COMPLEXITY REDUCTION OF MARKOV CHANNEL MODELS FOR WIRELESS NETWORKS USING GRAPH THEORY

Hassaan Khaliq Qureshi, Khurram Shahzad, Syed Ali Khayam
School of Electrical Engineering & Computer Science
National University of Sciences & Technology
(NUST) Rawalpindi, Pakistan.

Muttukrishnan Rajarajan, Veselin Rakocevic
School of Engineering and Mathematical Sciences
City University, London, UK.

ABSTRACT

Accurate simulation and analysis of wireless networks are inherently dependent on accurate models which are able to provide real-time channel characterization. High-order Markov chains are typically used to model errors and losses over wireless channels. However, complexity (i.e., the number of states) of a high-order Markov model increases exponentially with the memory-length of the underlying channel. In this paper, we present a novel graph-theoretic methodology that uses Hamiltonian circuits to reduce the complexity of a high-order Markov model to a desired state budget. Our trace-driven performance evaluations for real wireless local area network (WLAN) and wireless sensor network (WSN) channels demonstrate that the proposed Hamiltonian Model, while providing orders of magnitude reduction in complexity, renders an accuracy that is comparable to the Markov model and better than existing reduced state models.

I. INTRODUCTION

Due to a lack of available infrastructure to perform realistic wireless experiments, system-level simulations are used to evaluate the performance of emerging wireless protocols and services. An accurate model of the wireless channel is an important component of such simulation-based performance evaluation. In the past three decades, channel error modeling techniques have been used extensively to improve the design of communication channels and the protocols that operate on these channels [1]. Using an accurate channel model, one can simulate the channel and can gain insights into the channel's underlying behavior. More importantly, an accurate and low-complexity channel model can be used to tune critical parameters of network protocols and applications at design time and in real-time. Lastly, a low-complexity channel model also allows real-time channel characterization and prediction which is required by rate adaptive protocols and applications.

In the channel modeling context, stochastic models have gained significant research attention [2], [3], [4]. In particular, high-order Markov channel models have been shown to be quite accurate in modeling link layer bit-errors and packet losses [1], [2], [5], [6], [7]. Unfortunately the complexity of Markov models increases with their memory length and, consequently, the viability of using Markov models in resource-constrained wireless environment is severely limited. Thus, accurate approximations of high-order Markov channel models are needed for wireless environments.

Many models have been proposed in recent literature to reduce the complexity of high-order Markov chains [6]. While there exists a clear tradeoff between complexity and accuracy (lower the model complexity, lower the accuracy), existing low-complexity channel models (with the exception of the bipartite model [7]) reduce the channel model's complexity to a fixed level and therefore do not cater for the emerging heterogeneous communication devices; for instance, on a given channel, high-end wireless devices (e.g., desktop and laptop computers) can afford higher complexity channel models than low-end devices (e.g., PDAs and smart phones.) To cater for such device heterogeneity, we need channel models that can adapt their complexity to an arbitrary level in accordance with the resources available at a wireless device.

In this paper, we propose a new variable-complexity wireless channel model referred to as the Hamiltonian Model (HM). The HM reduces the complexity of high-order Markov channel models by identifying and aggregating Hamiltonian circuits present in the states of the Markov chain. Given a desired complexity budget in terms of the total number of Markov states, the proposed model identifies a Hamiltonian circuit in the Markov chain, finds cycles of the needed complexity, and then aggregates these cycles into odd and even states based on the number of total states present in those cycles.

The performance of the proposed model is compared with the Bipartite Model (BM) of [7]. The performance characteristics of both models are evaluated using a comprehensive dataset of actual traces collected in two different environments: (i) 802.11 MAC layer bit errors at 5.5 Mbps; (ii) 802.15.4 MAC layer bit errors at 250 Kbps. We compare

the models' accuracies by measuring their closeness to the actual wireless channel traces using an information-theoretic Kullback–Leibler-based divergence measure and by comparing the Cumulative Distribution Functions (CDF) of bit errors. Our results demonstrate that HM has significantly higher modeling accuracy than BM.

The rest of this paper is organized as follows. Section II describes the related work in this area. Section III provides the background that is required to understand the contribution of this paper. Section IV describes error trace collection on Wireless Sensor and Local Area Networks. Section V describes the proposed Hamiltonian model. Section VI outlines performance evaluation of the HM model and compares its performance with the BM model. Section VII summarizes key conclusions of this paper.

