Attack mitigation of hardware trojans for thermal sensing via micro-ring resonator in optical NoCs

Zhou, Jun; Li, Mengquan; Guo, Pengxing; Liu, Weichen

2021

Zhou, J., Li, M., Guo, P. & Liu, W. (2021). Attack mitigation of hardware trojans for thermal sensing via micro-ring resonator in optical NoCs. ACM Journal On Emerging Technologies in Computing Systems, 17(3), 31-. <https://dx.doi.org/10.1145/3433676>

<https://hdl.handle.net/10356/152430>

<https://doi.org/10.1145/3433676>

© 2021 Association for Computing Machinery. All rights reserved. This paper was published in ACM Journal on Emerging Technologies in Computing Systems and is made available with permission of Association for Computing Machinery.

Downloaded on 26 Apr 2023 10:00:19 SGT

Attack Mitigation of Hardware Trojans for Thermal Sensing via Micro-ring Resonator in Optical NoCs

JUN ZHOU

School of Computer Science and Engineering, Nanyang Technological University, Singapore
edo.zhou@ntu.edu.sg

MENGQUAN LI

School of Computer Science and Engineering, Nanyang Technological University, Singapore
College of Computer Science, Chongqing University, Chongqing, China
mengquan.li@ntu.edu.sg

PENGXING GUO

School of Computer Science and Engineering, Nanyang Technological University, Singapore
School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, China
N1805884A@c.ntu.edu.sg

WEICHEN LIU

School of Computer Science and Engineering, Nanyang Technological University, Singapore
liu@ntu.edu.sg

ABSTRACT

As an emerging role in new-generation on-chip communication, optical networks-on-chip (ONoCs) provide ultra-high bandwidth, low latency and low power dissipation for data transfers. However, the thermo-optic effects of the photonic devices have a great impact on the operating performance and reliability of ONoCs, where the thermal-aware control with accurate measurements, e.g., thermal sensing, is typically applied to alleviate it. Besides, the temperature-sensitive ONoCs are prone to be attacked by the hardware Trojans (HTs) covertly embedded in the counterfeit integrated circuits (ICs) from the malicious third-party vendors, leading to performance degradation, denial-of-service (DoS), or even permanent damages. In this paper, we focus on the tampering and snooping attacks during the thermal sensing via micro-ring resonator (MR) in ONoCs. Based on the provided work flow and attack model, a new structure of the anti-HT module is proposed to verify and protect the obtained data from the thermal sensor for attacks in its optical sampling and electronic transmission processes. In addition, we present the detection scheme based on the spiking neural networks (SNNs) to implement an accurate classification of the network security statuses for further high-level control. Evaluation results indicate that, with less than 1% extra area of a tile, our approach can significantly enhance the hardware security of thermal sensing for ONoC with trivial costs of up to 8.73%, 5.32% and 6.14% in average latency, execution time and energy consumption, respectively.

CCS CONCEPTS

• Security and privacy → Security in hardware; • Networks → Network on chip; • Computer systems organization → Interconnection architectures; • Hardware → Emerging technologies;

KEYWORDS

Hardware security, Thermal sensing, Network-on-chip, Micro-ring resonator, Trojan, Neural network

1 Introduction

With the rising of design complexity and manufacturing costs of the emerging many-processor systems, designers tend to utilize the devices or components from other vendors directly, namely, the third-party intellectual property (3P-IP), to fulfill the integration of the whole single facility [1]. However, these IPs can

probably introduce some security risks [1, 2]. For instance, malicious or counterfeit circuits with hardware Trojans (HTs) will snoop or tamper the critical and sensitive data in the interconnection systems of the multi-processor systems-on-chip (MPSoCs). In fact, the HTs within the compromising MPSoCs will result in data leakage, information manipulation, etc., inducing certain irreversible losses, such as the systemic abnormal functions, performance degradation, denial-of-service (DoS), or even permanent breakdown, which attracts more and more attention from both the academia and industry [1].

Many MPSoCs now use electronic networks-on-chip (ENoCs) for inter-processor data transfers. As a traditional network-based communications sub-system in integrated circuits (ICs), ENoC is a router-based packet switching interconnection among modules of SoC, applying the theories and methods of computer networks to on-chip communications and bringing notable improvements over conventional bus and crossbar architectures. Meanwhile, the recent progresses in silicon photonics also have accelerated the integration of optical components with the complementary metal-oxide-semiconductor (CMOS) circuits in SoCs [3]. As a promising promotion to ENoCs, the optical networks-on-chip (ONoCs), especially those employing the wavelength division multiplexing (WDM) technology [3, 4], can provide ultra-high bandwidth, low latency and power dissipation for modern many-processor systems. These advantages motivate the applications of ONoC for the inter-processor communications on chip.

However, both the communications in ENoC and ONoC need to face the challenges of overheating. The thermal issue impacts not only the network performance (e.g., communication latency, throughput, packet loss rate, and so on) but also the reliability of chip. Higher temperatures will cause longer propagation delay, slower circuits switching, larger leakage power, and optical bandwidth loss [4]. Hence, for ONoCs, which concern us more here, to reduce or eliminate the thermo-optic effects in interconnection has been one of the major research priorities. Specifically, temperature is a key indicator for the operation system (OS) to carry out the thermal-aware control of ONoCs with certain effective schemes, such as scheduling, mapping, and routing, etc., to avoid the hotspots and local overheating phenomena. The balance maintenance of on-chip temperature distribution guarantees the efficient communications of ONoCs.

Similar to an MPSoC with ENoCs, a many-processor system using ONoCs is also expected to adopt the 3P-IPs, which is vulnerable to security risks [2, 4]. The communication packets are probably to be snooped or tampered because of these unreliable IPs where HTs are embedded in ONoCs. On the other hand, as an effective measurement, the thermal sensing technique is widely utilized in the ONoCs to perform the thermal-aware control [4]. Similar to the communication packets, the sensing data obtained by thermal sensors may also suffer from leakage and manipulation during its sampling and transmission. Consequently, OS will be cheated and misled seriously when the inaccurate on-chip temperature distribution is gained from the sensors, leading to the misjudgment of OS and induced inappropriate operations. Especially, tampering attacks are the primary security threats to cause the performance degradation, DoS of the core components in ONoCs, or even heat damage of the whole chip [1, 5]. Hence, efficient and effective mechanisms need to be proposed to accurately detect these HT attacks and recover the normal communications of ONoCs.

In this paper, for the first time to our knowledge, we propose a comprehensive and effective HT-mitigation method for the thermal-aware control under snooping and tampering attacks for ONoCs using the accurate and low-power optical sensors [6]. Moreover, the neural network (NN) is utilized to address the detection of network security statuses by classification. The novel contributions of this paper are listed as follows:

- We present a thermal sensing flow of ONoCs and an attack model therein. In this model, tampering attacks can be triggered by the counterfeit circuits within the control logic of the thermal sensor, which is based on the silicon micro-ring resonator (MR).
- A new structure of the anti-HT module is proposed to verify the data from the thermal sensor for attacks in its sampling and transmission processes based on the mentioned work flow and attack model. Retransmission operations will be continuously applied until the countermeasures work, which will ensure the accuracy of the sampling data.
- For the designed anti-HT module, we also make use of the optics-related parameter as the symmetric key to encrypt the sensitive temperature information to prevent the data from being leaked during the delivery. Meanwhile, the guarantee for data integrity is also performed on the encrypted thermal value in case that it is directly manipulated or the encryption key is stolen or cracked.

- We use a classification scheme for the runtime detection of the security statuses of ONoCs through the spiking neural network (SNN) [7], which can identify and classify different patterns to carry out the relevant analyses. This monitoring mechanism conduces to guiding the OS to take further effective processing to avoid DoS, disorder of system function, etc.
- Evaluation and analysis show that the proposed mitigation approach provides the security assurance against HT attacks with trivial costs on latency, execution time, area and energy consumption, which indicates that our solution is attractive for the secure thermal-aware control of ONoCs.

The rest of the paper is organized as follows: Section 2 introduces the preliminaries and related works; Section 3 presents the work flow and attack model; Section 4 proposes the mitigation design of the HT attacks during thermal sensing of MPSoCs; Section 5 takes advantage of the trained SNN to implement the analysis of security statuses for the whole ONoC based on the proposed mitigation design. Experimental results are provided in Section 6, and Section 7 concludes this paper.

2 Preliminary and Related Work

2.1 Background

2.1.1 ONoC Architecture

As illustrated in Figure 1, a typical 2D-mesh ONoC is an optoelectronic hybrid network. Vertically on top of the layer of processor cores, i.e., processing elements (PEs), an optical network (O-net) provides the optical interconnects between optical routers (O-routers) for the inter-processor data transferring. And, an electronic network (E-net) consisting of electronic routers (E-routers) and metallic links lies in the bottom layer is to perform logical control. The photonic and electronic sub-networks together constitute an ONoC. Note that the extra metallic links between the optical router and network interface (NI) are exploited here to enhance the communication of control data in ONoC. Different from the traditional E-router, in general, an O-router consists of three main parts: an optical transmitter, an optical receiver and an optical switching network [4]. Besides, silicon MRs and optical waveguides perform the switching and transmission of optical signals, respectively, composing the optical switching network.

