Integrated Circuit Digital Rights Management Techniques Using Physical Level Characterization

Sheng Wei
Computer Science Department
University of California, Los Angeles (UCLA)
Los Angeles, CA 90095
shengwei@cs.ucla.edu

Farinaz Koushanfar
Electrical & Computer Engineering Department
Rice University
Houston, TX 77005
farinaz@rice.edu

Miodrag Potkonjak
Computer Science Department
University of California, Los Angeles (UCLA)
Los Angeles, CA 90095
miodrag@cs.ucla.edu

ABSTRACT

Digital rights management (DRM) of integrated circuits (ICs) is a crucially important task both economically and strategically. Several IC metering techniques have been proposed, but until now their effectiveness for royalty management has not been quantified. IC auditing is an important DRM step that goes beyond metering; it not only detects that a pirated IC has been produced but also determines the quantity of pirated ICs. Our strategic objective is to create a new intrinsic passive metering technique as well as the first IC auditing technique, and to maximize and quantify their effectiveness using statistical analysis and IC characterization techniques. Our main technical innovations include physical level gate characterization, a Bayesian approach for coincidence analysis, and an adaptation of animal counting techniques for IC production estimation. We evaluate the accuracy of the IC metering and auditing approach using simulations on a set of ISCAS benchmarks.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection—*Physical Security*

General Terms

Security

Keywords

Digital rights management, IC metering, IC auditing, physical level characterization, animal counting

1. INTRODUCTION

Digital rights management (DRM) [1] [2] of integrated circuits (ICs) has drawn a great deal of attention in the recent years with the fast growth of outsourcing in the IC industry. In the current model of IC manufacturing, the IC design companies deliver their designs to IC foundries without having any control over the manufacturing process. In this process, it is likely that an untrusted IC foundry fabricates a larger number of ICs than it was authorized to produce. Such misconduct of unauthorized IC production has become a crucial concern in the IC industry, with illegal copies of ICs costing the design companies billions of dollars annually.

IC metering approaches [1] [3] [4] [5] have been proposed to deter or detect the unauthorized IC production. More formally, IC metering or hardware metering refers to tools, methodologies, and protocols that enable post-fabrication tracking of the ICs. The metering approaches proposed thus far are classified into two categories, *passive* and *active*. In passive metering, the ICs are individually identified, either in terms of their functionality, or by other forms of unique identification. The identified ICs may be matched against their record in a pre-formed database that could reveal unregistered ICs or overbuilt ICs (in case of collisions). A beauty of passive metering is that the *intrinsic* process variation of the legacy chips, without any modifications, can be exploited to identify and to track each individual chip [5]. For a more comprehensive review of hardware metering, we refer the readers to recent surveys on the topic [6] [7]. In active metering, not only the ICs are specifically identified, but also parts of the IC functionality can be only accessed, locked (disabled), or unlocked (enabled) by the design house or IP owners by exploiting the design details that are not transferred to the foundry.

The existing IC metering approaches can detect intellectual property (IP) violations in IC manufacturing. However, no exact quantification of the number of chips that have been produced illegally was proposed, beyond the collision probabilities that were computed by variations of the Birthday paradox. This renders the royalty management and charging of the parties using the IP extremely challenging.

To address this issue, we propose the new concept of *IC auditing*, which aims to provide a quantified estimation of the number of chips produced and released to the IC market. Our strategic objective is to create a new intrinsic passive metering technique and the first IC auditing technique using a combination of statistical, majorization, and IC characterization techniques. In particular, we propose a new intrinsic passive metering scheme based on physical level IC characteristics, such as threshold voltage (V_{th}) and effective

DRM'11, October 21, 2011, Chicago, Illinois, USA.

channel length (L_{eff}). The reason why we use physical level properties is that they are more stable than the properties are used in manifestational tests (e.g., delay and power) in the sense that they do not depend on the environmental factors (e.g., temperatures) and thus, they can serve as stable IDs for the chips. We analyze the uniqueness of our generated IDs by conducting coincidence estimation, which verifies the uniqueness of IDs over a large number of chips.

Based on the IC metering and coincidence estimation, we propose a systematic way of conducting IC auditing, which is to predict the total number of chips produced and released to the market. Our auditing scheme is based on an animal counting model that predicts the total population from a partial sampling and labeling of the chips in the market. Our main contributions in this paper include the following:

- We propose a new intrinsic passive metering scheme based on physical level characterization which provides stable and unique IDs for the chips in the market without instrumentations to the IC during the design.
- We introduce a statistical method of estimating the coincidence among the IC physical characteristics based on Bayesian analysis and majorization techniques.
- We develop an IC auditing approach based on the classic animal counting model and statistical sampling.

The remainder of this paper is organized as follows. In Section 2, we summarize the existing research work regarding IC metering and digital rights management. Section 3 introduces the system models we use in this work. In Section 4, we introduce the overall flow of our IC metering and auditing scheme. The IC metering approach we are using is discussed in Section 5, followed in Section 6 by our method of estimating the coincidence in the IC metering process. Section 7 gives a complete solution for predicting the number of chips, based on the animal counting model. We show our simulation results for our IC metering and auditing scheme in Section 8. We conclude the paper in Section 9.

2. RELATED WORK

Metering and auditing have been recently studied in the area of WWW, such as for client counting for client/server management [16] and click fraud prevention [17]. Similarly, in the area of IC design and manufacturing, there are several active or passive IC metering schemes that have been proposed. Some of them require instrumentation in the design and manufacturing process, which are called *extrinsic* metering; the others utilize the existing IC characteristics for metering purpose without modifying the design flow, which are called *intrinsic* metering. Intrinsic IC metering methods are all passive.

2.1 Extrinsic IC Metering

Extrinsic IC metering introduces extra hardware/software components to the chips, in order to make a unique identification for each chip and use it to detect IP violation. Extrinsic IC metering methods maybe either active or passive. Fingerprinting schemes [8] [9] assign a unique fingerprint on each IP that the manufacturer is allowed to use. The manufacturer is supposed to use each IP once when producing the chips. Therefore, each chip would have a unique fingerprint compared to the other chips. Then, the design company can detect the IP violation by finding out the chips with the same unique fingerprint. Another extrinsic metering scheme [3] adds a small programmable component in each design which can be configured in a unique way for each chip during the manufacturing process. The foundry reports to the design company all the IDs of the manufactured chips. To detect IP violation, the design company would conduct a random sampling in the market and record the number of unreported chips. From a statistical analysis based on collision probabilities computed by the Birthday paradox, the number of unauthorized chips can be estimated.

