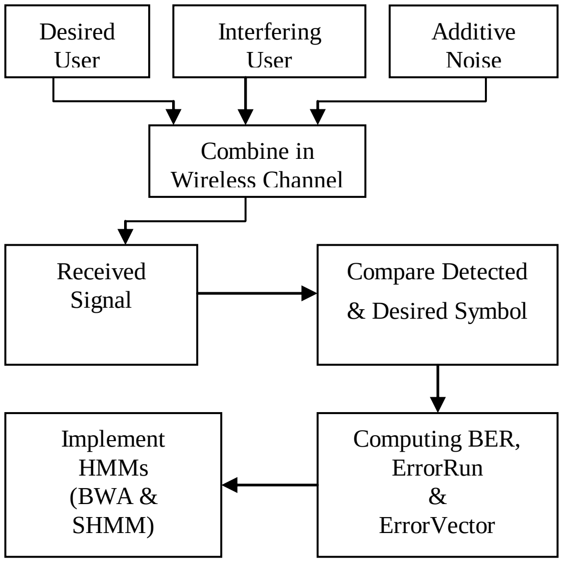

Figure 3: (CDMA Model)

The desired signal and all MAI (Multiple Access Interference) signals are assumed to be chip synchronized at the receiver. The detected symbols and the desired symbols are compared and then Error Run and Error Vector are generated. These are used for the parameter estimation of an HMM model that would identify the CDMA communication link with respect to the error sequence generated. The simulation results for the process have been discussed in the next section.

4. SIMULATION RESULTS

The error sequence generated by the CDMA model is used to fit HMM model using two approaches: the BWA algorithm based HMM and the Semi- Hidden Markov model. After the fitting, the new error sequence are regenerated by these estimated HMM models. The original error sequence and the regenerated error sequence are compared to establish the superiority, if any, of one method over the other. In the simulation, the spreading factor (SF) can be taken in form of $2^n - 1$, where n is an integer less than or equal to 12. The number of Interfering users has been varied from zero to one less than the spreading factor (i.e. SF- 1). In this particular work, a 3-state Markov model has been assumed. The simulation was performed with 100000 symbols processed for a particular value of the ratio of bit energy to noise PSD of 5 dB. Blocks of 1000 symbols are serially processed at a time for maintaining a reasonable computational efficiency. The Baum-Welch algorithm was tested upon by initial assumption of the following state transition matrix $[A_{in}]$, error generation matrix $[B_{in}]$ and $\pi_{i_{in}}$ as follows:

$$[A_{in}] = \begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.15 & 0.80 & 0.05 \\ 0.15 & 0.10 & 0.75 \end{bmatrix}$$

$$[B_{in}] = \begin{bmatrix} 0.90 & 0.85 & 0.80 \\ 0.10 & 0.15 & 0.20 \end{bmatrix}$$

$$\pi_{i_{in}} = \begin{bmatrix} \mathbf{0.40} & \mathbf{0.30} & \mathbf{0.30} \end{bmatrix}$$

While for the Semi Hidden Markov model no initial assumptions are required. The error sequence generated by the CDMA model is directly applied to the SHMM algorithm. The log likelihood function log ($\Pr[O|\Gamma]$) is determined in each of the iterations of the estimation algorithms to test and compare the error sequences generated by the CDMA model. The results are as shown in Figure 4 and Figure 5 for the BWA and SHMM respectively. The figures show the variation of the log-likelihood function values with the number of interferers (NoI). For both the BWA and SHMM the number of iteration is limited to 20, as the log-likelihood values become constant after few numbers of iterations. During the simulations, in general it is found that the log-likelihood value become more negative as the number of interferers is increases, due to increases in the probability of occurrence of error. Although the simulation results here have been illustrated for 14 interferers but the analysis is easily extendable for more number of interferers. The increase in the number of interferers increases the length of spreading factor and this increase the simulation time.

