様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　４ 年　 ５ 月 １３ 日現在

機関番号：３２６８９
研究種目：基盤研究(B)（一般）
研究期間：2019～2021
課題番号：１９Ｈ０４０８０

研究課題名（和文）機械学習による集積回路設計データ中のハードウェアトロイ検知

研究課題名（英文）Hardware-Trojan Detection for Integrated Circuit Design Data based on Machine Learning

研究代表者
戸川　望（Togawa, Nozomu）

早稲田大学・理工学術院・教授

研究者番号：３０２９８１６１

交付決定額（研究期間全体）：（直接経費）　　13,300,000円

研究成果の概要（和文）：　現在，集積回路設計・製造は低コスト化のため積極的に外注が利用され，外部の悪意ある設計・製造者により悪意ある回路を故意に侵入する「ハードウェアトロイ」が現実的な脅威として指摘されている．特に集積回路設計データに挿入されたハードウェアトロイは，軽微な設計データ改変で重大な事象を引き起こす可能性がある．ハードウェアトロイはその対策技術が開発されると，それを回避するハードウェアトロイが開発される「いたちごっこ」が続いているのが現状である．本研究では，集積回路設計データ中のハードウェアトロイの各種特徴量を積極的に学習することにより，既知・未知のハードウェアトロイを検知する技術を確立した．

研究成果の学術的意義や社会的意義
IoT機器の心臓部は集積回路によって構成されるが，現在の集積回路の開発プロセスは設計・製造コストを削減するため，積極的に外注を利用している．一方，集積回路のサプライチェーンは危険にさらされており，悪意ある設計者・製造者によって，悪意ある回路，すなわちハードウェアトロイの挿入が容易に実現され得る．国家安全，医療，宇宙航空などの分野で使用されるIoT機器にハードウェアトロイが含まれれば大きな危機を招くことになる．
本研究成果は，上記の問題に対して一定の答えを与えるものであり，ハードウェアトロイの危険性がないセキュア集積回路チップの実現により安全安心なIoT機器の構築に大きく寄与するものと考える．

研究成果の概要（英文）：　Recently, as Internet of Things (IoT) devices become widespread, the demand for embedded hardware devices has been increasing. In order to produce embedded hardware devices more inexpensively, the manufacturing bases have been internationalized, and several processes in the IC design and manufacturing steps have been outsourced to third-party vendors. Under the circumstances, a hardware Trojan, which is a malicious function circuit inserted into a hardware device, may be inserted into IC products by the malicious third-party vendors, and therefore the risk of hardware Trojans has arisen. In this research, we have developed a machine-learning-based hardware Trojan detection method to detect known and unknown hardware Trojans effectively and efficiently.

研究分野：集積システム設計

キーワード：ハードウェアトロイ　機械学習　レジスタトランスファレベル　ゲートレベル

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

IoT（Internet of Things）機器の心臓部は集積回路によって構成されるが，現在の集積回路の開発プロセスは設計・製造コストを削減するため，積極的に外注を利用している．例えば，安い賃金で労働している海外企業への設計の発注や，IP コア（IP は Intellectual Property の略）と呼ばれる既設計部品の利用が多用されている．外注によるリソースや購入した外部 IP コアの信頼性は，利用した海外企業等の性善説に基づく一方，集積回路のサプライチェーンは危険にさらされており，悪意ある設計者・製造者によって，悪意ある回路の挿入が容易に実現され得る．このような悪意ある回路部品は「ハードウェアトロイ」と呼ばれ，海外を中心にその危険性が大きく指摘されている．特に国家安全，医療，宇宙航空などの分野で使用される IoT 機器にハードウェアトロイが含まれれば，潜在的に大きな危機を招くことなる．

