# ニューラルネットワークをベースとした IoT デバイス指向型認証モジュールとその評価

野崎佑典[1]　吉川雅弥[1]

**概要：**人工知能技術は，幅広い分野で使用されており，近年ではエッジコンピューティングでの利用が期待されている．一方で，これらの IoT デバイスに対する不正な攻撃の危険性が報告されており，認証などのセキュリティ対策を施すことは非常に重要である．しかし，IoT デバイスでは利用できる回路規模に制約があるため，人工知能技術であるニューラルネットワーク（Neural Network：NN）や認証技術である Physical Unclonable Function (PUF)の専用ハードウェアをそれぞれ実装することは難しい．そこで本研究では，NN と認証技術を併用させた，新たな認証モジュール（PUF）を提案する．提案 PUF では，デバイス固有の製造ばらつきにより生じる NN の計算時間の違いを抽出し，これを認証に利用する．そして，Field Programmable Gate Array (FPGA)を用いた評価実験では，安定性や差異性，ユニーク性などの代表的な PUF の性能指標について検討し，提案 PUF の有効性について定量的に評価する．

# IoT Device Oriented Neural Network Based Authentication Module and Its Evaluation

YUSUKE NOZAKI[1]　MASAYA YOSHIKAWA[1]

## 1. はじめに

Artificial Intelligence（AI）技術は著しく発展しており，Internet of Things（IoT）などの組込み機器（エッジ）でも利用されている [1]．エッジコンピューティングでは，大量のデータを用いた解析は，計算能力の高いクラウド側のマシンを利用する．そして，即時性の高い判断が求められるエッジ側では，主に推論のみが行われる [1]．一方で，エッジで利用される IoT デバイスは，外部のネットワークと接続されているため，認証などのセキュリティ対策が重要である．認証が行われない場合，外部からの乗っ取りや不正な制御コマンドにより，深刻な事故が発生する危険性がある [2].

そのため，認証技術として Physical Unclonable Function（PUF）が注目されている [3]–[7]．PUF は半導体の製造ばらつきを認証に利用する技術である．半導体の製造ばらつきは人工的に制御することが難しいため，PUF 回路を物理的に複製することは困難である．一方で，エッジデバイスで推論と認証を実現するためには，ニューラルネットワーク（Neural Network：NN）と PUF の専用ハードウェアがそれぞれ必要となる．しかし，IoT デバイスでは利用できる回路規模に制約があるため，それぞれの専用ハードウェアを実装することは難しい．

そこで本研究では，IoT デバイスで利用される NN と認証機能を併用させた新たな認証モジュール（PUF）を提案する．提案 PUF では，NN と認証用回路を併用させることで，回路規模を削減する．具体的には，NN の入力層から出力層までの計算時間が，製造ばらつきによってわずかに異なることに着目し，これを認証に利用する．また，実際の実装においては，広く使用されている Field Programmable Gate Array（FPGA）を指向した NN の実装方式を導入する．具体的には，FPGA の Look Up Table (LUT)に適した 2 値の NN 実装方式を用いる．さらに，実デバイスによる評価実験により，提案 PUF の有効性について検証する．

## 2. 準備

### 2.1 Physical Unclonable Function

PUF は半導体の製造ばらつきを抽出し，そのデバイス固有の ID を生成する技術である．PUF の概要を図 1 に示す．図 1 に示すように，PUF はチャレンジと呼ばれる入力が与えられたとき，このチャレンジに対するレスポンスを出力

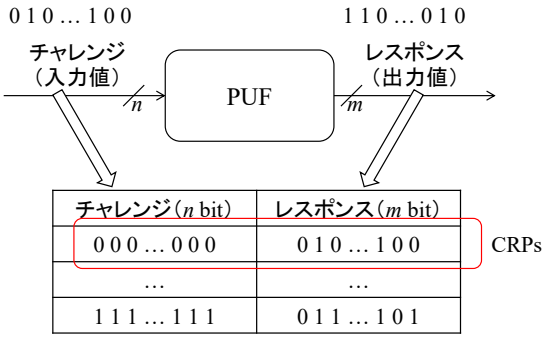



図 1　PUF の概要

Figure 1　Outline of PUF.

