Security Beyond CMOS: Fundamentals, Applications, and Roadmap

Fahim Rahman, Bicky Shakya, Xiaolin Xu, *Student Member, IEEE*, Domenic Forte, *Member, IEEE*, and Mark Tehranipoor, *Senior Member, IEEE*

***Abstract*— Hardware-oriented security and trust has traditionally relied on the dominant CMOS technology to develop security primitives and provide protection against different attacks and vulnerabilities. With CMOS nearly reaching its fundamental scaling limit and the shortcomings of current solutions, researchers are now looking to exploit emerging nanoelectronic devices for various security applications. In this paper, we discuss the unique features of three emerging nanoelectronic technologies, namely, phase-change memory, grapheme, and carbon nanotubes, and analyze how these features can aid in hardware security and trust. In addition, we present challenges and future research directions about how to effectively integrate emerging nanoscale devices into hardware security. We emphasize that an interdisciplinary initiative is needed for emerging technologies to reach their full potential in security and trust applications.**

***Index Terms*— Carbon nanotubes (CNTs), emerging nanoscale devices and technologies, hardware security and trust, phase-change memory (PCM), physical unclonable function (PUF), supply chain security, tamper detection, true random number generator (TRNG).**

I. Introduction

HARDWARE security has become an increasing concern in today's world, where securing software and protocols have become insufficient. The past few decades of research in this area have yielded many security primitives, such as physical unclonable functions (PUFs) and true random number generators (TRNGs), various defensive mechanisms, and other numerous applications to aid different aspects of hardware security [1]. However, many of these security strategies heavily rely on preexisting CMOS technologies, which are slowly saturating in development. Furthermore, new attack models and vulnerabilities are constantly emerging and cannot be adequately addressed by current CMOS technology with which the primitives/countermeasures have been developed.

More recently, nanoscale devices and technologies, such as phase-change memory (PCM), memristors, carbon nanotubes (CNTs), and grapheme, have emerged [2] with promising improvements in speed and performance over the conventional CMOS technologies. Being less mature than their CMOS-counterparts, they also provide some unusual features that may somehow limit their full-phase implementation for regular logic and memory applications. While much research into these devices has focused on their performance, power, and reliability, the opportunities they offer for security purposes have not been assessed adequately. Some of these devices have been recently adapted to design security primitives like PUFs and TRNGs. However, other equally important security issues, such as antitamper, counterfeit detection/avoidance, side-channel attacks, and reverse engineering, have not been fully explored in the context of emerging nanoscale devices.

In this perspective paper, we attempt to link emerging devices and their features with a wide variety of applications in security. To achieve this goal, we consider notable properties of emerging devices, such as PCM, CNT, and graphene, and point out how their distinct features can provide countermeasures to different threats and vulnerabilities beyond traditional CMOS-based security solutions. We also provide a security-oriented roadmap for these devices by discussing the challenges and limitations that hinder security requirements. We emphasize that hardware security solutions with these nanoscale devices are capable of providing effective solutions for new threats. However, we see that much effort is still needed to provide a holistic solution that balances both security and performance in ICs. In many cases, there is a lack of necessary stochastic models, designs, experimental demonstrations, and vulnerability analysis − from the security perspective − to offer a proper evaluation and implementation platform. We hope that this paper can serve as a guide for both device and circuit/system-level security groups in exploring new avenues of nanoelectronic security, and we urge that a multidisciplinary effort should be taken by both device and hardware security community. Since there is already some prior work investigating similar applications for memristors [3] and spintronic memory devices [4], and utilizing crossbar memory architectures for security applications for neuromorphic computing [5], along with NEMS structures [6] and camouflaging and polymorphic gates [7], we focus more on PCM and carbon-based structures (graphene and CNT) and their possible security applications in this paper. We also believe that our effort to build a bridge between the device-intrinsic unique features and designing hardware security primitives and countermeasures to existing attacks can be expanded in a similar fashion for any other emerging device with analogous qualities.

The rest of this paper is organized as follows. In Section II, we provide preliminaries to popular hardware security prim-

Manuscript received January 17, 2017; revised June 19, 2017; accepted August 1, 2017. Date of publication September 28, 2017; date of current version November 22, 2017. This work was supported by the Air Force Office of Scientific Research MURI Grant under Award FA9550-14-1-0351. *(Corresponding author: Fahim Rahman.)*

The authors are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: fahim034@ufl.edu; bshakya@ufl.edu; xiaolinxu@ece.ufl.edu; dforte@ece.ufl.edu; tehranipoor@ece.ufl.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2017.2742943

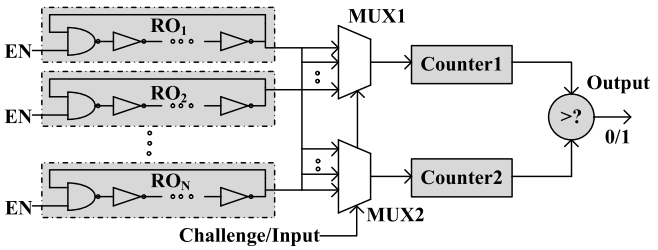


Fig. 1. RO PUF.

itives, attacks, and countermeasures. Section III briefly introduces PCM, graphene, and CNT with device properties and unique features that potentially lead to different security applications. In Sections IV and V, several hardware-oriented security applications leveraging PCM, graphene, and CNTs, respectively, are discussed. Related challenges that need to be overcome to use these nanodevices for security applications are also presented in this section. In Section VI, we highlight some open questions and future research directions in the emerging field of nanoscale security. Section VII concludes this paper.

II. Hardware Security Preliminaries

A good degree of security and trust can be ensured by employing circuits and sensors that leverage inherent properties of the hardware to design security primitives, and develop countermeasures to different attacks. In this section, we discuss some popular security primitives and some hardware-based attacks with related countermeasures.

A. Physical Unclonable Function

PUFs have been proposed in [8] and [9]. The fundamental idea behind a PUF is to create a cryptographically secure one-way function that produces a unique and unclonable output (response) to a given input (challenge). As the name suggests, a PUF can generate keys by leveraging *inherent physical variations from manufacturing processes*. Thus, identical (by design and lithography) integrated circuits manufactured by the same fabrication facility and process can generate different challenge-response pairs (CRPs) (or cryptographic keys) as there always exist small but nondeterministic variations in the manufacturing process. For example, a ring-oscillator (RO-) based PUF (Fig. 1) generates a physically unclonable response by comparing the frequencies of two identically designed ROs, as each of them oscillates at a slightly different frequency due to intrinsic process variations. This RO-PUF architecture has gained much popularity, since it does not require a rigorous design and can be implemented in both ASIC and reconfigurable platforms [10].

Till date, several security applications or protocols based on PUFs have been proposed, such as key generation for encryption IC identification and authentication and hardware metering [10], [11]. Since a PUF can produce responses on the fly, it offers an inexpensive, nonvolatile, and tamper-resistant alternative to conventional approaches, which store the keys in a nonvolatile memory (NVM). A PUF generally utilizes the uncontrollable analog behavior due to manufacturing process variation, and hence compromising the generated key should ideally be impossible without invasive attacks [12].

Among PUF quality metrics, the most populars are *uniqueness*, *randomness or uniformity*, and *reliability*. *Uniqueness* measures the distinctive quality of PUFs within a PUF-set by calculating the interchip Hamming distance (inter-HD) among all the instances for the same challenges. An inter-HD of 50% produces the maximum ideal uniqueness, meaning that any two PUFs within a given group have 50% different responses for the same given challenge on average. *Randomness or uniformity* stands for the unpredictability of a PUF showing if it has any bias or measurable trend in the generated key. An ideal PUF should maintain a good diffusive property, i.e., changing one bit in the challenge should randomly and equally flip the response bits. Such properties are crucial as a qualitatively poor PUF may be prone to different modeling and machine learning attacks [13], [14]. *Reliability* of a PUF assesses its capability to generate the same CRPs across different environmental conditions and over time by measuring average intrachip HD over different operating conditions and/or times. Ideally, a PUF should always maintain the same CRPs resulting in zero-bit error rate (i.e., 0% intra-HD) throughout its operational lifetime.

It is crucial that a PUF maintains high quality for reliable cryptographic applications. *Environmental (temporal) variations*, such as power supply and temperature variations and *aging* lead to PUF performance degradation and reliability issues. For example, the power supply noise and temperature variations have negative, though temporary, effect on the analog behavior (such as drain current, delay, leakage current, and so on) of the transistors in CMOS platform by impacting bias point, threshold voltage (V_{th}), mobility (μ), and other critical parameter, respectively, hence making the PUF performance less robust [15], [16]. *Aging*, on the other hand, creates permanent degradation to the critical parameters due to *bias temperature instability*, *hot carrier injection*, *time-dependent dielectric breakdown*, and *electromigration* (EM), and thus has significant impact on the PUF reliability with time [17]. PUFs suffering from reliability issues may produce an error up to 25% in the key generated for the cryptographic applications [18]. Researchers have proposed error correcting codes (ECCs) and other novel algorithms and architectures to improve PUF reliability [19]. However, they mostly result in a high area and power overhead, and are not suitable for lightweight applications. Also, they may potentially introduce security issues such as the secret key leakage [20]. It is worth noting that researchers mostly rely on algorithmic and architectural solutions to enhance PUF reliability till date, and the search continues to find a proper solution using inherent device properties.