Extrinsic metering approaches can detect IP violations but they require a high instrumentation to either the design or manufacturing process. It complicates the IC design process and increases the cost of each chip. Also, there are still security concerns in this scheme, because the design company do not have control over the manufacturing process, it is possible that untrusted manufacturers modify the assigned fingerprint or ID and compromise the IP protection scheme.

2.2 Intrinsic IC Metering

Intrinsic IC metering approaches do not interrupt the IC design and manufacturing process. Instead, they characterize the existing properties of the chips and assign a unique ID obtained from the characterization results of each chip. The IDs are used in the same way as in the extrinsic metering scheme. [1] proposed a CAD-based intrinsic passive IC metering approach. It characterizes each gate of an IC in terms of its delay on critical path. Because of the existence of process variation, the delay values of the gates are different even if they are from the same design. Therefore, the delay value can be used as a unique ID of the IC. Paper [5] proposed a nondestructive approach for gate-level characterization and a hardware metering protocol based on the characteristics. They analyze the probability of collision of IDs in presence of intra- and inter-chip correlations.

The intrinsic metering approaches avoid the instrumentation to the IC design and manufacturing process and are still able to generate unique IDs for the chips. However, they would require high accuracy in the gate-level characterization results. Also, the existing approaches did not provide quantified solutions in terms of the number of chips that a foundry may have produced.

3. PRELIMINARIES

In this section, we introduce the system models we use in the IC metering and auditing process, including power/delay models and process variation model.

3.1 Power and Delay Models

In our IC metering and auditing approach, we use leakage power ($P_{leakage}$), switching power($P_{switching}$), and delay ($Delay$) as the conventional structural test parameters of an IC, which provide a manifestation of the physical properties such as effective channel length (L_{eff}) and threshold voltage (V_{th}). We call these the *manifestational* properties. This subsection introduces the power and delay models we use in this paper, which are obtained from [10].

There are typically two possible sources for power dissipation on an IC. One is from gate switching, in which the ICs dissipate power by charging the load capacitances wire and gates. The other source is leakage power, where even if the gates do not switch, they dissipates power due to the leakage current. Equation (1) is the gate-level subthreshold leakage power model [10], where L is effective channel length, V_{th} is threshold voltage, and T is temperature. The other parameters are considered as constants in the discussion of this paper and can be found in paper [10].

$$P_{leakage} = 2 \cdot n \cdot \mu \cdot C_{ox} \cdot \frac{W}{L} \cdot (\frac{kT}{q})^2 \cdot D \cdot V_{dd} \cdot e^{\frac{\sigma \cdot V_{dd} - V_{th}}{n \cdot (kT/q)}} \quad (1)$$

We note that the leakage power has a nonlinear (exponential) relation with the temperature T, which provides us with a method to vary the leakage power. In particular, if we apply a set of input vectors to the circuit that switch a subset of gates, the gates can be heated up and the heat will be transferred to other gates on the circuit, which causes the temperatures of the gates to vary over time. In this way, we can condition the temperatures of the circuit and utilize the exponential relation between temperature and leakage power for our IC metering purpose.

Equation (2) describes the gate-level switching power model [20], where the switching power is dependent on switching probability α, load capacitance C_L, gate width W, gate length L, and supply voltage V_{dd}.

$$P_{switching} = \alpha \cdot C_L \cdot W \cdot L \cdot V_{dd}^2 \quad (2)$$

Equation (3) shows the gate-level delay model [10] that depends on L and V_{th} in a non-linear manner.

$$Delay = \frac{k_{tp} \cdot k_{fit} \cdot L^2}{2 \cdot n \cdot \mu \cdot \phi_t^2} \cdot \frac{V_{dd}}{(ln(e^{\frac{(1+\sigma)V_{dd} - V_{th}}{2 \cdot n \cdot \phi_t}} + 1))^2} \cdot \frac{\gamma_i \cdot W_i + W_{i+1}}{W_i} \quad (3)$$

where subscripts i and $i+1$ represent the driver and load gates, respectively; γ is the ratio of gate parasitic to input capacitance; and k_{tp} and k_{fit} are delay-fitting parameter and model-fitting parameter, respectively.

The power and delay models connect the manifestational properties with the physical level properties. We employ these models in our physical level characterization approach under the consideration that gate-level physical properties (e.g., L and V_{th}) would naturally serve as a unique ID of an IC because of process variation.

3.2 Process Variation Model

Process variation (PV) in IC manufacturing is the deviation of IC parameter values from nominal specifications, due to the nature of the manufacturing process [11] [12] [13]. PV causes major variations in gate-level physical properties such as L_{eff} and V_{th}, which are two major sources of PV. In the discussion of this paper, we follow the quad-tree model presented in paper [14] for the variation of L_{eff} (ΔL). In particular, ΔL is distributed into multiple levels where there are different number of grids allocated on each level. The grids on each level are assigned variation values that follow a Gaussian distribution. Then, the total ΔL can be calculated as the sum of variation values on each level of the grids to which the corresponding gate belongs. Equation (4) shows the total variation in the effective channel length of gate j, where ΔL_{ij} is the variation in the ith level grid to which gate j belongs, and μ_i and σ_i are parameters of the Gaussian distribution at level i.

$$\Delta L_j = \sum_i \Delta L_{ij}, \qquad where \quad \Delta L_{ij} \sim N(\mu_i, \sigma_i) \quad (4)$$

For V_{th}, we use the model proposed in [15], where the distribution of V_{th} is obtained by the simulation study of random dopant. V_{th} in this model is fit into a Gaussian distribution, where the parameters are determined by the dopant number and dopant position.

4. IC METERING AND AUDITING MODEL

Our IC auditing approach is based on physical level characterization and IC process characterization, by which we identify the physical level properties of each gate accurately and use them as the IDs. Next, we conduct IC counting using sampling and labeling based techniques. Fig. 1 shows the overall flow of our IC metering and auditing model. We first sample a set of chips and conduct physical level IC characterizations. In particular, we take in the power/delay measurements of the gates and use Equations (1) to (3) to formulate a nonlinear equation with variables V_{th} and L_{eff}. By conditioning the temperatures of the IC and characterizing the power and delay values at the manifestation-level, we can formulate a system of nonlinear equations. The solution to the system of nonlinear equations would provide us with estimated values of V_{th} and L_{eff} for each gate on the sampled chips. After that, we begin our process characterization step, in which our goal is to generate the PV model for all the chips that the foundry may have produced. We use the PV model discussed in Section 3 in this step and characterize the distribution parameters (mean and variance) by using a fitting algorithm. Then, we use the calculated physical level PV distribution in the coincidence estimation process, where we estimate the probability of coincidence, i.e., the probability of false positives that we identify a chip as of our design but indeed it is not, as well as the probability of false negatives that we identify our chips as from other designs by mistake. Based on the coincidence estimation, we conduct IC counting by sampling and labeling the chips in the market to obtain a prediction of the number of manufactured chips.