1 名城大学
Meijo University

する回路である．PUF ではこの操作を繰り返し行い，チャレンジとレスポンスのペア（Challenge and Response Pairs：CRPs）を収集し，データベース上に登録する．そして，認証時には予め収集した CRPs と，取得した CRPs が一致するかによって認証を行う．また，これまでにアービターPUF [3]やリングオシレータ PUF (RO PUF) [4]，SRAM PUF [5] など数多くの PUF が提案されている．

### 2.2 ニューラルネットワーク

NN は，人の脳内で行われている計算を数式でモデル化したものである [8]．NN は主に，入力層，中間層，出力層で構成する．また，NN では主に学習と推論が行われる．学習では，NN に対して入力が与えられ，NN を構成する各ニューロンの重みの更新が行われる．そして推論では，ある入力に対する出力を推定する．エッジコンピューティングにおいては，大量のデータを利用した学習は計算能力の高いクラウド側のマシンを使用する．そして，エッジ側では主に推論のみを行う．

また，これまでに NN のハードウェア実装に関する研究はいくつか報告されている [9][10]．一方で，これらの実装において，PUF などの認証技術を併用させた研究は筆者らの知る限り報告されていない．

## 3. 提案手法

### 3.1 提案手法の概要

本研究では，NN と認証技術を併用させた新たな認証モジュール（PUF）を提案する．提案 PUF の概要を図 2 に示す．提案 PUF では，システムのディペンダビリティを向上させるため，推論に用いる NN を二重化回路として実装する．具体的には，同じ数のニューロン，層，重みを与えた NN を 2 つ実装する（NN A と NN B）．このとき，2 つの NN（NN A と NN B）は同一の層，重みであるため，同じ入力（チャレンジ）を与えた場合，計算結果は等しくなり，その計算時間は理想的には同じになる．しかし，実際には半導体の製造ばらつきにより，2 つの NN の計算時間には

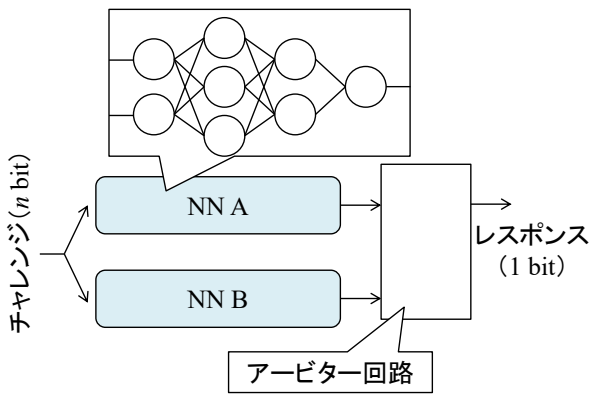



図 2　提案 PUF の概要

Figure 2　Outline of the proposed PUF.

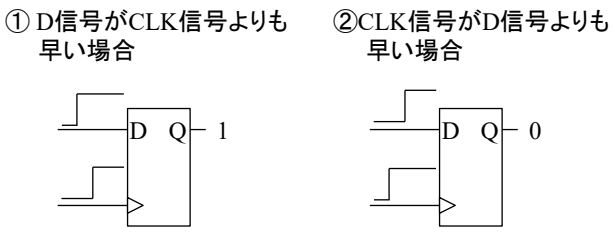



図 3　アービター回路でのレスポンス生成

Figure 3　Response generation in arbiter circuit.

