早稲田大学大学院　基幹理工学研究科

博　士　論　文　概　要

論　文　題　目

ハードウェアトロイの検出および無効化
に関する研究

Hardware Trojan Detection and
Invalidation Methods

申　請　者

大屋	優
Masaru	OYA

情報通信・情報理工専攻　情報システム設計研究

2018 年 12 月

近年，デジタル回路設計はコスト削減のために集積回路 (IC)の設計・製造工程を外部に委託するようになってきた．ICの設計・製造が自社内で完結しないようになり，第三者が関わるようになってきたことで，ハードウェアトロイと呼ばれるICに悪意のある機能を持つ回路を挿入される可能性が出てきた．例えば，2008年のIEEE Spectrumによると，2007年9月のシリアへの核関連施設へのイスラエルの空爆が成功したのは，遠隔操作による監視レーダーを停止する「キルスイッチ」が発動したからだと報告されている．他にも，匿名でアメリカ合衆国国防総省に接触した，あるヨーロッパのチップメーカによると，マイクロプロセッサにリモートで発動するキルスイッチを作成することに成功したと書かれている．ICにハードウェアトロイを挿入されることは，特に国防面においてセキュリティ上の懸念となっている．

ハードウェアトロイはICの全ての設計・製造工程において挿入可能ではあるものの，製造工程よりも設計工程の方がハードウェアトロイの製作・挿入の技術的な障壁が低く，容易に挿入できることは知られている．なぜなら，デジタル回路設計の設計工程ではソフトウェアによってICの設計が行われるためである．設計工程でハードウェアトロイを挿入する利点は，ソフトウェアの改竄で悪意のある機能を実現することができるため，チップとして残らず正規のICの機能の一つとして紛れてしまい検出が困難になる点である．一方，製造工程でハードウェアトロイを挿入することを考えると，悪意のある機能を行うチップが形として残ってしまうという欠点がある．このような背景により，設計工程におけるハードウェアトロイに着目した研究開発の重要性と緊急性が高まっているのが現状である．

本論文では，設計工程のハードウェアトロイの検出および無効化を目的とした手法を提案する．本論文で提案する手法により，設計工程におけるICに挿入されたハードウェアトロイの検出および無効化を達成し，IC自体の安全性検証に貢献する．

本論文は6章から構成される．

第1章では本論文の研究背景と目的を説明する．

第2章では，ゲートレベルネットリストを対象としたスコアに基づくハードウェアトロイ識別手法を提案する．これは，ハードウェアトロイに含まれるネット（トロイネットと呼ぶ）の特徴に注目し，トロイネットを検出することでハードウェアトロイを検出する手法である．まず，与えられたゲートレベルのネットリストに含まれる全てのネットにスコアを与える．次に，トロイ要素という3つのパラメータを最大のスコアを持つネットに設定する．最後に，3つのトロイ要素の値とそれらの閾値を比較することで，トロイネットの可能性が最も高いトロイネットを検出する．実際に，Trust-HUBのAbstraction Gate Levelで公開されている全てのゲートレベルネットリストに対して，ハードウェアトロイの有無を識別することに成功した．提案手法を適用する際に要する時間は高々数時間程度である．

第3章では，ゲートレベルネットリストの危険性を表現する指標を提案する．HT rankを設計するにあたり，全てのゲートレベルのTrust-HUBベンチマークを解析し，いくつかのトロイネットの特徴を発見する．解析結果に基づいて，3種類のトロイポイントを開発する．1個目はキャラクタリスティックポイント（C-ポイント），2個目はスケールポイント（S-ポイント），3個目はロケーションポイント（L-ポイント）である．これらのポイントを全てのネットに割り当てることで，ゲートレベルネットリスト内の最大のトロイポイントを持つネット（最大トロイポイントネット）を検出する．最大トロイネットポイントはネットリスト内で最もトロイネットであると疑われるネット (疑トロイネット)であり，最大トロイポイントがHT rankとなる．HT rankの値が高いほど，与えられたネットリストはハードウェアトロイが挿入されている可能性が高いことを示す．一方HT rankの値が低いほど，与えられたネットリストはハードウェアトロイが挿入されていない可能性が高いことを示す．HT rankは他にも設計者にハードウェアトロイと疑われる部分の位置を提供する．実際に HT rankは全てのTrust-HUB, ISCAS85，ISCAS89，ITC99のゲートレベルネットリストに加え，いくつかのOpenCoresゲートレベルネットリスト，そしてハードウェアトロイの挿入されているAESと挿入されていないAESに対して，ハードウェアトロイの有無を分類することに成功した．