Figure 1: A typical 2D-mesh ONoC

2.1.2 Optical Thermal Sensor

The silicon MRs, functioning as wavelength-selective optical switches, are essential to implement high-level routing policies in ONoCs. A single MR consists of one ring and two straight waveguides. In Figure 2(a) and 2(b), specifically, there are two different MR designs as illustrated, the parallel switching element (PSE) design, and the crossing switching element (CSE) design. Ideally, the optical signals from the *Input* port will be delivered to the *Through* port if the MR is configured to be switched *OFF*. Or else, the optical signals will get resonated into the ring and delivered to the *Drop* port if their wavelengths are in the range of resonant wavelength, while the signals whose wavelengths are beyond the range will be filtered then. Also, a direction change will be achieved by the optical signal through the *Drop* port, where the PSEs turn the optical signal by 180°, while the CSEs turn it by 90°.

Figure 2: Micro-ring resonators (MRs)

Here, take the PSE for example, the cascaded MRs have been widely used as the basic optical filtering element (BOFE) at the receiver to demultiplex the WDM wavelengths (i.e., λ_i) before optical signals reach the photo-detectors (PDs), as demonstrated in Figure 3. Besides, the wavelength selectivity of MR lends itself naturally for WDM operations, which multiplies the bandwidth of optical interconnects [4]. With denser channel spacing than the normal WDM, dense wavelength division multiplexing (DWDM) is used to increase bandwidth over the existing fiber networks via combining and transmitting multiple signals simultaneously at different wavelengths on the same fiber. Generally, a typical DWDM system would use 40 channels with 100 GHz spacing or 80 channels at 50 GHz spacing [4]. In this paper, we design the corresponding security mechanisms for the platform of DWDM-based ONoCs.

Figure 3: Cascaded MRs for basic optical filtering element (BOFE)

On the other hand, MR is highly sensitive to the temperature variations (TVs) [8]. [9] systematically modeled and quantitatively analyzed the thermal effects of MRs. [10] surveyed the thermal effects on MR-based devices. [11] concluded the loss contributions of a ring resonator. [12] presented an analytical model to quantify the resonator losses at room temperature without considering its physical structure and waveguide cleavage facets. These studies provide the foundation for the adopted thermal sensing technique in this paper. As shown in Figure 4(a) [13, 14], the resonant wavelength of an MR red-shifts with increasing temperatures and blue-shifts with decreasing ones. While Figure 4(b) shows the simulation results of relationship between the ambient temperature and power losses of MRs in [13, 14]. In general, the optical power loss of an MR increases monotonically with the temperature's increasing over a large range from 300 K to 400 K.

The undesired mismatch between resonant wavelength and signal wavelength will lead to, because of TVs, additional optical power loss (indicated as ΔP , the same below) that can be converted to corresponding temperature data within a certain idle PE of MPSoCs [8]. Thus, the inherent thermal sensitivity of MR makes itself an ideal choice of on-chip thermal sensing. And, we can easily measure the received optical power using PDs in O-router, then get the loss value. Compared with the traditional electric thermal sensors, MR-based type has several favorable features, such as compact size, robustness against mechanical shock, immunity to electro-magnetic interference, and well-modeled temperature dependence [15]. Though the fabrication-induced process variation (PV), as well as the device-level wavelength tuning mechanisms, will introduce the challenges on accuracy and reliability of thermal sensing via MRs [13, 16, 17], it is sufficient for OS to obtain the temperature distribution on chip to execute the suitable thermal-aware control for the communications of ONoCs. Specifically, [15] started to use an MR-based sensor to implement thermal sensing. Then, [16] and [17] firstly utilized MRs to carry out the on-chip thermal sensing with resolving the challenges in accuracy and reliability which arise from PVs and local wavelength tuning.

Figure 4: The characteristics of MRs change with temperature

2.2 Related Work and Motivation

Because of the embedded malicious circuits with HTs, both ENoCs and ONoCs suffer from the security issues nowadays. As surveyed in [1] and [2], previous research works on hardware security in NoCs focused on various attacks & defenses patterns.

So far, the majority of existing security methods have been proposed for ENoCs, which can be classified as follows: 1) **Detection and correction.** [18] gave us the packet-security (P-Sec) scheme, a packet validation technique to protect compromised ENoC from the fault injection, side-channel attacks, and covert HT communications by merging two error detection schemes, namely, algebraic manipulation detection (AMD) and cyclic redundancy check (CRC). Also, with the security points-of-interest (SPOI) and its vulnerabilities defined first, [19] provided the secure model checkers (SMCs), a real-time solution for the control logic verification and functional correctness in the routers to detect the HT-induced DoS attacks and improve reliability in a compromised network. 2) **Authentication.** [20] presented a progressive authentication scheme for the packets in third-party ENoC, where a packet is tagged at the source node for data integrity checking through all nodes on the communication path till the destination. This scheme was designed to reduce the bandwidth overhead to avoid potential conflicts in networks with tampering attacks. 3) **Routing.** [21] secured ENoCs against timing attacks with the non-interference based adaptive routing (NIBR). The network traffics are executed separately in different security domains to improve performance and prevent data leakage. [22] firstly proposed an HT model that applies DoS attacks by misrouting the packets, which can cause deadlock and consequent ENoC performance degradation, and then presented a secure routing algorithm that provides a runtime HT detection and avoiding scheme. 4) **Shuffling scheme.** [23] exploited error control approaches for the transient and permanent HTs on ENoC links at runtime with link shuffling and isolation techniques, respectively. Error correction code (ECC) is used to detect the attacks at the input and output ports of network routers. For [24], the authors designed an HT that can compromise the crucial details of packets to induce performance degradation and DoS attacks in ENoC. A mitigation method based on bit shuffling mechanism inside the routers with a key extracted from input message was proposed to limit the adverse effects of HTs. 5) **Obfuscation.** [25] presented a new lightweight target-activated sequential payload (TASP) HT model that performs packet inspection and injects faults to create DoS attacks, then proposed a heuristic threat detection model to classify faults and discover HTs within the compromised links, and several switch-to-switch link obfuscation methods to avoid triggering of HTs to continue using the links for minimal overhead instead of rerouting the packets. In [26], state obfuscation was exploited to facilitate HT detection in the NIs of ENoCs. The authors firstly added key bits to the finite state machine (FSM) of the NI control unit and created dummy states to increase the difficulty for attack execution, then examine the illegal states and illegal state transitions induced by a wrong key to detect the occurrence of HTs. Besides, the authors of [27] and [28] combined the dynamic permutation and flit integrity check method to form a collaborative method to mitigate the potential inside-router HTs inserted by the disloyal member in the design house or the third-party system integration company. 6) **Firewall or security module.** Authors of [29] presented a hardware firewall to secure the data

transactions between the ENoC routers. This security mechanism is capable of detecting and preventing attacks, such as unauthorized accesses that aim at stealing or corrupting sensitive data, and DoS that aim at degrading the system performances. Two functionalities are considered for the access control and verification of the occupation time of the physical links between routers. In [30], the extended ENoC firewall module can check the physical addresses and reject untrusted requests from the central processing unit (CPU) to on-chip memory, thus protecting the legitimate processes running within a multi-core SoC from the injection of malicious data or instructions to shared memory. Also, [31] described a low-overhead snooping invalidation module (SIM) in NI to prevent the malicious data replication by HTs in ENoCs, then devised a snooping detection module, i.e., THANOS, to detect malicious applications that utilize such HTs. What's more, [32] proposed Prometheus, a drop-in solution inside the NI that provides protection from the spoofing, DoS, and eavesdropping attacks in wireless NoC (WNoC, still a kind of ENoC). 7) **Mixed scheme.** [33] proposed an integrated runtime solution for both security and fault tolerance of SoCs based on the field-programmable gate array (FPGA) through digital signatures, live monitoring, partial reconfiguration, and adaptive routing supported by an ENoC called X-Network that can reduce the ratio of required routers versus PEs to facilitate security designs and fault tolerance. In [34], authors offered Fort-NoCs, a series of techniques including data scrambling, packet certification, and node obfuscation, to disable covert backdoor activation and reduce the chance of encountering the side-channel attacks by “clouding” the information obtained by the attackers for defense against the compromised ENoC in MPSoC.

Meanwhile, the specialized security approaches designed for ONoCs are still at an early stage. Since the traditional ONoC is an optoelectronic hybrid network, several security mechanisms proposed for ENoC can still be referred and applied. Whereas, the photonic components are also under the potential malicious attacks from different HTs. In [35] and [36], the authors presented a new framework using key-based authentication signatures along with the architecture-level enhancements to protect the packets in communications of the ONoCs from data-snooping HTs, which can be categorized as the mixed scheme.