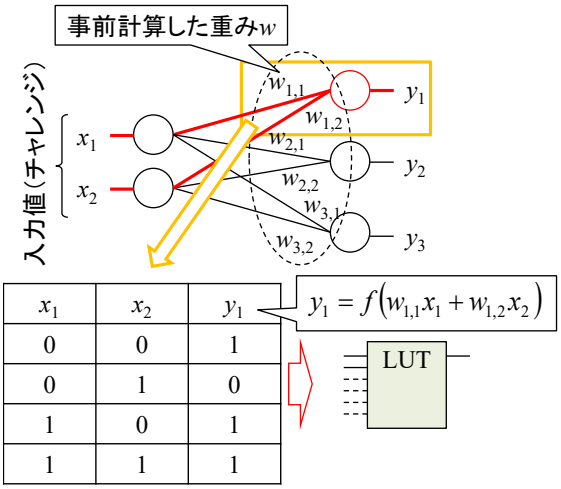
差が生じる．提案 PUF ではこの計算時間の違いに着目し，これをレスポンスの生成に利用する．具体的には，2 つの NN の出力をアービター回路に接続し，計算時間の差を抽出する．アービター回路に D-FF を利用する場合の例を図 3 に示す．図 3 に示すように，D に入力される信号の方が早い場合，出力（レスポンス）は 1 に，CLK に入力される信号の方が早い場合，レスポンスは 0 となる．

以上のように提案 PUF では，NN の計算時間の差をレスポンスとして利用することで，NN と PUF の機能を併用させており，回路規模を削減することができる．

### 3.2 LUT を指向した 2 値化 NN 実装方式

本研究では，FPGA の LUT の構造を考慮した 2 値の NN 実装を導入する．提案 PUF では，識別器としての NN をハードウェアで実装する．そのために，NN を構成する各ニューロンをそれぞれ FPGA の LUT を用いて実装する．このとき，実装する NN の各ニューロンの重みは，Chainer [11] などのソフトウェアで予め事前計算して算出したものを利用する．

実装方式の概要を図 4 に示す．まず実装する NN において，入力層では値の入力のみを行うため，中間層から実装する．この計算式を式(1)に示す．ただし，$x$ はニューロンに対する入力，$y$ はニューロンの出力，$w_i$ はニューロンの重み，$f(x)$は活性化関数である．




| $x_1$ | $x_2$ | $y_1$ |
|---|---|---|
| 0 | 0 | 1 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |
| 1 | 1 | 1 |

図 4　LUT を指向した 2 値化 NN 実装方式

Figure 4　LUT oriented NN implementation method.

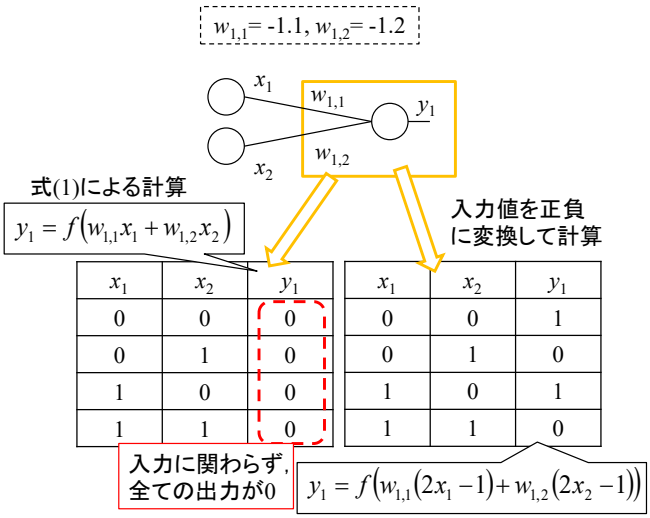



図 5　正負を考慮した実装

Figure 5　Implementation considering the value of positive and negative.

$$y_i = f\left(\sum_{j=1}^{n} w_{i,j} x_j\right) \tag{1}$$

また，$f(x)$は以下の式(2)による計算を行う．

$$f(x) = \begin{cases} 1 \text{ for } x \geq 0 \\ 0 \text{ for } x < 0 \end{cases} \tag{2}$$

