Optik 158 (2018) 855–860

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

# Thermal images based Hardware Trojan detection through differential temperature matrix

Jingxin Zhong*, Jianye Wang

*Air and Missile Defense College, Air Force Engineering University, China*

ARTICLE INFO

*Article history:*
Received 28 November 2017
Accepted 28 December 2017

*Keywords:*
Thermal images
Hardware Trojan
Detection
Differential temperature matrix

ABSTRACT

Insertions of Hardware Trojan result in unexpected harm to integrated circuits, and will change the characteristics of integrated circuits, such as thermal characteristics. In this paper, we proposed a Hardware Trojan detection method based on differential temperature matrix analysis. The matrix includes 750 differential thermal images data which were obtained from two differences between the testing chips and the golden chip, and the matrix shows each pixel's differential temperature value from thermal image and provides a novel insight to investigate the impact of Trojan. During experiment procedure, many programmable heating modules were distributed in chips to interfere with Trojan detection. Experimental results show that Trojan insertion could increase the differential temperature, and expand the density distribution of differential temperature at the same time. The proposed method can achieve effective and accurate detection result, Trojan with only a logical gate size has been detected and located in two proportions of logic units occupied between Trojan and programmable heating modules, the proportions are 0.45% and 0.29%, respectively.



## 1. Instruction

With the rapid development of the semiconductor design and fabrication technology, integrated circuits (ICs) have become widely used in military and civil applications. However, malicious modification of hardware during design or fabrication has emerged as a major security concern, such malicious circuit modification (referred to as Hardware Trojan) can affect the functions of IC, potentially with disastrous influences in safety-critical applications [1]. Therefore, devising effective methods for Hardware Trojan (HT) detection has become essential.

There are many ways of dealing with the risk of HTs. These ways normally include three categories: prevention, detection and localization. Prevention aims to never even permit HTs insertion. Detection is determines whether a hardware system contains a HT. Location is the regional detection of a HT, and the premise is at least one HT has been detected.

Since 2007, when the first side channel HT detection paper published [2], researchers have devoted much attention to devising HT detection methods in recent years. Up to now, methods of HT detection can be divided in three major categories: reverse engineering, logic tests and side-channel analysis [3]. Reverse engineering based specification comparison detection [4] is to remove IC chips' packages, split inner layers, take photos of layouts and wirings, and then compare images with original specifications. Reverse engineering is effective for any kind of HTs detection, but it isn't suitable for large numbers

* Corresponding author.
*E-mail addresses:* zhjx_08@163.com (J. Zhong), w_jianye@msn.com (J. Wang).

https://doi.org/10.1016/j.ijleo.2017.12.145
0030-4026/

of ICs testing: Splitting inner layers and taking photos of layouts and wires can cost much time and money, and the chips under test can't be used any more. Logic testing is the attempt to activate Trojans by applying test vectors and comparing the responses with the legitimate results. However logic testing has a great disadvantage: it is impractical to enumerate all states of a circuit with the numerous logic states and HTs are always activated under rare conditions [5]. As an effective method for HTs detection, side channel analysis was used to detect HTs through examining the abnormal signals from IC's physical parameters, such as delay [6], voltage [7], current [8], power [9], thermal [10–11] and electromagnetic emanation [12]. Thermal is a sensitive parameter and widely used for characterizing the performance of circuits. In 2014, Nowroz utilized thermal and power maps in HTs detection for the first time [10].

In this paper, we propose a novel HTs detection approach based on the differential temperature (DT) matrix which were obtained from two differences between the testing chips' and the golden chips' thermal images. The DT matrix shows each pixel's DT value on the thermal image varies with time. The accumulations of abnormal thermal signals caused by HTs in time enable the HTs can be easily detected and feature higher detection accuracy.

## 2. Theory of thermal images differential based hardware Trojan detection

Hardware Trojans are malicious alterations or insertions of extra circuitry to integrated circuits for malicious using, such as obstruct a system or intercept its confidentiality. When an IC starts working, the operation temperature in IC will rise up and stabilized in a few minutes. HT's insertion will bring extra thermal causing the temperature rising quicker than the IC without HT insertion. Fig. 1 shows thermal images on the same kinds of chips with or without a HT in the same moment $t_0$, and as Fig. 1(b) shows, a HT has inserted in point **P**.