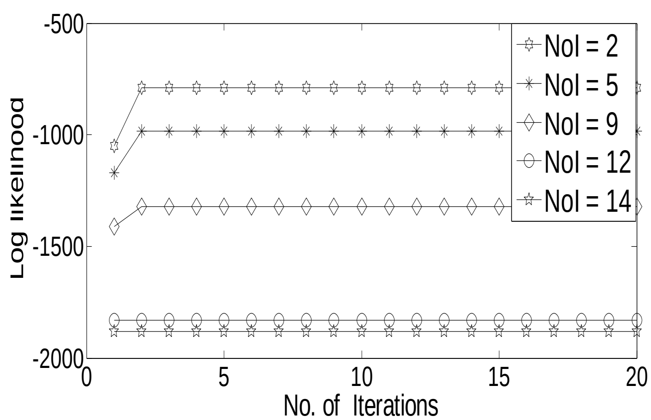

Figure 4: (Log-likelihood-BWA)

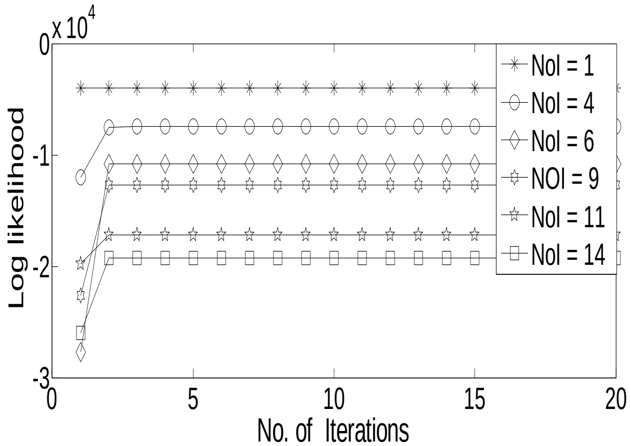

Figure 5: (Log-likelihood-SHMM)

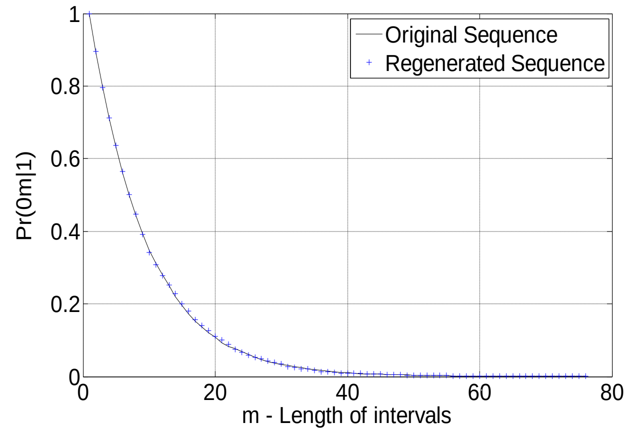

Figure 7: (Pr (0^m|1) vs. m -SHMM)

In Figure 6 and Figure 7, the two error sequences for the same input conditions to the CDMA model are compared for Pr (0^m|1) vs. length of interval (m) for the BWA and SHMM respectively. The notation (0^m|1) denotes the occurrence of m or more consecutive error free transmission followed by one error. The error sequence regenerated by the SHMM algorithm has close match with the original sequence. In the BWA algorithm, the error sequences have a relatively poorer match. The mean square error between the original error sequence and regenerated error sequences is calculated for the different number of interferers and is depicted in Figure 8 and Figure 9 for BWA and SHMM respectively. The order of mean square error for the BWA is 10^{-2} while for the SHMM the order of mean square error is 10^{-4}, which is very smaller than BWA case. Both the approaches are efficient, as the values of mean square error are very less. But, the SHMM algorithm gives better match between the two error sequences.

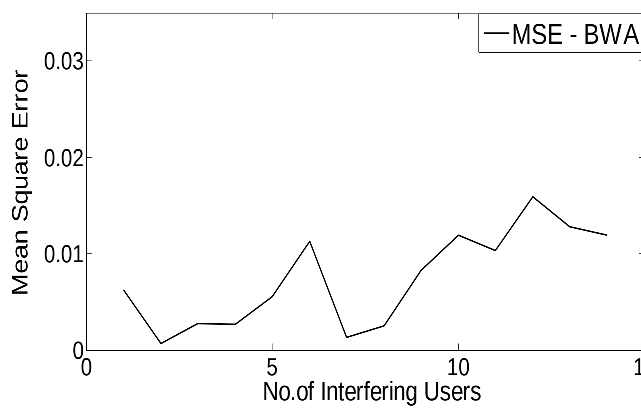

Figure 8: (MSE curve - BWA)