以上の計算により得られた 2 値の計算結果から真理値表を作成し，LUT にニューロンを実装する．

ここで 3 層目以降（中間層の 2 層目以降）では，上記の計算方法で NN を実装する場合，入力値に関わらず全ての出力が 0 となる場合がある．具体的に，図 5 を用いて説明する．図 5 は出力層における計算例である．各重み $w$ をそれぞれ，$w_{1,1}$ = -1.1，$w_{1,2}$ = -1.2 とすると，各出力は図 5 に示すように入力に関わらず全て 0 となる．そこで，提案実装手法では，「0」「1」の 2 値の入力値を「-1」「+1」の正負に変換し，計算を行う．このように変換処理を導入することで，LUT において 2 値での実装を行う場合でも，計算結果を反映させることができる．この計算式を式(3)に示す．

$$y_i = f\left(\sum_{j=1}^{n} w_{i,j}(2x_j - 1)\right) \tag{3}$$

## 4.　評価実験

### 4.1　実装ニューラルネットワーク

ここでは，本研究で実装する NN について説明する．本研究では, 提案 PUF として 16bit 入力に対する XOR 演算を行う NN を実装する．この NN は，2 入力 XOR 演算を行う NN を 1 つの基本ユニットとして構成する. まず, 提案 PUF を構成する基本ユニットを図 6 に示す. 図 6 に示すように，この NN は 4 層の NN であり，入力層，2 層の中間層，出力層で構成する．

次に，この基本ユニットの LUT を用いた実装について，図 7 を用いて説明する．図 7 に示すように，入力層を除いた各ニューロンは, それぞれ 1 つの LUT を用いて実装する．ここで中間層の 1 層目のニューロンはそれぞれ 2 入力であるため，LUT では 2 入力 1 出力の論理で実装する．

また，中間層の 2 層目と出力層のニューロンはそれぞれ 3 入力であるため，LUT では 3 入力 1 出力の論理で実装する．

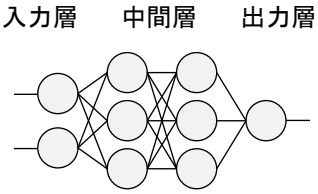



図 6　提案 PUF を構成する基本ユニット

Figure 6　Basic unit of the proposed PUF.

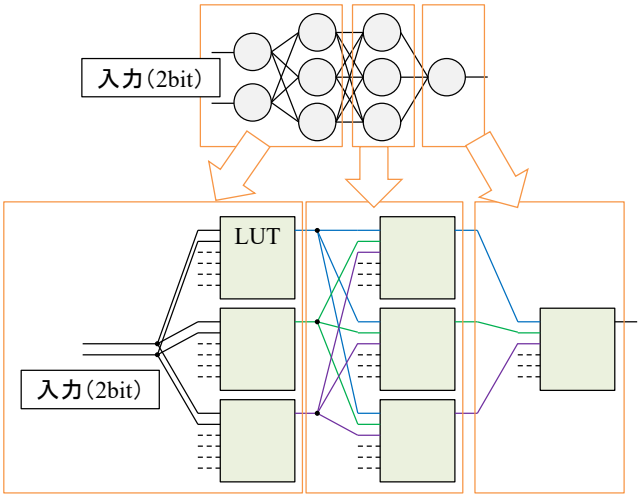



図 7　基本ユニットの LUT での実装

Figure 7　Implementation of the basic unit using LUT.

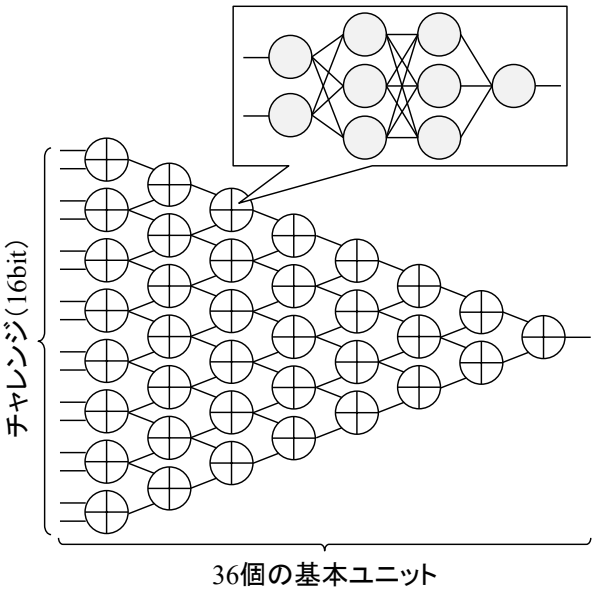



図 8　実装するニューラルネットワーク
Figure 8　Implemented neural network.