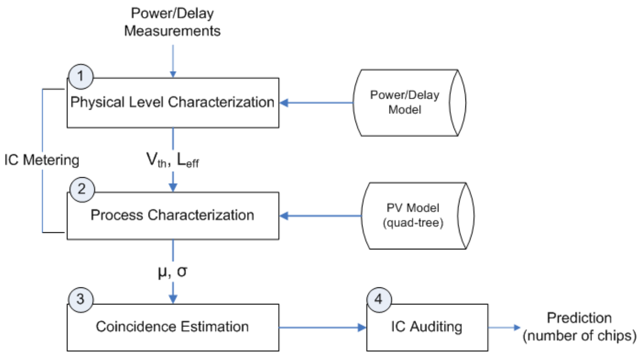


Figure 1: Overall flow of our IC metering and auditing scheme.

5. IC METERING BASED ON PHYSICAL LEVEL CHARACTERIZATION

5.1 Overview

In order to predict the number of chips produced by the foundry, we must first uniquely identify an IC and distinguish it from other ICs. We call this process IC metering or IC identification. Our goal in IC metering is to characterize the physical level properties of the sampled chips and quantify the PV model for all the manufactured chips. By accomplishing this, we can have an accurate method to uniquely identify the chips, as well as a global statistical view of all the chips in the market.

In IC metering, we take into account both the manifestational test properties (e.g., power and delay) and physical device properties (e.g., threshold voltage and effective channel length). From Equations (1) to (3) we know that the conventional gate level manifestational properties are impacted by many variables, which make the property values unstable and unpredictable. For example, the temperature (T) impacts leakage power exponentially, which means that the leakage power would have a large variation when the temperature varies due to IC activities or environmental factors. Therefore, unless one is in very controlled settings, the manifestational properties are not appropriate for the purpose of IC identification and, therefore, we consider using physical level properties as the IDs for the chips.

Our flow of IC metering is shown in Fig. 2. We first conduct gate-level characterization to determine the power/delay of each gate on the sampled chips, which is done by solving a system linear equations using linear programming. Then, we conduct physical level characterization to calculate the V_{th} and L_{eff} of each gate, based on the manifestational properties and the models shown in Equations (1) to (3). This is a nonlinear programming process since the models of power and delay are nonlinear with V_{th} and L_{eff}. Finally, we conduct process characterization to determine the parameter values in the PV model of all the manufactured chips.

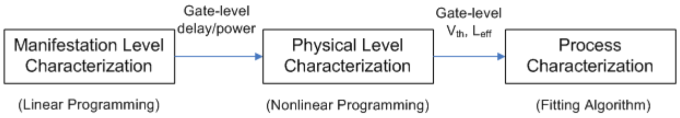


Figure 2: Flow of IC metering.

5.2 Manifestation-level Characterization

We use the GLC method proposed in [18] and [19] to characterize the manifestational properties in the presence of PV. In the GLC method, the power and delay models are expressed in a linear format assuming that the variation of all the physical level properties is represented by a single PV scaling factor K. Then, the value of K can be obtained by solving a system of linear equations. For example, Equation (5) shows the linear model of using leakage power.

$$\tilde{p_j} = e_{sj} + e_{rj} + \sum_{\forall gate\ i=1,\ldots,n} K_{ij}\ s_i \qquad (5)$$

where $\tilde{p_j}$ is the full-chip leakage power at input state j; s_i is the PV scaling factor of gate i; K_{ij} is the nominal leakage power for the gate at input state j, which is dependent on the constant parameters in Equation (1) and the input states; and e_{sj} and e_{rj} are systematic and random measurement errors, respectively. We can obtain a system of linear equations by varying the primary input vectors and measuring the leakage power of the entire circuit for each of them.

5.3 Physical Level Characterization based on Thermal Conditioning

From the characterization results from the manifestational properties, we are able to formulate a nonlinear equation based on Equation (1) in the following format:

$$P_{leakage} = \frac{A}{L} \cdot T^2 \cdot e^{\frac{C-V_{th}}{BT}} \qquad (6)$$

where L and V_{th} are the two variables that we are characterizing. A, B and C are transistor level parameters in the leakage power model that we assume as constant values.

Equation (6) provides us with a nonlinear equation that relates L_{eff} and V_{th} to the manifestational properties (leakage power). We can obtain the leakage power value from the characterization as discussed in Section 5.2. However, with only one nonlinear equation, we are not able to solve two variables L_{eff} and V_{th}. Therefore, we must find a way to add additional variations to the leakage power model, so that a system of nonlinear equations can be obtained. We achieve this goal by varying the temperatures of the circuit using thermal conditioning. As discussed in Section 3, leakage power has an exponential relation with temperature T, and we can use thermal conditioning to control the temperatures and obtain multiple leakage power nonlinear equations for each single gate. By applying different T to the IC and repeat the manifestational property characterization in terms of leakage power, we can formulate a system of nonlinear equations. Then, we solve the nonlinear equations using a nonlinear program solver and obtain characterization results for V_{th} and L_{eff}.

5.4 Process Characterization

In process characterization we aim to find out the quantified PV model parameters as discussed in Section 3 for all the manufactured chips. In particular, for the quad-tree model of L_{eff}, we characterize the Gaussian distribution parameters μ_i and σ_i for all the ΔL_i. For the model of V_{th}, we find out the mean and variance in the Gaussian distribution.

For the V_{th} distribution, we can refer to a Gaussian fitting tool that can provide distribution parameters (mean and variance). Then, we use the obtained parameters as the estimation of those for the entire chip population. For the quad-tree model of L_{eff}, the problem becomes more complicated because it is a sum of multiple Gaussian distributions on multiple levels, and there is no direct way to break down the compound distribution and obtain parameter values for each single distribution. In order to solve the problem, we develop a decomposition algorithm and use a divide-and-conquer approach to keep fitting the sampled L_{eff} (compound distribution) to individual distributions. The objective in this process is to fit the individual distributions to Gaussian distributions as accurate as possible, i.e., optimize the approximation error provided by the Gaussian fitting tool for each individual distribution. Our solution is based on the fact that a Gaussian distribution is infinitely divisible, i.e., a Gaussian distribution X with mean μ and variance σ can be decomposed to multiple Gaussian distributions X_i with mean of μ_i and variance of σ_i, where the following equations hold:

$$\sum_i \mu_i = \mu \tag{7}$$

$$\sum_i \sigma_i^2 = \sigma^2 \tag{8}$$

Based on this divisibility feature of Gaussian distribution, we design a decomposition algorithm of process characterization. We start from the highest level (root) of the quad-tree and conduct a breadth-first search of the tree. At each node, we guess and verify the constant component of the leaf node with the requirement that the remainder obtained by subtracting this constant component from the L_{eff} value should follows a Gaussian distribution, which is the L_{eff} value of the lower level nodes of the current node.