According to the current state-of-the-art research status, in this paper, the security issues in ONoCs are our main concerns here. *Up to now, in addition, nearly all the relevant works are concentrated on the attacks and defenses of sensitive data packets transmitted in the networks.* But similarly, the real-time distribution of on-chip temperature we care about is also prone to be “obfuscated” because of its tight relationship with the network control [4]. *In fact, however, there is scarcely any research work aiming at this field.* It is indeed significantly necessary to design an effective and targeted mitigation approach for the HT attacks in thermal-aware control of ONoCs due to the great adverse impacts of thermo-optic effects on the performance and reliability of photonic devices. Moreover, as the relevant attacks from HTs may occur in either the optical or the electronic components of an ONoC, we need to alleviate the negative influences induced in the respective parts separately. In the meanwhile, how to grasp the overall network security statuses is also a critical issue for OS to implement the appropriate runtime action to maintain the normal operation of MPSoCs, which is actually an essential HT-mitigation measure.

3 Work Flow and Attack Model

A thermal sensing flow within ONoC and our attack model are described and demonstrated in this section to introduce the processes of optical sampling and electronic transmission of the thermal sensing data, and the corresponding attack scenarios created by the triggered HTs therein.

3.1 Thermal Sensing Flow

As no specific work flow of the thermal sensing with MR-based sensor has been provided before in previous researches of this domain [13-17], we present one in this section in detail for ONoC of the classical 2D-mesh topology in Figure 5. For the normal communications, generally, the *optical circuit switch* is applied in ONoC for routing. An electronic “*path-setup*” packet will be delivered to E-router from NI while communications are in need between two PEs. When this packet arrives, the E-router will inject it to the E-net, and attempt to establish an optical path for the later data transmission. Specifically, according to the routing information contained in the packet header, the E-routers internally mark the associated resources in ONoC as “reserved”

after receiving the feedbacks from the destination nodes, and then activate the corresponding basic optical switching elements (BOSEs) in the connected O-routers to establish an optical path. Along the optical path, bulk data can be sent end-to-end without halfway buffering, arbitration and other processing.

Figure 5: An example of thermal sensing flow in traditional 2D-mesh ONoC

For our scenario of thermal sensing, the O-router is a key device to obtain the temperature-related data. Generally, an MR-based thermal sensor is comprised of three core components: a multi-wavelength laser source, a BOFE, and multiple PDs, as shown in Figure 6. The MRs in the BOFE filter out their own designated resonant wavelengths from the input ones, demultiplexing the W wavelengths before the optical signals reach respective PDs. For a single MR, we can derive a predicted temperature value based on the temperature dependence of MRs formulated in [13] and [16], which well model the relationship between the thermal-induced ΔP of an MR and its ambient temperature. Actually, the temperature values cannot be gained directly through the MR-based sensors. The relevant control data including ΔP of MR will be transferred to the corresponding E-router from the *sampling node* through NI's packetization with the mentioned extra metallic link, as shown in Figure 5. Afterwards, in the E-net, the router will route the received data to the *collection node* that is an "idle" PE to carry out the temperature calculation and centralized control, since the access latency through the shared memory of MPSoCs is extremely high and not appropriate for the real-time measurements [13]. This PE is designated by OS for the load balancing of ONoCs, which is the criteria to specify the *collection node* dynamically and periodically. In our design, the ΔP transmission is based on E-net because the temperature sampling should not affect the normal optical communications.

3.2 Attack Model

Prior related works [18-36] seldom focused on the security of control information, but that of data packets of networks. According to the thermal sensing flow in ONoCs mentioned above, some serious kinds of HT attacks can be triggered by the embedded counterfeit circuits. Generally, for the thermal-aware process, there are two chief potential HT attack scenarios during this flow in the compromised O-net and E-net, namely, attacks during the optical sampling and electronic transmission of thermal value.

Firstly, we introduce the attack scenario within the O-net. As described in [13, 16], the thermal sensitivity of MRs can be abstracted as Equation (1), which describes the relationship between the ΔP value of an *ON*-state MR at the *Drop* port and its ambient temperature, T .

$$T = T_0 + [(\sqrt{\alpha \cdot c^{\Delta P}} - \beta + \Delta\lambda) / \rho_{MR}] \quad (1)$$

where the temperature-dependent resonant wavelength shift parameter of the MR, i.e., ρ_{MR} , is a constant as well as α , β and c according to the specific MR, T_0 denotes the nominal operating temperature (typically the indoor temperature, i.e., 25°C), and $\Delta\lambda$ represents the known change of resonant wavelength due to the PV of MR. The related calculations for T will be carried out in the idle PE where all the ΔP -loaded packets arrive eventually. Here, because the MR-based sensor we adopt is applied to implement thermal sensing without participation in optical communications, no consideration is needed for local wavelength tuning. Generally, eight MRs are contained in one thermal sensor for the accurate temperature value via average calculation

[13-17]. Thus, as demonstrated in Figure 5, the sampling sensing data including all the ΔP values are collected and delivered from the *sampling node* to *collection node* through the E-routers on the path in red.

By applying the modulation voltage (MV) on an MR, its resonance wavelength will be changed, which improves the flexibility and robustness of MR for DWDM operations. Note that, the non-linear relationship between ΔP and MV from the electronic control logic was provided in [37]. However, in Figure 6, the HTs embedded in control logic will manipulate the actual value of MV, resulting in the incorrect ΔP , then finally the wrong temperature value. $V_1, V_2, V_3, \dots, V_w$ are the output MVs generated from the control logic for each MR in the sensor [38], which face the common malicious threats from voltage variations [39]. In fact, they can be easily manipulated by utilizing a direct current (DC) chopper [40] or even a simple DC voltage regulator (e.g., voltage divider) [39] at a negligible overhead for just a series of extra resistors, diodes, and capacitors, which are easy to deploy while hard to notice for testers. Once the chopper or regulator is triggered by an internal activation code, the original MV values will be tampered.

Figure 6: An MR-based thermal sensor with HTs for ONoCs

On the other hand, the malicious attacks for electronic transmission of ΔP values within E-net are similar to those modeled in the previous relevant works in the NoCs [18-36]. After all, as one sort of control data, the temperature information is also prone to be under data leakage and manipulation as mentioned [41, 42], causing negative impacts on the performance and reliability of ONoCs. In general, the correctness, integrity, and privacy are the three most critical aspects of data security we need to take into account here.

It is important to note that the HTs within the network routers may not be noticed by the testers for their 1.56%~4.62% extra area overhead [20, 34]. However, the SMCs in [19] could probably detect this type of HTs with the control logic violations of routers. We thus concentrate on the HT attacks that are harder to be detected. For the NoC based on the wormhole routing, NI is the core module to perform packetization and depacketization amongst PE and O/E-router with just one connection to them, respectively. It is easier for the attackers to snoop or tamper data through the relatively simpler compromised circuits without interfering with the basic functionality of NI at a smaller area overhead of 1.3% [31].

When the malicious logic, namely, the HT, is embedded into the packetizer of NI and triggered internally from the adjacent depacketizer, as shown in Figure 7 below, it can modify the header flit of the ΔP packet to duplicate the packets to transfer to a preset counterfeit destination PE in which the suspicious tasks execute. This type of attack behavior will lead to data leakages for the sensitive information that could be obtained by the attackers to deduce the temperature distribution of the whole chip. Furthermore, another type of attack is to directly tamper the related control data in the ΔP packet, such as the coordinates of the O-router where the certain thermal sensor locates, and also the specific values of ΔP . The OS will be spoofed and misled by the calculated temperature information due to both of the incorrect values.

Besides, according to the various behaviors of actual activation signals, the relevant counterfeit circuits of HTs in O-net and E-net can be triggered to perform the persistent attacks, intermittent attacks or mixed attacks of the former two. Meanwhile, the attack types could also be either static or dynamic during the operations of MPSoCs. These attack characteristics will bring different effects to the collected ΔP values, and make the caused damages more difficult to be detected and mitigated. Efficient response measures need to be proposed to find out which ΔP values are hijacked or corrupted, and protect them properly by carrying out the specific and targeted counteracting measures in each sub-network of ONoC.

4 Mitigation of the HT Attacks in ONoC

In this section, based on the work flow and attack model presented above, we propose the attack mitigation methods for HTs in the optical sampling and electronic transmission of ΔP , respectively. However, the two scenarios are not irrelevant, but still have close ties in the thermal-aware control of the whole optoelectronic hybrid network. They constitute a complete mitigation approach for the malicious HTs in ONoCs.

Figure 7: Attack model and mitigation of the malicious circuits in ONoCs

Considering the key role of NI in the thermal sensing flow of ONoC, it is a suitable module to execute the anti-HT mechanism, while not adding extra overhead to the on-chip interconnections. The relevant anti-HT functions are integrated together to constitute a whole module for the mitigation approach, i.e., MHT, which is embedded in NI, as illustrated in Figure 7, where the HTs within MR control logic and packetizer of NI are also demonstrated. Note that, the network data is transmitted between O/E-router and NI in packet (P) format, while ΔP info will be sent in flit (F) format from E-router to the NI of the *collection node*, as it is delivered via wormhole routing in E-net. MHT is mainly consisted of the three functional parts: optic-checker sub-module for optical sampling, en-/decryption and hash function sub-modules for electronic transmission, which will be described in detail as follows. Some auxiliary logics are also included in MHT, such as the feedback logic that is in charge of providing the related feedbacks for reference of resistance to HTs.