そして，この基本ユニットを用いて，提案 PUF を構成する NN（16bit 入力の XOR）を実装する．実装する NN を図 8 に示す．図 8 に示すように，実装する NN は 36 個の基本ユニットを用いて構成する．実際には，提案 PUF は図 8 の NN を 2 つ実装し，2 つの NN の計算時間の差をレスポンスに使用する．そのため，合計で 72 個の基本ユニットを実装する．

### 4.2 実験環境

評価実験では，提案 PUF を FPGA ボード（SASEBO-GII）[12] に実装した．実験環境を図 9 と表 1 に示す．また，重みの事前計算には Chainer [11]を使用した．実装に関して，提案 PUF のフロアプランを図 10 に示す．図 10 の左側は実装した提案 PUF の全体図を，図 10 の右側は最終段における基本ユニットの拡大図をそれぞれ示している．

次に，PUF の性能評価では，3 種類の SASEBO-GII ボード（ボード A，ボード B，ボード C）にそれぞれ提案 PUF を実装し，代表的な性能指標 [13]である安定性，差異性，ユニーク性について評価した．またそれぞれの性能指標は，同じデバイスに対して同じチャレンジを入力したときの ID 間のハミング距離（Same Challenge Intra Hamming Distance：SC Intra-HD），同じデバイスに対して異なるチャレンジを入力したときの ID 間のハミング距離（Different Challenge Intra-HD：DC Intra-HD），異なるデバイスに対して同じチャレンジを入力したときの ID 間のハミング距離（Same Challenge Inter-HD：SC Inter-HD）で評価することができる [14]．PUF の性能指標の概要を以下に示す．

- 安定性：同一の PUF に同じチャンレジを与えたとき，同じレスポンスを出力する性質．SC Intra-HD が 0 に近い程，安定性が高い．
- 差異性：同じ PUF に対して異なるチャレンジを入力したとき，異なるレスポンスを出力する性質．DC Intra-HD が ID 長の半分に近い程，差異性が高い．
- ユニーク性：異なる PUF に対して同じチャレンジを与えたとき，異なるレスポンスを出力する性質．SC Inter-HD が ID 長の半分に近い程，ユニーク性が高い．

実験では，128bit のレスポンスを 1 つの ID として使用した（ID 長は 128bit）．また，同じチャレンジに対する ID を 100 個，異なるチャレンジに対する ID を 100 個用意し，合計で 200 個の ID（CRPs）を実験に使用した．

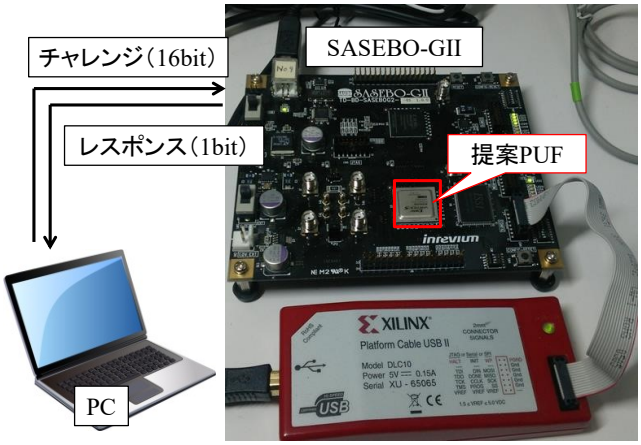



図 9　評価システム
Figure 9　Evaluation system.