B. True Random Number Generators

A TRNG is a security primitive that is widely used in security and cryptographic applications to generate session keys, one-time pads, random seeds, nonces, challenges to PUFs, and so on [23], [25]. It typically generates a random digital bitstream with high uncertainty, or *entropy*. Following

the definition of Shannon entropy [24], the equal probability of generating 0 or 1 produces the highest entropy. To generate an output that is "truly" random and not dependent to previous state/bit, a TRNG needs to rely on *device intrinsic electrical and/or thermal noise* that is inherently nondeterministic.

A typical random number generator consists of an entropy source, entropy extraction/sampling unit, and, in most cases, a cryptographic conditioning unit. The entropy source is the focal point of a TRNG, as the output quality highly depends on the raw entropy coming from this entropy source. For a TRNG, such sources are analog in nature and include random telegraph noise (RTN) found in scaled transistors, power supply noise, radioactive decay, latch metastability, jitter in ROs, and so on [22], [25]. On the other hand, a software-based random number generator—mostly known as a pseudorandom number generator (PRNG)—relies on algorithms and produces high throughput with lightweight implementations, although the output is deterministic for a very long sequence with a fixed seed.

In a typical intrinsic variation-based TRNG, the analog entropy source is sampled using the entropy extraction/sampling unit, e.g., a latch sampling an RO signal or a voltage comparator producing a digital output from a comparison of the RTN-prone signal to a reference voltage. More often than not, the problem with entropy sources is that although they might be "intuitively random," statistical test, such as the NIST Test Suite [26], DieHARD [27], run on the output of the TRNG show a certain level of bias and predictability, especially under different environmental conditions and deterministic manufacturing variations. To combat this, cryptographic hash functions, von Neumann corrector, and stream ciphers are employed to manipulate the raw output of the TRNGs to ensure the uniformity and statistical randomness. Also, additional tuners and processing blocks may be employed to control the TRNG quality and throughput [27].

Although fast and lightweight, a PRNG is not secure because its next state can be predicted from the current state if someone gains access to the design and seed. In communication and cryptographic applications, a predictable RNG can expose the sensitive data to the adversary, and it is, in such cases, not "truly random" anymore. On the other hand, physical variation-based RNGs are also prone to different physical attacks that tend to bias the outcome of the generator by altering the operational conditions and exploiting the dependence of physical entropy on different runtime parameters [21].

C. Design-for-Antitamper

Design for antitamper plays a crucial role in the security of silicon chips used for critical applications. Antitamper designs, for example, protect cryptographic keys and sensitive user data from being stolen, prevent unauthorized access fraudulent use, and denial of service attacks, protect intellectual properties (IPs), and prevent cloning. However, adversaries like insiders, outsiders, or funded organizations can carry on such attacks in different ways, for example, by probing, reverse engineering, remote attacks, and so on, that are invasive, semi-invasive, or noninvasive in nature. Prevention of such attacks naturally needs a thorough understanding of the threat model, cost, and possibilities, and require protection mechanisms and regular security evaluation [28].

Attacks that involve physical tampering, such as microprobing, may be invasive or semi-invasive in nature. Also, such physical attacks can allow the adversary to penetrate and probe from the top of the chip (frontside attack) for an easier access to critical nets and bus, or penetrate from the bottom (backside attack) providing a direct access to the active layer, i.e., transistors. Such attacks may involve grinding, polishing, and decapsulation of IC packages and chips, delayering, imaging, thermal cycling, and microprobing into selective areas of the die with or without power/frequency-based assisted attacks from small to large scale. The protection against such tampering attacks can, in a broader sense, be classified into two categories [28].

1) *Tamper-evident design* allows the authorized user to check whether a chip or system has gone through any physical tampering with mechanical and optical instruments. It, however, cannot prevent the attacker from accessing the system and stealing secret key/data.
2) *Tamper-resistant/tamper-responsive design* provides a higher level of security with the capability to "respond" to unauthorized physical access, for example, by erasing the secret key or by shutting down the system for good, based on available tamper-sensing mechanisms.

The security based on antitamper trivially relies on sophisticated tamper-sensing mechanisms that may utilize device-inherent properties or circuit-level solutions. Till date, these techniques have largely depended on creating power-net-based active shields, mechanical force sensors, and light-based sensors. The key idea behind it is that whenever the adversary tries to delayer and mill through the chip or does optical imaging, the active net and surrounding sensors get triggered and the key is erased [30]. However, a major challenge is that active power net can be bypassed easily by state-of-the-art focused ion beam (FIB) [31] attacks and microprobing attacks [29]. The use of nano/micro-elctromechanical systems (NEMS/MEMS)-based mechanical sensors and the optical sensors may also be futile, since the exerted small mechanical force may not activate the sensors placed in the die and powerful tools can do imaging outside the operational bandwidth of the optical sensors. Also, many of such schemes do not offer protections against silicon backside attacks, which can extend to permanent circuit modifications and data extraction [29].

D. Design-for-Anticounterfeit

Counterfeit ICs can be categorized into recycled, remarked, cloned, overproduced, or defective, with each type of counterfeit posing a unique threat to global semiconductor supply chain [32]. For example, overproduced and cloned chips manufactured by untrusted foundries/assemblies may cause legal issues and loss of profit for legitimate chip designers. On the other hand, chips that are recycled, remarked, or defective pose an even riskier threat, as they can crash critical infrastructures like military, transportation, and health-care systems.

To combat the threat of counterfeit ICs, various prevention and detection mechanisms have been proposed [32]. Detection mechanisms for counterfeit ICs usually involve the identification of the defects produced by counterfeiting. This includes physical inspection or electrical characterization to check for anomalies, such as black-topping or physical defects caused by sanding and changes in the threshold voltage degradation of the transistors. Physical inspection is, however, largely limited to recycled and remarked types.

Prevention mechanisms usually involve the design of sensors to detect IC usage in the field [33] or metering techniques to prevent unauthorized production by untrusted foundry [34]. Currently, the challenges for detection techniques include the need for "golden" samples for comparison (which might not always be available), and the lack of automation, high test cost, and long time required, especially for physical inspection methods. For prevention techniques, the challenge is the versatility and coverage of the proposed solutions. For example, the CDIR sensor, proposed in [33] only combat recycled ICs, whereas metering techniques proposed in [34] only control overproduction. Another limitation is that the solutions focusing on design-for-anticounterfeit incur design-level changes and additional overhead and, thus, are not applicable for legacy designs. Furthermore, the majority of design approaches focus on digital components, but analog and discrete components are just as often counterfeited, if not more. Most importantly, these schemes cannot guarantee the trust and security in the electronic component supply chain. This makes it more difficult for the countermeasures against counterfeiting to further provide a secure and trustable supply chain. The supply chain hardware integrity for electronics defense program from Defense Advanced Research Projects Agency (DARPA) [35] is currently developing miniature dielets that can be inserted into an IC package and then read in a contactless manner to detect cloned and remarked ICs. With nanoscale devices, there might exist opportunities to miniaturize and/or find new modes of developing counterfeit detection mechanisms for supply chain traceability.

E. Antireverse Engineering

Reverse engineering is an emerging threat to the semiconductor industry to possibly endanger the IP rights of designers. With specialized companies, e.g., TechInsights [36], reverse engineering has become popular for the proof of IP infringement. However, an adversary or untrusted party, by obtaining full knowledge of an IP/circuit design, can illegally use the derived information for piracy/counterfeiting, along with identifying vulnerabilities in the IC and develop low-cost noninvasive attacks. To combat this threat, several solutions have been proposed at different levels of abstraction: IP or software level, gate level, chip level, and PCB level. Circuit information in the form of gate-level netlists or HDL codes can be reverse-engineered by adversaries to steal information or in worst cases, to redesign with a hardware Trojan inserted. For chip-level reverse engineering, an adversary may sequentially grind away layers of an integrated circuit to image and gain layer-by-layer information of a circuit and work his/her way back to a netlist, at which point the original design has been compromised and can be used for illicit production by an adversary. To combat this, integrated circuits are shielded with power supply nets [37] or layers of chemical shields [38] to deter reverse engineering. Additionally, integrated circuit camouflaging [39] and split manufacturing [40] have been proposed to prevent an untrusted foundry from reverse engineering and obtain the functional netlist the design. Besides, PCB antireverse engineering solutions have also been proposed recently based on locked permutation of wires via CPLDs [41]. The outstanding challenge with most of these solutions is the overheads (area, timing, and power) associated with them. In the case of split manufacturing, a significant change in the design flow is also needed.

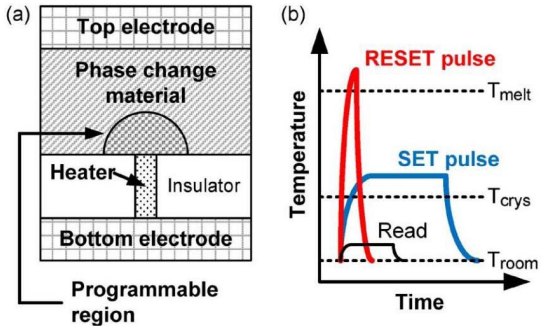


Fig. 2. (a) PCM "Mushroom" cell. (b) Program/read pulses [42].

III. Emerging Nanodevices With Unique Features

In the recent era, post-CMOS devices such as PCM, memristors, CNTs, and graphene, have shown promising potential with respect to speed and performance improvement over the traditional CMOS devices. Unfortunately, these devices are still immature and possess behavior that may limit their usage to traditional high-speed logic and memory applications. However, these same behaviors and properties might be leveraged for different hardware-based security applications which, in many cases, are less considered, in comparison to performance and reliability-focused analysis and applications. In this section, we discuss some promising properties of the emerging devices, namely, PCM, CNTs, and graphene.