II. RELATED WORK

Some previous studies have proposed methods to reduce the complexity of high-order Markov models [2], [6], [7]. In [6], guidelines were proposed to accurately model Markov based wireless channels and a constant complexity probabilistic model was proposed. In [8], authors showed empirically that low complexity hierarchical and hidden Markov models cannot characterize the bit error processes and proposes to employ high-order Markov chains for accurate channel characterization. These studies, however, resulted in models with fixed, non-scalable complexities.

Since complexity and accuracy of a model generally exhibit a direct proportionality relationship, we argue that, in view of the heterogeneity of contemporary wireless devices, a model should be able to scale its complexity in accordance with the complexity that can be afforded at a wireless device. More specifically, given a complexity budget (for instance, in terms of number of model states,) a scalable channel modeling algorithm should be able to produce a channel model to satisfy that budget. We are only aware of two studies ([2] and [7]) which approach the wireless channel modeling problem in this light.

Chen and Rao [2], [3] used the lumpability framework to reduce the order of a Markov channel model. However, the lumpability conditions place very stringent constraints on the transition probabilities of a Markov chain. These constraints are generally not satisfied by real-life wireless channel models. Therefore, we do not compare performance of our proposed technique with lumped Markov chains.

Willig [7] proposed a scalable-complexity Bipartite Model for wireless channels. The bipartite model uses the notion of a binary indicator sequence and divides the sequence according to its burst order. Based on the burst order, the image of good and bad burst Probability Mass Functions (PMFs) are divided into burst intervals. A transition matrix

is then computed for transiting between the burst intervals. We compare the performance of our proposed model with the BM.

III. BACKGROUND AND NOTATION

In this section we present brief background about the binary nature of traces and then most commonly used Markov chains of order K and the important notations used in this paper.

REPRESENATION OF BINARY WIRELESS TRACES

Traces collected over a wireless medium generally represent two states. One state is the good state and the other state is the bad state or the lossy state. Hence we can characterize wireless traces generally as a binary time series $\{x(n)\}_{n=1}^{l}$, where $x(n) \in \{0,1\}$ and l is the length of the error trace.

Without loss of generality, throughout this paper we use zero to represent an error-free bit and one for a bit in error. The sequence of these bits form a alternating bursts of zeros and ones. If the burst consists of number of zeros then we refer to it as a *good burst* and if the burst consists of one's then we call it a *bad burst*. The trace can hence be represented as pairs of good and bad bursts: $(N_1, G_1), (N_2, G_2), ..., (N_n, G_n)$, where N_n and G_n represent the length of the nth good and the bad bursts, respectively [1]. Many channel modeling studies have showed that this binary representation is suitable for representing channel traces [2], [7], [8], [9].

K^{TH} ORDER MARKOV CHAINS

A Markov chain of memory K is a discrete time random process whose probabilities for going to future states at a given present state are independent of the past states. For a memory length of K, the Markov chain comprises of 2^K possible combinations of K consecutive bits. If we have a set of states K which consists of $S = \{S_1, S_2, S_3, ..., S_n\}$, then the process starts in one of these states and moves successively from one state to another. If we define S_i as a current state then for moving to next state S_j, the probability will be denoted by P_{ij}. The probabilities are called transition probabilities and are computed by sliding bit by bit a K bit memory window over the data [10]. In this paper, Markov chain corresponding to even (odd) decimal numbers is referred as even (odd) states. Using the above notation, an example 3-rd order Markov chain is shown in Figure 1; only transitions to even states are shown. If the
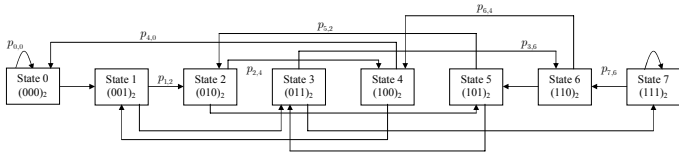
Figure 1: A 3-rd order (memory length=3) Markov chain.

Markov chain is in an even state, the last received bit (i.e., the least significant bit position in the memory window) must be error-free. Similarly, a Markov chain in the odd states implies that the last bit was corrupted. Due to the binary nature of the underlying wireless bit-error process, each Markov chain state can transit to only two other states [11]. This is due to the Markov chain definition in which the memory-window at each time instance is left-shifted by one bit and a one or a zero bit is added to the least-significant bit position. Thus from state S, a Markov chain can transit either to even state $(2S)\bmod 2^K$ or to odd state $(2S+1)\bmod 2^K$. Since the sum of all transitions from a Markov state must sum to one, for any state S we have $P_{(2S+1)\bmod 2^K} = 1 - P_{(2S)\bmod 2^K}$. It should also be emphasized that once a corrupted bit is received, a K-th order Markov chain will return to state 0 (i.e., the no error state) only from state 2^{K-1} after K transitions; see in Figure 1 that at state $2^{3-1} = 4$, the Markov chain wraps around to state 0.