4.1 Mitigation for HTs in Optical Sampling

The optic-checker sub-module within MHT here is applied for the checking and correction of the tampered ΔP value transferred from O-router to NI in ONoCs. Actually, ΔP value is mainly affected by MV, TV, and PV of MR together [13, 16, 43]. Note that, normally, the eight MRs in an optical thermal sensor have different fabrication parameters, i.e., the PV of individual MR is different from each other. As PV is a known value to us, we can calculate its corresponding “impact” on ΔP , i.e., ΔP^{PV} , through Equation (2) [16, 44, 45].

$$\Delta P^{PV} = 10 \lg [(\Delta \lambda^2 + \beta) / \alpha] \quad (2)$$

where α , β and $\Delta\lambda$ represent the same parameters with those in Equation (1). And, on account of almost the same temperature change for all MRs in one sensor, the impact that ΔP of each MR obtains under TV is also regarded to be the same [16]. Besides, the possibility that all the MRs are under attack simultaneously in one thermal sensor is extremely low. For instance, even if the activation probability of the HT in an MR's control logic circuit can reach 50% or so, the probability of the scenario that the HTs in all eight MRs' control circuits are triggered is less than 0.4% indeed. However, in fact, all of these control logics are still potentially under the tampering attacks simultaneously no matter how small the probability is. For explanation, in one sensor, we mark a specific ΔP value as the standard one, namely, ΔP_S , which will be explained in detail later in this sub-section. Meanwhile, ΔP_A is used to indicate the ΔP value of the MR under a severer attack. We mark the ΔP data of relevant MRs in Equation (3) and (4), respectively [13, 16, 43, 46].

$$\Delta P_S = \Delta P_S^{MV} + \Delta P_S^{TV} + \Delta P_S^{PV} \quad (3)$$

$$\Delta P_A = \Delta P_A^{MV} + \Delta P_A^{TV} + \Delta P_A^{PV} \quad (4)$$

where ΔP^{MV} , ΔP^{TV} , and ΔP^{PV} indicate the specific ΔP values affected by MV, TV and PV. As declared above, ΔP_S^{PV} and ΔP_A^{PV} are both available. $\Delta P'$ can be worked out if we set it as $(\Delta P - \Delta P^{PV})$, namely, $(\Delta P^{MV} + \Delta P^{TV})$. Above all, if all the $\Delta P'$ values of the MRs are equal, then no HT is embedded or activated, or the attack effect for each MR is the same. Once there is at least one $\Delta P'$ value that is different from $\Delta P_S'$, it is indicated that this sensor must be under attack. In addition, the TV for each MR within one thermal sensor is assumed to be the same according to our scenario, thus, $\Delta P_S^{TV} = \Delta P_A^{TV}$. Therefore, these ΔP^{TV} values will be eliminated through the operation of subtraction between ΔP_S and ΔP_A . Then the difference between ΔP_S and ΔP_A values will be only dependent on the ΔP^{MV} -based expression. Moreover, according to the relationship between the transmissivity at modulation wavelength and MV of the MR-based sensor [37], here we fit an approximate function as Equation (5) through certain functions [47] using MATLAB Curve Fitting Toolbox [48] and a convincing photonic integrated circuits simulator INTERCONNECT of Lumerical [49].

$$f(x) = [(e^{ax-b})/(e^{cx-d} + 1)] + q \quad (5)$$

where a , b , c , d , and q are the constant parameters. The curve of the function in Equation (5) is described in Figure 8 for the ideal MR without PV. For an actual MR, all the constant parameters need to be regulated accordingly. Also, by reason of $\Delta P^{MV} = -10 \lg[f(x)]$ [13, 16], we can calculate the malicious variation of MV, then inform the corresponding control logic to modify the output modulation voltage to an appropriate value for sensor to retransmit the correct data by tolerating the attacks inside, such as adjusting the input voltage of the control logic or directly isolating the affected circuits to recover the original MV. Whether the tampering behavior is a persistent, intermittent or mixed type, our strategy checks out the ΔP values when receiving them from PDs of the BOFE. The retransmission will not end until the effective modifications are applied to the control logics circuits for MRs. For instance, the number of retransmissions depends on the specific data-sampling rate and operating frequency of the proportional-integral-derivative (PID) controller in modulation controller for the relatively simple voltage-adjusting method [17, 50, 51].

Figure 8: Transmissivity at modulation wavelength over MV for an ideal MR

Note that, because all the ΔP^{TV} values are equal and all the ΔP^{PV} values are known for eight MRs in each sampling of a sensor, we choose the ΔP with the lowest specific $\Delta P'$ value as ΔP_S in one sensor. As mentioned before, ΔP^{MV} is the only possibly attacked variable among the eight different ΔP values, and the value “1” should be normally obtained by the calculation of the transmissivity at modulation wavelength as closely as possible so as to resonate the particular optical signals, while letting ΔP^{MV} be closer to “0”. In other words, generally, ΔP^{MV} of a certain MR should be “0” if its MV is not manipulated. Here, we set $\Delta P_S^{MV}=0$ for the calculation of the modified MV values of specific MRs possessing ΔP_A values. Even if the HTs within all of the control logics of MRs are activated rarely, causing that any ΔP^{MV} value of the eight MRs cannot be equal to “0” indeed, we still need to adopt the lowest ΔP^{MV} value to be the standard one. It is because this value is the closest one to “0”, which can guarantee the most approximate accuracy of the temperature distribution of the whole chip for coordination and control of OS.

4.2 Mitigation for HTs in Electronic Transmission

After the thermal sensing process, MHT within NI checks the correction result of the sampling information from O-router. And then NI will continue to send these data to the *collection node* through the connected E-router. As discussed in the attack model of Section 3, the thermal sensing data is probably to be snooped and tampered before the package’s injection to the network from NI.

Taking the unique features of an MR into account, PV is an ideal attribute for en-/decryption methods to protect the sampling data from leakage, which is implemented in the encryption & decryption sub-module of MHT. However, PV is also a complex attribute of MR, including length, width, thickness, etc. Therefore, it is needed to comprehensively consider the impacts caused by all these aspects to gain a PV-induced effect. We adopt a defined concept in Section 3, i.e., $\Delta\lambda$ (the change of resonant wavelength due to the PV of MR), as the basis of encryption key. $\Delta\lambda$ is a known constant according to a specific MR, and can be converted to a 64-bit digital signal through a dithering-signal-based in-situ method [35]. As the data flit size of the network is considered to be typically set as 128 bits in our mechanism, this digital signal needs to be appended twice to generate a 128-bit encryption key to match the flit size. And then, the exclusive OR (XOR) operation is applied and performed between this key and the data flit, as demonstrated in Equation (6).

$$E_{Temp} = O_{Temp} \oplus K_{\Delta\lambda} \quad (6)$$

where E_{Temp} and O_{Temp} are the encrypted and original sampling data flits, and $K_{\Delta\lambda}$ is the $\Delta\lambda$ -based encryption key. Here, all the $K_{\Delta\lambda}$ values involved are stored in each PE node of the ONoC by default in our mechanism. Specifically, the current *collection node* collects the encrypted ΔP data from the different *sampling nodes* at a certain frequency [13, 16], and decrypt them with the corresponding PV-related key, which also performs in MHT of NI after receiving the thermal information.

As the snooping attack duplicates the sampling packets and changes the destination address in the header flit, the amount of data transmitted in the network will be increased, which will potentially cause the traffic contentions. Different from the normal communication data, the real-time thermal sensing packets will be sent only once during an optical sampling period, e.g., 40 ns consisting of 30 ns for sampling operation and 10 ns for interval time [17]. As a result, MHT can easily distinguish the duplicated packet with the modified header flit, which includes the incorrect destination address and misleads the NI to inject the related flits from the “*source node*” to a compromised node where malicious tasks may be carried out. And then, the duplicated packets will be discarded directly and the corresponding flits will not enter the flit queue of packetizer in NI, leading to the higher traffic and improved performance of the E-net. In comparison, the duplicated packets in the snooping attack for normal network communications are much more difficult to identify, requiring the specialized invalidation mechanisms [31].