IoT 機器に組み込まれたハードウェアトロイに関して，米国では国防総省高等研究所を中心に「集積回路の信頼性確保プログラム」を開始している．実際，IoT 機器中のハードウェアトロイの実例として，特定メーカの PC に遠隔から侵入できるバックドア回路が含まれると指摘され，国家機関で当該メーカ製 PC の使用が禁止された事例もある．アイロンに不正チップが発見された事例も報告されている．対象アイロンは WiFi 経由でウィルスを仕掛ける回路が含まれていたと言う．また特定の国で製造されたチップにハードウェアトロイが挿入され別の国に出荷されたと報告されている．これらの実例は IoT 機器に挿入されたハードウェアトロイが現実の脅威として存在することを強く示唆する．

集積回路の設計・製造は仕様から始まり，いくつかの工程を経て最終製品として出荷される．製造工程でハードウェアトロイを挿入するには回路の物理的構造そのものを改変することが必要であり，ハードウェアトロイの可能性・危険性は限定的である．一方，設計工程でハードウェアトロイを挿入するには，電子的に作成された設計データのごく一部を改変するだけで，極めて重大な攻撃が可能となる．実際，集積回路の設計データ中の 0.01%程度の改変で，内部秘密情報を漏洩したり制御装置を誤動作させたりすることも可能である．

製造工程では 1 つのマスクパターンから多くの集積回路チップを製造するのに対し，設計工程ではただ一つの設計データを設計する．つまり製造工程では多数の集積回路チップの一部だけにハードウェアトロイを侵入した不正チップを作り込むことができるのに対し，設計工程ではただ一つの設計データにハードウェアトロイを侵入し不正設計データとなる．これは，製造工程では原理的に「正チップ」と「不正チップ」との双方が存在し，「正チップと異なるチップを発見」することは比較的容易であるのに対して，設計工程では「ただ一つの設計データが不正かどうかを判断する」ことが目標となる．しかも我が国の半導体設計製造メーカは内製の設計ツールの開発研究がほぼなく，米国製設計ツールや東アジア諸国の安価な外部 IP コア・セルライブラリを使わざるを得ない状況にある．我が国こそが設計工程でハードウェアトロイ侵入の危機に直面していると言っていい．

つまり，我が国の集積回路のハードウェアトロイ検知では，設計データ中のハードウェアトロイ検知が特に重要であるが，これは極めて困難な問題点がある．すなわち，原理的に「正データ」は存在せず，ただ一つの設計データにハードウェアトロイが侵入しているかを判断する技術が必要となる．

これまで，我々の研究においてベンチマークレベルで既知ハードウェアトロイについて，これらが設計データ中に含まれていた場合，正しく検知する技術を研究開発してきたが，こうした技術が開発されると「未知」ハードウェアトロイの出現が危ぶまれる．実際，ハードウェアトロイ検知手法に対して，それに対応したハードウェアトロイの開発等の事例が報告され，今後も「いたちごっご」が繰り返される恐れがある．ハードウェアトロイの開発は「いたちごっこ」であり，いかに既知ハードウェアトロイを学習し，未知ハードウェアトロイを検知するかが大きな問題となる．

２．研究の目的

本研究では，レジスタトランスファレベル・論理レベル等の集積回路設計データを対象に，機械学習によるハードウェアトロイの「学習」を利用しハードウェアトロイ識別器を進化，未知ハードウェアトロイを含む設計データ（未知設計データ）について，高い精度で未知設計データ中の「各信号線のトロイ／非トロイを識別」する技術を確立する．

ハードウェアトロイの学習では「ハードウェアトロイ回路の特徴量として何を学習するか」が最大のポイントとなる．研究代表者らはこれまでハードウェアトロイ回路に含まれる信号線を精査し，これらをデータベースとして蓄積しているが，これらの知見のもとハードウェアトロイ識別器ならびにその特徴量を最適化することで，目的を達成するものとする．

３．研究の方法

(1) ハードウェアトロイ

一般にレジスタトランスファレベルの回路設計データは，ハードウェア記述言語（HDL）によって記述され，これが EDA ツール（Electronic Design Automation）によってゲートレベルのネットリストに変換される．ゲートレベルのネットリストとは，論理和，論理積，フリップフロッ