IV. DATA COLLECTION

To perform realistic performance evaluation over operational channels, we collected a comprehensive dataset of wireless error traces over two different channels: 1) an 802.15.4 WSN channel, and 2) an 802.11b WLAN channel. All traces were collected at the MAC layer after physical layer processing; MAC layer channels are referred to as *residual channels* in prior literature [4], [6]. This section describes the data collection and some preliminary trace statistics.

802.15.4 DATA COLLECTION

We used Crossbow's Micaz motes [12] to collect residual bit-error traces over wireless sensor networks. These motes operate on the ISM frequency band of 2.4 GHz and support a peak data rate of 250 Kbps. Sensor motes were running the open-source TinyOS operating system [13]. We modified the source code of TinyOS applications to disable the MAC layer checksum feature at the receiver. Hence, corrupted packets were not dropped in a receiver's kernel, and were passed to a data logging application. The application logged all packets on an attached computer through the serial port. We collected the traces at four different locations or setups. These setups are named accord-
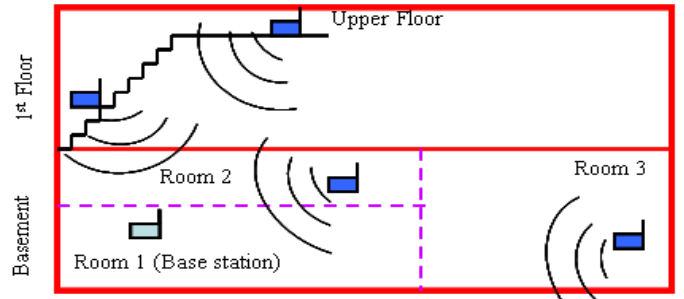

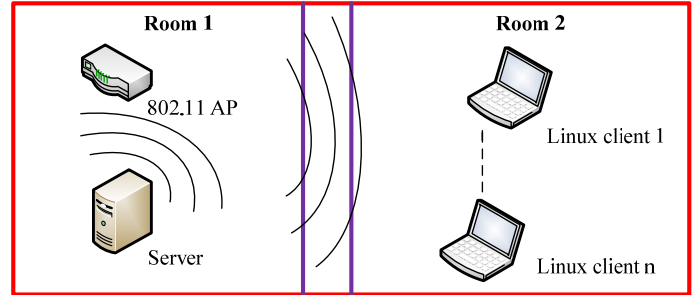Figure 2: Setup for 802.15.4 bit error traces.

Figure 3: Setup for 802.11b bit error traces.

ing to their geographical location as shown in Figure 2. The light-shaded mote in Figure 2 is the base station which received and logged data, while the remaining motes are sending motes which transmitted packets with predefined contents to the base station. In each experiment, one sender transmitted unicast data to the base station and the other senders were inactive; i.e., in each trace collection, there was no channel contention and collisions as there was only a single sender and a single receiver. While performing experiments, motes were kept stationary. In order to collect traces with varying error behaviors, the distance between the motes and the base station was varied from 5 to 12 meters. The senders transmitted 20-byte fixed-sized frames at a rate of 10 frames per second.

We first performed the experiments by having a direct line of sight (LoS) between the sender and the base station but the error rates observed in those experiments were too low to warrant further analysis. Therefore, throughout this paper, we focus on non-LoS traces. The average number of frames per trace was approximately 31,000 frames. Thus, the average length of each trace was approximately 5 million bits [4]. For evaluation of HM and BM we used five traces for each setup and a total of 20 traces for all setups. While we collected more actual wireless sensor network traces but the trends observed in 20 traces (5 per setup) are representative of the trends that we observed in other traces.

802.11 DATA COLLECTION

For 802.11b traces, we repeated the same process as the 802.15.4 traces using the topology shown in Figure 3 and

	Setup	BER
802.15.4	Room2	0.00085689
	Room3	0.01519663
	Stairs	0.00737452
	Upper floor	0.00735198
802.11b	Location 1,2,3	0.003067968

a) Markov chain of order $K = 4$.

b) Hamiltonian circuit.

c) Hamiltonian circuit arranged by vertices traversed

d) Hamiltonian circuit reduced to $K = 3$.