Whereas, the encrypted sampling data is still exposed to the possible risks of data manipulation. This will probably result in the failure of decryption at the *collection node*, which needs to send the retransmission request to the corresponding *sampling node*. In addition, the OS will be deceived if the encryption key $K_{\Delta\lambda}$ is stolen/cracked (e.g., under certain extra side-channel attacks) and the ΔP value is forged by the malicious third party. In order to avoid the scenarios above, we present a scheme to check the data integrity via the hash algorithm, as an effective authentication solution, to map data of arbitrary size to a hash of a fixed size for

one-way operation. The hash function sub-module (efficient message-digest algorithm 5, namely, MD5 [52] is employed here) in the MHT is to attach an extra 128-bit (i.e., 1 flit) hash value within the sensing packet. While the received packet is decrypted after its arrival, MHT will authenticate the data to check whether the original information is tampered or not by employing the same hash algorithm on the decrypted packet and comparing the obtained result with the attached hash value. If they are not matched, the ΔP value should be covertly modified owing to the uniqueness of the specific hash value. Regardless of being tampered directly upon the encrypted data or indirectly with the stolen/cracked keys as ΔP is, the *collection node* will send the request packets back to the *sampling nodes* to retransmit the new real-time thermal sensing data, for the rate of sampling (e.g., 25,000 kSa/s or so [17]) is sufficient to support this operation.

Considering the relevant encryption keys are probably leaked out when there are some mismatches in the comparison during the data-integrity checking, the newly generated keys need to be adopted for the sampling data against the tampering attacks for better robustness and flexibility of the encryption. The mentioned hash function will continue to be used to update the encryption keys via Equation (7) as follows.

$$K_{\Delta\lambda}' = \text{Hash}(K_{\Delta\lambda}) \quad (7)$$

where $K_{\Delta\lambda}'$ is the updated key, which is still a 128-bit value to be employed in Equation (6) instead of $K_{\Delta\lambda}$. Meanwhile, the keys in *collection node* will also be updated accordingly and kept consistent for decryption operation. In conclusion, similar to the mitigation scheme in compromised O-net, whether the snooping and tampering attacks in E-net are persistent, intermittent or mixed types, our strategy can alleviate the impacts induced by the malicious attacks of HTs embedded in the NIs with continuous runtime countermeasures.

5 Analysis for Security Status of ONoC

Here, we briefly summarize the proposed HT-mitigation scheme for ONoCs for a clear demonstration in Table 1 as below. The attack behaviors can be thwarted by performing corresponding checking or correction operations as discussed. However, the detection of security statuses of ONoC is highly essential to be applied to report to OS for the subsequent high-level control to prevent the further negative impacts of the malicious attacks. Taking into consideration of the various attack scenarios from HTs in O-net and E-net, the specific information of retransmission requests, such as the sending node and time, turns out to be eligible as the key indicators for network security analyses. In this paper, we utilize these indicators to effectively perform the detection of security statuses of the whole network, which could be transformed as an abstract recognition issue for the spatial and temporal components of data indeed.

Table 1: The brief summary of the proposed mitigation scheme

Location	Work Flow	Attack Type	Mitigation Scheme
O-net	Optical sampling	Tampering	Theoretical checking & correction
E-net	Electronic transmission	Snooping & Tampering	Symmetric encryption & Integrity checking

In view of that the artificial neural network (ANN) [53] is a traditional and efficient tool to identify and classify different patterns, which can be exactly utilized to carry out the analyses of network security statuses. In particular, as the third generation of ANN, SNN has the special ability to recognize the spatial and temporal components in event data. The event data that behaves as the spike, in principle, is aggregated through several synapses and accumulated by a neuron. That is, neurons detect particular features of event data via groups of synapses tuned by the synaptic plasticity [7, 54]. When the accumulative effect (i.e., the membrane potential) of the neuron reaches or exceeds a certain threshold value, this neuron will “fire”, and output a spike event. Subsequently, the neuron will stop firing when the membrane potential falls below the threshold. This process enables the spatio-temporal events to be propagated by groups of neurons. Note that the membrane potential of the neuron can rise or fall depending on the stimulation level of signals from other connected neurons via the synapses. Hence, in our case, the detection issue can be well analyzed through SNN to intuitively classify the security statuses of ONoC as “normal” patterns or “abnormal” patterns.

Figure 9: SNN-based detection for security status of ONoCs

As shown in Figure 9, for SNN, the input neurons are connected to one hidden layer for interpretation and ultimately two output neurons. In one tile that consists of PE, NI, routers and relevant links [55], the detection method with SNN explores the connection of synapses to each O/E-router-linked port of the NI. The synapses “read” the pulses of digital data that is transferred through the two ports. The retransmission request signals act the role of input spikes here, including the spatial (2D array of the components in ONoC) and temporal (a serial exchanges through the involved links) features. It is necessary to note that, there would be common retransmission requests during the presented thermal-aware work flow because of the existing contentions, noise interference, transit faults or other possible reasons within both O-net and E-net. We need to detect the abnormal patterns which present the different spatial and temporal features, namely, for a specific port, there are more retransmission request signals in a certain time period than there should normally be, which can signify the security issues. For each output neuron, similarly, there is also a threshold which is set to finally judge the network security status is normal or abnormal. As discussed above, if this threshold is reached or exceeded, respective patterns will be detected. Both the to-be-classified normal and abnormal patterns need to be well distinguished from each other by updating the associated parameters for SNN training accordingly. Furthermore, the synaptic plasticity allows the neural network to adapt to changes in event data. Hence, the training and test datasets are obtained by applying a variety of traffic profiles in different communication scenarios to ensure the relatively real situations for higher-level control operations executed by the OS for ONoCs. Different from the known training samples, test samples need to follow the principle of “the attacker has the element of surprise” for SNN [1]. In addition, since the detection of network security statuses is based on a specific time range during the thermal sensing, the relevant data (i.e., the information of retransmission requests, etc.) can be collected by a certain idle PE of ONoC to implement the respective trained SNN therein, using the access of on-chip shared memory without strict real-time demands, which also creates fewer adverse impacts on the runtime thermal-aware control of networks. For load balancing of the whole on-chip network, the criteria to specify the idle PE here is similar to that of the *collection node*, which is described in Section 3.1. In accordance with the related configurations and settings, multi-level notifications or alerts will be raised to the OS, which informs the users and takes preventive measures for the possible serious DoS, performance degradation, malfunction, chip destruction, and so on.

6 Evaluation and Analysis

6.1 Evaluation Setup

We modified a full system simulator Gem5 [56] and a cycle-accurate NoC simulator Booksim [57] to model the ONoC of $m \times n$ 2D-mesh topology. And, each of the tile [55] consists of an x86 processor core based on out-of-order (OoO) execution and an SRAM cache bank in 22 nm process technology. The detailed system configurations are described in Table 2 for the evaluations of performance (including transmission latency and execution time) and overhead (including area footprint and energy consumption). Benchmark suite of PARSECv2.1 [58] is used with Gem5 to emulate the proposed HT-mitigation approach with real applications. We also validated the experimental platform in terms of the overhead on area and energy according to the estimated results obtained from the NoC simulator and DSENT tool [59].

As most of the attack model and network architecture of related works are widely different with ours, we present a comprehensive and fair comparison between the baseline ONoC (without any security mechanisms) with its variants and the proposed mitigation method for investigation. Besides, the types of the unexpected attack (i.e., persistent, intermittent and mixed) are set to be launched randomly from the activated HTs.

Table 2: Full-system simulation configurations

Properties	Specific Configurations
Processing Element	m×n processor cores, 2GHz OoO, 4 issues, x86, 32KB, 2-way I-cache/D-cache (1-cycle latency)
L2 Cache	Private, 256KB, 4-way, 8-cycle latency
ONoC Topology	m×n 2D-mesh, XY routing
Memory	4GB DRAM, 1 rank, 1 channel, 8 banks, 90-cycle latency
Coherence Protocol	MOESI protocol
O-router	DWDM supported, 8 MRs per thermal sensor, Cygnus architecture [60]
E-router	3 pipeline stages, wormhole flow control, 8-flit depth buffer, 2 virtual channels

6.2 Performance Study

6.2.1 Transmission Latency

Transmission latency is a key indicator for the performance analysis of the HT-mitigation approaches for ONoCs, where we employ the PARSEC benchmark suite as the workloads in an 8×8 mesh network. Note that, in view of the HT-mitigation operations implemented in the MHT of NI, we measure this latency existed between the respective NI of the source node and destination node for fair and reasonable evaluation. Here, Figure 10 below presents the average latency of packet transmission normalized to the latency of respective corresponding baseline networks with different ratios of activated HTs (i.e., 5%, 10%, 15%, and 20%, the common cases in practice [1]). HTs are embedded in the relevant circuits by default as assumed here. For the comparison, the performances of secure ONoCs are a bit lower due to the en-/decryption operations, checking for accuracy and integrity of ΔP values, and retransmission, etc.

Figure 10: Transmission latency comparison with different ratios of triggered HTs

Specifically, for the high-contented applications, such as *canneal* and *freqmine*, the increment of packet latency is relatively large. But, for the low-contented applications, for instance, *streamcluster*, the increment is moderate. In conclusion, the higher the ratio of activated HTs is, the higher the extra packet latency will be introduced. The latency rises 9.89% maximally and 8.73% on average with 20% of HTs triggered, which

is acceptable for ONoC under such a high-ratio scenario to implement the security-enhanced thermal-aware communications at a relatively low cost. The performance will degrade significantly if the ratio of triggered HTs rises continuously, which will be detected/classified and reported to OS to take the next actions, such as the logical isolation and physical replacement for related modules or devices under attack.