A. Phase-Change Memory

PCM is an emerging nanoscale device that enables nonvolatile storage with high density and fast read/write operations. PCM is primarily based on chalcogenide materials, such as $Ge_2Sb_2Te_5$ (GST), and their transition to and from an amorphous (high resistance) phase and a crystalline (low resistance) phase. The difference in resistance between these phases is typically on the order of 10^2–10^4 [42]. To "reset" a PCM cell, a high-current pulse applied over a short duration ($\sim$ 50 ns) melts the GST by localized heating. It is then cooled rapidly, forming an amorphous plug that creates a high resistance between the electrodes of the PCM cell. For the "set" operation, a moderate current pulse with a longer duration ($\sim$120 ns) is applied to melt the GST, which is then

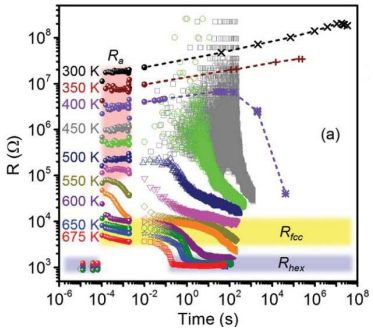


Fig. 3. Resistance drift in amorphized PCM cells at various temperatures [43].

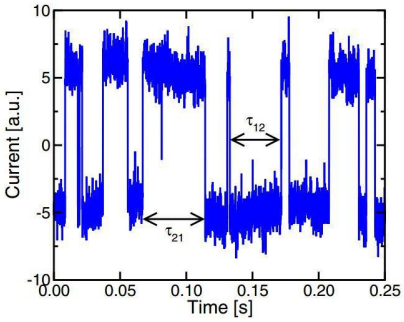


Fig. 4. RTN in a PCM cell at room temperature [45].

cooled down slowly for crystallization. A small read voltage determines whether the cell is in the amorphous state (considered as logic "0") or the crystalline state (considered as logic "1"). Additionally, a PCM cell can have a variety of geometries, e.g., mushroom cell structures (Fig. 2), μ-trench, line cell, and so on, with each geometry exhibiting different current requirements, scalability, and thermal properties [42].

We now identify a few features that are inherent in and, in some cases, exclusive to PCM devices.

1) *Programming Variability:* Stochastic programming variability in a PCM cell changes its resistance to a relatively moderate resistance window, not to a specific predefined resistance. For example, given two PCM cells, a reset operation with the same reset pulse yields two close but different resistance values. In both cases, the exact resistance is defined stochastically by the thermal properties and geometrical dimensions of the cell, such as GST layer thickness and bottom electrode contact diameter, that vary from cell to cell due to manufacturing process variation [44].
2) *Resistance Drift:* Resistance drift is a phenomenon whereby an amorphized PCM cell may have an increase or "drift" in resistance over time [43], and eventually change to crystalline phase with a drastic decrease in resistance (Fig. 3). This phenomenon is attributed to different amorphization/crystallization mechanisms of the cell, the cell structure and possible intermediate states. Though it is commonly considered a problematic issue for data retention and reliable operation, it may be useful for security applications.

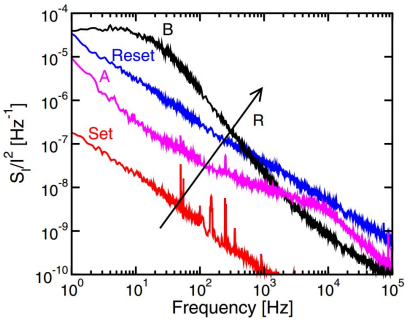


Fig. 5. Normalized PSD of set, reset, and intermediate states shows $1/f$-noise-like behavior for PCM devices [45].

3) *Random Telegraph Noise and* $1/f$ *Noise:* PCM displays RTN similar to ultrascale CMOS devices (Fig. 4) [45]. With this, the device current fluctuates randomly between several discrete stages within a broad range of timeframe. It causes short-term resistance fluctuations in PCM devices with the power spectral density varying with multiple parameters, such as cell contact area, temperature, and applied voltage. Also, it exhibits $1/f$ noise in the normalized power spectral density with Lorentzian components (Fig. 5) [45]. Such RTN and $1/f$ noises are great sources of variability, nondeterministic in nature, and inherent to the device itself and, hence, can act as intrinsic entropy sources for TRNGs.
4) *Multibit Storage Per Cell and Variability:* PCM also supports multilevel cell (MLC) operation, where the resistance window between the amorphous and crystalline states is used to store multiple bits in a single PCM cell. This is similar to flash-based multibit operation; however, a more unstable resistance window will create more variability for the multibit operation.
5) *Initial Forming Step:* PCMs sometimes require an initial "forming step." The resistance of a newly manufactured PCM cell in amorphous phase is much higher than the usual amorphous resistance of a reset PCM cell. Thus, to "form the device," a higher initial programming pulse is required. Note, however, that most PCM devices today are optimized with respect to the interface between the heater area and the chalcogenide material, thereby removing the necessity to "form" a device [46].

B. Graphene and Carbon Nanotube Electronics

Graphene and CNT-based electronics have emerged as lucrative alternatives to conventional CMOS-based digital applications as well as non-logic analog circuitry and sensors [2].

The main advantages that graphene and CNTs have over conventional silicon-based designs and architectures arise from their unique physical structures and associated energy-band diagrams. Graphene of a large area is a sheet of carbon atoms that are sp^2-bonded in a honeycomb lattice, whereas a CNT can be visualized as a seamless cylinder by rolling up graphene. Such structures create interesting energy-band diagrams, and hence electronic states that determine their

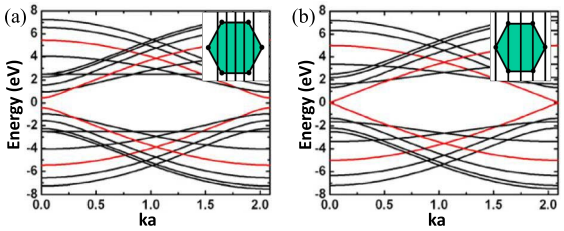


Fig. 6. Illustration of band-structures for (a) semiconducting and (b) metallic CNTs. Allowed wavevector lines are shown in respective insets [47].

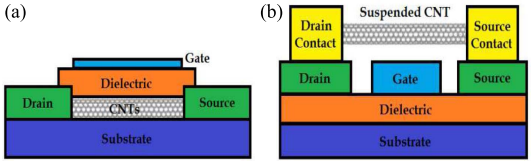


Fig. 7. (a) Top-gated CNTFETs. (b) Suspended-channel CNTFETs [48].

fundamental properties. For example, quantization of the electronic states in graphene results into subbands passing through the corner points (also known as *K-points* in reciprocal space) of the Brillouin zone, and thus it shows no energy bandgap (E_g), i.e., exhibits (semi)metallic properties for large area graphene. Patterning the graphene into nanoscale ribbon can increase the bandgap to offer semiconducting behavior, and thus can potentially be used as the channel material in graphene nanoribbon field-effect transistors (GFET). On the other hand, a CNT has quantization of the electronic states in the circumferential direction, with subbands having sets of 1-D dispersion relations and are determined by the periodic boundary condition around the circumference of the CNT. Thus, the generated subbands for a CNT may or may not pass through the K points, making it metallic or semiconductive, respectively (Fig. 6) and can be expressed as:

$$E_g(eV) \approx \frac{0.7}{d_{\mathrm{CNT}}(\mathrm{nm})} \tag{1}$$

where d_{CNT} is the diameter of the tube [2].

Additionally, CNT provides 1-D ballistic transportation for electrons and holes. A field-effect transistor that uses CNTs as channels (CNTFET) requires a low electric field in comparison to silicon-based transistors and can be used to create nano-CMOS architectures with robust performance. Hence, high mobility CNTFETs can be used in high-speed and ultralow power logic and RF applications [2].

1) *Metallic/Semiconducting Behavior:* As stated, it is uncertain whether the generated device would possess metallic or semiconducting behavior. For example, as given in (1), the bandgap energy of CNT may vary greatly due to the process variation and mismatch in the tube diameter [47]. Hence, the CNTFET performance may also vary in terms of the device electrical properties, such as current density and operating bias point, based on the inherent characteristics.
2) *Electrical Variability:* Different transistor architectures have been proposed using GNR and CNTs as channel materials (Fig. 7) to design high-mobility transistors [2], [48]. Since the property of these GFETs and CNTFETs greatly depends on the channel-GNR and CNT properties (e.g., semiconducting or metallic, and so on), length and patterning, drain/source-contact, CNT numbers and placements, and numerous other factors, the inherent sources of variability are quite large and largely manufacturing process-dependent. Also, due to ambipolar conduction, a CNTFET would not remain in the OFF state if the gate bias is swept too far unlike the ideal CMOSFET operation.
3) *Channel Sensitivity:* The channel material in GFETs and CNTFETs is highly sensitive to external excitation, which causes unwanted variations in transistor performance. Such excitation may arise from mobility variation due to operating conditions (such as exerted electric field and temperature), channel contamination, physical deformation in channel nanotubes, quality of passivation (i.e., whether the channel is contaminated or not), by photons, and other phenomena. Hence, most efforts have been focused on controlling the channel quality. Moreover, researchers have also leveraged this high sensitivity for many nanoscale sensor applications, since these effects can be translated into digital data for sensing. MEMS/NEMS architectures involving CNT structures and printable graphene electronics can also be utilized for various applications beyond conventional logics [49].
4) *Flexibility and Printability:* Solution-processable graphene sheets can be used for bulk scale printing, for example, using ink-jet printers, on both hard and flexible substrates to create transparent and functional electronic circuits that can potentially work as a processing block with an appropriately designed interface. Fig. 8 shows a simple structure of a single transistor constructed using printable graphene via ink-jet printing [50]. Here, graphene works as the channel material and can offer the similar functionality as that of a conventional CMOS transistor.