Figure 4: An example of complexity reduction using the Hamiltonian Model.

AP (access point) was operating in Distributed Coordination Function (DCF) mode and three wireless stations communicating in the infrastructure network configuration. One of the stations was operating as the server and the remaining two as multicast clients. All wireless stations were Linux boxes using Dlink wireless cards with Prism2 chipset device drivers [14]. The server was stationary and transmitted a continuous stream of predetermined patterns to the multicast clients. Traces were generated for each bitrates at different stationary client positions with and without LoS. It was observed that, with clear LoS, the error rate at all bitrates was extremely low. Such excellent performance deemed further LoS study inconsequential. Hence, both clients were positioned in a separate room across two walls in order to simulate a more realistic business/classroom/home-network wireless setup and forced to transmit non Line of Sight frames of 512-bytes at a physical layer data rate of 5.5Mbps. The average length of each trace was approximately 6 million bits. We collected 3 traces at different locations for 5.5Mbps. We also calculated the traces at 11 Mbps but due their non-Markovian behavior we do not use those traces in this paper.

TRACE BIT ERROR RATE (BER) STATISTICS

TABLE I shows the average bit error rates of 802.15.4 and 802.11b original traces, respectively. It can be seen that highest bit error rate was observed for Room 3 as in this case the sender and receiver were at the farthest distance from each other. Room 2 has the lowest bit error rate as in this case the sender and receiver were at the closest distance from each other. The bit error rate of Upper floor and stairs are similar because of the small distance variation between these setups. In case of 802.11b where the receiver's were kept stationary, the average bit error rate observed at different locations was 0.003.

V. THE HAMILTONIAN WIRELESS CHANNEL MODEL

Let a K-th order Markov chain be represented as a K vertex connected digraph $G = (V, E)$ with positive edge wei-
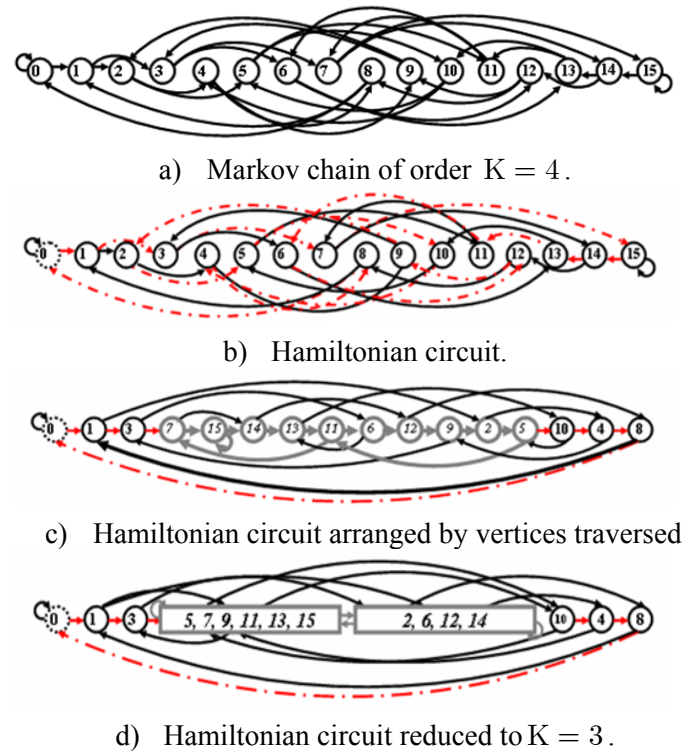
ghts. Markov chain of any order forms a connected graph as long as both the two state transitions probability in each Markov state are greater than zero. We observed that Markov chains exhibit many interesting graph-theoretic properties which can be used to reduce the complexity associated with higher order Markov chains and those properties can be used to develop a scalable model. A graph in which we can traverse each vertex or node exactly once forms a Hamiltonian circuit. A Hamiltonian circuit can be identified in Markov chains of any arbitrary order. By using this property, we can further arrange the states according to the nodes traversed, which gives a easy method for aggregating states that comprise the Hamiltonian circuit. Since a cycle of arbitrary length can be identified in the Markov digraph, the states of the circuit can be aggregated to a desired state budget. These characteristics are a consequence of the Markov chain construction and are therefore present at all orders of the Markov chains.