6.2.2 Execution Time

Execution time is another important indicator of performance evaluation for different real applications run in NoCs. Compared with the global average latency of packets, execution time reflects the network performance more directly and intuitively for each application. Figure 11 illustrates the benchmark execution time that is also normalized to the execution time of respective corresponding baseline networks with different ratios of activated HTs (i.e., 5%, 10%, 15%, and 20%). Similar trends of transmission latency can also be observed in execution time but with a smaller degree of increment. Still take the HT-activation ratio of 20% as an example, the related metric rises 6.48% maximally and 5.32% on average, which is affordable for ONoCs under the severe attacks. Real applications can avoid the significant restrictions on their execution with the security assurance of thermal sensing in MPSoCs.

Figure 11: Execution time comparison with different ratios of triggered HTs

6.3 Overhead Analysis

6.3.1 Area Footprint

An MR-based thermal sensor occupies 0.0123 mm^2 of silicon area that is similarly sized to most of CMOS-compatible electric ones compared in [14] and [17]. Furthermore, the area of MHT is $317.94 \mu\text{m}^2$ with an extra footprint increment of 6.22% for an NI module via the simulation.

Table 3: Area footprint of different security-enhanced mechanisms for NoCs

Categories	Security Design	Area
Detection and Correction	P-Sec [18]	$300\sim12850 \mu\text{m}^2$
	SMCs [19]	$230 \mu\text{m}^2$
Shuffling Scheme	Link shuffling [23]	$126.72\sim2152.8 \mu\text{m}^2$
Obfuscation	TASP [25]	$498.33 \mu\text{m}^2$
	State obfuscation [26]	$614.5 \mu\text{m}^2$
	Harden Design [27]	$57514.3 \mu\text{m}^2$
Firewall or Security Module	Prometheus [32]	$7000\sim85000 \mu\text{m}^2$
Mixed Scheme	Fort-NoCs [34]	$4917 \mu\text{m}^2$
	Proposed MHT	$317.94 \mu\text{m}^2$

Compared with the promising security enhancement modules in the typical state-of-the-art progresses, as shown in Table 3, our MHT still possesses an obvious advantage on area footprint with consideration of the

security capability. Moreover, for some of the referred designs, the area values within a specific range show that there is more than one implementation for similar protection capabilities. Generally, the implementation of a smaller footprint indicates the weaker security assurance [18, 23, 32]. As the NI's area occupies less than 10% of a tile of NoCs [55], the increase of the tile area will be smaller than 1%, which is negligible indeed to perform the proposed mitigation approach.

6.3.2 Energy Consumption

The energy consumption of our optical MR-based sensor is just 627.71 pJ/Sa [17], which is much lower than those of electric sensors. In Figure 12, it is the comparison on energy consumption, which is normalized to the energy consumption of respective corresponding mesh-based baseline ONoCs of different scales (i.e., 6×6 , 8×8 , 10×10 , and 12×12) with activated HTs at the ratio of 15%. For a relatively large-sized network, for instance, it is indicated that the energy consumption rises 7.35% maximally and 6.14% on average in the 12×12 mesh ONoC. For the scalability study of our approach, the power growth rates are still in a reasonable scope with an overall apparent convergence trend as shown in the figure. Though the absolute number of activated HTs increases, which leads to more possible overhead on extra control and data transfers because of the related checking, en-/decryption and retransmission operations, while the overall energy consumption (both the static and dynamic portions) will rise more rapidly when the network size reaches a certain level according to the related systemic configurations.

Figure 12: Energy consumption comparison in networks at different scales

6.4 Evaluation on SNN-based Detection

We gather 60 datasets (12,000 samples for each) for SNN training/test (7:3) via different traffic models (such as Uniform, Transpose, Hotspot, etc. [57]) with the oblivious and adaptive routings under HT-activation ratio of 10%. Each router-linked port of NI is connected to one of the total 32 neurons of input layer for 4×4 mesh ONoC, respectively. As utilized in several neuroscience applications, NeuCube is used to create and visualize our SNN [61]. The built SNN consists of 1000 neurons, and has been trained with the normal and abnormal patterns for offline unsupervised learning. As a media-processing program of PARSEC, x264 application is employed to be implemented in HT-embedded ONoC with classical DyAD routing [62] (a hybrid algorithm combined with order and adaptive types). As a result, the SNN can detect the normal patterns unmistakably for the test samples of various attacks, while the abnormal patterns are properly classified at the accuracy of 90.07% on average for those samples under different scenarios as shown below.

Specifically, in Table 4, we divide each attack scenario into three sub-aspects, such as the duration type, specific duration of sample length, and location of the network. It is indicated from the experimental results that the attack duration is the key factor of detection of the network security statuses, while the start time of attacks within the lifetime of the execution of an application is not. Note that, the detection accuracy within E-net is relatively lower than that in O-net or the whole network for the same attack duration. This is because the retransmission request signals will also be sent from the destinations when there are packet losses, which is more common for relatively long-distance transmission in E-net.

Table 4: The accuracy of SNN detection for abnormal patterns

#	Attack Scenarios (for all activated HT carriers)			Accuracy
	Duration Type	Specific Duration	Location	
1	Fixed attack duration	50% of sample length	Within both O-net and E-net	93.25%
2	Incremental attack duration	30%~50% of sample length	Within both O-net and E-net	86.83%
3	Fixed attack duration	50% of sample length	Only within O-net	94.31%
4	Incremental attack duration	30%~50% of sample length	Only within O-net	88.04%
5	Fixed attack duration	50% of sample length	Only within E-net	92.28%
6	Incremental attack duration	30%~50% of sample length	Only within E-net	85.68%
7	Fixed attack duration	50% of sample length at different start times	Within both O-net and E-net	93.27%
8	Incremental attack duration	30%~50% of sample length at different start times	Within both O-net and E-net	86.76%
9	Fixed attack duration	50% of sample length at different start times	Only within O-net	94.43%
10	Incremental attack duration	30%~50% of sample length at different start times	Only within O-net	88.02%
11	Fixed attack duration	50% of sample length at different start times	Only within E-net	92.35%
12	Incremental attack duration	30%~50% of sample length at different start times	Only within E-net	85.64%
Avg.	-			90.07%

7 Conclusion and Future Work

An HT is a malicious modification of the circuitry of an integrated circuit. It can also be introduced to the on-chip optoelectronic interconnections by the integration of counterfeit circuits (e.g., 3P-IPs) in MPSoCs with ONoC exploited. In view of the potential attacks aiming at the thermal-aware control of ONoCs from the malicious HTs, we thus propose a complete mitigation scheme to alleviate the negative impacts, while adopting an SNN-based detection to relatively accurately classify the normal and abnormal statuses of the network. Evaluation results indicate that, with less than 1% extra area of a tile, our approach is an attractive security-enhanced solution for thermal sensing in ONoCs with trivial costs of up to 8.73%, 5.32% and 6.14% in average latency, execution time and energy consumption, respectively.

As many attacks for NoCs, especially those for the photonic devices in ONoCs, are probably much more complicated and diversified than those described in our attack model in the future, such as hybrid attacks on both the communication and control data within the whole ONoC, we will continue to be concentrated on the reliable, scalable and lightweight security enhancements for the networks of different topologies.

ACKNOWLEDGMENTS

This work is partially supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 2 (MOE2019-T2-001-071) and Tier 1 (MOE2019-T1-001-072), and Nanyang Technological University, Singapore, under its NAP (M4082282) and SUG (M4082087).