As shown in Fig. 1(a), the temperature of point **P** ($T_P$) can be express as:

$$T_P = T_0 + T_{\Delta t0} + e_{t0} \tag{1}$$

$T_0$ is the initial temperature on the chip, $T_{\Delta t0}$ is the temperature accumulation in the process of the circuits operation, $e_{t0}$ is the noises of process and measurement in moment $t_0$. Similarly, the temperature of point **P** ($T'_P$) in Fig. 1(b) can be express as:

$$T'_P = T'_0 + T'_{\Delta t0} + e'_{t0} + T_{Trojan} \tag{2}$$

In order to highlight the temperature caused by HT, two differences method has been used in this paper. Two differences include two parts: self-difference and mutual-difference.

Self-difference is all the thermal image data from chip will be subtracted by the initial mean value ($T_{0\_avr}$) which from the same chip's first thermal image, the purpose is to eliminate the impact of the different initial temperature between $T_0$ and $T'_0$. As shown in Formulas (3) and (4), $T_{P\_tmp}$ and $T'_{P\_tmp}$ are temporary variables.

$$T_{P\_tmp} = T_P - T_{0\_avr} = \Delta T_0 + T_{\Delta t0} + e_{t0} \tag{3}$$

$$T'_{P\_tmp} = T'_P - T'_{0\_avr} = \Delta T'_0 + T'_{\Delta t0} + e'_{t0} + T_{Trojan} \tag{4}$$