第4章では，回路の動的な振る舞いから定常状態を学習することでハードウェアトロイを検出する手法を提案する．この手法は短時間ランダムシミュレーションを通じ，回路の信号遷移の定常状態を学習することでハードウェアトロイを検出する．多くのハードウェアトロイは特別な条件を満たした時にのみ動作するように設計されており，短時間シミュレーションでランダムにテストパターンを与えてもハードウェアトロイは動作しないままである．この状態の信号遷移を定常状態として学習し，長時間ランダムシミュレーションによりハードウェアトロイを動作させることで，ハードウェアトロイを検出することができる．ハードウェアトロイは短時間ランダムシミュレーション中は動作しないで隠ぺいされたままである．これはほとんどのハードウェアトロイが特別な条件を満たした時にのみ動作するよう設計されているからである．したがって，各疑トロイネットにおける定常信号遷移状態を短時間ランダムシミュレーションによって得る．定常信号遷移状態は一定のクロックサイクル間の4次元ベクタによって定義される．4次元ベクタとは，0を保持し続けるクロックサイクル数の合計，1を保持し続けるクロックサイクル数の合計，クロックの立ち上がり数の合計，クロックの立ち下がり数の合計である．ハードウェアトロイは短時間ランダムシミュレーションでは動作しないが，長時間のシミュレーションでは動作する可能性が極めて高い．それは，ハードウェアトロイは稀に動作するように設計されているが，長時間のシミュレーションを行うことでハードウェアトロイが動作することが期待でき

るからである．そこで，各疑トロイネットに対して長時間のシミュレーションを行い，信号遷移ベクタの集合を得る．定常信号遷移ベクタと長時間シミュレーションで得た信号遷移ベクタを比較し，定常信号遷移ベクタと異なる場合は異常状態が存在するため，ハードウェアトロイと判断する．加えて，異常状態が検出された場合は更に分析することで，ハードウェアトロイの機能を推測する．ハードウェアトロイのタイプはトリガとペイロードによってそれぞれ二種類のタイプに分類できる．トリガによる分類は，組み合わせ回路をトリガとするハードウェアトロイと順序回路をトリガとするハードウェアトロイである．ペイロードによる分類は，観測可型トロイと観測不可型トロイである．提案手法は組み合わせ回路・順序回路をトリガとしたハードウェアトロイと観測可型トロイ・観測不可型トロイの全てを検出することができる．実際にTrust-HUBベンチマークに対して，組み合わせ回路をトリガとしたハードウェアトロイ，順序回路をトリガとしたハードウェアトロイ，観測可型トロイ，観測不可型トロイを検出し，ハードウェアトロイの機能を推測することに成功した．

第5章では，ハードウェアトロイの機能を無効化する内部トロイ認証を提案する．本手法は認証を通じて信頼性の低いネットリストに挿入されたハードウェアトロイの機能を無効化する．提案手法のアプローチは，信頼性の低いICを認証モードとノーマルモードで動作させる．認証モードでは，埋め込まれたトロイ認証回路が疑トロイネットの一定クロック間のビットフリッピング回数を監視し，本当のトロイか否かを識別し，それらが本当にトロイであるか否かを判断する．もし認証条件を満足したら，疑トロイネットはノーマルネットと判断するため有効化させる．一方，トロイネットと判断した場合は無効化し，トロイの機能をマスキングする．これにより，ハードウェアトロイが挿入されている信頼性の低いネットリストをノーマルモードでも安全に動作させることができる．頻繁に動作するハードウェアトロイは回路検証で容易に検出されてしまうため，賢いハードウェアトロイは滅多に動作しないように設計されている．このハードウェアトロイの特徴に基づき，提案した内部トロイ認証技術は認証条件により，ノーマルネットとトロイネットの相違を認識し，本当のトロイネットのみを無効化する．このアプローチにより信頼性の低いネットリストにハードウェアトロイが存在しても，安全に動作させることができる．本章の実験では，いくつかのTrust-HUBベンチマークをトレーニングセットとし，認証条件を最適化する．適切な認証条件を設定することにより，信頼性の低いネットリストにあるトロイネットのみを無効化することができる．一般的な128ビットのAESに内部トロイ認証回路を埋め込み，例えハードウェアトロイがあろうとも安全かつ正常に動作することに成功した．面積オーバーヘッドは0.79%だけであり，遅延オーバーヘッドは生じなかった．