6. COINCIDENCE ESTIMATION

An important and difficult step in IC auditing is to be able to distinguish each chip from the others. Due to possible measurement and characterization errors in the IC metering process, there are possibilities of false positives and false negatives. The former means that we count chips that are not of our design as ours, and the latter means the opposite. Our goal in coincidence estimation is to measure the probabilities of false positives and false negatives, so that we can estimate their impacts on the accuracy of IC auditing.

We develop a Bayesian-based approach to calculate the probability of coincidence when only a single gate on each chip is considered. Then, we employ a majorization technique to conduct worst case analysis, which assumes that all the gates on the circuit have the same probability of coincidence as the gate with the largest possible probability. From this analysis, we obtain an upper bound of the probability of coincidence and use it for analyzing the impact on IC auditing.

6.1 Bayesian-based coincidence analysis

Since our IC metering is based on the characterization results of L_{eff} and V_{th}, there is a possibility that two gates that are not from the same chip would have the same measured L_{eff} and V_{th} due to either measurement errors or characterization errors.

Our IC auditing process works in the following way. We take a sample of chip from the market and conduct IC metering to obtain L_{eff} and V_{th}, and we label this chip according to the L_{eff} and V_{th} values. Then, we put this chip back to the market and continue to take other samples. When we take the next sample in, it is possible that it is not the same chip as the previous labeled chip but we characterize them as similar L_{eff} and V_{th} (i.e., false positives), or it is the same chip as the previous labeled chip, but we have different L_{eff} and V_{th} measurements (i.e., false negatives).

We employ Bayesian-based probability analysis [21] to calculate the probability of coincidence. Taking the false positive case as an example, we have the following Bayesian-based calculation:

$$P(H|D) = \frac{P(D|H) \cdot P(H)}{P(D)} \tag{9}$$

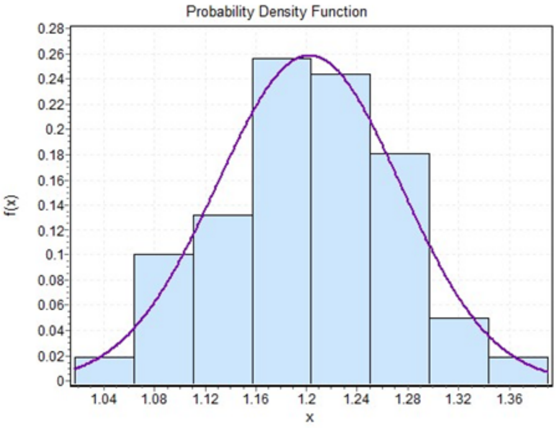


Figure 3: Probability that a gate has coincidence with other gates in terms of L_{eff} (benchmark C432 with 160 gates; mean value of L_{eff} is 1.2).

where H is the event that a gate matches with at least one other gate according to either L_{eff} or V_{th} measurement. D represents the event that we have a certain set of L_{eff} or V_{th} measurements for N sampled chips. Therefore, $P(H|D)$ represents that the probability that a gate matches with other gates under the condition that we have that certain set of measurements; and $P(D|H)$ is the probability of having the certain set of measurements under the condition that the gate matches with some other gates. We assume that $P(D|H)/P(D)$ forms a normalization constant that does not vary with the variation of D. We calculate $P(H)$ in the following way by using the rationale in the well known birthday paradox problem [25]:

$$P(H) = 1 - \prod_1^N (1 - P_i) \tag{10}$$

where P_i is the probability that a certain gate i matches with one another gate j. The value of P_i depends on the position of the L_{eff} or V_{th} in the whole distribution. Fig. 3 shows our simulation results of P_i in terms of L_{eff}. We can see that the gates with L_{eff} around the mean value ($L_{eff} = 1.2$ in the figure) of the distribution have a relatively large P_i.

Putting it all together, we have the following estimate for $P(H|D)$:

$$P(H|D) \propto 1 - \prod_1^N (1 - P_i) \tag{11}$$

We use two approaches for determining whether two chips match with each other: (i) an extreme method, in which we claim two chips are identical as long as there is an overlap between the distributions of their measured values; and (ii) a threshold approach, in which only when the overlap between two distributions exceeds a threshold value do we assume they are identical. We conduct simulations on ISCAS benchmarks for the false negatives and false positives in each approach. The results are shown in Table 1 and discussed in details in Section 8.

6.2 Majorization and Worst Case Analysis

As mentioned in the previous subsection, the probability P_i varies over the absolute value of L_{eff} or V_{th}, and it reaches the highest value if the sampled chip is at the mean value of the entire distribution. In order to conduct coincidence estimation considering all gates on a chip, we must take into account the variations of P_i for all the gates. In our coincidence estimation process, we approximate the value of each P_i using a majorization technique. In other words, we use the highest possible P_i (that of the mean value L_{eff} or V_{th}) to represent all the P_i values. In this way, we indeed overestimate the probability of coincidence and aim to obtain an upper bound value for the worst case analysis.

6.3 Summary of Coincidence Estimation

From the results in Table 1 that we will discuss in Section 8 in details, we can conclude that the worst case probability of coincidence is small enough to hold a large number of chips (e.g., in millions), and the probabilities of false positives and false negatives are close to zero. This conclusion enables us to assume that all the chips are distinguishable from each other and we can label them uniquely without overlaps. This is important in the next step of our IC auditing process, because the sampling and re-sampling in the IC auditing approach are based on replacement.

7. IC AUDITING USING SAMPLING

Based on the IC metering with near-zero false positives and false negatives, we are able to conduct IC auditing using a sampling approach. Our IC auditing scheme is based on the animal counting techniques proposed in the statistical field [22] [23]. The main idea is to predict the total population of a kind of animals by capturing and recapturing samples. In this section, we show how our IC auditing problem is adapted to the animal counting model and how we solve the IC counting problem based on the model.