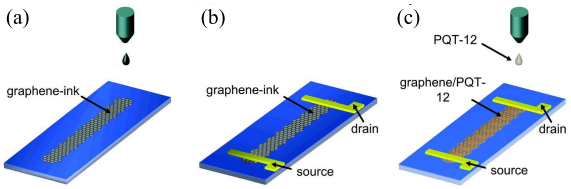


Fig. 8. Printable graphene electronics. (a) Ink on Si/SiO_2 to define channel. (b) Cr–Au pads define the source and drain contacts. (c) Layer of PQT-12 is printed on top to define gate [50].

We see that emerging nanoscale devices, such as PCM, graphene, and CNT, show inherent distinct properties that are not always prominently present in traditional CMOS devices. From a hardware security point of view, these features can be leveraged for different applications, for example, to build security primitives, such as PUFs and TRNGs, using the physical variability of the devices, to ensure supply chain security using printed electronics, to thwart physical tampering and reverse engineering attacks using device-level

TABLE I
UNIQUE FEATURES OF PCM AND CNT/GRAPHENE-DEVICES FOR HARDWARE SECURITY APPLICATIONS

Security Applications	Ideal Device Intrinsic Features	Nano-device Potentials for Security Applications	
PUFs	Non-deterministic Physical Variability due to Manufacturing Process Variations	PCM	Stochastic Resistance Variability due to Geometric Variation MLC Operation with Non-deterministic Resistance Window
		CNTs / Graphene	Unpredictable Bandgap and Metallic/ Semiconducting Properties Channel Variability due to Manufacturing Process
TRNGs	Random Variability and Device Intrinsic Noise	PCM	RTN and $1/f-$Noise
		CNTs / Graphene	Stochastic Physical and Electrical Variability
Design for Anti-Tamper & Anti-RE	Sensitivity to Environmental Variations and Physical and Chemical Impacts	PCM	Initial Forming Step
		CNTs / Graphene	Channel Sensitivity due to External Physical and Chemical Impact
Design for Anti-Counterfeit	Predictive Degradation (Aging and Usage) Printability for secure supply chain	PCM	Resistance Drift & Data Retention Failure
		CNTs / Graphene	Degradation Over Aging & Printable Electronics

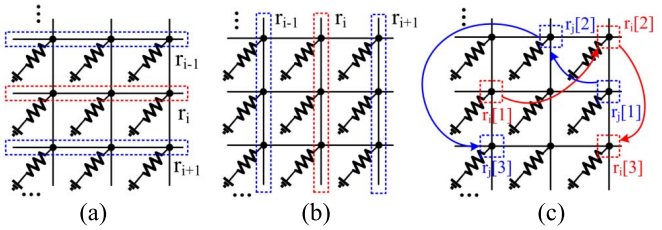


Fig. 9. Crossbar architecture for PCM CRP generation. (a) Horizontal, (b) vertical, and (c) random challenge selection [44].

inherent sensing mechanisms. In Table I, we highlight the ideal device-intrinsic features for different hardware security primitives and countermeasures as well as summarize the unique properties (of PCM and graphene/CNT-devices) that can be potentially leveraged for security applications. It should also be noted that PCM, graphene, and CNT-based designs are not traditionally sought for designing security primitives and countermeasures against hardware attacks. Hence, these devices have potentials to offer more, in terms of security, by suitable exploitation of the distinct inherent features with appropriate designs and architectures. In Sections IV and V, we elaborate the summarized idea given in Table I to analyze the potential hardware security applications driven by these emerging devices.

IV. SECURITY APPLICATIONS USING PCM

In this section, we discuss how PCM and its unique properties can be used to design potential security primitives as well as provide solutions to different attacks.

A. PCM for Physical Unclonable Functions

In this section, we will look at some PCM features that may have potential in creating or improving PUFs, while also reviewing prior work that has employed PCMs to create different types of PUFs. Furthermore, we also point out existing challenges for PCM-PUFs.

1) PCM-Based PUFs: Zhang *et al.* [44], [51] have used the variations between PCM cells in an array to generate keys (Fig. 9). Here, the leveraged intrinsic feature is that *PCM cells cannot be programmed in a deterministic fashion*. For example, given two PCM cells, a reset operation on them with the same reset pulse will yield two different resistances, where each resistance is defined stochastically by the geometrical properties of the specific PCM cells. With this, two PCM cells can be invoked by a challenge C, and a key can be generated by a simple comparison. Another advantage of this approach is that the PCM response generated also depends on the specific programming pulse used, as the current magnitude of the pulse changes the amorphous resistance. Thus, a different programming pulse would then yield another fresh set of CRPs, leading to a reconfigurable PUF, a concept introduced in [52]. However, this approach is not without its drawbacks, as the authors point out the need for postprocessing, using a logarithmic amplifier [51] and more elaborate ECCs [44], to remove bias.

2) PUFs Based on MLC Operation: Kursawe *et al.* [52] proposed the idea of PCM-based PUF using MLC feature. In this approach, a PCM resistance window is divided into N logical states and r_N windows within each logical state. Upon programming a PCM cell, a measurement device should not only be able to tell which logical state the PCM cell holds, but also what resistance window the PCM cell is in. Here, the position of the PCM cell in the resistance windows is dictated by process variations within a PCM cell. With this ability, a PCM cell may be programmed and then read out with fine granularity to obtain responses as a PUF.

3) Reliability Concerns Regarding PCM-Based PUFs: An important point to be noted for PCM is that crystallization and amorphization is a thermally activated process [42]. Thus, the impact of temperature and environmental variations on PCM-PUF key reliability would be a concern. Resistance drift, a phenomenon that is known to worsen with higher temperatures, is another issue for PCM. While the effect of resistance drift is limited in single-level cells, they are severe for MLCs, making it difficult to practically implement MLC based PCM PUFs. In addition to external temperature, the relatively high amorphization temperature can cause neighboring PCM cells to be disturbed, when an adjacent PCM cell is programmed [53] causing a "failure" in an adjacent PCM cell. Further, CRPs generated with PCM resistances may change over time as thermal disturbance changes over time. Moreover, since thermal disturbance is a stochastic process and depends on the frequency of set/reset operations too, the stability of CRPs generated by a PCM-based PUF could be affected. Although countermeasures for both resistance drift and thermal disturbance have been proposed in [53] and [54],

these solutions need to be reconsidered in the scope of security primitives, instead of memory structures, where the overheads for area/power might be very different. Furthermore, helper data algorithms and ECCs need to be analyzed for PCM-PUF-based implementations, as the area advantage provided by high-density PCM might be countered by the high area overhead of the postprocessing required.

Finally, we also mention withstanding issues and opportunities with PCM as a traditional NVM for key storage. Although PUFs offer a tremendous advantage over NVMs for secure key storage, the latter is still widespread in smart cards, embedded systems, and cryptographic key storage. Traditional NVMs, such as flash, have an array of vulnerabilities to data remanence attacks [55] and imaging attacks [56]. With PCM, these vulnerabilities are yet to be assessed. Intuitively, we could point out a few advantages of using PCM as the NVM memory for security. For example, a PCM-based memory would be inherently immune to EM-based attacks for key extraction (since amorphization/crystallization is purely thermal processes). Also, they could be more immune to data remanence attacks, compared to SRAM, as the set/reset operation changes the physical characteristics of the PCM cell, leaving behind little to no evidence of the previous state of the cell. However, to date, no experimental analysis of such features of PCM has been analyzed for security applications.

B. PCM for True Random Number Generators

PCM offers multiple sources for physical entropy. First, PCM offers RTN, also terms as burst noise, that is a phenomenon commonly exploited for high-quality random number generation. Also, RTN, exhibited as short-term resistance fluctuation in PCM, gets more prominent below the 90-nm regime, as the percolation path that 'averages out' the RTN gets shorter [45]. Furthermore, the work in [57] shows significant low-frequency RTN in 90-nm PCM devices, observed on programming a PCM cell to an intermediate resistance state of 150 $k\Omega$ at a low bias voltage of 0.2 V. Although RTN does exist in PCM devices, its suitability for random number generation is yet to be assessed, as questions regarding robustness and statistical randomness of generated bits for these PCM devices still need to be answered. In addition to RTN, the amorphous phase of GST in a PCM cell could possess a good source of entropy. Since amorphization is an intrinsically random phenomenon, the amorphization resistance reached by a PCM cell will vary stochastically from cycle-to-cycle, which could potentially be used as a source of true random numbers.

