# Modeling Switching Activity Using Cascaded Bayesian Networks for Correlated Input Streams *

Sanjukta Bhanja and N. Ranganathan
Center for Microelectronics Research
Department of Computer Science and Engineering
University of South Florida, Tampa, FL-33620
bhanja, ranganat@csee.usf.edu

## Abstract

*We represent switching activity in VLSI circuits using a graphical probabilistic model based on Cascaded Bayesian Networks (CBN's). We develop an elegant method for maintaining probabilistic consistency in the interfacing boundaries across the CBN's during the inference process using a tree-dependent (TD) probability distribution function. A tree-dependent (TD) distribution is an approximation of the true joint probability function over the switching variables, with the constraint that the underlying Bayesian network representation is a tree. The tree approximation of the true joint probability function can be arrived at using a Maximum Weight Spanning Tree (MWST) built using pairwise mutual information between switchings at two signal lines. Further, we also develop a TD distribution based method to model correlations among the primary inputs which is critical for accuracy in Bayesian modeling of switching activity. Experimental results for ISCAS circuits are presented to illustrate the efficacy of the proposed methods.*

## 1 Introduction

Switching activity of a node is affected by factors such as the connectivity of the circuit, the input statistics, the correlation among nodes (or lines), the gate type, and the gate delays, thus making the estimation process a complex procedure. It is well-known that switching activity depends on temporal, spatial, and spatio-temporal correlations exhibited by the signals (could be internal nodes or primary inputs or state lines) [4, 5, 6, 3].

In this work, we propose a a new switching probability model for combinational circuits based on the concept of **Cascaded-Bayesian Networks** (CBN), capturing complex conditional dependencies over a set of random variables. While single Bayesian Network representation is guaranteed to give us accurate estimates [2], the inference process in establishing these estimates is NP-hard. This forces us, given the available computing constraints, to represent combinational circuits using Cascaded BNs (CBN). A large circuit is partitioned into smaller one, and each partition is represented as a Bayesian Network (BN). Consecutive BN's are cascaded through a Tree-Dependent distribution. In the Bayesian network model discussed in [2], the partitioning was done in an ad-hoc manner without maintaining spatial correlations across the boundary nodes. Moreover, we address the problem of modeling the switching activity among *correlated* primary inputs which is essentially the same as maintaining correlations across the CBN. We can elegantly address both these problems with solutions based on the tree-dependent (TD) probability function. This tree-dependent representation offers a good trade off between the need for accurate representation of correlations and computational efficiency. We use the tree-dependent (TD) distribution at the BN boundaries, as well as, at the primary inputs. Since the tree-dependent (TD) distribution can also be represented as a Bayesian Network, we can fuse this approximate tree representation over the primary nodes with the accurate BN-based representation over the internal nodes of the combinational circuit as well as fuse the TD between adjacent BNs to capture correlations in the boundary nodes. Results show that mean error, standard deviation of error and maximum error are lower, indicating that CBN is indeed superior model which estimates switching activity accurately and uniformly over all the switching nodes.

## 2 Cascaded BN Modeling

Switching correlations among the primary input nodes can affect switching activity estimates across the whole circuits. An ideal way to model the input switching would be, given a training set of input line transition, to learn an exact switching model in terms of a joint proba-

*This work is being supported in part by a grant from the National Science Foundation Grant No. CCR-0098103



Proceedings of the 2002 IEEE International Conference on Computer Design: VLSI in Computers and Processors (ICCD'02)
1063-6404/02 $17.00 © 2002 IEEE

bility function over the input lines, which, of course, can also be represented by a Bayesian network (BN). This learned BN then would be coupled with the BN representation of the combinational circuit. However, Learning exact BN is NP-hard. Hence, we resort to TD based approximate modeling as a practical compromise of the accuracy of representation of dependencies and computational costs in terms of time and storage. The BN coupling problem between the cascaded BNs can be seen as an instance of the primary input modeling problem, where the "inputs" are not the primary input lines but nodes in the previous BN. Here too we can use TD distribution over the boundary nodes and cascade it with adjacent BN and form an elegant cascade structure.

**Definition 1:** Any tree-dependent distribution $P^t(x)$ can be defined as a Markov field relative to the tree $t$ which can be written as the product of $n-1$ pair-wise conditional probability distributions,

$$p^t(x) = \prod_i p(x_i|x_{j(i)}) \tag{1}$$

where $X_{j(i)}$ is the designated parent of $X_i$ in some orientation of the tree $t$. The root node $X_1$ is chosen arbitrarily without any parents and $P(x_1|x_0) = P(x_1)$. Apart from the memory requirement, only second order statistics are needed to construct the tree.