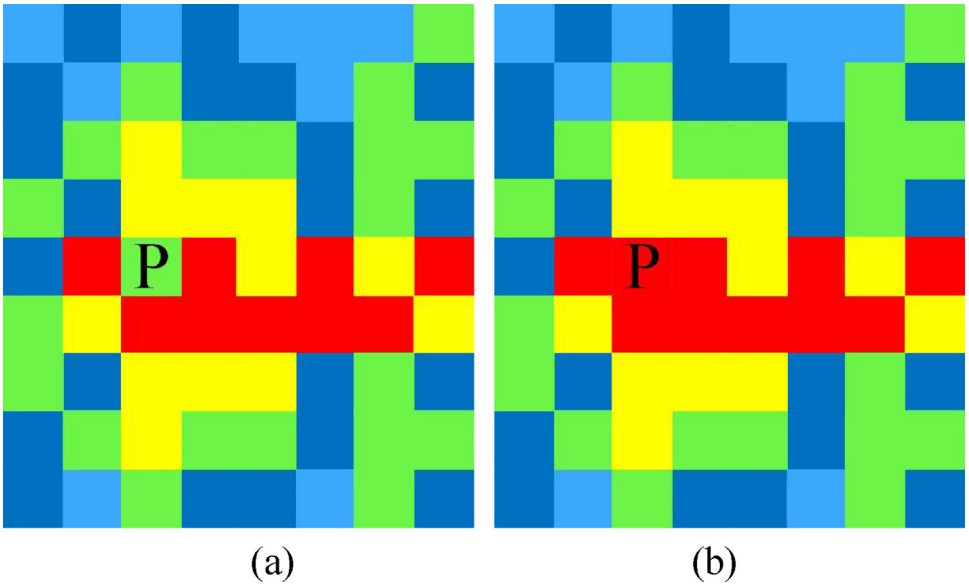


**Fig. 1.** Thermal images on chip in moment $t_0$. (a) Chip without HT; (b) Chip with HT.

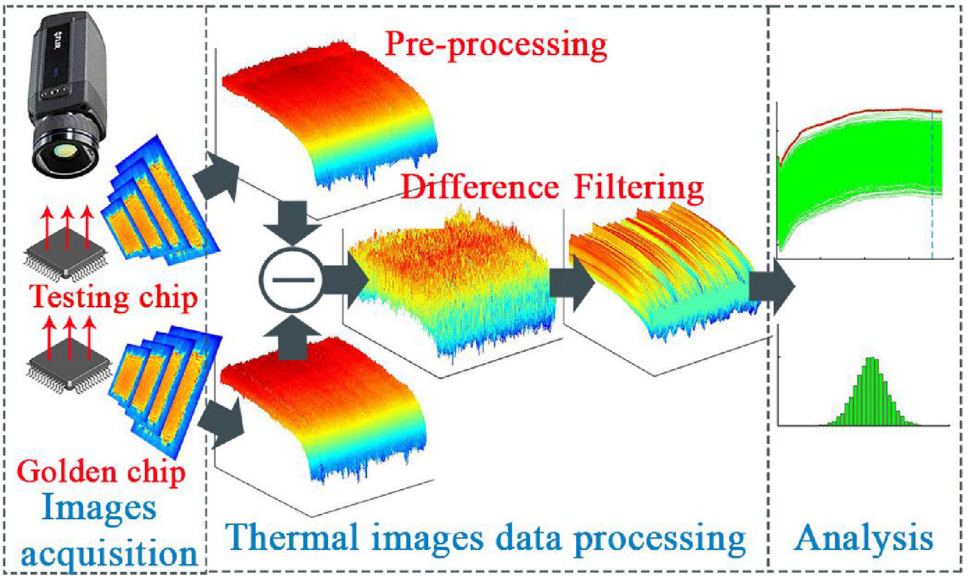


**Fig. 2.** Process of proposed HT detection method.

Mutual-difference is the comparison between chips with HT and chips without HT, as shown Formula (5). This procedure can highlight the influence of HTs and obtain the DT matrix.

$$T'_{P\_tmp} - T_{P\_tmp} = (\Delta T'_0 - \Delta T_0) + (T'_{\Delta t0} - T_{\Delta t0}) + (e'_{t0} - e_{t0}) + T_{Trojan} \tag{5}$$

As the testing chips with same circuits, and testing under the same condition, so $\Delta T'_0$ and $\Delta T_0$ can be considered equal approximately, and the same as $T'_{\Delta t0}$ and $T_{\Delta t0}$. So the temperature caused by HT can be express as:

$$T_{Trojan} = (T'_{P\_tmp} - T_{P\_tmp}) - (e'_{t0} - e_{t0}) \tag{6}$$

As the $e_{t0}$ and $e'_{t0}$ are white noises, so the temperature data caused by HT $T_{Trojan}$ can be obtained after filtering from temperature data $(T'_{P\_tmp} - T_{P\_tmp})$.

## 3. Thermal images based hardware Trojan detection methodology

As shown in Fig. 2, the proposed HTs detection method in this paper includes three parts: thermal images acquisition, thermal images data processing and results analysis.

### 3.1. Thermal images acquisition

The FLIR A645 thermal infrared imager is used for thermal images acquisition. The resolution of FLIR A645 is 640*480 pixels, and each pixel data presented by temperature value. During experiment procedure, FLIR A645 takes 750 thermal images in 30 s. Thermal images obtained from the thermal infrared imager have a large area of void for HTs detection, as shown in the left side in Fig. 3. Therefore, the thermal image data in the areas surrounded by the white square loop were investigated in this paper.

### 3.2. Thermal images data processing

Thermal images data processing is the core of this paper, and includes three parts: data pre-processing, data difference and Kalman filtering.

In order to show the relationship between temperature and time, data pre-processing accumulates all the useful pixels data from the thermal images in a new matrix. And the new matrix consists of 750 rows, and each row is made up of all pixels data from one thermal image. As shown in the right side of Fig. 3, the three dimensional diagram of the new matrix and each section of the diagram represents a thermal image.

Data difference includes two parts: self-difference and mutual-difference. Self-difference is all the thermal image data from chip will be subtracted by the initial mean value which from the same chip's first thermal image. Mutual-difference is the comparison between testing chips and golden chip to highlight the influence of HTs and obtain the DT matrix.

The thermal images data from FLIR A645 contains a lot of white noise and brings seriously influence on HTs detection. Kalman filter can effectively filter the white noise. During the filtering, the value of each pixel variations during the testing