表 1　実験環境
Table 1　Experimental condition.

| | |
|---|---|
| PUF | 提案 PUF |
| NN の推論機能 | XOR 演算 |
| チャンレンジ | 16bit |
| レスポンス | 1bit |
| FPGA ボード | SASEBO-GII |
| FPGA | Virtex-5 XC5VLX30 |
| 開発ツール | Xilinx ISE Design Suite 14.7 |
| 配置配線 | Xilinx PlanAhead v14.7 |
| PUF ID (CRPs) の数 | 200 |
| 重みの事前計算 | Chainer [11] |

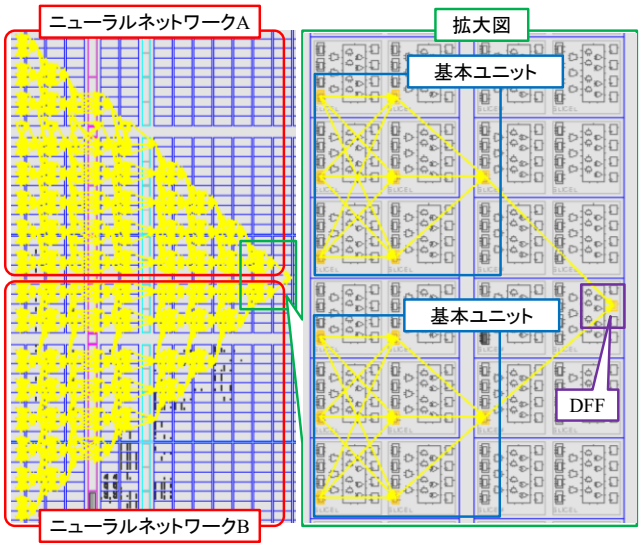



図 10　フロアプラン

Figure 10　Floorplan.

### 4.3 実験結果

まず，提案 PUF の安定性，差異性，ユニーク性について評価した．各ボードでの実験結果を図 11 と図 12 に示す．図の横軸はハミング距離を，縦軸はその出現頻度を示している．また，各実験結果の平均の結果を表 2 と表 3 に示す．安定性については，図 11 と表 2 に示すように SC Inter-HD は 0 に近く，安定性が高いことが確認できる．また，差異性については，DC Intra-HD が ID 長の半分である 64 に近い値となっており，差異性が高いことがわかる．最後にユニーク性については，図 12 と表 3 に示すように，平均して約 43 であることが確認できる．

表 2　SC Intra-HD と DC Intra-HD の平均

Table 2　Average of SC Intra-HD and DC Intra-HD.

| | ボード A | ボード B | ボード C |
|---|---|---|---|
| SC Intra-HD | 6.454 | 5.740 | 6.010 |
| DC Intra-HD | 61.29 | 64.78 | 59.44 |

表 3　SC Intra-HD と DC Intra-HD の平均

Table 3　Average of SC Intra-HD and DC Intra-HD.

| | ボード A | ボード B | ボード C |
|---|---|---|---|
| ボード A | — | 44.73 | 40.43 |
| ボード B | 44.73 | — | 44.72 |
| ボード C | 40.43 | 44.72 | — |

次に認証に関して，提案 PUF における本人拒否率（False Rejection Rate : FRR）と他人受入率（False Acceptance Rate : FAR）について検証した．ここで，FRR は本物のデバイスを誤って偽物と判定する確率，FAR は偽物のデバイスを誤って本物と判定する確率である．これらの FRR と FAR は，SC Intra-HD と SC Inter-HD を用いて評価することができる [14]．具体的には，SC Intra-HD の分布に SC Inter-HD が重なっている割合を FRR，SC Inter-HD の分布に SC Intra-HD が重なっている割合を FAR として算出することができる．各ボードにおける SC Intra-HD と SC Inter-HD をまとめた結

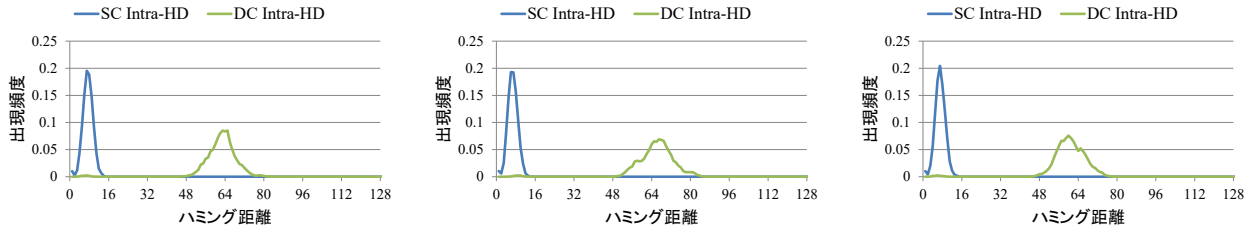



図 11　安定性と差異性の評価（左からボード A，ボード B，ボード C）

Figure 11　Experimental results for steadiness and diffuseness in each device.