REFERENCES

- [1] Zhao Huang, Quan Wang and Pengfei Yang, “Hardware Trojan: research progress and new trends on key problems,” in *Chinese Journal of Computers*, 42(5), 2019, pp. 993-1017. DOI: <http://dx.chinadoi.cn/10.11897/SP.J.1016.2019.00993>.
- [2] Luka Daoud, “Secure network-on-chip architectures for MPSoC: overview and challenges,” in

- Proceedings of IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2018, pp. 542-543. DOI: <https://doi.org/10.1109/mwscas.2018.8623831>.
- [3] David A. B. Miller, "Device requirements for optical interconnects to silicon chips," in *Proceedings of IEEE*, 97(7), 2009, pp. 1166-1185. DOI: <https://doi.org/10.1109/jproc.2009.2014298>.
 - [4] Keren Bergman, Luca P. Carloni, Aleksandr Biberman, Johnnie Chan and Gilbert Hendry, "Photonic network-on-chip design," 1st edition, Vol. 68 of *Integrated Circuits and Systems*, Springer, 2014. DOI: <https://doi.org/10.1007/978-1-4419-9335-9>.
 - [5] Luka Daoud, "Secure network-on-chip against black hole and tampering attacks," a thesis for the degree of Ph.D. in *Boise State University, USA*, 2020.
 - [6] Jun Zhou, Mengquan Li, Pengxing Guo and Weichen Liu, "Mitigation of tampering attacks for MR-based thermal sensing in optical NoCs," in *Proceedings of IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2020, pp. 554-559. DOI: <https://doi.org/10.1109/ISVLSI49217.2020.00009>.
 - [7] Jilles Vreeken, "Spiking neural networks, an introduction," in *Computing*, vol. 7, no. 3, 2002, pp. 1-5.
 - [8] Hyun-Tae Kim and Miao Yu, "Cascaded ring resonator-based temperature sensor with simultaneously enhanced sensitivity and range," in *Optics Express*, 24(9), 2016, pp.9501-9510. DOI: <https://doi.org/10.1364/oe.24.009501>.
 - [9] Yaoyao Ye, Jiang Xu, Xiaowen Wu, Wei Zhang, Xuan Wang, Mahdi Nikdast, Zhehui Wang and Weichen Liu, "System-level modeling and analysis of thermal effects in optical networks-on-chip," in *IEEE Transactions on Very Large Scale Integration Systems (TVLSI)* 21(2), 2012, pp. 292-305. DOI: <https://doi.org/10.1109/tvlsi.2012.2185524>.
 - [10] Kishore Padmaraju and Keren Bergman, "Resolving the thermal challenges for silicon microring resonator devices," in *Nanophotonics* 3(4-5), 2014, pp. 269-281. DOI: <https://doi.org/10.1515/nanoph-2013-0013>.
 - [11] Wim Bogaerts, Peter De Heyn, Thomas Van Vaerenbergh, Katrien De Vos, Shankar Kumar Selvaraja, Tom Claes, Pieter Dumon, Peter Bienstman, Dries Van Thourhout and Roel Baets, "Silicon microring resonators," in *Laser and Photonics Reviews*, 6(1), 2012, pp. 47-73. DOI: <https://doi.org/10.1002/lpor.201100017>.
 - [12] Shijun Xiao, Maroof H Khan, Hao Shen and Minghao Qi, "Modeling and measurement of losses in silicon-on-insulator resonators and bends," in *Optics Express*, 15(17), 2007, pp. 10553-10561. DOI: <https://doi.org/10.1364/OE.15.010553>.
 - [13] Mengquan Li, Weichen Liu, Nan Guan, Yiyuan Xie and Yaoyao Ye, "Hardware-software collaborative thermal sensing in optical network-on-chip based manycore systems," in *ACM Transactions on Embedded Computing Systems (TECS)*, 18(6), No. 118, 2019, pp. 1-24. DOI: <https://doi.org/10.1145/3362099>.
 - [14] Weichen Liu, Guiyu Tian, Mengquan Li, "Autonomous temperature sensing for optical network-on-chip," in *Elsevier Journal of Systems Architecture (JSA)*, 102, No. 101650, 2020. DOI: <https://doi.org/10.1016/j.sysarc.2019.101650>.
 - [15] Weichen Liu, Peng Wang, Mengquan Li, Yiyuan Xie and Nan Guan, "Quantitative modeling of thermo-optic effects in optical networks-on-chip," in *Proceedings of ACM/IEEE Great Lakes Symposium on VLSI (GLSVLSI)*, 2017, pp. 263-268. DOI: <https://doi.org/10.1145/3060403.3060457>.
 - [16] Weichen Liu, Mengquan Li, Wanli Chang, Chunhua Xiao, Yiyuan Xie, Nan Guan and Lei Jiang, "Thermal sensing using micro-ring resonators in optical network-on-chip," in *Proceedings of ACM/IEEE Design, Automation & Test in Europe (DATE)*, 2019, pp. 1611-1614. DOI: <https://doi.org/10.23919/DATE.2019.8714844>.
 - [17] Mengquan Li, Jun Zhou and Weichen Liu, "Lightweight thermal monitoring for optical networks-on-chip via router reuse," in *Proceedings of ACM/IEEE Design, Automation & Test in Europe (DATE)*, 2020, pp. 406-411. DOI: <http://doi.org/10.23919/DATE48585.2020.9116351>.
 - [18] Travis Boraten and Avinash Karanth Kodi, "Packet security with path sensitization for NoCs," in *Proceedings of ACM/IEEE Design, Automation & Test in Europe (DATE)*, 2016, pp. 1136-1139. DOI: https://doi.org/10.3850/9783981537079_0180.
 - [19] Travis Boraten, Dominic DiTomaso and Avinash Karanth Kodi, "Secure model checkers for network-on-chip (NoC) architectures," in *Proceedings of ACM/IEEE Great Lakes Symposium on VLSI (GLSVLSI)*, 2016, pp. 45-50. DOI: <https://doi.org/10.1145/2902961.2903032>.
 - [20] Mubashir Hussain, "Runtime detection of hardware Trojan in untrusted network-on-chip," a thesis for the degree of Ph.D. in *University of New South Wales (UNSW), Australia*, 2018.
 - [21] Travis H. Boraten and Avinash K. Kodi, "Securing NoCs against timing attacks with non-interference

- based adaptive routing,” in *Proceedings of ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, 2018, No. 14, pp. 1-8. DOI: <https://doi.org/10.1109/NOCS.2018.8512166>.
- [22] Luka Daoud and Nader Rafla, “Routing aware and runtime detection for infected network-on-chip routers,” in *Proceedings of IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2018, pp. 775-778. DOI: <https://doi.org/10.1109/MWSCAS.2018.8623972>.
- [23] Qiaoyan Yu and Jonathan Frey, “Exploiting error control approaches for hardware Trojans on network-on-chip links,” in *Proceedings of IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFTS)*, 2013, pp. 266-271. DOI: <https://doi.org/10.1109/DFT.2013.6653617>.
- [24] Manoj Kumar J.Y.V., Ayas Kanta Swain, Sudeendra Kumar, Sauvagya Ranjan Sahoo and Kamalakanta Mahapatra, “Run time mitigation of performance degradation hardware Trojan attacks in NoC,” in *Proceedings of IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2018, pp. 738-743. DOI: <https://doi.org/10.1109/ISVLSI.2018.00139>.
- [25] Travis Boraten and Avinash Karanth Kodi, “Mitigation of denial of service attack with hardware Trojans in NoC architectures,” in *Proceedings of IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, 2016, pp. 1091-1100. DOI: <https://doi.org/10.1109/IPDPS.2016.59>.
- [26] Jonathan Frey and Qiaoyan Yu, “Exploiting state obfuscation to detect hardware Trojans in NoC network interfaces,” in *Proceedings of IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2015, pp. 1-4. DOI: <https://doi.org/10.1109/MWSCAS.2015.7282167>.
- [27] Jonathan Frey and Qiaoyan Yu, “A hardened network-on-chip design using runtime hardware Trojan mitigation methods,” in *Elsevier Integration, the VLSI Journal*, 56, 2017, pp. 15-31. DOI: <https://doi.org/10.1016/j.vlsi.2016.06.008>.
- [28] Musharraf Hussain, Naveed Khan Baloach and Aamir Anwar, “A non-invasive approach to detect and mitigate hardware Trojan in network on chip,” in *International Journal of Computer Science and Mobile Computing (IJCSMC)*, 8(7), 2019, pp. 188-202.
- [29] Ahmed Ben Achballah, Slim Ben Othman and Slim Ben Saoud, “Toward on hardware firewalling of networks-on-chip based systems,” in *Proceedings of IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, 2017, pp. 7-13. DOI: <https://doi.org/10.1109/ASET.2017.7983658>.
- [30] Miltos D. Grammatikakis, Polydoros Petrakis, Antonis Papagrigoriou, George Kornaros and Marcello Coppola, “High-level security services based on a hardware NoC firewall module,” in *Proceedings of International Workshop on Intelligent Solutions in Embedded Systems (WISES)*, 2015, pp. 73-78.
- [31] Venkata Yaswanth Raparti and Sudeep Pasricha, “Lightweight mitigation of hardware Trojan attacks in NoC-based manycore computing,” in *Proceedings of ACM/IEEE Design Automation Conference (DAC)*, 2019, No. 48, pp. 1-6. DOI: <https://doi.org/10.1145/3316781.3317851>.
- [32] Brian Lebednik, Sergi Abadal, Hyoukjun Kwon and Tushar Krishna, “Architecting a secure wireless network-on-chip,” in *Proceedings of ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, 2018, No. 16, pp. 1-8. DOI: <https://doi.org/10.1109/NOCS.2018.8512168>.
- [33] Taimour Wehbe and Xiaofang Wang, “Secure and dependable NoC-connected systems on an FPGA chip,” in *IEEE Transaction on Reliability*, 65(4), 2016, pp. 1852-1863. DOI: <https://doi.org/10.1109/TR.2016.2606883>.
- [34] Dean Michael Ancajas, Koushik Chakraborty and Sanghamitra Roy, “Fort-NoCs: mitigating the threat of a compromised NoC,” in *Proceedings of ACM/IEEE Design Automation Conference (DAC)*, 2014, No. 158, pp. 1-6. DOI: <https://doi.org/10.1145/2593069.2593144>.
- [35] Sudeep Pasricha, Sai Vineel Reddy Chittamuru, Ishan G Thakkar and Varun Bhat, “Securing photonic NoC architectures from hardware Trojans,” in *Proceedings of ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, 2018, No. 15, pp. 1-8. DOI: <https://doi.org/10.1109/NOCS.2018.8512167>.
- [36] Sai Vineel Reddy Chittamuru, Ishan G Thakkar, Varun Bhat and Sudeep Pasricha, “SOTERIA: exploiting process variations to enhance hardware security with photonic NoC architectures,” in *Proceedings of ACM/IEEE Design Automation Conference (DAC)*, 2018, No. 81, pp. 1-6. DOI: <https://doi.org/10.1109/DAC.2018.8465939>.
- [37] Ali Ayazi, Tom Baehr-Jones, Yang Liu, Andy Eu-Jin Lim and Michael Hochberg, “Linearity of silicon ring modulators for analog optical links,” in *Optics Express*, 20(12), 2012, pp. 13115-13122. DOI: <https://doi.org/10.1364/OE.20.013115>.
- [38] Zhongqi Li, Amer Qouneh, Madhura Joshi, Wangyuan Zhang, Xin Fu and Tao Li, “Aurora: a cross-layer solution for thermally resilient photonic network-on-chip,” in *IEEE Transactions on Very Large*