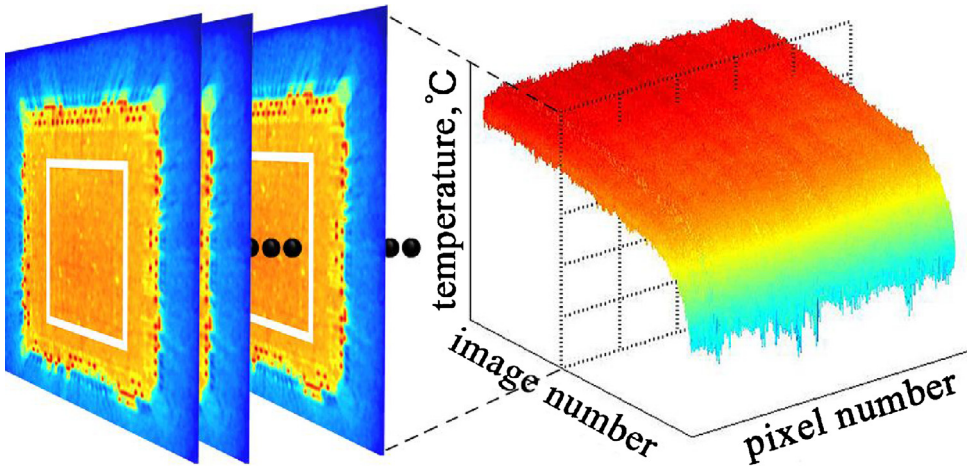


**Fig. 3.** Sketch of thermal image data pre-processing.

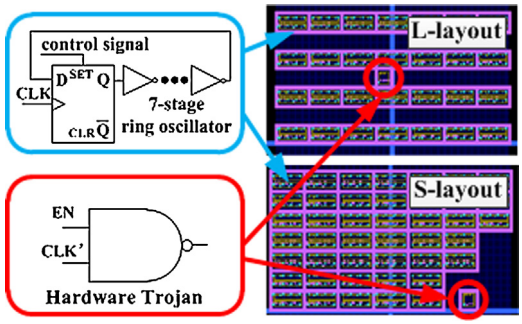


**Fig. 4.** Programmable heating modules and two layout maps of FPGA.

will be considered as a single point to filter, leading to accurately describe the change of each pixel's temperature as a function of time.

## 4. Authentication experimental setup

The testing chips are Xilinx Spartan3E FPGAs, and each testing chip consists of many programmable heating modules and with or without HT. The programmable heating module is a power programmable ring oscillator, as shown in upper left side of Fig. 4, the frequency (from 50 to 200 MHz) of clock (CLK) signal can control the output power of this module, and this module will replace the real circuits to generate heat for cover up the heat generated by HT. HT is instead of a NAND gate, and drived by CLK' signal which also can control the output power of HT, as shown in lower left side of Fig. 4.

Finally, two layout methods were adopted in the experiment, marked as Line Layout (L-layout) and Surface Layout (S-layout), as shown in right side of Fig. 4. The proportions of logic units occupied by HT and programmable heating modules are 0.45% and 0.29%, respectively. In order to verify the effectiveness of the proposed HT detection method, each programmable heating module is driving by different frequency CLK signal, and the HT is only driving by 100 MHz CLK' signal.

## 5. Results and analysis

Two Spartan3E FPGAs have been tested, one was tested as golden chip, and the other was the testing chip and has been tested with or without HT. Fig. 5 shows the experiment results of DT values distribution after Kalman filtering.

Fig. 5(a), (c), (e) and (g) shows the *DT* distribution from 750 thermal images on HT-free chips and HT-insertion chips. Fig. 5(b), (d), (f) and (h) shows the *DT* density distribution from 700th differential thermal image (the *DT* resolution is 0.01 °). Fig. 5(a) and (e) shows that the *DT* values are relatively stable on HT-free chips, and most *DT* are near 0 °. And from Fig. 5(c) and (g), the *DT* values present an obvious rise trend on HT-insertion chips, and most of them are more than 0 °. Fig. 5(b) and (f) shows a high *DT* density distribution within −0.1 to 0 on HT-free chips, and the maximum probability *DT* value has more than 2000 pixels. Compared with Fig. 5(b) (f), (d) and (h) show wider distribution range (almost from 0 to 0.3) and lower distribution density about the *DT* values on HT-insertion chips, and the maximum probability *DT* value is only about 1000 pixels. As a conclusion, the existence of a HT will increase the *DT* values and expand the range of *DT* distribution.