プ，マルチプレクサなどの論理ゲートと論理ゲートを接続する信号線（ネットと呼ばれる，以降ネットと信号線は同じ意味で用いる）から構成される．ここではゲートレベルのネットリストに対し，不正回路としてハードウェアトロイが挿入された場合を想定し，ゲートレベルのネットリスト中からいかにハードウェアトロイを検知するかに焦点を当てる．特に，ネットリストを構成する各ネットを，(1)ハードウェアトロイを構成するネット（トロイネットと呼ぶ）あるいは(2)ハードウェアトロイを構成するネット以外のネット（ノーマルネットと呼ぶ）に分類することを考える．

ハードウェアトロイは，図 1 に示すように，トリガ回路とペイロード回路から構成される．トリガ回路とは，ハードウェアトロイを発現するための条件を構成するものであり，極めて稀な条件によって，後段のペイロード回路をアクティベートする．ペイロード回路は，ハードウェアトロイの機能を発現するものであり，例えば，機密情報の漏洩，余分な電力増加等の不正な動作を引き起こす．

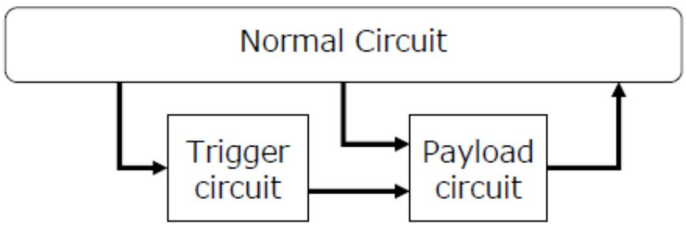

図 1：ハードウェアトロイの構成

(2) ハードウェアトロイの特徴量

上述のようにハードウェアトロイは，極めて稀な条件のもとに動作を行うものが多く，さまざまな構造上の特徴を持つ．例えば，条件を構成するために多段の論理積あるいは論理和ゲートが接続され，不自然な信号線の集中等が見られる．こうした条件のもと，本研究では，ハードウェアトロイを構成する信号線（ネットnと呼ぶ）の特徴量として，まず以下の 11 個の特徴量を見出した（表 1）．

表 1：特徴量 1

番号	特徴量	意味
1	fan_in_4	ネットnの入力側 4 段手前に接続される論理ゲートの数
2	fan_in_5	ネットnの入力側 5 段手前に接続される論理ゲートの数
3	in_flipflop_4	ネットnの入力側 4 段手前に接続されるフリップフロップの数
4	out_flipflop_3	ネットnの出力側 3 段手前に接続されるフリップフロップの数
5	out_flipflop_4	ネットnの出力側 4 段手前に接続されるフリップフロップの数
6	in_loop_4	ネットnの入力側 4 段それぞれでループを構成する数
7	out_loop_5	ネットnの出力側 5 段それぞれでループを構成する数
8	in_nearest_pin	ネットnから最も近いプライマリ入力の段数
9	out_nearest_pout	ネットnから最も近いプライマリ出力の段数
10	out_nearest_flipflop	ネットnから出力側で最も近いフリップフロップの段数
11	out_nearest_multiplexer	ネットnから出力側で最も近いマルチプレクサの段数

表 1 の 11 種類の特徴量は，対象ネットの入力側や出力側のネットを参照することで抽出される．ここでさらにハードウェアトロイのトリガ回路に着目すると，数十本のネットが一本のネットに集約されるピラミッド型の構造を持っていることが判明した．そこで，表 1 に加えて fan_in_uxdy なる 25 個の特徴量を提案した．fan_in_uxdy は，対象ネットより出力側 x 段目のゲートの出力側のネットから見て入力側 y 段目のゲートに接続するネットの数で定義される．x と y は変数を表しており，既存の多数のハードウェアトロイのトリガ回路が最大 5 段であることから，x と y の値を 1 から 5 の間で変化させて組み合わせた 25 個の特徴量となる．