- Scale Integration Systems(TVLSI)*, 23(1), 2015, pp. 170-183. DOI: <https://doi.org/10.1109/TVLSI.2014.2300477>.
- [39] Young Sil Lee, Non Thirananant, HyeongRag Kim, JungBok Jo and HoonJae Lee, "Fault attacks by using voltage and temperature variations: an investigation and analysis of experimental environment," in *Information Science and Applications, Springer*, 2015, pp. 483-490. DOI: http://doi.org/10.1007/978-3-662-46578-3_57.
- [40] Ali Adam, A. Elnady and Amer Ghias, "A novel multilevel DC chopper supplying DC motor," in *Proceedings of International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, 2016, pp. 1-5. DOI: <https://doi.org/10.1109/ICEDSA.2016.7818494>.
- [41] Yazhou Tu, Sara Rampazzi, Bin Hao, Angel Rodriguez, Kevin Fu and Xiali Hei, "Trick or heat?: manipulating critical temperature-based control systems using rectification attacks," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2019, pp. 2301-2315. DOI: <http://doi.org/10.1145/3319535.3354195>.
- [42] Kishore Padmaraju, Dylan F. Logan, Xiaoliang Zhu, Jason J. Ackert, Andrew P. Knights and Keren Bergman, "Integrated thermal stabilization of a microring modulator," in *Optics Express*, 21(12), 2013, pp. 1-10. DOI: <http://doi.org/10.1364/OE.21.014342>.
- [43] Wim Bogaerts, Peter De Heyn, Thomas Van Vaerenbergh, Katrien DeVos, Shankar Kumar Selvaraja, Tom Claes, Pieter Dumon, Peter Bienstman, Dries Van Thourhout and Roel Baets, "Silicon microring resonators," in *Laser Photonics Reviews, Wiley*, 6(1), 2012, pp. 47-73. DOI: <http://doi.org/10.1002/lpor.201100017>.
- [44] Yi Xu, Jun Yang and Rami Melhem, "Tolerating process variations in nanophotonic on-chip networks," in *Proceedings of ACM/IEEE International Symposium on Computer Architecture (ISCA)*, 2012, pp. 142-152. DOI: <https://doi.org/10.1109/ISCA.2012.6237013>.
- [45] Pengxing Guo, Weigang Hou, Lei Guo, Wei Sun, Chuang Liu, Hainan Bao, Luan H. K. Duong and Weichen Liu, "Fault-tolerant routing mechanism in 3D optical network-on-chip based on node reuse," in *IEEE Trans. on Parallel and Distributed Systems*, 31(3), 2020, pp.547-564. DOI: <http://doi.org/10.1109/TPDS.2019.2939240>.
- [46] Christopher Nitta, Matthew Farrens and Venkatesh Akella, "Addressing system-level trimming issues in on-chip nanophotonic networks," in *IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2011, pp. 122-131. DOI: <http://doi.org/10.1109/HPCA.2011.5749722>.
- [47] Jun Han and Claudio Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," in *Proceedings of International Work-Conference on Artificial Neural Networks (IWANN)*, 1995, pp. 195-201. DOI: https://doi.org/10.1007/3-540-59497-3_175.
- [48] MATLAB Curve Fitting Toolbox, <https://www.mathworks.com/products/curvefitting.html>.
- [49] INTERCONNECT simulator, <https://www.lumerical.com/products/interconnect/>.
- [50] Zhao Wang, "Silicon micro-ring resonator modulator for inter/intra-data centre applications," a thesis for the degree of Ph.D. in *Mcmaster University, Canada*, 2017.
- [51] Mohammed F. Tolba, BahaaAlDeen M. AboAlNaga, Lobna A. Said, Ahmed H. Madian and Ahmed G. Radwan, "Fractional order integrator/differentiator: FPGA implementation and FOPID controller application," in *International Journal of Electronics and Communications (AEÜ)*, 98, 2019, pp. 220-229. DOI: <https://doi.org/10.1016/j.aeue.2018.10.007>.
- [52] Ronald Linn Rivest, "The MD5 message digest algorithm," in *Requests for Comments (RFC)*, 1321, 1992. DOI: <https://doi.org/10.17487/RFC1321>.
- [53] Mohamad H. Hassoun, "Fundamentals of artificial neural networks," 1st edition, *The MIT Press*, 1995.
- [54] Kyle Madden, Jim Harkin, Liam McDaid and Chris Nugent, "Adding security to networks-on-chip using neural networks," in *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2018, pp. 1299-1306. DOI: <https://doi.org/10.1109/SSCI.2018.8628832>.
- [55] Sriram R. Vangal, Jason Howard, Gregory Ruhl, Saurabh Dighe, Howard Wilson, James Tschanz, David Finan, Arvind Singh, Tiju Jacob, Shailendra Jain, Vasantha Erraguntla, Clark Roberts, Yatin Hoskote, Nitin Borkar and Shekhar Borkar, "An 80-tile sub-100-w teraflops processor in 65-nm CMOS," in *IEEE Journal of Solid-State Circuits*, 43(1), 2008, pp. 29-41. DOI: <https://doi.org/10.1109/JSSC.2007.910957>.
- [56] Nathan Binkert, Bradford Beckmann, Gabriel Black, Steven K. Reinhardt, Ali Saidi, Arkaprava Basu, Joel Hestness, Derek R. Hower, Tushar Krishna, Somayeh Sardashti, Rathijit Sen, Korey Sewell, Muhammad Shoaib, Nilay Vaish, Mark D. Hill and David A. Wood, "The gem5 simulator," in *SIGARCH Computer Architecture News*, 39(2), 2011, pp. 1-7. DOI: <https://doi.org/10.1145/2024716.2024718>.
- [57] Nan Jiang, Daniel U. Becker, George Michelogiannakis, James Balfour, Brian Towles, D. E. Shaw, John

- Kim and William J. Dally, "A detailed and flexible cycle-accurate network-on-chip simulator," in *Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, 2013, pp. 86-96. DOI: <https://doi.org/10.1109/ISPASS.2013.6557149>.
- [58] Christian Bienia, Sanjeev Kumar, Jaswinder Pal Singh and Kai Li, "The PARSEC benchmark suite: characterization and architectural implications," in *Proceedings of International Conference on Parallel Architectures and Compilation Techniques (PACT)*, 2008, pp. 72-81. DOI: <https://doi.org/10.1145/1454115.1454128>.
- [59] Chen Sun, Chia-Hsin Owen Chen, George Kurian, Lan Wei, Jason Miller, Anant Agarwal, Li-Shiuan Peh and Vladimir Stojanovic, "DSENT - a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *Proceedings of ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, 2012, pp. 201-210. DOI: <https://doi.org/10.1109/NOCS.2012.31>.
- [60] Ruiqiang Ji, Jiang Xu and Lin Yang, "Five-port optical router based on microring switches for photonic networks-on-chip," in *IEEE Photonics Technology Letters*, Vol. 25, No. 5, 2013, pp. 492-495. DOI: <https://doi.org/10.1109/LPT.2013.2243427>.
- [61] Nikola K. Kasabov, "NeuCube: A spiking neural network architecture for mapping, learning and understanding of spatio-temporal brain data," in *Neural Networks*, vol. 52, 2014, pp. 62-76. DOI: <https://doi.org/10.1016/j.neunet.2014.01.006>.
- [62] Jingcao Hu and R. Marculescu, "DyAD - smart routing for networks-on-chip," in *Proceedings of ACM/IEEE Design Automation Conference (DAC)*, 2004, pp. 260-263. DOI: <https://doi.org/10.1145/996566.996638>.