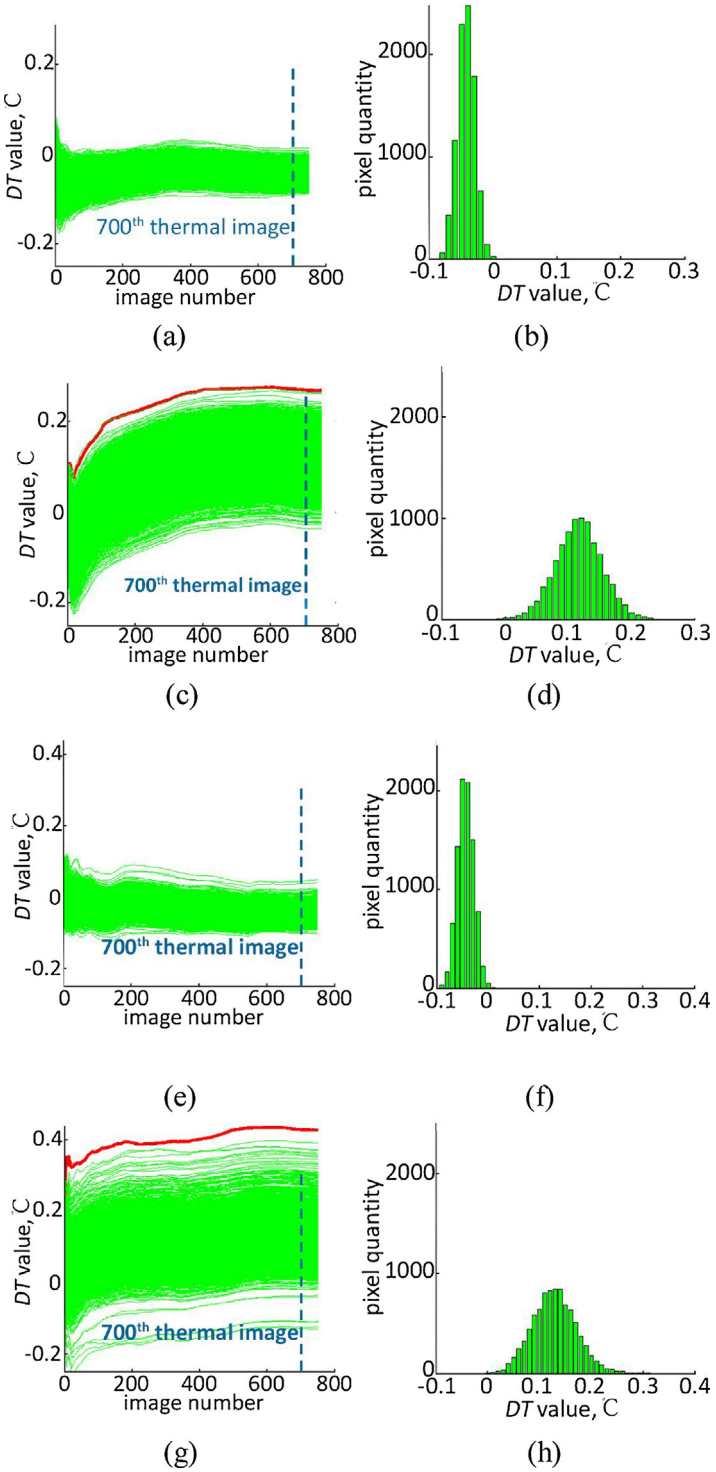


**Fig. 5.** DT values distribution after Kalman filtering. (a) DT distribution on L-layout HT-free chip; (b) DT density distribution from 700th differential thermal image in (a); (c) DT distribution on L-layout HT-insertion chip; (d) DT density distribution from 700th differential thermal image in (c); (e) DT distribution on S-layout HT-free chip; (f) DT density distribution from 700th differential thermal image in (e); (g) DT distribution on S-layout HT- insertion chip; (h) DT density distribution from 700th differential thermal image in (g).

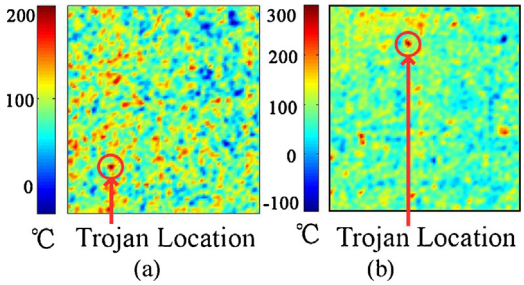


**Fig. 6.** Location analysis of HT on two layout methods chips. (a) Location of HT on L-layout chip; (b) Location of HT on S-layout chip.