Moreover, the Hamiltonian circuit formed in the Markov digraph clearly identifies the good and bad nodes which remain separated during state aggregation. This is a very important property because it is generally undesirable to merge good (even) and bad (odd) states together [7]. After Hamiltonian state aggregation, the merged states probabilities are aggregated and normalized into one aggregate state of the low-complexity model. This graph-theoretic realization helps us in reducing the complexity in a finite time

Model	Setup	Number of States	
		16	256
Hamiltonain_802.15.4	Room2	0.001315	0.000416
	Room3	0.022036	0.021816
	Stairs	0.010695	0.011524
	Upper Floor	0.010517	0.014039
Hamiltonian_802.11b	Location 1,2,3	0.004328	0.005284
Bipartite_802.15.4	Room2	0.0009345	0.0540207
	Room3	0.0032156	0.0725722
	Stairs	0.0017525	0.0453318
	Upper Floor	0.0080524	0.1279515
Bipartite_802.11b	Location 1,2,3	0.033962	0.030718

and in a simple and easy manner. In the following discussion, we provide an example of applying the proposed Hamiltonian Model based state aggregation on a Markov chain of order $K = 4$.

A Markov chain of order $K = 4$ is shown in Figure 4[a]. If we start from state zero and start traversing all the other vertices, the order of traversed vertices in the consequent Hamiltonian circuit will be: 0-1-3-7-15-14-13-11-6-12-9-2-5-10-4-8-0. This is depicted with dotted line or edges in Figure 4[b]. After finding a Hamiltonian cycle we observed that there are n cycles in the graph. Based on the cycles still left in the graph we applied our algorithm to work in such a way to find a cycle exactly aggregating the graph into $K = 3$ states. This procedure is shown in Figure 4[c]. We first found a cycle consisting of 5 vertices (7, 5, 14, 13, 11,) and then again we found another edge coming from vertex 5 towards vertex 11 which was already merged. Hence, we actually looked for a cycles consisting of ten nodes and then we separated the cycle into odd and consecutive even nodes. Figure 4[d] shows the final transformation from $K = 4$ to $K = 3$. We then normalized the probabilities of the states aggregated in odd and even parts. We first compressed our model from $2^9 = 512$ to $2^8 = 256$ and then to $2^7 = 128$ and so on up to $2^3 = 8$. We applied the algorithm on the graph as depicted in the figures and at each state we generated artificial traces generated by our model. It is also possible to find the Hamiltonian circuit by traversing vertices different than the traversed vertices shown in Figure 4. For that the model remains the same and is still applicable in reducing the needed complexity.

VI. PERFORMANCE EVALUATION

In this section, we use the error traces to compare the performance of the Hamiltonian Model with the Bipartite Model using Bit Error Rate (BER), Kullback-Leibler Divergence (KLD) and bit error distributions.
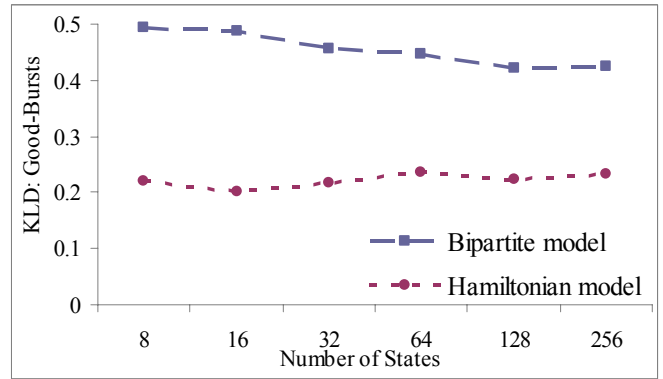

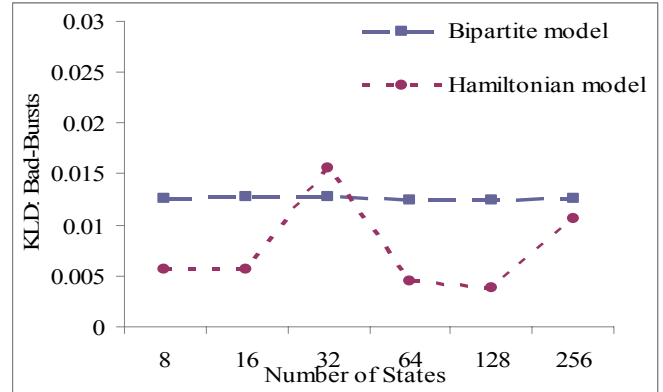Figure 5: Resister average divergence of good bursts versus complexity for the 802.15.4 bit error process.