(3) 機械学習モデルの構成

表 1 の 11 個の特徴量と，25 個の fan_in_uxdy 特徴量とを合わせて 36 個の特徴量を用いた学習と識別を考える．予備的な実験を多数通して，ハードウェアトロイの識別には，アンサンブル学習が効果的であることが分かった．そこで機械学習モデルとして，ランダムフォレストを用い

ることにする．ランダムフォレストのハイパーパラメータは，弱学習器の数を 200，不純度の計算を交差エントロピー，木の深さの最大を 13，最小分割数を 2 とした．また，以下の成果に示すように，ハードウェアトロイを構成するトロイネットの数は，回路全体のノーマルネットの数に比べ非常に少ない．そこで，重複する特徴量をもつ学習データを除外した上で，ノーマルネットとトロイネットの数が同等になるように重み付けし，トロイネットを複数回学習することとした．

またネットリストを構成するネットの識別には，交差検証を適用する．今回，以下の結果に示す 24 種類のネットリストを対象に，一個抜き交差検証で識別実験をする．つまり，24 種類のベンチマークのうち，1 種類を未知のネットリストとして識別テストの対象とし，残りの 23 種類をトロイネットの場所が既知のネットリストとして学習する．全ネットリストが識別テストの対象となるよう検証を繰り返して平均のスコアを評価することとした．

表 2：識別結果

分類するネットリスト	TN	FP	FN	TP	TPR	TNR
RS232-T1000	308	1	0	10	100.0	99.68
RS232-T1100	308	1	0	11	100.0	99.68
RS232-T1200	310	0	0	13	100.0	100.0
RS232-T1300	309	0	0	7	100.0	100.0
RS232-T1400	306	0	0	12	100.0	100.0
RS232-T1500	309	2	0	11	100.0	99.36
RS232-T1600	311	1	3	6	66.67	99.68
s15850-T100	2419	1	16	10	38.46	99.96
s35932-T100	6409	0	12	1	7.69	100.0
s35932-T200	6405	0	11	1	8.33	100.0
s35932-T300	6405	0	12	25	67.57	100.0
s38417-T100	5799	0	10	1	9.09	100.0
s38417-T200	5802	0	10	1	9.09	100.0
s38417-T300	5801	0	0	44	100.0	100.0
s38584-T100	7330	14	17	1	5.56	99.81
s38584-T200	7327	17	104	22	17.46	99.77
s38584-T300	7311	34	1134	9	0.79	99.54
EthernetMAC10GE-T700	102969	0	0	12	100.0	100.0
EthernetMAC10GE-T710	102969	0	0	12	100.0	100.0
EthernetMAC10GE-T720	102969	0	0	12	100.0	100.0
EthernetMAC10GE-T730	102969	0	3	9	75.0	100.0
B19-T100	70626	23	0	96	100.0	99.97
B19-T200	70626	23	0	96	100.0	99.97
wb_conmax-T100	22185	1	10	1	9.09	100.0
Average	–	–	–	–	63.1	99.9

4．研究成果

評価結果を表 2 に示す．ここで，True Negative（TN）は正しくノーマルネットとして分類されたノーマルネットの数を表す．False Positive（FP）は誤ってトロイネットとして分類されたノーマルネットの数を表す．False Negative（FN）は誤ってノーマルネットとして分類されたトロイネットの数を表す．True Positive（TP）は正しくトロイネットとして分類されたトロイネットの数を表す．True Positive Rate（TPR）はトロイネットのうち，正しくトロイネットとして分類されたものの割合を示し，TP/(TP+FN) で定義される．True Negative Rate（TNR）はノーマルネットのうち，正しくノーマルネットとして分類されたものの割合を示し，TN/(TN+FP) 定義される．識別実験の結果，平均 TPR が 63.1%，平均 TNR が 99.9%なる結果得た．12 個のネットリストにおいて TPR が 100%を達成し，13 個のネットリストにおいて TNR が 100% を達成した．