The maximum *DT* value was probably resulted from HTs, as the red bold line represents in Fig. 5(c) and (g), and this value can be used for HT's locating. In order to highlight the existence of HT, all the *DT* values from750 differential thermal images on HT-insertion chips were superimposed together and restored to thermal images, as shown in Fig. 6, the location of the HTs are inside the red circle. Affected by the shooting angle of thermal imager, the location of HTs may be different from the layout maps.

## 6. Conclusion

HTs insertions will change the thermal characteristics of ICs. In this paper, we proposed a more accurate HT detection method based on DT matrix through thermal image. The matrix combines all the pixels' DT values on the thermal images vary with time and makes detection results more intuitive. Results show HT has been detected and located in two proportion of logic units occupied by HT and programmable heating modules (0.45% and 0.29% respectively), and HT was only a logical gate size. The purpose of the experiment is to verify the effectiveness of the detection method, therefore the diversity of HTs and simulation circuits were not considered in the experimental setup, and the future works will design more circuits to analysis and modify the locating method.

## References

[1] R.S. Chakraborty, S. Narasimhan, S. Bhunia, Hardware Trojan: threats and emerging solutions, in: IEEE International High Level Design Validation and Test Workshop Conference, 2009, pp. 166–171.
[2] D. Agrawal, S. Baktir, D. Karakoyunlu, Trojan detection using IC fingerprinting, in: IEEE Symposium on Security and Privacy Conference, 2007, pp. 296–310.
[3] B. Zhou, W. Zhang, S. Thambipillai, Cost-efficient acceleration of hardware Trojan detection through fan-out cone analysis and weighted random pattern technique, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 35 (2016) 792–805.
[4] P. Subramanyan, N. Tsiskaridze, K. Pasricha, D. Reisman, Reverse engineering digital circuits using functional analysis, in: Design Automation and Test in Europe Conference and Exhibition, 2013, pp. 1277–1280.
[5] R.S. Chakraborty, S. Pagliarini, J. Mathew, S.R. Rajendran, A flexible online checking technique to enhance hardware Trojan horse detectability by reliability analysis, IEEE Trans. Emerg. Top. Comput. 5 (2017) 260–270.
[6] D. Rai, J. Lach, Performance of delay-based trojan detection techniques under parameter variations, in: IEEE International Workshop on Hardware-Oriented Security and Trust Conference, 2009, pp. 58–65.
[7] S. Kelly, X. Zhang, M. Tehranipoor, Detecting hardware Trojan using on-chip sensors in an ASIC design, J. Electron. Test. 31 (2015) 11–26.
[8] F. Koushanfar, A. Mirhoseini, A unified framework for multi-modal submodular integrated circuits Trojan detection, IEEE Trans. Inf. Forensics Secur. 6 (2011) 162–174.
[9] Q. Sui, Z.K. Wu, J. Li, S.Q. Li, A detection method of hardware Trojan based on two-dimension calibration, in: 2nd IEEE International Conference on Computer and Communications, 2016, pp. 2795–2799.
[10] A.N. Nowroz, K.Q. Hu, F. Koushanfar, Novel techniques for high-sensitivity hardware Trojan detection using thermal and power maps, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 33 (2014) 1792–1805.
[11] C.B. Bao, D. Forte, A. Srivastava, Temperature tracking: toward robust run-time detection of hardware Trojan, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 34 (2015) 1577–1585.
[12] J. He, Y. Zhao, X. Guo, Y. Jin, Hardware Trojan detection through chip-free electromagnetic side-channel statistical analysis, IEEE Trans. Very Large Scale Integr. Syst. 25 (2017) 2939–2948.

**Jingxin Zhong** received the B.S. degree in radar engineering from Air Force Engineering University, Xi'an, China, in 2012, and M.S. degree in electronic science and technology from Air Force Engineering University, in 2015. He is currently pursuing the Ph.D degree in electronic science and technology from Air Force Engineering University. His current research interests include trusted circuit design, hardware Trojan detection, and IC security.

**Jianye Wang** received his B.S. degree in Radio Engineering from Shaanxi University of Science and Technology, China, in 1984, and the M.S. degree in Electrical Engineering from Nanjing University of Science and Technology, Nanjing, China, in 1990. He was the visiting scholar in University of Tennessee, Tennessee, USA, from 2005 to 2006. He joined the Air Force Engineering University in 1999 and he is now a professor. His research interests are VLSI design, electronic design automation and IC security.