7.1 Animal Counting Model

The animal counting problem was first studied for estimating the dynamic of biological populations. One of the widely used approaches is the capture-recapture method [22] [23], in which samples are taken and labeled at periodic intervals. Then, the total population can be predicted from the number of captured, and more importantly, recaptured animals in each sample. For example, in the fish counting problem discussed in [22], the following information is recorded for each sample i: (i) the total number of fish (t_i); (ii) the number of new fish (d_i); and (iii) the number of recaptures (r_i). Next, the probability of obtaining such a sample can be calculated by using binomial distribution:

$$p_i = \binom{t_i}{r_i} (\frac{M_i}{N})^{r_i} (1 - \frac{M_i}{N})^{d_i} \tag{12}$$

where N is the total number of fish, and M_i is the number of labeled fish when the ith sample is drawn. Assuming all the samples are taken randomly and independently, the probability of obtaining n samples with specific t_i, d_i, and r_i is the product of p_i: $P = \prod_{i=1}^{n} p_i$. Then, by using maximum likelihood analysis, paper [22] gives the equation that holds for N and M_i:

$$\sum_{i=1}^{n} \frac{d_i M_i}{N - M_i} = \sum_{i=1}^{n} r_i; \tag{13}$$

Paper [22] solves the equation and gives an approximation solution of N as the following:

$$N = \frac{\sum_{i=1}^{n} t_i M_i}{\sum_{i=1}^{n} r_i} \tag{14}$$

Equation (14) indicates that the predicted number of fish is a function of t_i, M_i, and r_i. All of these parameters can be obtained easily from the sampling and labeling process.

7.2 IC Auditing

Our IC auditing problem is similar with the animal counting problem, in both required inputs and outputs. However, we must analyze the assumptions behind the problem and verify that our IC counting problem still makes the assumptions hold. We note that the fundamental assumptions that are required by the animal counting model include the following: (i) there must be a method to uniquely label the captured samples; and (ii) the sampling model must be with replacement so that the captured samples can be recaptured, which provides an indicator on how large the total population is.

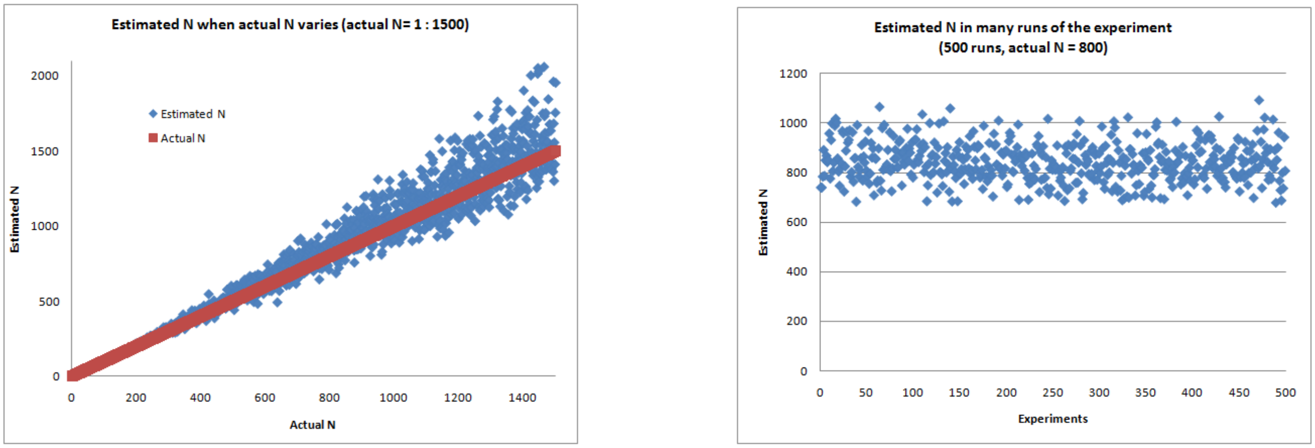
From the discussion in Section 6, the first assumption holds because the probability of coincidence becomes extremely small when we consider all the gates on the chip. For the second assumption, we make our IC auditing process spread into the IC marketing period. In other words, we collect IC samples periodically and put them back into the market after each sampling period. This would make our auditing process long, but it is doable. Furthermore, the number of samples can be adjusted according to the required accuracy of the prediction results. Section 8 gives an analysis of the prediction accuracy in terms of the number of samples taken, which can serve as a reference of how many samples are needed and an estimation of how long the entire IC auditing process would take.

Based on the above analysis, we apply the animal counting technique to our IC auditing problem. We use the same symbols of t_i, d_i, and r_i as in Section 7.1 for chips. The number of chips can be predicted by Equation (14).

7.3 Post-processing

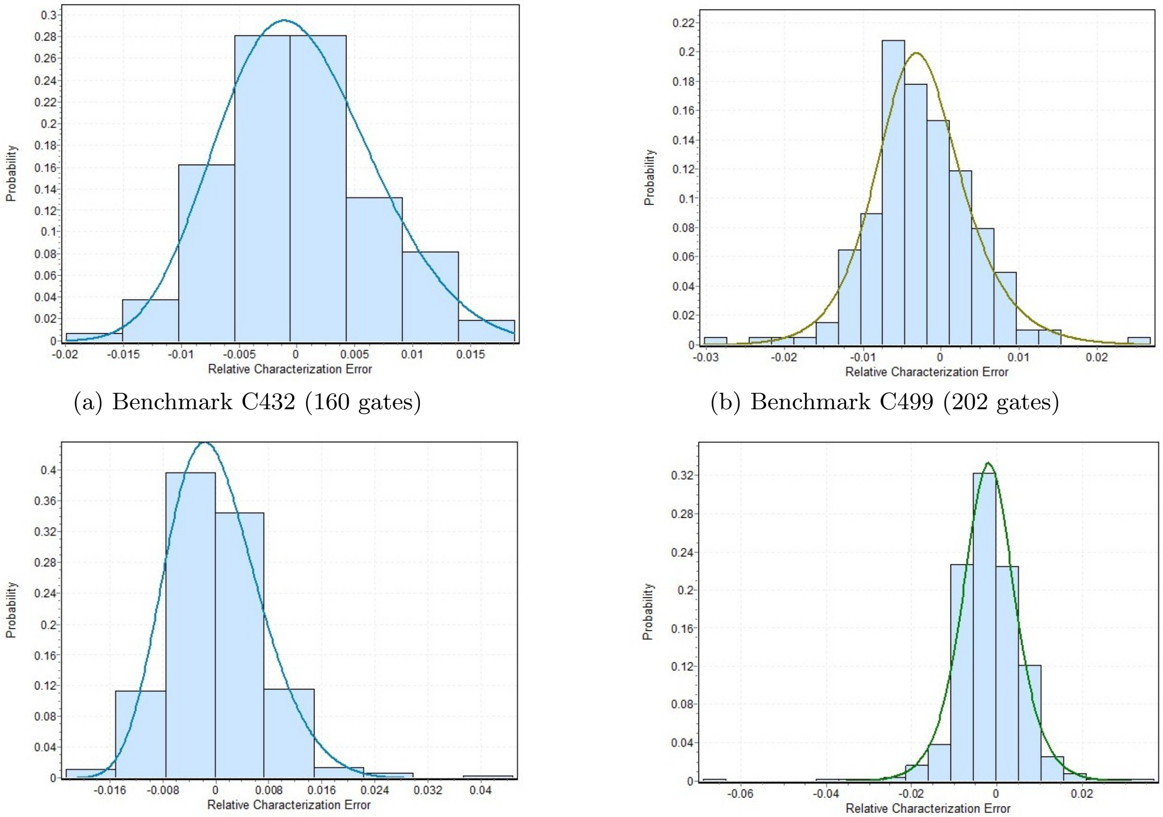
The accuracy of the prediction results can be impacted by many factors, such as the degree of independence of the samples and the approximation errors in the animal counting model. In order to improve the accuracy of IC auditing, we employ a statistical method, namely maximum likelihood estimation to post-process the data after many runs of the sampling experiments have been conducted. Then, we apply goodness-of-fit tests [24] on the data from each run, and estimate the statistical distribution of the predicted results over different runs. According to the distribution that each result follows, we obtain its approximate density function, i.e., $p(N)$, and set our estimated value of N to be the one that maximizes the likelihood function:

$$\tilde{N} = argmax_N \; log \; p(N) \tag{15}$$



(a) (b)

Figure 4: Validation of our IC auditing scheme: (a) on known sets of chips; N varies from 1 to 1500; (b) on 500 runs of the analytical simulation; N is fixed to 800.



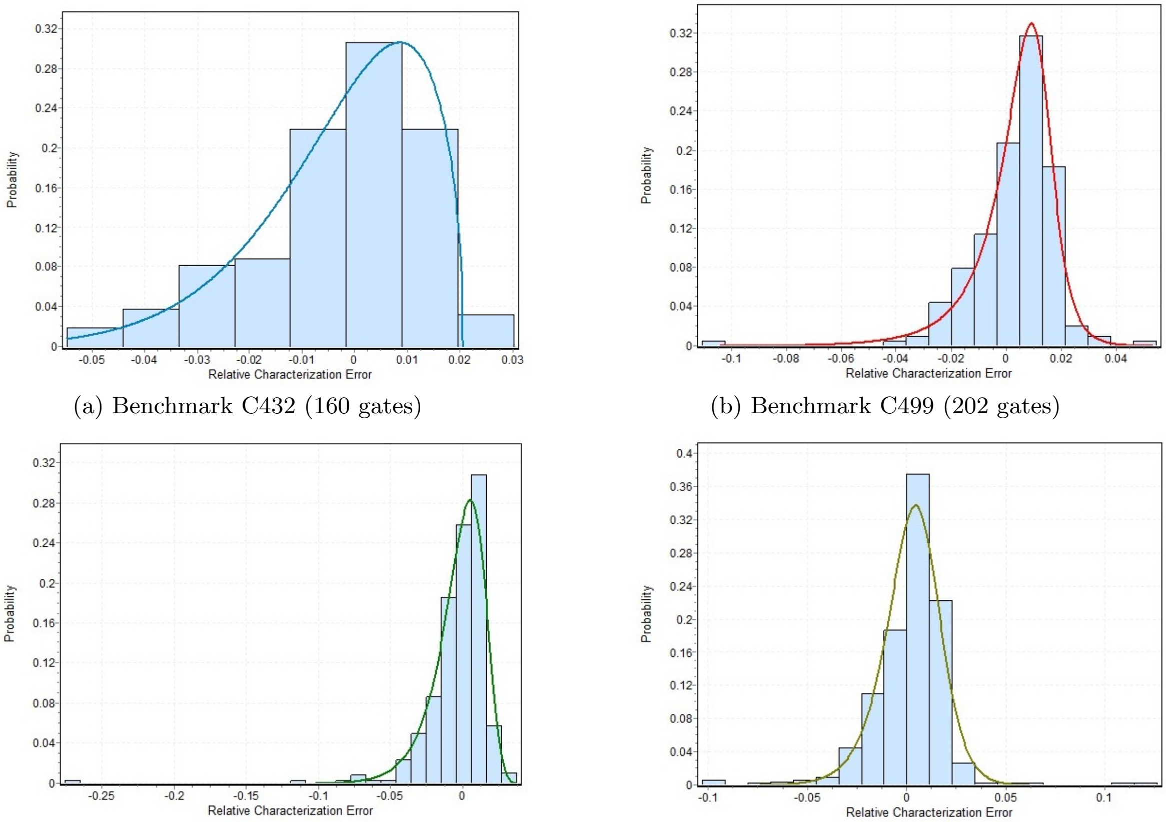
(a) Benchmark C432 (160 gates)

(b) Benchmark C499 (202 gates)

(c) Benchmark C880 (383 gates)

(d) Benchmark C1355 (546 gates)

Figure 5: Accuracy of L_{eff} characterization on a set of ISCAS benchmarks.



(a) Benchmark C432 (160 gates)

(b) Benchmark C499 (202 gates)

(c) Benchmark C880 (383 gates)

(d) Benchmark C1355 (546 gates)

Figure 6: Accuracy of V_{th} characterization on a set of ISCAS benchmarks

Table 1: Accuracy of Coincidence Estimation

Benchmark	Relative GLC Error (%)	Extreme Method		Threshold Method	
		False Positive (%)	False Negative (%))	False Positive (%)	False Negative (%)
C432	-2.0 ~ +1.5	6.9	0	5.5	64
C499	-2.0 ~ +3.0	9.8	0	7.8	64
C880	-6.0 ~ +4.0	19.0	0	15.4	64
C1355	-6.0 ~ +2.0	15.4	0	12.4	64
C1908	-3.2 ~ +3.2	12.4	0	10.0	64
C2670	-3.0 ~ +3.0	11.6	0	9.4	64
C3540	-3.0 ~ +3.0	11.6	0	9.4	64
C5315	-3.0 ~ +3.0	11.6	0	9.4	64
C6288	-2.0 ~ +3.0	9.8	0	7.8	64
C7552	-3.2 ~ +2.4	11.6	0	8.6	64

7.4 Validation

We can validate our prediction results in two ways. One is to experiment it directly on a known set of chips. By comparing the actual number of chips and our predicted results, a conclusion can be drawn on how accurate our prediction model is. Fig. 4(a) shows one of validation results, in which we apply our IC auditing approach to unknown sets of chips with up to 1500 chips. For each set of chips, we plot and compare our prediction results with the actual number of chips. We observe from Fig. 4(a) that the estimated N is close to the actual N, but the distance between them is increasing as the actual N grows. This is due to the fixed number of samples and sample sizes, which are not enough when the population is large. We analyze this problem in more details in Section 8 by simulating a varying number of samples in the IC auditing process.