C. PCM for Tamper Detection

One feature that could be used for tamper detection using PCMs is to use its initial "forming step". The resistance of a newly manufactured PCM cell in amorphous phase is much higher than the usual amorphous resistance of a reset PCM cell. Thus, to "form the device," a higher initial programming pulse is required [46]. This can be used to check whether a PCM cell is "fresh" or has been tampered with, as a quick check of the amorphous resistance of new PCM cells (provided that the correct resistance value is known), or a count of the number of pulses required to crystallize the cell (for example, 20 pulses being required instead of 5 if the cell is new) can help to detect any tampering attempt on new PCM cells. However, most PCM devices today are optimized with regard to the interface between the heater area and the chalcogenide material, removing the necessity to "form" a device before using it [46]. Thus, to utilize this tampering-detection feature, PCM with older heater architectures might have to be used. In addition, self-powered light sensors, coupled with PCM as an NVM, can be used for effective tamper resistance. As illustrated in [58], an energy-harvesting photovoltaic sensor is coupled along with a portion of the PCM memory, possibly storing secret keys and highly reactive materials deposited as metal multilayers (e.g., $Si + 2B$ and $Cu + Pd$). When an invasive attack is attempted, the current pulse generated by the sensor can ignite the reactive material, causing heat generation to set/reset the PCM cell, which effectively "destroys" its information content. It opens up the opportunity to create a tamper-responsive secure memory architecture, or use PCM cells as stand-alone tamper-evident sensors without any further cost regarding integration.

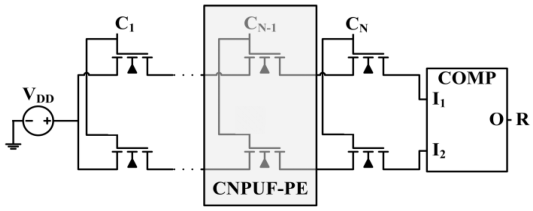


Fig. 10. CNPUF proposed in [61]. Characteristics of CNPUF parallel element vary (metallic or semiconducting) due to process variation.

D. PCM for Anticounterfeiting

The phenomena of resistance drift in PCM cells could potentially be used to design passive aging sensors for roughly detecting if a chip has been out of the supply chain before (and for how long). Although the drift feature is passive and ideal for aging sensors, there are several hurdles to overcome. The PCM cells must be isolated and protected against any form of set/reset operations, as its resistance values can be reverted. Furthermore, the amount of time that can be detected is highly subjective to device geometry and stochastic material properties of the PCM cell. Alternatively, data retention failure can also be used to detect arbitrary durations of time, similar to the SRAM-based data decay strategy presented in [59]. This is possible in PCM as the gradual process of seed crystal nucleation and formation of percolation paths causes a PCM cell to crystallize (permanently) and fail, while the resistance continually drops with the gradual crystallization. However, such a failure mechanism is too slow to be practical at room temperature (10 years at 85 °C for complete crystallization) [60].

V. Carbon Nanotubes and Graphene for Hardware Security Applications

In this section, we highlight how some unique properties of graphene and CNTs can be potentially leveraged for several

hardware-oriented security applications and countermeasures against different attacks.

A. CNT/CNTFET-Based Physical Unclonable Functions

As discussed in Section III-B, the inherent random variations in a GFET/CNTFET can be exploited to generate PUF-based signatures. Konigsmark *et al.* [61] proposed a CNT-based PUF (namely, CNPUF, see Fig. 10). It relies on the fact that the lack of chirality control in the manufacturing process yields metallic CNTs over semiconducting CNTs in a nondeterministic way. Utilizing the characteristic variation between semiconducting and metallic properties of CNTs can lead to distinguishable, but random, states since the OFF current for semiconducting CNTs is considerably lower than that for metallic CNTs. Simulated results of CNPUF show reduced area and power footprint, and higher robustness against environmental variations with respect to selected CMOS-PUFs.

As presented in [62], CNTs can also be used to provide the interconnection in a cross-bar architecture. The physical unclonable property comes, in this case, from the fact that a well-designed (i.e., optimal) architecture would allow uncontrolled placement, or self-assembly, of CNTs into random locations resulting into different short/open nodes for the crossbar architectures varying from device to device, as shown in Fig. 11. This would also allow a lightweight implementation of PUF leveraging the uncontrollability of CNT manufacturing and placement.

However, major barriers to evaluating graphene and CNT-based PUF architectures must be overcome, since we lack proper and reliable models incorporating such stochastic behaviors as well as predicting the impact of environmental variations and aging. It may also offer degraded performance and lower reliability due to poor quality channel formation and contamination. Furthermore, mass production of such architectures still lacks technological maturity, and integration scheme with CMOS platform still needs thorough investigation.

B. CNTs for True Random Number Generation

Random variations that occur due to a CNTFET's channel-tubes' chirality, placement, spacing, and dimensions, as well as other physical variations can be exploited as entropy sources for TRNGs. For example, a metastable RO [63] implemented with CNTFETs may produce high entropy due to numerous sources of variations. However, for CNTFETs as well as PCM, digital extraction of entropy from such an inherent phenomenon is challenging and may be biased by the extraction circuitry due to a lack of resolution and operational limitations.

C. Graphene-Based Printable Electronics for Supply Chain Security

Graphene-based printable electronics exhibit high potential in electronic supply chain security. Following the technique discussed in Section III, active or passive electronic components can be printed on the top of an insulating material, such as the IC package material. The main advantage of printing electronics over conventional logic circuitry is that the circuit does not need to be fabricated on the die during manufacturing; rather it can be printed onto the package as well. This means that when chips are returned from the untrusted foundry, the IP owner can "print" circuits, on the chip package, that can generate digital fingerprints for identification and tracking to ensure the security of the product in the supply chain. Such a printed circuit can potentially make a touch-and-go solution for chip authentication, and to some extent, also make a counterfeit and tamper-evident architecture, since any polishing of package for recycled and remarked chips, or delayering, will destroy the printed circuit on the package. Note, however, a key concern of such a technique would be designing the necessary interface and supporting circuits and architectures. A power harvesting technique can allow the printed circuit to be implemented as a passive element, and only be activated while participating in an authentication process. Scalability, complexity, and area/power footprints of the printed circuits are also major concerns. Detail of such digital fingerprint circuitry and interface is still under investigation.

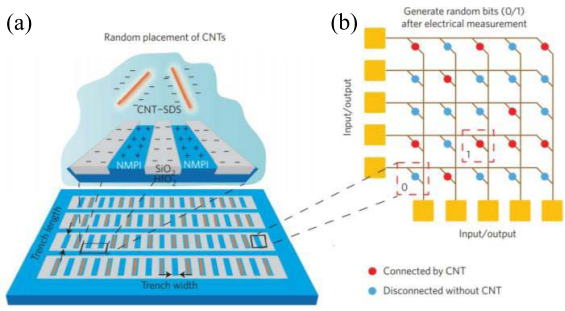


Fig. 11. Self-assembled CNT-based PUF. (a) Chemical self-assembly of CNTs. (b) Randomly connected 2-D CNT array in a 5×5 crossbar [62].

D. Graphene and CNT Nanosensors Combating Invasive/Semi-Invasive Attacks

Graphene and CNTs offer a large variety of sensors and MEMS/NEMS-based actuators. CNTs exhibit unique variation in properties, for example, the quality to conduct current through the channel of a CNTFET, due to variation in exerted mechanical force on the device, chemical/biochemical adsorption, or even optical exposure. Naturally, it makes graphene and CNT-based nanodevices good candidates for sensor applications [64]. Such sensors and actuators can be used to create a shield around the critical components (e.g., cryptomodule, secure data bus, and so on) of the circuit to prevent physical tampering and eavesdropping [30], as well as reverse engineering. As discussed in Section II-C, an antitamper design can leverage sensors that capture any kind of unwanted and unauthorized activity inside the chip to detect and resist physical attacks, such as delayering, imaging, probing, and milling. Keeping that in mind, we can utilize graphene and CNT-based mechanical, optical, and chemical sensors to extend security to invasive and semi-invasive attacks.

1) Mechanical Pressure Sensors: CNTs offer several MEMS/NEMS structures, as well as floating gate CNTFET structures that work as mechanical pressure (or force) sensors. In such cases, the electrical properties of the CNTFET such as

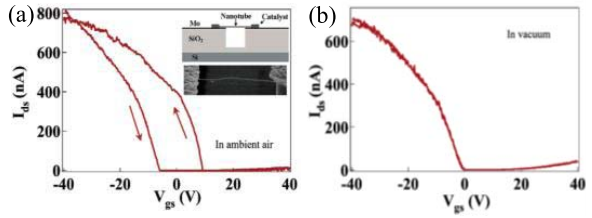


Fig. 12. I_{DS}–V_{GS} curve for a CNTFET exposed to ambient (a) air and (b) vacuum [68].

carrier mobility within the channel, or resonant frequency of a cantilever structure, change because of physical deformation due to exerted physical force [64]. Such a pressure/force sensor can be used in a chip to detect physical force while delayering and polishing.

2) Optical Sensors: Imaging is one of the key steps in invasive/semi-invasive attacks, and hence optical/image sensors are necessary to combat such physical attacks. Graphene photodetectors and single wall CNT optical sensors provide high sensitivity in a broad range of optical and near infrared wavelength [65], [66]. As a security component in ICs, these sensors will trigger an alert flag in case of exposure to light while delayering, milling, or probing, and will erase secret key or data. Hence, it would become difficult to attack a cryptographic module surrounded by such optical sensors without any loss of data.