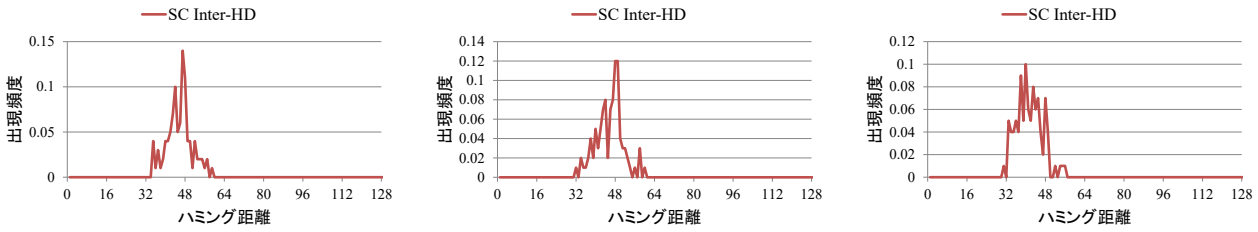



図 12　ユニーク性の評価（左からボード AB 間，ボード BC 間，ボード AC 間）

Figure 12　Experimental results for uniqueness in each device.

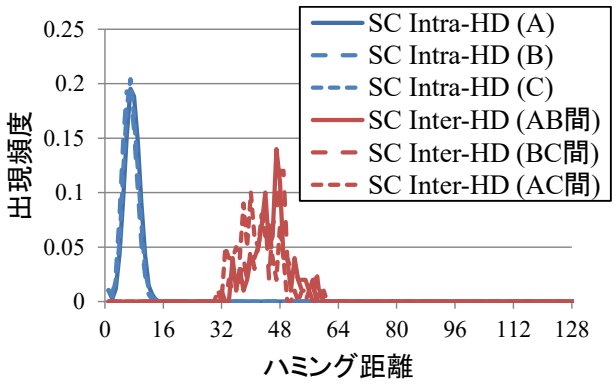



図 13　SC Intra-HD と SC Inter-HD の比較

Figure 13　Comparison of SC Intra-HD and SC Inter-HD.

表 4　回路規模の比較

Table 4　Comparison of circuit area.

| | LUT 数 | レジスタ数 |
|---|---|---|
| 提案 PUF | 532 | 17 |
| NN とアービターPUF | 598 | 21 |
| NN と RO PUF | 2,893 | 90 |

果を図 13 に示す．図 13 の青線は SC Intra-HD を，赤線は SC Inter-HD をそれぞれ示している．図 13 に示すように，提案 PUF では SC Intra-HD と SC Inter-HD の分布が重なり合う箇所は見られず，FRR と FAR は低いことが分かる．また，分布の境界であるハミング距離 18 を閾値とすることで，安定して認証を行うことができると考えられる．

最後に，回路規模について比較した．この比較では，提案 PUF と，NN とアービターPUF [3]，NN と RO PUF [4]をそれぞれ FPGA 実装した．ここで，アービターPUF は 16 段セレクタユニットのものを，RO PUF は 256 個の RO をもつものを実装した．比較結果を表 4 に示す．表 4 に示すように，提案 PUF は他の PUF と比較して，小回路規模で実装可能であり，有効であることが分かる．これは提案 PUF では，NN と PUF の機能を，回路を併用させることで実現しているためだと考えられる．