Another method for validation is to conduct statistical analysis on multiple runs of the sampling process. In particular, we repeat the experiment many times and compare the results of each run in terms of the variance of the predicted results. If the variance is within a small enough range, it indicates that our prediction model converges and is stable. Fig. 4(b) shows our validation results based on this method. We repeat the experiment 500 times for a known set (800) of chips. The plotted results of predicted number of chips indicate that they are within the range of 25% of the actual number of chips.

8. SIMULATION RESULTS

We simulate our IC metering and auditing schemes on a set of ISCAS benchmarks. We use leakage power as the manifestational test properties in the simulation because every gate on the circuit has leakage power regardless of its activities. This provides us with more variabilities in metering and labeling the gates.

8.1 IC Metering

We use the manifestation-level characterization results as well as thermal conditioning to formulate a system nonlinear equations. In this simulation, we use 20 nonlinear equations (temperatures) per gate and obtain V_{th} and L_{eff} for each gate by solving the system of nonlinear equations. We solve the nonlinear equations by using the Gauss-Newton method provided by Matlab. The PV model we use in the simulation is the quad-tree model as discussed in Section 3.

Fig. 5 and Fig. 6 show the accuracy of our characterization results for L_{eff} and V_{th}, respectively. In each benchmark, we characterize each gate and compare the characterization results with the actual values to calculate the accuracy of our characterization. We plot the relative characterization errors for all gates in histograms and fit them into a distribution as shown in the curves. We can see from the curves that we have less than 1% of average errors and less than 5% of maximum errors except for few outliers. We consider these error distributions in the next steps where we conduct coincidence estimation and IC auditing.

8.2 Coincidence Estimation

We perform coincidence estimation on the same set of ISCAS benchmarks and characterize the probabilities of false positives and false negatives when using both the extreme method and the threshold method (with a 20% threshold). Table 1 shows the results when considering one single gate on each chip. The extreme method gives zero false negatives and false positives from 6.9% to 19.0%, while the threshold method has lower false positives from 5.5% to 15.4% and constant false negative values depending on the threshold value.

Given the coincidence estimation for having only one single gate considered on each chip, we can calculate the probabilities of false negatives and false positives that consider all gates on the chip by using Equation (11). We find that the probability of coincidence becomes extremely small (e.g., 10^{-95} for benchmark C432) because of the large number (e.g., at least 160) of gates in the benchmark circuits. Considering the fact that there are many more (in millions or more) gates on a single chip in modern IC design, the probability of coincidence is very low even if there are huge number of chips in the market.

8.3 IC Auditing

In our IC auditing simulation, we evaluate the IC counting model in terms of the prediction accuracy. Also, we estimate the impact of the sampling parameters, such as the number of samples and the sample size, as well as the impact of the total number of chips on the prediction accuracy.

8.3.1 Prediction Error vs. Number of Chips

We simulate our IC auditing scheme on different numbers of chips in order to find out how the total number of chips would impact the prediction accuracy. In Fig. 7 we show the relative prediction errors when the number of chips varies from 1 to 2000, the number of samples is fixed at 20, and the size of each sample is 20. We observe that the relative prediction error becomes higher as the number of chips increases, but it is always below 15%. Also, we observe that the variance of the prediction error grows as the number of chips increases. This is due to the insufficient samples compared to the total number of chips. We will discuss the impact of the number of samples later in Section 8.3.2. Also note that the results in Fig. 7 are obtained without post-processing, i.e., each experiment is conducted only once, which is another reason why the variance of the prediction error increases.

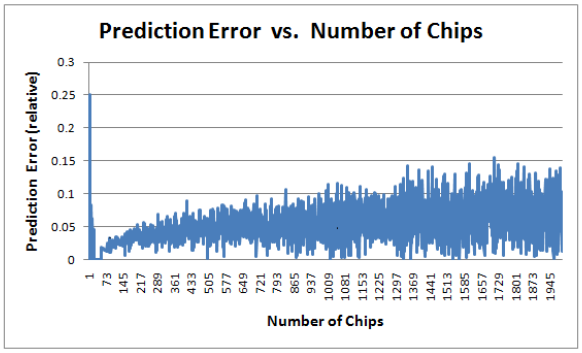


Figure 7: IC auditing results: prediction error vs. total number of chips (the number of samples is fixed to 20; the sample sizes are fixed to 20; the total number of chips varies from 1 to 2000; and no post-processing of the prediction results is performed).

Table 2 shows our simulation results on a large number of chips (up to 100 million). In this set of simulation, we set the sample rate (the ratio between the number of sampled chips and the total number of chips) as 0.1%, 0.5%, or 1.0% of the total number of chips. Also, we repeat each experiment 100 times and conduct MLE post-processing towards the collected results. We observe that the estimation error decreases as the increase of the number of chips with the same sample rate. Also, a sample rate of 0.5% can provide us with estimation errors below 5% for large numbers of chips (10^7 or 10^8).

Table 2: IC auditing on large numbers of ICs.

Number of ICs	Total Sample Rate	Estimation Error
10^6	0.5%	15.0%
10^6	1%	3.77%
10^7	0.1%	29.9%
10^7	0.5%	2.16%
10^8	0.1%	6.34%
10^8	0.5%	5.04%

8.3.2 Prediction Error vs. Number of Samples

We find in our simulation results that the number of samples taken plays an important role in the eventual prediction accuracy. In order to find out more about the impact of the number of samples, we perform simulations on 1600 chips while varying the number of samples from 10 to 50, with 20 chips in each sample. We show the results in Fig. 8. We can observe that the prediction accuracy keeps improving as the number of samples increases. This verifies our intuition that the prediction becomes more accurate with more information from the samples.

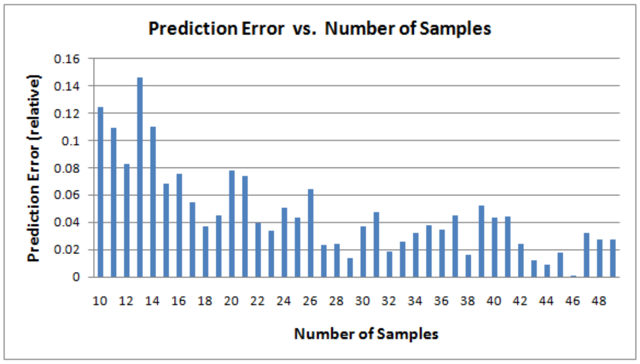


Figure 8: IC auditing results: prediction error vs. number of samples (the number of chips is 1600; the number of samples varies from 10 to 50; and the sample sizes are fixed to 20 chips.)