3) Chemical Sensors: Researchers have proposed several chemical and biochemical sensors using CNTs to provide high selectivity and sensitivity to detect chemical/biochemical materials and their amount [67], [68]. For example, Fig. 12 shows the change in the electrical properties (i.e., I–V characteristics) of a suspended CNTFET exposed to ambient air and vacuum. Placement of such a CNTFET acting as chemical sensors within the die can potentially detect chemical activities occurring while delayering or polishing, i.e., when the chip (sensor) is exposed to air. As shown, typical CNT-based sensors tend to drive different drain currents, since the associated chemicals and their amount change the electrical properties (e.g., conductance, carrier mobility, threshold voltage, and so on) of the channel. Hence, by converting the current into different logic levels, these sensors can trigger alert flags and erase valuable data while under corresponding invasive/semi-invasive attacks.

It should be noted that the major difficulty regarding graphene and CNT electronics is the integration with conventional CMOS platform for high processing yield. For example, the difficulties to overcome regarding CNT-based applications are the directional placement of individual CNTs on a substrate, high yield with desired electronic properties, and incorporating individual CNTs with proper contacts for a dense architecture. For graphene, a major barrier lies in creating a proper bandgap for using CNT transistors as digital switches. Unfortunately, it also compromises the device's carrier mobility. Additionally, a graphene and CNT-based primitive may itself be a part of a graphene-based platform. Such a graphene/CNT-based monolithic implementation can be prepared with solution-based chemical processing techniques that may not have the bulk-manufacturing capability [69], [70]. On the other hand, integration of graphene/CNT-based sensors and security primitives to CMOS-platform also poses additional challenges in terms of cost, yield, and large-scale manufacturability. Nevertheless, the advancement of 3-D integration techniques offers lucrative potentials for integrating CMOS and CNT-devices in memory and logic domain [71]. Furthermore, with the help of the state-of-the-art technologies in nanoscale regime, e.g., with an FIB that can operate in sub-10-nm region [31], it is much easier to put such a circuitry or sensor architecture in places for low-volume, selective, and critical applications. Hence, graphene and CNT-based architectures certainly exhibit strong candidacy in terms of hardware security-based applications, even though they may lack conventional logic applications.

VI. Emerging Devices for Security: The Road Ahead

In Sections IV and V, our discussion mainly focused on the unique features of emerging nanoscale devices, such as PCM and graphene/CNTs, targeting hardware security applications. As we see, these devices offer various intrinsic properties that enable us to establish different security primitives, such as PUFs and TRNGs, and provide countermeasures to several hardware-oriented attacks, such as tampering and counterfeiting. However, exploiting such nanoelectronic logic and memory devices for hardware security applications creates an additional set of challenges and calls for well-guided research from both the device and the hardware security community.

The notable challenges faced by researchers in this domain arise from some of the fundamental limitations of these nanodevices and their targeted applications. Although it is difficult to point out and solve each and every possible challenge, we note some of the major obstacles that hinder the ideal security-oriented applications of the emerging devices.

1) The primary challenge for using the emerging logic and memory devices for hardware security applications is that the design and architecture of these devices, like all others, do not take security as one of the fundamental focal points like performance-oriented features such as speed, power, and reliability. Device and fabrication parameters relevant to security are usually not determined or modeled beforehand to maximize the quality of resulting security primitives. The security versus performance tradeoff is not properly established either [72].
2) Security properties of such devices are not well established among the cross-community researchers. Additionally, the metrics available for hardware-oriented security assessments are limited and focus only on the circuits' and systems' output-level appraisal (e.g., PUF metrics, such as uniqueness, assess only the quality of the PUF output bitstream and do not provide any judgment on the device's intrinsic security quality). Lack of understanding of potential device features and metrics for assessment poses major difficulties for building security primitives using emerging devices.

Notable Challenges for Nanodevice-based Security Applications

Challenges		Device and Security Applications					
		PUF / TRNG		Anti-Tampering		Anti-Counterfeit	
		PCM	CNTs / Graphene	PCM	CNTs / Graphene	PCM	CNTs / Graphene
Device Modeling & Metrics	Statistical Models & Metrics for Variability	✓	✓	✓	✓	✓	✓
	Process Variation	✓	✓	✓	✓	✓	✓
	Environmental Variations & Aging	✓	✓				✓
	Measurable Entropy	✓	✓				
	Resistance Drift	✓				✓	
Circuit & Design Optimization	Cycle-to-Cycle Stochasticity	✓		✓		✓	
	Entropy Extraction	✓	✓				
	Efficient Post-processing	✓	✓		✓		✓
	Device Forming Step			✓			
	Data Retention loss					✓	
Physical Demonstration	Platform Compatibility		✓		✓		✓
	Integrate Analog Sensors into Package/Die				✓		✓
	Sensor Printability				✓		✓
	Noise Amplification	✓	✓				
	Initial Device Resistance	✓				✓	
Resistance to Attacks (Invasive / Non-Invasive)	Metrics to Assess Attack Vulnerability	✓	✓	✓	✓	✓	✓
	Passive Sensor Operation				✓		✓
	Can sensor be bypassed?				✓		✓
	Side Channel Leakage via Remanence	✓		✓			

Fig. 13. Withstanding challenges for developing emerging nanodevice-based hardware security primitives and countermeasures for attacks.

3) Most of the hardware security primitives and countermeasures still rely on existing mature technology, especially CMOS devices. The primary reason behind it is that the mature CMOS devices offer predictable behavior and enable the designers to envision the outcome of the primitive to some extent. However, some of the properties exploitable for security primitives, such as PUFs and TRNGs, may not actually be available in a more mature technology and hence do not necessarily offer better quality in terms of security.
4) Much of the proposed ideas for hardware security primitives and countermeasures presented in this paper and other related literature are yet to be experimentally validated and integrated as part of a security-enabling system. As a result, it is difficult to assess the yield and cost of the emerging technology-based security primitives for bulk manufacturing.

Some of the currently withstanding challenges (and, consequently, potential research directions) related to security-oriented applications of PCM and graphene/CNTs are summarized in Fig. 13. We categorize these major challenges from the following perspectives and highlight the impacted primitives and countermeasures.

1) *Device Characterization and Modeling:* The lack of proper statistical models for device variability and stochastic processes leads to difficulties in entropy estimation along with runtime and aging-based reliability analysis for PCM and CNT/graphene-based PUFs and TRNGs. The existing device models mostly target traditional logic/memory operations and do not necessarily incorporate security-oriented unique features and variabilities. Additionally, a phenomenon like resistance drift in PCM devices also affects the robustness of the PUF and must be incorporated into the model. Building such an exhaustive device variability model for emerging devices is highly challenging due to enormous quantity of characterization, physical modeling, and statistical data processing keeping both traditional logic/memory operations and nontrivial security applications in mind.
2) *Circuit-Level Architecture and Design Optimization:* As stated previously, CNTFETs and PCM cells are mostly modeled, analyzed, and optimized keeping logic and memory applications in mind, whereas hardware security applications may require different properties. Hence, modification is needed to develop security-optimized designs and architecture beyond classic applications. Such security-focused designs require efficient extraction of device-level entropy and other security-beneficiary properties. It is also necessary to obtain versatile compositions of emerging devices allowing maximum signal (or noise, as needed) amplification and minimum postprocessing.
3) *Physical Demonstration:* Physical demonstrations of proposed ideas are quite challenging, yet highly anticipated. Integration of emerging devices to an existing platform with security in mind allows assessing a number of security benefits. It can also work as a guidance for weighing security versus performance tradeoff for a given cost and yield rate for respective security applications. Evaluating foundry capabilities and advanced technologies for large-scale production should be another key point in this regard.
4) *Resistance to Attacks:* A poor resistance to additional vulnerabilities and attacks may highly degrade the security impact of emerging-device-based primitives and countermeasures. Developing novel metrics for evaluating different attack vulnerabilities and assessing the limitations against possible emerging attacks are, therefore, very important for using these emerging nanoelectronic devices to gain maximum security.

It is crucial to overcome these withstanding challenges for ideal unified applications focusing on both performance and security. To do so, the following aspects should be taken into consideration as a high-level guidance for potential research directions.

1) *Security Evaluation:* For security-oriented evaluation of emerging nanodevices such as PCM and graphene/CNTs, necessary metrics are needed to quantify device intrinsic and extrinsic parameters such

as variability and entropy, vulnerability to tampering, device-level sensor gain, and so on. These metrics are sought specifically for device-level security (and quality) assessment, as opposed to trivial system/output/circuit-level metrics that are currently prevalent. Unlike the traditional performance-oriented evaluation techniques, e.g., assessment of CMOS-logic devices using the I–V characteristics, device-level security metrics would focus on the unique security-oriented properties for different applications such as PUFs and TRNGs. This will allow the designers to balance the tradeoffs among devices, designs, and architectures to build targeted security primitives and provide countermeasures to different hardware vulnerabilities and attacks in a more efficient way, and maintain performance and security requirements as necessary.

2) *Design and Modeling:* Emerging devices may be designed as needed for particular security applications. For example, the PCM-cells to be used as tamper-evident sensors may be designed to require forming steps. Additionally, proper device-level models are highly required for these security-oriented device designs as such models work as the very first building blocks for designing, simulating, and analyzing the resistance of the security primitives and countermeasures to different hardware attacks. These models allow the designers to exploit inherent properties and variabilities of the device and assess the security outcome that may not be evident from trivial designs and models. For this, good statistical models capturing security features of devices, such as sources of entropy, process variation, changes to parameters by tampering/environmental variations, the impact of aging, and so on, are highly sought.
3) *Integration and Demonstration:* Many of the emerging devices, e.g., CNTFETs, still lack technological maturity and bring several challenges associated with the design and integration to CMOS platforms. Also, composition and integration of other required blocks for holistic operation bring about many difficulties and need further investigation. For example, it is quite challenging, at the present state, to integrate tamper-detection sensors based on graphene/CNTs onto actual IC die or package while maintaining the robustness of their detection capabilities (responses) across different environmental conditions. We also note that these devices are not traditionally designed for security applications, and challenges like manufacturing issues, environmental sensitivity, and low reliability still need to be overcome. Therefore, such devices may not always provide highly lucrative and readily available security solutions with their current status. However, a more hardware security-oriented composition of such devices can harness the intrinsic properties more fittingly to further open up and enhance possible security applications. Furthermore, it needs to be made sure that these nanodevice-based designs are not vulnerable to any existing or emerging attacks that can disable or bypass them (such as FIB-based invasive attacks), or if they leak sensitive information via side channel. Demonstration of the proposed applications is, therefore, of extreme need.