Figure 6: Resister average divergence of bad bursts versus complexity for the 802.15.4 bit error process.

Bit Error Rate (BER)

We calculated the BER of 802.15.4 and 802.11b synthetic traces generated by Hamiltonian Model and the Bipartite Model. TABLE 2 shows the Average BER per setup of Hamiltonian Model traces parameterized from actual traces at $K = 4$ and $K = 8$. It also shows the BER of Bipartite Model for varying number of K-states. Comparing with TABLE I, it can be observed that BER of Hamiltonian Model (HM) traces are closer to the actual traces' BER than the Bipartite Model (BM). Overall, the inaccuracy in BER estimates shows the opposite trends in HM and BM. For the HM, the inaccuracy decreases with an increase in the number of states. The BM, on the other hand, incurs more inaccuracy for higher number of states. Thus for the BM introducing more state does not necessarily increase the accuracy of the model. Typically, an increase in the complexity of a model (e.g., with an increase in the order of a Markov chain) causes the accuracy of the model to improve. However, for the HM model we did not observe this proportionality trend between complexity and accuracy. On the contrary, after empirically evaluating the HM model's accuracy for varying state merging orders, we ob-
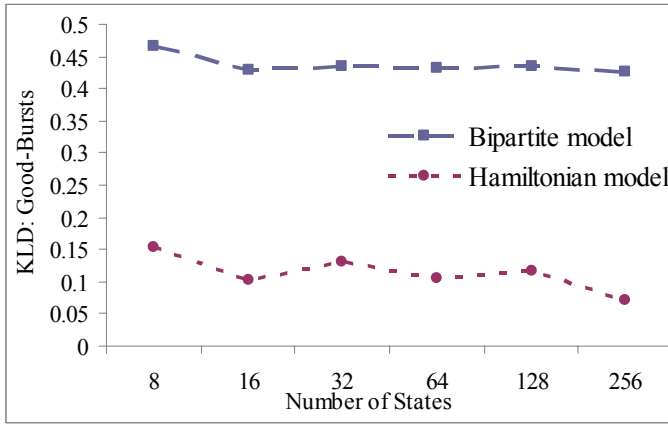
Figure 7: Resister average divergence of good bursts versus complexity for the 802.11b bit error process.
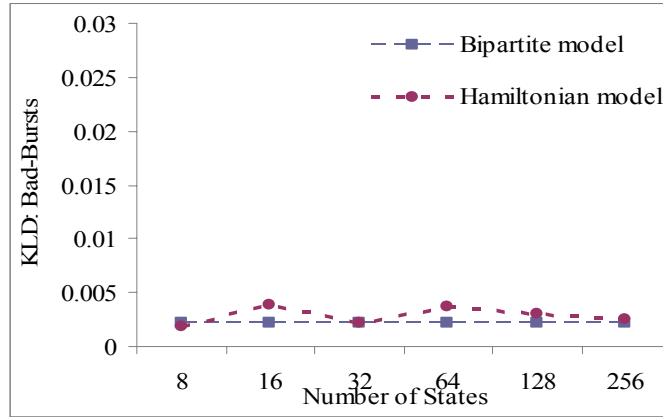

Figure 9: Good burst length distribution for 802.15.4 bit error traces at $K = 5$.

Figure 8: Resister average divergence of bad bursts versus complexity for the 802.11b bit error process.
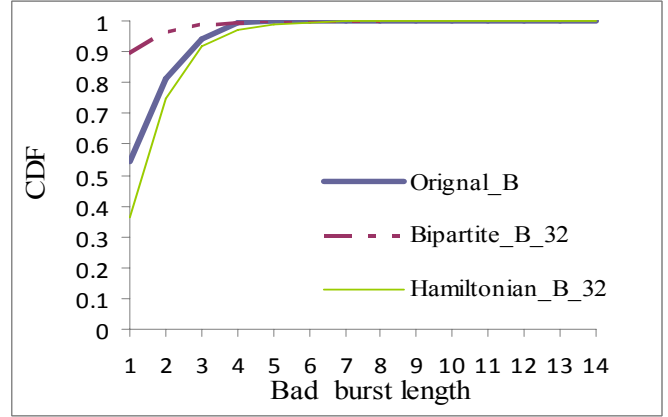
Figure 10: Bad burst length distribution for 802.15.4 bit error traces at $K = 5$.

served that a particular HM order provides better accuracy than orders above and below it.