Our goal is to construct a tree over $n$ variables, representing the input nodes, that is the closest representation of the underlying joint probability function over the $n$ variables. Hence, out of all the spanning tree over the $n$ variables that can be constructed, we have to select the one which preserve the correlations to a maximum level. Now, we have two subgoals: 1) To choose the best conditional probabilities between the parent and the child nodes in the tree given a fixed tree $t$ such that $P^t$ is the best approximation of $P$. This distribution is called the **projection of $P$ on $t$ ($P^t_P$)** . And, 2) to choose a tree from a set of all the spanning trees over the nodes such that it would make the projection $P$ on this tree $P^t_P$ closest to $P$. We will use the two following theorems to arrive at a tree structure [1].

**Theorem 1:** The projection of $P$ on $t$ is characterized by the equality

$$P^t_P(x_i|x_{j(i)}) = P(x_i|x_{j(i)}) \tag{2}$$

Proof in [1].

This implies that the conditional probabilities for a branch a tree has to coincide with that computed from $P$ will produce the best **projection of $P$ on $t$ ($P^t_P$)**. The optimal tree structure is chosen by minimizing Kullbuck-Leibler cross-entropy measure.

**Theorem 2:** The Kullbuck-Leibler distance measure is minimized by projecting $P$ on any maximum weight spanning tree (MWST) where the weight of the branch $(X_i, X_j)$ is defined by the information measure between them

$$I(X_i, X_j) = \sum_{x_i, x_j} P(x_i, x_j) \log \left( \frac{P(x_i, x_j)}{P(x_i)P(x_j)} \right) \tag{3}$$

Proof in [1].

Hence, by the above definitions and theorems, a tree-dependent (TD) distribution is an optimal (best) approximation of the true joint probability function over the switching variables, with the constraint that the underlying Bayesian network representation is a tree. The tree structure controls the computational complexity. The tree approximation of the true joint probability function can be arrived at using a Maximum Weight Spanning Tree (MWST) built obtained by the pairwise mutual information between switchings at two signal lines [1]. Using a tree-structured representation ensures that storage proportional to $(r-1)r(n-1) + r - 1$ [1] is used where $r$ is the number of states (in our case $r$ = 4) and $n$ is the number of variables of the primary inputs which is much less than $r^n$ which would be needed for a complete representation. Moreover, by the above algorithms, we ensure that at least the pairwise correlations are captured effectively and propagated to the internal nodes.

## 3 Experimental Results and Conclusion

We mapped several ISCAS circuits to their corresponding Cascaded Bayesian Networks representation. The conditional probabilities are pre-determined by the type of gate connecting the parents and the child. We used HUGIN's Bayesian Network tool [7] for compiling the junction tree and propagating the probabilities. We also performed in-house zero-delay logic simulation providing "ground truth" estimates of switching for computation of the errors on each signal. In our experiments, we want to establish that the tree-dependent distribution works accurately for highly correlated input streams. We use 16 bit counters to generate highly correlated sequences for the experiments. These experiments are carried with in DELL PC with WINDOWS operating system running at 750 MHz. In Table 1, we present the time for estimation as well as time for input modeling for several circuits under highly correlated input streams. We observed that maximum cost of input modeling in terms of time is less than 1.32 sec.

We provide a comparison with BN models (without capturing spatial correlation in the boundary nodes and in the primary inputs) versus the new CBN with TD model in



Proceedings of the 2002 IEEE International Conference on Computer Design: VLSI in Computers and Processors (ICCD'02)
1063-6404/02 $17.00 © 2002 IEEE

| Circuits | Time (BN Inference)(s) | Time (input modeling) (s) |
|---|---|---|
| c17 | 0.00 | 0.00 |
| c432 | 0.33 | 1.31 |
| c499 | 0.21 | 1.26 |
| c880 | 0.25 | 1.26 |
| c1355 | 1.04 | 1.27 |
| c1908 | 1.87 | 1.26 |
| c3540 | 3.47 | 1.32 |
| c6288 | 8.68 | 1.27 |

**Table 1. Estimation time for CBN Models.**

| | BN Model | | | CBN Model | | |
|---|---|---|---|---|---|---|
| Circuits | $\mu$ | $\sigma$ | $Max$ | $\mu$ | $\sigma$ | $Max$ |
| c17 | | | | 0.000 | 0.000 | 0.00 |
| c432 | 0.003 | 0.023 | 0.191 | 0.001 | 0.011 | 0.141 |
| c499 | 0.0002 | 0.023 | 0.172 | 0.000 | 0.017 | 0.125 |
| c880 | 0.000 | 0.002 | 0.033 | 0.000 | 0.002 | 0.031 |
| c1355 | 0.007 | 0.039 | 0.191 | 0.001 | 0.011 | 0.126 |
| c1908 | 0.006 | 0.027 | 0.388 | 0.005 | 0.019 | 0.122 |
| c3540 | 0.000 | 0.001 | 0.083 | 0.000 | 0.001 | 0.054 |
| c6288 | 0.009 | 0.034 | 0.450 | 0.001 | 0.027 | 0.350 |