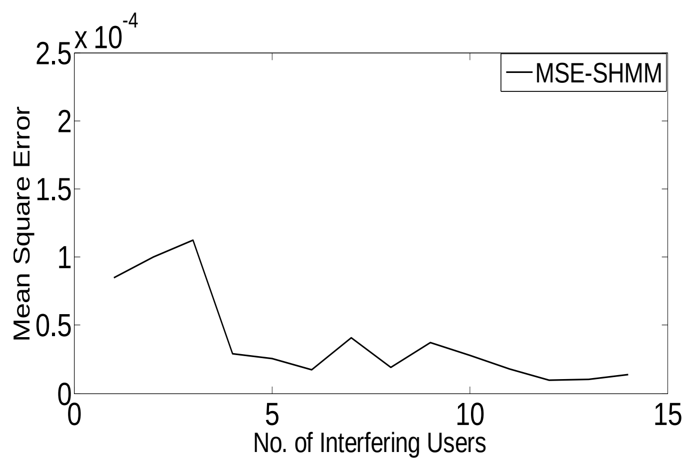

Figure 9: (MSE curve - SHMM)

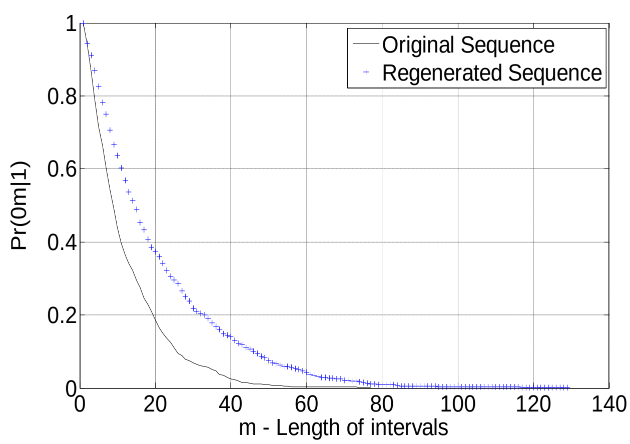

Figure 6: (Pr (0^m|1) vs. m -BWA)

The effect of the number of symbols on the MSE is shown in the Table 1 and also in Figure 10 & Figure 11. The number of symbols is varied from 10000 to 100000 with the interval of 10000. Table 1 shows the variation of MSE with respect to the no. of symbols for both BWA and SHMM. It is observed from that for the BWA algorithm, the order of MSE is 10^{-2} , while for SHMM it is 10^{-5}. After a number of simulations, it is found that for BWA the MSE is varying in between 0.003 to 0.025, while for SHMM it is in between 8.07×10^{-6} to 3.5×10^{-4} . As the order of MSE is lesser in SHMM, its performance is superior irrespective of number of symbols taken for the simulation in comparison to BWA.

Table 1 Comparison between BWA & SHMM

(In terms of Number of Symbols and MSE)

Number of Symbols	MSE - BWA	MSE - SHMM
10000	1.53×10^{-2}	5.99×10^{-5}
20000	1.98×10^{-2}	3.45×10^{-5}
30000	3.69×10^{-3}	6.75×10^{-5}
40000	1.09×10^{-2}	7.72×10^{-5}
50000	2.00×10^{-2}	2.13×10^{-5}
60000	1.15×10^{-2}	5.98×10^{-5}
70000	8.56×10^{-3}	3.55×10^{-5}
80000	2.49×10^{-2}	2.55×10^{-5}
90000	1.69×10^{-2}	7.96×10^{-6}
100000	1.41×10^{-2}	2.58×10^{-5}

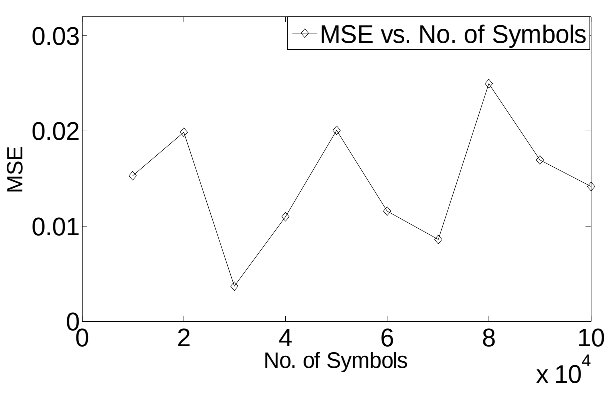

Figure 10: (MSE vs. No. of Symbol – BWA)

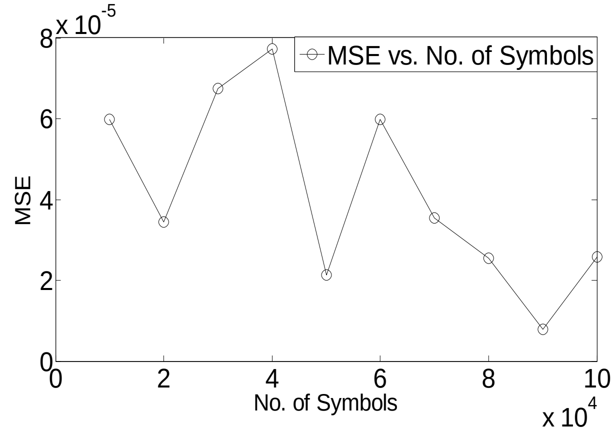

Figure 11: (MSE vs. No. of Symbol – SHMM)