第6章では，本論文を総括し，今後の課題を示す．

早稲田大学　博士（工学）　学位申請　研究業績書

氏 名　大屋　優　　印

（2018 年　9 月　現在）

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
a. 論文：査読付き	○ [1] **M. Oya**, M. Yanagisawa and N. Togawa,“Hardware Trojan Detection and Classification based on Logic Testing utilizing Steady State Learning,”IEICE Transactions on Fundamentals, vol. E101-A, no. 12, pp.1--12, Dec. 2018. ○ [2] **M. Oya**, N. Yamashita, T. Okamura, Y. Tsunoo, S. Goto, M. Yanagisawa and N. Togawa,“Hardware-Trojans Rank: Quantitative Evaluation of Security Threats at Gate-Level Netlists by Pattern Matching,”IEICE Transactions on Fundamentals, vol. E99-A, no. 12, pp. 2335--2347, Dec, 2016. ○ [3] **M. Oya**, Y. Shi, N. Yamashita, T. Okamura, Y. Tsunoo, S. Goto, M. Yanagisawa and N. Togawa,“Hardware-Trojans Identifying Method based on Trojan Net Scoring at Gate-Level Netlists,”IEICE Transactions on Fundamentals, vol. E98-A, no. 12, pp. 2537--2546, Dec, 2015.
c. 講演：査読付き国際会議	○ [1] **M. Oya**, Yanagisawa and N. Togawa,“Hardware trojan detection and classification based on steady state learning,” in Proc. International Symposium on On-Line Testing and Robust System Design (IOLTS), pp. 215--220, Jul, 2017. ○ [2] **M. Oya**, Yanagisawa and N. Togawa,“Redesign for untrusted gate-level netlists,” in Proc. International Symposium on On-Line Testing and Robust System Design (IOLTS), pp. 219--220, Jul, 2016. [3] **M. Oya**, K. Hasegawa, Yanagisawa and N. Togawa,“Hardware trojans detection based on steady state learning,” in Proc. Design Automation Conference (DAC) Work in Progress, Jun, 2016. ○ [4] **M. Oya,** Y. Shi, M. Yanagisawa and N. Togawa,“In-situ trojan authentication for invalidating hardware-trojan functions,” in Proc. International Symposium on Quality Electronic Design (ISQED), pp. 152--157, Mar, 2016. [5] **M. Oya**, Y. Shi, M. Yanagisawa and N. Togawa,“Hardware-trojan ranking at gate-level netlists based on trojan net features,” in Proc. Design Automation Conference (DAC) Work in Progress, Jun, 2015. ○ [6] **M. Oya**, Y. Shi, M. Yanagisawa and N. Togawa,“A score-based classification method for identifying hardware-trojans at gate-level netlists,” in Proc. Design, Automation and Test in Europe (DATE), pp. 465--470, Mar, 2015. ○ [7] **M. Oya**, Y. Atobe, Y. Shi, M. Yanagisawa and N. Togawa,“Secure scan design using improved random order and its evaluations,”in Proc. Asia Pacific Conference on Circuits and Systems (APCCAS), pp. 555-558, Nov, 2014. [8] K. Hasegawa, **M. Oya**, Yanagisawa and N. Togawa,“Hardware Trojans classification for gate-level netlists based on machine learning,” in Proc. International Symposium on On-Line Testing and Robust System Design (IOLTS), pp. 203--206, Jul, 2016.