**Table 2. Estimation Errors for BN and CBN models.**

Table 2 for highly correlated input streams. We achieved considerable improvement in terms mean, standard deviation and maximum errors. It is obvious that with TD, the mean is always lower than that without TD-based models. Standard deviation of error which signifies the diversity of error estimates are 1.5 to 2 times smaller with TD based coupling than the naive one. For all the circuits, maximum error is reduced by significant amount; for some, we achieved 3 times improvements. The reduction in standard deviation and the maximum errors signifies that the estimation based on CBN models are not only more accurate but also more uniform. The error distribution (for all the nodes having errors higher than 0.01) with CBN and BN models for both low and highly correlated inputs for benchmark c1355 are shown in Figures 1 and Figures 2(CBN in yellow (or light) and BN in red (or dark)). It is obvious that with TD based CBN modeling, we achieve very low error spread compared to naive BN models. Hence TD based CBN modeling is essential for accurate and uniform switching activity estimation for all the nodes internal to the circuits as well as for modeling correlated inputs.

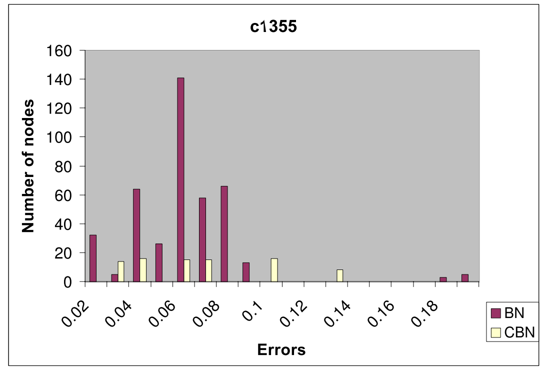


**Figure 1. Error distribution for c1355 for input streams with low correlation. (BN: dark or red; CBN: light or yellow)**

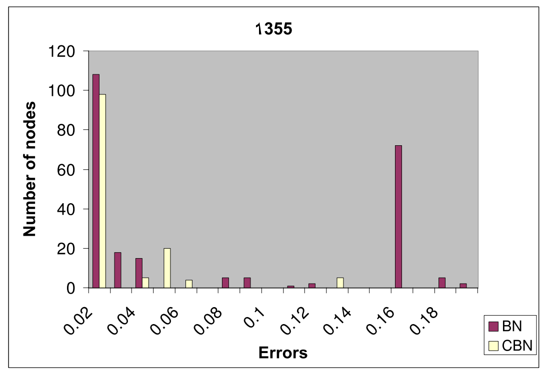


**Figure 2. Error distribution for c1355 for input streams with high correlation. (BN: dark or red; CBN: light or yellow)**

## References


[1] C. K. Chow, C. N. Liu, "Approximating Discrete Probability Distributions with Dependence Trees", *IEEE Transaction on Info. Theory*, vol. 14, pp.462-467, 1968.

[2] S. Bhanja, N. Ranganathan, "Dependency Preserving Probabilistic Modeling of Switching Activity using Bayesian Networks", *38th Design Automation Conference*, pp.209-214, 2001.

[3] A. Ghosh, S. Devadas, K. Keutzer, and J. White, "Estimation of Average Switching Activity in Combinational and Sequential Circuits", *Proceedings of the 29th Design Automation Conference*, pp.253-259, June 1992.

[4] R. Marculescu, D. Marculescu, and M. Pedram, "Probabilistic Modeling of Dependencies During Switching Activity Analysis", *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, vol 17-2, pp.73-83, February 1998.

[5] F. N. Najm, "Transition Density: A New Measure of Activity in Digital Circuits", *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, vol 12-2, pp.310-323, February 1993.

[6] P. Schneider, and U. Schlichtmann, "Decomposition of Boolean Functions for Low Power Based on a New Power Estimation Technique", *Proc. 1994 Int'l Workshop on Low Power Design*, pp.123-128, April 1994.

[7] HUGIN URL `http://www.hugin.com/`





Proceedings of the 2002 IEEE International Conference on Computer Design: VLSI in Computers and Processors (ICCD'02)
1063-6404/02 $17.00 © 2002 **IEEE**