While the BER estimates give us an overall picture of the accuracy of a model, for more elaborate performance comparison, we need to compare the burst distributions of the HM and BM. To this end, we compare the good- and bad-bursts' distributions of the two models in the next two sections.

Kullback-Leibler Divergence (KLD) of Good- and Bad-Bursts

Entropy is a measure of the average number of bits required to represent all outcomes of a probability distribution. The Kullback-Leibler divergence quantifies the difference in the entropies of two probability distributions [15]. The KLD divergence quantifies the source-coding-like overhead incurred by employing a model instead of the actual source. For two probabilities distributions p and q defined over a common alphabet Ψ, the KL divergence is defined as:

$$D(p \mid\mid q) = \sum_{x} p(x)\log_2 \frac{p(x)}{q(x)}.$$

The KLD has two shortcomings: 1) non symmetry, 2) it requires the two distributions to be continuous with respect to each other. Therefore, instead of Kullback-Leibler we used the KL based Resistor-Average (R) divergence measure defined as [15]:

$$\frac{1}{R(p,q)} = \frac{1}{D(p \mid\mid q)} + \frac{1}{D(q \mid\mid p)}.$$

For accuracy evaluation of HM and BM, we compare the R divergence of good- and bad-bursts distributions derived from actual traces and the models. The R divergence observed in Figure 5, Figure 6, Figure 7 and Figure 8 at different states demonstrate that the HM shows results demonstrating very small R values for good bursts and outperforms the BM. The results for R divergence of 802.15.4 HM traces and BM are demonstrated in Figure 5. It shows consistently better performance with decreasing order of K and renders good behavior with increasing state compression. Figure 6 shows the R divergence of bad-bursts for 802.15.4 traces generated from HM and BM. For bad bursts, it demonstrates very small R values elaborating similarity with actual traces when compared with BM and showing a slightly higher R divergence at
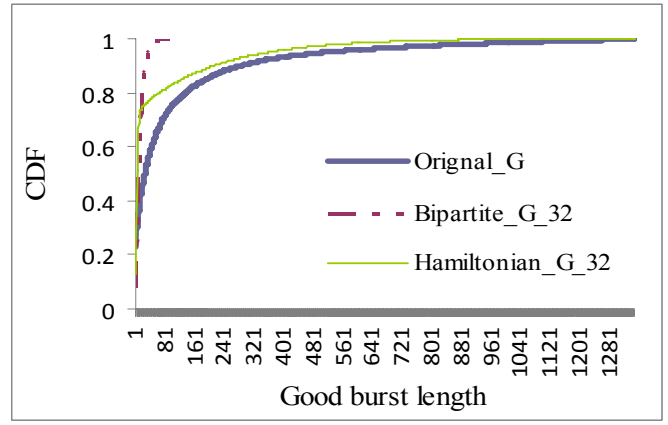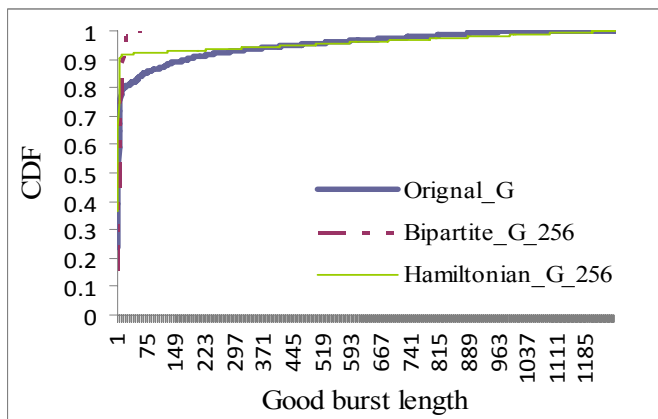
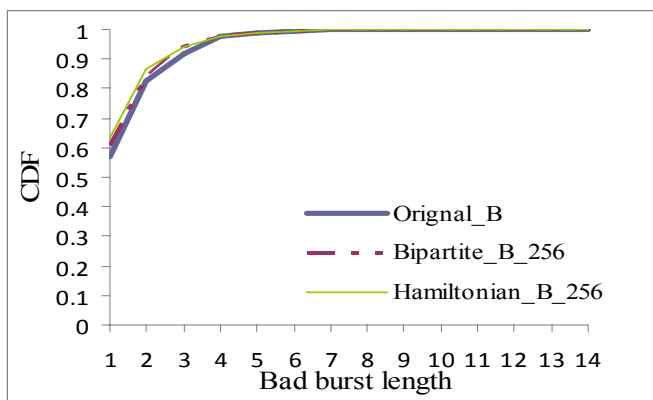Figure 11: Good burst length distribution for 802.11b bit error traces at $K = 8$.