8.3.3 Prediction Error vs. Size of Samples

We further investigate the possible impact of the sample sizes by conducting a set of simulations on 1600 chips, with a fixed number of samples (e.g., 20) and varied sample sizes (e.g., from 10 to 50). We show the results in Fig. 9, where there are no improvements in the prediction accuracy as we increase the sample sizes.

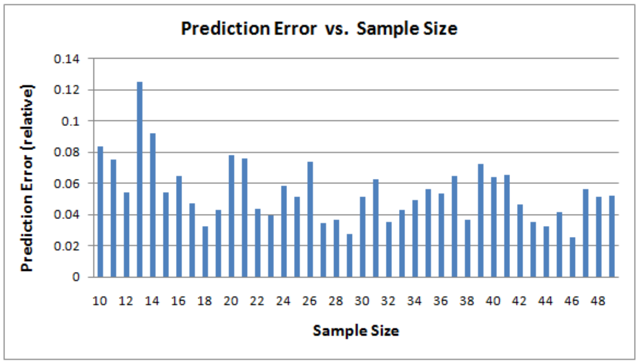


Figure 9: IC auditing results: prediction error vs. sample sizes (the number of chips is 1600; the number of samples is fixed to 20; and the sample sizes vary from 10 to 50.)

9. CONCLUSION

We have developed a new intrinsic passive IC metering scheme based on physical level IC characterization, which provides a new labeling method to identify different chips. Our estimation of coincidence proves the uniqueness of the IC labeling for large numbers of chips. Based on the IC metering scheme, we audit the number of manufactured chips in the market by employing a statistical method, namely the animal counting model. We evaluate the accuracy of our IC metering and monitoring schemes on several ISCAS benchmarks, and the simulation results show that both IC metering and auditing are accurate even for small circuits and huge numbers of chips in the market.

10. REFERENCES

[1] F. Koushanfar, M. Potkonjak. CAD-based Security, Cryptography, and Digital Rights Management. DAC 2007, pp. 268-269.

[2] Y. Alkabani, F. Koushanfar, M. Potkonjak. Remote Activation of ICs for Piracy Prevention and Digital Right Management. ICCAD 2007, pp. 674-677.

[3] F. Koushanfar, G. Qu, M. Potkonjak. Intellectual Property Metering. Information Hiding 2001, pp. 81-95.

[4] Y. Alkabani, F. Koushanfar. Active hardware metering for intellectual property protection and security. USENIX Security Symposium, 2007, pp. 291-306.

[5] Y. Alkabani, F. Koushanfar, N. Kiyavash, M. Potkonjak. Trusted Integrated Circuits: A Nondestructive Hidden Characteristics Extraction Approach. Information Hiding 2008, pp. 102-117.

[6] F. Koushanfar. Integrated Circuits Metering for Piracy Protection and Digital Rights Management: An Overview. GLSVLSI 2011, pp. 449-454.

[7] F. Koushanfar. Hardware Metering: A Survey. Book Chapter, in: “Introduction to Hardware Security and Trust”, M. Tehranipoor and C. Wang (eds.), Springer, 2011.

[8] A. Caldwell, H. Choi, A. Kahng, S. Mantik, M. Potkonjak, G. Qu, J. Wong. Effective Iterative Techniques for Fingerprinting Design IP. IEEE Transactions on CAD, Vol. 23, No. 2, 2004. pp. 208-215.

[9] A. Kahng, D. Kirovski, S. Mantik. M. Potkonjak, J.L.

Wong. Copy Detection for Intellectual Property Protection of VLSI Designs. ICCAD 1999, pp. 600-604.

[10] D. Markovic, C. Wang, L. Alarcon, T. Liu, J. Rabaey. Ultralow-Power Design in Near-Threshold Region, Proceedings of the IEEE, Vol. 98, No. 2, 2010. pp. 237-252.

[11] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, V. De. Parameter Variations and Impact on Circuits and Microarchitecture. DAC 2003, pp. 338-342.

[12] B. Cheng, S. Roya, A. Browna, C. Millara, A. Asenov. Evaluation of statistical variability in 32 and 22 nm technology generation LSTP MOSFETs. Solid-State Electronics, Vol. 53, 2009. pp. 767-772

[13] S. Roy and A. Asenov, Where do the dopants go? Science, Vol. 309, No. 5733, pp. 388-390, 2005.

[14] B. Cline, K. Chopra, D. Blaauw, Y. Cao. Analysis and Modeling of CD Variation for Statistical Static Timing. ICCAD 2006, pp. 60-66.

[15] A. Asenov. Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1 um MOSFET's: A 3-D Atomistic Simulation Study. IEEE Transactions on Electron Devices, Vol. 45, No. 12, 1998, pp. 2505-2513.

[16] M. Naor and B. Pinkas. Secure and Efficient Metering. EUROCRYPT 1998, pp. 576-590.

[17] M. Liedtke. Google to Pay $90M in 'Click Fraud' Case. Washington Post Magazine, March 9, 2006.

[18] S. Wei, S. Meguerdichian, M. Potkonjak. Gate-level characterization: foundations and hardware security applications. DAC 2010, pp. 222-227.

[19] S. Wei, M. Potkonjak. Scalable Segmentation-Based Malicious Circuitry Detection and Diagnosis. ICCAD 2010, pp. 483-486.

[20] A. Srivastava. Statistical Analysis and Optimization for VLSI: Timing and Power. Springer, 2005.

[21] M. Mitzenmacher, E. Upfal. Probability and Computing: Randomized Algorithms and Probabilistic Analysis. Cambridge, 2005.

[22] Z. Schnabel. The Estimation of Total Fish Population of a Lake, The American Mathematical Monthly, Vol. 45, No. 6, 1938, pp. 348-352.

[23] A. Chao. Estimating the Population Size for Capture-Recapture Data with Unequal Catchability. Biometrics, Vol. 43, No. 4, 1987. pp. 783-791.

[24] P. Lewis, E. Orav. Simulation Methodology for Statisticians, Operations Analysts, and Engineers (Volome I), Chapman & Hall/CRC, Dec. 1988.

[25] P. Flajolet, D. Gardy, L. Thimonier. Birthday Paradox, Coupon Collectors, Caching Algorithms and Self-organizing search. Discrete Applied Mathematics, Vol. 39, No. 3, 1992. pp. 207-229.