Toward this threefold approach, hardware security researchers could contribute to metrics, while device researchers could use those metrics to guide the design and modeling of devices. This clearly points to the need for a multidisciplinary effort in this field. It should be noted that these challenges and possible research directions are not only limited to PCM and graphene/CNT-based devices, rather they can be expanded to any emerging nanoelectronic logic and memory devices, such as Si-nanowire FET, memristive, and spintronic memory devices, without any loss of generality.

VII. Conclusion

In this paper, we have identified a plethora of features inherent in emerging nanoscale technologies and show how they potentially create new opportunities for establishing hardware security. These features enable applications ranging from new PUF/TRNG mechanisms, and supply chain security using graphene-solution-based printable electronics to a variety of sensors capable of detecting different modes of tampering. We also point out the withstanding challenges to establishing efficient and robust primitives using the emerging nanodevice and address what needs to be done for overcoming these obstacles. We urge that the research and scientific community from both device and hardware security domain should work together to ensure the best possible outcome.

References

[1] M. Tehranipoor and C. Wang, *Introduction to Hardware Security and Trust*. New York, NY, USA: Springer, 2011.
[2] A. Chen, J. Hutchby, V. Zhirnov, and G. Bourianoff, *Emerging Nanoelectronic Devices*. Hoboken, NJ, USA: Wiley, 2014.
[3] J. Rajendran *et al.*, "Nano meets security: Exploring nanoelectronic devices for security applications," *Proc. IEEE*, vol. 103, no. 5, pp. 829–849, May 2015.
[4] S. Ghosh, "Spintronics and security: Prospects, vulnerabilities, attack models, and preventions," *Proc. IEEE*, vol. 104, no. 10, pp. 1864–1893, Oct. 2016.
[5] B. Liu, C. Yang, H. Li, Y. Chen, Q. Wu, and M. Barnell, "Security of neuromorphic systems: Challenges and solutions," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2016, pp. 1326–1329.
[6] C. K. H. Suresh *et al.*, "A comparative security analysis of current and emerging technologies," *IEEE Micro*, vol. 36, no. 5, pp. 50–61, Sep. 2016.
[7] Y. Bi *et al.*, "Emerging technology-based design of primitives for hardware security," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 13, no. 1, Dec. 2016, Art. no. 3.
[8] B. Gassend, D. Clarke, M. van Dijk, and S. Devadas, "Silicon physical random functions," in *Proc. 9th ACM Conf. Comput. Commun. Secur. (CCS)*, New York, NY, USA, Nov. 2002, pp. 148–160.
[9] R. Pappu, B. Recht, J. Taylor, and N. Gershenfeld, "Physical one-way functions," *Science*, vol. 297, no. 5589, pp. 2026–2030, Sep. 2002.
[10] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proc. 44th Annu. Design Autom. Conf.*, 2007, pp. 9–14.
[11] S. Katzenbeisser *et al.*, "PUFs: Myth, fact or busted? A security evaluation of physically unclonable functions (PUFs) cast in silicon," in *Cryptographic Hardware and Embedded Systems—CHES*. Berlin, Germany: Springer, 2012, pp. 283–301.
[12] C. Helfmeier, C. Boit, D. Nedospasov, and J.-P. Seifert, "Cloning physically unclonable functions," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST)*, Jun. 2013, pp. 1–6.
[13] X. Xu and W. Burleson, "Hybrid side-channel/machine-learning attacks on PUFs: A new threat?" in *Proc. Conf. Design, Autom. Test Eur.*, Mar. 2014, p. 349.

[14] U. Rührmair *et al.*, "PUF modeling attacks on simulated and silicon data," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 11, pp. 1876–1891, Nov. 2013.

[15] T. Rahman, F. Rahman, D. Forte, and M. Tehranipoor, "An aging-resistant RO-PUF for reliable key generation," *IEEE Trans. Emerg. Topics Comput.*, vol. 4, no. 3, pp. 335–348, Jul./Sep. 2016.

[16] L. Lin, S. Srivathsa, D. K. Krishnappa, P. Shabadi, and W. Burleson, "Design and validation of arbiter-based PUFs for sub-45-nm low-power security applications," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1394–1403, Aug. 2012.

[17] S. Han, J. Choung, B.-S. Kim, B. H. Lee, H. Choi, and J. Kim, "Statistical aging analysis with process variation consideration," in *Proc. Int. Conf. Comput.-Aided Design*, Nov. 2011, pp. 412–419.

[18] M. Bhargava and K. Mai, "An efficient reliable PUF-based cryptographic key generator in 65 nm CMOS," in *Proc. Conf. Design, Autom. Test Eur.*, Mar. 2014, p. 70.

[19] J. Delvaux, D. Gu, D. Schellekens, and I. Verbauwhede, "Helper data algorithms for PUF-based key generation: Overview and analysis," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 34, no. 6, pp. 889–902, Jun. 2015.

[20] J. Delvaux and I. Verbauwhede, "Attacking PUF-based pattern matching key generators via helper data manipulation," in *Topics in Cryptology*. Cham, Switzerland: Springer, 2014, pp. 106–131.

[21] A. T. Markettos and S. W. Moore, "The frequency injection attack on ring-oscillator-based true random number generators," in *Cryptographic Hardware and Embedded Systems*, vol. 5747. Berlin, Germany: Springer, 2009, pp. 317–331.

[22] M. T. Rahman *et al.*, "Ti-trng: Technology independent true random number generator," in *Proc. 51st Annu. Design Auto. Conf. (DAC)*, New York, NY, USA, 2014, pp. 179:1–179:6.

[23] K. Yang, D. Blaauw, and D. Sylvester, "A robust −40 to 120 °C all-digital true random number generator in 40 nm CMOS," in *Proc. IEEE Symp. VLSI Circuits (VLSI Circuits)*, Jun. 2015, pp. C248–C249.

[24] C. E. Shannon, "The mathematical theory of information," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423 and 623–656, Jul./Oct. 1948.

[25] M. Stipčević and Ç. K. Koç, "True random number generators," in *Open Problems in Mathematics and Computational Science*. Cham, Switzerland: Springer, 2014, pp. 275–315.

[26] A. Rukhin *et al.*, "A statistical test suite for random and pseudorandom number generators for cryptographic applications," NIST, McLean, VA, USA, Tech. Rep. 800-22, 2001.

[27] G. Marsaglia. (1996). *Diehard: A Battery of Tests of Randomness*. [Online]. Available: http://www.stat.fsu.edu/pub/diehard

[28] S. P. Skorobogatov, "Semi-invasive attacks: A new approach to hardware security analysis," Ph.D. dissertation, Comput. Lab., Univ. Cambridge, Cambridge, U.K., 2005.

[29] C. Helfmeier, D. Nedospasov, C. Tarnovsky, J. S. Krissler, C. Boit, and J.-P. Seifert, "Breaking and entering through the silicon," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2013, pp. 733–744.

[30] D. Shahrjerdi *et al.*, "Shielding and securing integrated circuits with sensors," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Design (ICCAD)*, Nov. 2014, pp. 170–174.

[31] *ORION NanoFab—Helium Ion Microscope (HIM)*. Accessed: Jun. 19, 2017. [Online]. Available: http://www.zeiss.com/microscopy/en_us/products/multiple-ion-beam/orion-%nanofab-for-materials.html

[32] M. Tehranipoor, U. Guin, and D. Forte, *Counterfeit Integrated Circuits*. Cham, Switzerland: Springer, 2015, pp. 15–36.

[33] U. Guin, X. Zhang, D. Forte, and M. Tehranipoor, "Low-cost on-chip structures for combating die and IC recycling," in *Proc. 51st Annu. Design Autom. Conf.*, 2014, pp. 1–6.

[34] Y. M. Alkabani and F. Koushanfar, "Active hardware metering for intellectual property protection and security," in *Proc. USENIX Secur.*, Boston, MA, USA, 2007, pp. 291–306.

[35] K. Bernstein. *Supply Chain Hardware Integrity for Electronics Defense (SHIELD)*. [Online]. Available: https://www.darpa.mil/program/supply-chain-hardware-integrity-for-electronics-defense

[36] *Techinsights*. Accessed: Jun. 19, 2017. [Online]. Available: http://www.techinsights.com/

[37] C. Tarnovsky, "Hacking the smartcard chip," Black Hat DC, Arlington, VA, USA, Tech. Rep., 2010. [Online]. Available: http://www.blackhat.com/html/bh-dc-10/bh-dc-10-archives.html#Tarnovsky

[38] R. N. Das, V. R. Markovich, J. J. McNamara, Jr., and M. D. Poliks, "Anti-tamper microchip package based on thermal nanofluids or fluids," U.S. Patent 8 288 857, Oct. 16, 2012.