5. CONCLUSIONS

The performance comparison of the computationally efficient HMM Discrete Channel Model has been evaluated for a DS-CDMA link in this work. The simulation includes the effects of AWGN, Multipath and Multiple Access Interference to generate the error sequence in developing the models. The performances of two popular Markov models namely the Baum Welch Algorithm based HMM and Semi Hidden Markov Model have been specifically compared. Validation includes a comparison of the run-length statistic for both the original error sequence and the regenerated error sequence produced by the estimated models. Although both the approaches are reasonably efficient, but testing through several simulations it is observed that SHMM is in all aspects superior to BWA in developing a more accurate model. The length of number of symbols processed at a time does not affect the accuracy with which the channel can be predicted by both the methods. The implementation and validation of the HMM models shows the capability of learning the stochastic mapping of the discrete-time channel input to the output for a CDMA link operating in a typical multipath fading channel. This clearly opens further scope of investigations in applications related to dynamic channel modeling applications in error control coding and handoff algorithms.

6. REFERENCES

[1] Rappaport, T. S. 2004. Wireless Communications. Second Edition, 2004, 472-487

[2] Khan, M. R. H. et al. 2006. “Comparative Analysis of CDMA based wireless Communication under Radio Propagation Environment”, Proceedings of IEEE 1st International Conference on Wireless Broadband and Ultra Wideband Communications (AusWireless’06), Sydney, Australia ,March 13-16,2006, 116-120.

[3] Balaban, P., Li, D., and Turin, W. 1998. "Performance evaluation of the reverse link of a CDMA cellular system using simulation," in Proc. IEEE GLOBECOM'98, 1998, 231-236.

[4] Balaban, P., Li, D., and Turin, W. 1999. “Performance Evaluation of the Forward Link of a CDMA Cellular System", IEEE Miltary Communication Conference (MILCOM), Vol. 1, 1999, 162-166.
[5] Tranter W.H., et al. 2004. Principles of Communication Systems Simulation with Wireless Applications, New Jersey, Prentice Hall, Professional Technical Reference, 529-718.
[6] Gilbert, E. N. 1960. “Capacity of a Burst-Noise Channel”, Bell System Technical Journal, Vol. 39, September 1960, 1253–1266.
[7] Fritchman, B. D. 1967. ‘‘A Binary Characterization Using Partitioned Markov Chains,’’ IEEE Transactions on Information Theory, Vol. IT-13, No. 2, April 1967, 221–227.
[8] Arauz, J., Krishnamurthy, P. 2002. “A study of different partitioning schemes in first order Markovian models for Rayleigh fading channels”, The 5th IEEE International Symposium on Wireless Personal Multimedia Communications, 2002, 277–281.
[9] Rabiner, L. R. 1989. “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”. Proc. IEEE 77 (1989), 257-285.
[10] Sadeghi, P., et al. 2008. "Finite- State Markov Modeling of Fading Channels", IEEE Signal Processing Magazine, September 2008, 57-80.
[11] Sivaprakasam, S., Sharnmugan, K. S. 1995. "An Equivalent Markov Model for Burst Errors in Digital Channels", IEEE Transactions on Communications, Vol. 43, No. 2/3/4, February/March/April 1995, 1347-1351.
[12] Turin, W., van Nobelen, R. 1998. “Hidden Markov modeling of flat fading channels”, IEEE J. Selected Areas in Communications 16, 1998, 1809–1817.
[13] Beverly, A., Shanmugan, K. S. 1998. “Hidden Markov models for burst errors in GSM and DECT channels”, Proceedings of the IEEE Global Telecommunications Conference (Globe-com’98), Sydney, Australia, 1998, 3692–3698.
[14] Kuczynski, P. et al. 2004 “Hidden Markov Modeling of Error Patterns and Soft Outputs for Simulation of Wideband CDMA Transmission Systems”, Int. J. Electron. Commun. (AEU), (2004), 1-12.
[15] Zhi-Jin, Z. et al. 2007. “Discrete Channel .Modeling based on genetic algorithm and simulated annealing for training hidden Markov model”, Chinese Physics, Vol 16 No 6, June 2007, 1619-1623.