Figure 12: Bad burst length distribution for 802.11b bit error traces at $K = 8$.

$K = 5$ because of high BER at $K = 6$. The slight variations in bad burst R divergence values can be removed by averaging over more traces. Due to space constraint, we do not report these results. R divergences of good and bad bursts for the 802.11b traces are shown in Figure 7 and Figure 8. It can be clearly observed that at decreasing order of K the HM has very small R divergence in the good bursts case. On the other hand, the good-bursts distributions of the BM diverge quite significantly from actual traces. For bad-bursts, the R divergence of the HM is slightly higher than the BM. Nevertheless, both models are able to capture the bad-bursts behavior quite accurately.

Cumulative Distribution Function (CDF) of Good and Bad Bursts

We derive CDFs of good and bad bursts from actual network traces and traces artificially synthesized by our model and the BM. For evaluation we took one trace per setup from each of the model at different memory length.

The cumulative distribution function for good and bad burst length at $K = 5$ is shown in Figure 9 and Figure 10 respectively. The CDF of the HM clearly follows the CDF of actual 802.15.4 traces for both the good and bad burst distributions. The CDF for $K = 8$ is also plotted for good

and bad burst lengths and is shown in Figure 11 and. Figure 12. The CDF of the HM clearly follows the CDF of actual 802.11b bit error traces in both the good and bad burst length distributions respectively. However, the CDF of the BM renders same behavior following the CDF of actual 802.11b bit error traces in the case of bad burst.

VII. CONCLUSION

In this paper, we presented a novel algorithm to reduce the complexity of high-order Markov channel models to a desired state budget. We performed state aggregation by identifying Hamiltonian Circuits in the Markov digraph. States in the circuit were then aggregated to a given and arbitrary state budget. We demonstrated that the HM provides orders of magnitude reduction in complexity and renders very accurate performance.

ACKNOWLEDGMENT

REFERENCES

[1] S. A. Khayam, H. Radha, S. Aviyente, and J. R. Deller, Jr., "Markov and multifractal wavelet models for wireless MAC-to-MAC channels," *Elsevier Performance Evaluation*, vol. 64, no. 4, pp. 298−314, May 2007.
[2] A.M. Chen and R.R. Rao, "Wireless Channel Models – Coping with Complexity," *Wireless Multimedia Network Technologies*, Kluwer Academic Publishers, 271-288, 1999.
[3] A.M. Chen and R.R. Rao, "On tractable wireless channel models," *IEEE PIMRC*, Sep. 1998.
[4] A. Iqbal and S.A. Khayam, "Improving WSN Simulation and Analysis Accuracy Using Two-Tier Channel Models," *ICC*, May 2008.
[5] A. Konrad, B.Y. Zhao, A.D. Joseph, and R. Ludwig, "A Markov based channel model algorithm for wireless networks," *ACM WINET*, (9), 189-199, 2003.
[6] S.A. Khayam and H. Radha "Constant-Complexity Models for Wireless Channels," *IEEE Infocom*, April 2006.
[7] A. Willig. "A new class of Packet- and Bit-Level Models for Wireless Channels," *TKN-02-009, Telecommunication Networks Group*, Technical University Berlin, June 2002.
[8] S.A. Khayam and H. Radha, "Linear-Complexity models for wireless MAC-to-MAC channels," *ACM WINET*, (11)5, 543-555. 2005.
[9] S.A. Khayam and H. Radha, "Markov-based modeling of wireless local area networks," *ACM MSWIM*, Sep. 2003.
[10] C.M. Grinstead and J.L. Snell, "Introduction to Probability: Second Revised Edition," AMS, 1997.
[11] S.A Khayam and H. Radha, "On the Impact of Ignoring Markovian channel Memory on the Analysis of Wireless Systems," *IEEE International Conference on Communications* (ICC), June 2007.
[12] Crossbow Homepage, http://www.xbow.com.
[13] TinyOS Homepage, http://www.tinyos.net/.
[14] Linux-wlan Home Page. *http://www.linux-wlan.org/.*
[15] T. Cover and J. A. Thomas, *Elements of Information Theory*, Second Edition. ISBN: 0-471-24195-4.