[39] J. Rajendran, M. Sam, O. Sinanoglu, and R. Karri, "Security analysis of integrated circuit camouflaging," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2013, pp. 709–720.

[40] K. Vaidyanathan *et al.*, "Efficient and secure intellectual property (IP) design with split fabrication," in *Proc. IEEE Int. Symp. Hardw.-Oriented Secur. Trust (HOST)*, May 2014, pp. 13–18.

[41] Z. Guo, M. Tehranipoor, D. Forte, and J. Di, "Investigation of obfuscation-based anti-reverse engineering for printed circuit boards," in *Proc. 52nd Annu. Design Autom. Conf.*, Jun. 2015, p. 114.

[42] H.-S. P. Wong *et al.*, "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, Dec. 2010.

[43] F. Dirisaglik *et al.*, "High speed, high temperature electrical characterization of phase change materials: Metastable phases, crystallization dynamics, and resistance drift," *Nanoscale*, vol. 7, no. 40, pp. 16625–16630, 2015.

[44] L. Zhang *et al.*, "Exploiting process variations and programming sensitivity of phase change memory for reconfigurable physical unclonable functions," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 921–932, Jun. 2014.

[45] D. Fugazza, D. Ielmini, S. Lavizzari, and A. L. Lacaita, "Random telegraph signal noise in phase change memory devices," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, May 2010, pp. 743–749.

[46] A. Pirovano *et al.*, "Reliability study of phase-change nonvolatile memories," *IEEE Trans. Device Mater. Rel.*, vol. 4, no. 3, pp. 422–427, Sep. 2004.

[47] M. P. Anantram and F. Léonard, "Physics of carbon nanotube electronic devices," *Rep. Prog. Phys.*, vol. 69, no. 3, p. 507, 2006.

[48] R. Vargas-Bernal and G. Herrera-Pérez, "Carbon nanotube- and graphene based devices, circuits and sensors for VLSI design," in *VLSI Design*. Rijeka, Croatia: InTech, 2012.

[49] B. Zhang and T. Cui, "An ultrasensitive and low-cost graphene sensor based on layer-by-layer nano self-assembly," *Appl. Phys. Lett.*, vol. 98, no. 7, p. 073116, 2011.

[50] F. Torrisi *et al.*, "Inkjet-printed graphene electronics," *ACS Nano*, vol. 6, no. 4, pp. 2992–3006, 2012.

[51] L. Zhang, Z. H. Kong, and C.-H. Chang, "PCKGen: A phase change memory based cryptographic key generator," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2013, pp. 1444–1447.

[52] K. Kursawe *et al.*, "Reconfigurable physical unclonable functions—Enabling technology for tamper-resistant storage," in *Proc. IEEE Int. Workshop Hardw.-Oriented Secur. Trust (HOST)*, Jul. 2009, pp. 22–29.

[53] S. Kim *et al.*, "Thermal disturbance and its impact on reliability of phase-change memory studied by the micro-thermal stage," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, May 2010, pp. 99–103.

[54] W. Zhang and T. Li, "Helmet: A resistance drift resilient architecture for multi-level cell phase change memory system," in *Proc. IEEE/IFIP 41st Int. Conf. Dependable Syst. Netw. (DSN)*, Jun. 2011, pp. 197–208.

[55] S. Skorobogatov, "Data remanence in flash memory devices," in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst.*, 2005, pp. 339–353.

[56] S. Skorobogatov, "Using optical emission analysis for estimating contribution to power analysis," in *Proc. IEEE Workshop Fault Diagnosis Tolerance Cryptograph. (FDTC)*, Sep. 2009, pp. 111–119.

[57] G. F. Close *et al.*, "Device, circuit and system-level analysis of noise in multi-bit phase-change memory," in *IEDM Tech. Dig.*, Dec. 2010, pp. 25–29.

[58] J. O. Chu *et al.*, "Integrated circuit tamper detection and response," U.S. Patent 8 861 728, Oct. 14, 2014.

[59] A. Rahmati *et al.*, "TARDIS: Time and remanence decay in SRAM to implement secure protocols on embedded devices without clocks," in *Proc. 21st USENIX Conf. Secur. Symp.*, 2012, p. 36.

[60] U. Russo, D. Ielmini, A. Redaelli, and A. L. Lacaita, "Intrinsic data retention in nanoscaled phase-change memories—Part I: Monte Carlo model for crystallization and percolation," *IEEE Trans. Electron Devices*, vol. 53, no. 12, pp. 3032–3039, Dec. 2006.

[61] S. T. C. Konigsmark *et al.*, "CNPUF: A carbon nanotube-based physically unclonable function for secure low-energy hardware design," in *Proc. IEEE Design Autom. Conf. (ASP-DAC)*, Jan. 2014, pp. 73–78.

[62] Z. Hu *et al.*, "Physically unclonable cryptographic primitives using self-assembled carbon nanotubes," *Nature Nanotechnol.*, vol. 11, no. 6, pp. 559–565, 2016.

[63] I. Vasyltsov, E. Hambardzumyan, Y.-S. Kim, and B. Karpinskyy, "Fast digital TRNG based on metastable ring oscillator," in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst.*, 2008, pp. 164–180.

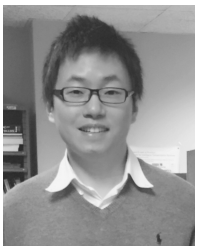
[64] C. Hierold *et al.*, "Nano electromechanical sensors based on carbon nanotubes," *Sens. Actuators A, Phys.*, vol. 136, no. 1, pp. 51–61, 2007.
[65] Y. Zhang *et al.*, "Broadband high photoresponse from pure monolayer graphene photodetector," *Nature Commun.*, vol. 4, p. 1811, May 2013.
[66] P. W. Barone *et al.*, "Near-infrared optical sensors based on single-walled carbon nanotubes," *Nature Mater.*, vol. 4, no. 1, pp. 86–92, 2005.
[67] P. Bondavalli *et al.*, "Carbon nanotubes based transistors as gas sensors: State of the art and critical review," *Sens. Actuators B, Chem.*, vol. 140, no. 1, pp. 304–318, 2009.
[68] W. Kim *et al.*, "Hysteresis caused by water molecules in carbon nanotube field-effect transistors," *Nano Lett.*, vol. 3, no. 2, pp. 193–198, 2003.
[69] M. M. A. Rafique *et al.*, "Production of carbon nanotubes by different routes—A review," *J. Encapsulation Adsorption Sci.*, vol. 1, no. 2, p. 29, 2011.
[70] Y. L. Zhong, Z. Tian, G. P. Simon, and D. Li, "Scalable production of graphene via wet chemistry: Progress and challenges," *Mater. Today*, vol. 18, no. 2, pp. 73–78, 2015.
[71] M. M. Shulaker *et al.*, "Monolithic 3D integration of logic and memory: Carbon nanotube FETs, resistive RAM, and silicon FETs," in *IEDM Tech. Dig.*, Dec. 2014, pp. 24–27.
[72] F. Rahman, D. Forte, and M. Tehranipoor, "Reliability vs. security: Challenges and opportunities for developing reliable and secure integrated circuits," in *Proc. IEEE Int. Rel. Phys. Symp.*, Sep. 2016, pp. 4C-6-1–4C-6-10.

Fahim Rahman is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA.

His current research interests include hardware security and trust and his main specialties include low-cost physical unclonable functions, and security aspects of emerging nanoelectronic devices.

Bicky Shakya received the B.Sc. degree (Hons.) in electrical engineering from Trinity College, Hartford, CT, USA, in 2014. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA.

His current research interests include hardware security and trust, techniques for semiconductor IP protection, and nanoscale security/integration challenges.



Xiaolin Xu (S'15) received the B.S. and M.S. degrees in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, China, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts at Amherst, Amherst, MA, USA.

His current research interests include physical unclonable functions, logic obfuscation, embedded systems, and hardware security

Domenic Forte (S'09–M'13) received the B.S. degree in electrical engineering from Manhattan College, Riverdale, NY, USA, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2010 and 2013, respectively.

He is currently an Assistant Professor with the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA. His current research interests include the investigation of hardware security primitives, hardware Trojan detection and prevention, electronics supply chain security, and antireverse engineering.

Dr. Forte was a recipient of the NSF Career Grant and the Young Investigator Award by the Army Research Office. He is currently serving as an Associate Editor for the *Journal of Hardware and Systems Security* and serving on the organizing committees of HOST and AsianHOST.

Mark Tehranipoor (S'02–M'04–SM'07) received the Ph.D. degree from the University of Texas at Dallas, Richardson, TX, USA, in 2004.

He is currently the Intel Charles E. Young Preeminence Endowed Professor in cybersecurity with the University of Florida, Gainesville, FL, USA. His current research interests include hardware security and trust, supply chain security, and VLSI design, test, and reliability. He has authored six books, 11 book chapters, and over 300 journal articles.

Dr. Tehranipoor is a recipient of several best paper awards as well as the 2008 IEEE Computer Society Meritorious Service Award, the 2012 IEEE CS Outstanding Contribution, the 2009 NSF CAREER Award, and the 2014 MURI Award. He is currently serving as an Associate Editor of the *Journal of Electronic Testing: Theory and Applications*, the *Journal of Low Power Electronics*, the IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION, and the *ACM Transactions on Design Automation of Electronic Systems* and as a Co-Director of the Florida Institute for Cybersecurity Research.