## 5. まとめ

本研究では，エッジ向けの NN と認証機能を併用させた新たな PUF を提案した．提案 PUF では，製造ばらつきにより生じる NN の計算時間の違いに着目し，これを ID の生成に利用する．そして，NN の回路と PUF 回路を併用させることで，小回路規模での実装を実現させる．また，FPGA を用いた実験では，PUF の性能に関して，安定性と差異性，ユニーク性について評価し，回路規模に関してはアービターPUF と RO PUF との比較を行い，提案 PUF の有効性を実証した．

今後は，NN の層を変化させた場合や，異なる学習パターンの NN での評価などを行う予定である．

## 参考文献


[1] Yu, W., Liang, F., He, X., Hatcher, W. G., Lu, C., Lin, J. and Yang, X.: A Survey on the Edge Computing for the Internet of Thingsm, *IEEE Access*, Vol.6, pp.6900–6919, (2018).

[2] Bitdefender: DDoS attack by massive IoT botnet takes down Krebs on Security, HOT for security powered by Bitdefender, available from <https://hotforsecurity.bitdefender.com/blog/ddos-attack-by-massive-iot-botnet-takes-down-krebs-on-security-16742.html> (参照 2018-05-14)

[3] Lee, J. W., Lim, D., Gassend, B., Suh, G. E., Dijk, M. V. and Debadas, S.: A Technique to Build a Secret Key in Integrated Circuits for Identification and Authentication Applications, Proc. *IEEE VLSI Circuits Symposium*, pp.176–179, (2004).

[4] Suh, G. E. and Devadas, S.: Physical Unclonable Functions for Device Authentication and Secret Key Generation, *Proc. 44th ACM/IEEE Design Automation Conference* (*DAC'07*), pp.9–14, (2007).

[5] Guajardo, J., Kumar, S. S., Schrijen, G. J. and Tuyls, P.: FPGA Intrinsic PUFs and Their Use for IP Protection, *Proc. 9th Int. Workshop on Cryptographic Hardware and Embedded Systems* (*CHES 2007*), LNCS 4272, pp.63–80, Springer-Verlag, (2007).

[6] Hori, Y., Kang, H., Katashita, T. and Satoh, A.: Pseudo-LFSR PUF: A compact, efficient and reliable physical unclonable function, *Proc. 7th Int. Conf. ReConFigurable Computing and FPGAs* (*ReConFig'11*), pp.223–228, (2011).

[7] Shimizu, K. and Suzuki, D.: Glitch PUF: Extracting Information from Usually Unwanted Glitches, *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, Vol.E95-A, No.1, pp.223–233, (2012).

[8] Zhang, G. P.: Neural Networks for Classification: A Survey, *IEEE Trans. Systems, Man, and Cybernetics*, Vol.30, No.4, pp.451–462, (2000).

[9] Zhao, W., Fu, H., Luk, W., Yu, T., Wang, S., Feng, B., Ma, Y. and Yang, G.: F-CNN: An FPGA-based Framework for Training Convolutional Neural Networks, *Proc. IEEE 27th Int. Conf. Application-specific Systems, Architectures and Processors* (*ASAP 2016*), pp.107–114, (2016).

[10] Nakahara, H., Yonekawa, H., Sasao, T., Iwamoto H. and Motomura, M.: A memory-based realization of a binarized deep convolutional neural network, *Proc. Int. Conf. Field-Programmable Technology* (*FPT 2016*), pp.277–280, (2016).

[11] Chainer, available from <https://chainer.org/> (参照 2018-05-14).

[12] SASEBO-GII, available from <http://satoh.cs.uec.ac.jp/SASEBO/ja/board/sasebo-g2.html> (参照 2018-05-14)

[13] Hori, Y., Yoshida, T., Katashita, T. and Satoh, A.: Quantitative and Statistical Performance Evaluation of Arbiter Physical Unclonable Functions on FPGAs, *Proc. 6th Int. Conf. ReConFigurable Computing and FPGAs* (*ReConFig'10*), pp.298–303, (2010).

[14] Hori, Y., Katashita, T., Kang, H., Satoh, A., Kawamura, S. and Kobara, K.: Evaluation of Physical Unclonable Functions for 28-nm Process Field-Programmable Gate Arrays, *Journal of Information Processing*, Vol.22, No.2, pp.344–356, (2014).