早稲田大学　博士（工学）　学位申請　研究業績書

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
学会発表	[1] **大屋優**, 史又華, 柳澤政生, 戸川望,``悪意ある機能を無効化する内部ハードウェアトロイ認証,'' 電子情報通信学会, 信学技報, VLD2015--124, vol. 115, no. 465, pp. 79--84, 那覇市, 2016 年 3 月 1 日.  [2] **大屋優**, 史又華, 山下哲孝, 岡村利彦, 角尾幸保, 柳澤政生, 戸川望,``回路の動的な振る舞いから定常状態を学習することでハードウェアトロイを検出する手法,'' 情報処理学会, SCIS2016, pp. 1--6, 熊本市, 2016 年 1 月 20 日. [3] **大屋優**, 史又華, 山下哲孝, 岡村利彦, 角尾幸保, 柳澤政生, 戸川望,``ゲートレベルネットリストの脆弱性を表現する指標,'' 電子情報通信学会, 信学技報, VLD2015--59, vol. 115, no. 338, pp. 141--146, 長崎市, 2015 年 12 月 3 日. [4] **大屋優**, 史又華, 柳澤政生, 戸川望, ゲートレベルネットリストを対象としたスコアに基づくハードウェアトロイ識別手法, 電子情報通信学会, 信学技報, VLD2014--182, vol. 114, no. 476, pp. 165--170, 那覇市, 2015 年 3 月 4 日. [5] **大屋優**, 史又華, 柳澤政生, 戸川望, トロイネットの特徴に基づくハードウェアトロイ検出手法, 電子情報通信学会, 信学技報, VLD2014--137, vol. 114, no. 426, pp. 157--162, 横浜市, 2015 年 1 月 30 日. [6] **大屋優**, 史又華, 柳澤政生, 戸川望, ハードウェアトロイに含まれるネットに着目したハードウェアトロイ検出手法, 電子情報通信学会, 信学技報, VLD2014--91, vol. 114, no. 328, pp. 135--140, 別府市, 2014 年 11 月 26 日. [7] **大屋優**, 史又華, 柳澤政生, 戸川望, 改良ランダムオーダースキャンによるセキュアスキャン設計とその評価, 電子情報通信学会, 信学技報, VLD2013--141, vol. 113, no. 454, pp. 43--48, 那覇市, 2014 年 3 月 3 日. [8] 長谷川健人, **大屋優**, 史又華, 柳澤政生, 戸川望,``SVM を利用したネットリストの特徴に基づくハードウェアトロイ識別,'' 電子情報通信学会, 信学技報, VLD2015--58, vol. 115, no. 338, pp. 135--140, 長崎市, 2015 年 12 月 3 日.

早稲田大学　博士（工学）　学位申請　研究業績書

種 類 別	題名、　発表・発行掲載誌名、　発表・発行年月、　連名者（申請者含む）
e.その他 (外部資金)	[1] 科学技術振興機構, **大屋優**(代表者), ``デジタル回路設計における耐ハードウェアトロイ設計仕様の研究開発'', 戦略的創造研究推進事業(ACT-I) 情報と未来, 2018--2019 年度, 3000 千円.  [2] 日本学術振興会, **大屋優**(代表者), ``ハードウェアの不正動作検出および個人情報漏えい対策に関する研究'', 日本学術振興会特別研究員 (DC2), 2018--2019 年度, 1900 千円.
(招待講演)	[1] 戸川望, **大屋優**, ``SoC 設計を取り巻くセキュリティリスクと取り組み,'' CDN Live Japan, 2018 年 7 月 20 日
(受賞)	[1] 特に優れた業績による返還免除（半額免除), 日本学生支援機構, May. 2018. [2] SLDM 優秀論文賞, 情報処理学会 システムと LSI の設計技術研究会, DA シンポジウム 2016, Sep. 2016. [3] SLDM 優秀発表学生賞, 情報処理学会 システムと LSI の設計技術研究会, DA シンポジウム 2016, Sep. 2016. [4] 特に優れた業績による返還免除（全額免除), 日本学生支援機構, May. 2016. [5] テレコムシステム技術学生賞, 電気通信普及財団, Mar. 2016. [6] 専攻賞, 早稲田大学基幹理工学研究科情報理工・情報通信専攻, Mar. 2016. [7] CPSY 研究会優秀若手講演賞, 電子情報通信学会コンピュータシステム研究会, SWoPP2015, Aug. 2015. [8] デザインガイア・ポスター賞, 電子情報通信学会・情報処理学会, デザインガイア 2014, Nov. 2014.
(特許)	[1] （発明者）戸川　望, **大屋　優**,（出願人）学校法人早稲田大学,“集積回路の正常化方法、正常化回路、及び集積回路,”特願 2016-27994,（出願日）2016 年 2 月 17 日. [2] （発明者）戸川　望, **大屋　優**,（出願人）学校法人早稲田大学,“ハードウェアトロイの検出方法, ハードウェアトロイの検出プログラム, およびハードウェアトロイの検出装置,”PCT/JP2015/082223,（出願日）2015 年 11 月 17 日.