評価結果から特筆すべきは，FP の値が極めて小さく，全信号線のうち FP となる信号線は 0.5%未満である．本研究によってトロイネットと判定されたもの（TP）は，ほぼ真のトロイネットとなっている．またハードウェアトロイを構成する回路から，少なくとも 1 つ以上のトロイネットを検出している．

５．主な発表論文等

〔雑誌論文〕　計12件（うち査読付論文　12件／うち国際共著　0件／うちオープンアクセス　1件）

１．著者名 Tatsuki Kurihara and Nozomu Togawa	４．巻 E105.A
２．論文標題 Hardware-Trojan Detection based on the Structural Features of Trojan Circuits Using Random Forests	５．発行年 2022年
３．雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	６．最初と最後の頁 NA
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2021EAP1091	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Kazunari Takasaki, Ryoichi Kida, and Nozomu Togawa	４．巻 E104.A
２．論文標題 An Anomalous Behavior Detection Method Utilizing Extracted Application-Specific Power Behaviors	５．発行年 2021年
３．雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	６．最初と最後の頁 1555～1565
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2020KEP0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Kento Hasegawa, Tomotaka Inoue, and Nozomu Togawa	４．巻 E104.A
２．論文標題 A Two-Stage Hardware Trojan Detection Method Considering the Trojan Probability of Neighbor Nets	５．発行年 2021年
３．雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	６．最初と最後の頁 1516～1525
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2020KEP0005	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 Tatsuki Kurihara and Nozomu Togawa	４．巻 NA
２．論文標題 Hardware-Trojan Classification based on the Structure of Trigger Circuits Utilizing Random Forests	５．発行年 2021年
３．雑誌名 2021 IEEE 27th International Symposium on On-Line Testing and Robust System Design (IOLTS)	６．最初と最後の頁 NA
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IOLTS52814.2021.9486700	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Kazunari Takasaki, Ryoichi Kida, and Nozomu Togawa	4．巻 NA
2．論文標題 An Anomalous Behavior Detection Method Based on Power Analysis Utilizing Steady State Power Waveform Predicted by LSTM	5．発行年 2021年
3．雑誌名 2021 IEEE 27th International Symposium on On-Line Testing and Robust System Design (IOLTS)	6．最初と最後の頁 NA
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IOLTS52814.2021.9486706	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Kento Hasegawa, Masao Yanagisawa, and Nozomu Togawa	4．巻 E103.D
2．論文標題 Trojan-Net Classification for Gate-Level Hardware Design Utilizing Boundary Net Structures	5．発行年 2020年
3．雑誌名 IEICE Transactions on Information and Systems	6．最初と最後の頁 1618～1622
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2019ICL0003	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Makoto Nishizawa, Kento Hasegawa, and Nozomu Togawa	4．巻 E103.A
2．論文標題 A Capacitance Measurement Device for Running Hardware Devices and Its Evaluations	5．発行年 2020年
3．雑誌名 IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences	6．最初と最後の頁 1018～1027
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2019KEP0005	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Kurihara Tatsuki, Hasegawa Kento, and Togawa Nozomu	4．巻 NA
2．論文標題 Evaluation on Hardware-Trojan Detection at Gate-Level IP Cores Utilizing Machine Learning Methods	5．発行年 2020年
3．雑誌名 2020 IEEE 26th International Symposium on On-Line Testing and Robust System Design (IOLTS)	6．最初と最後の頁 NA
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IOLTS50870.2020.9159740	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Kazunari Takasaki, Kento Hasegawa, Ryoichi Kida, and Nozomu Togawa	4．巻 NA
2．論文標題 An Anomalous Behavior Detection Method for IoT Devices by Extracting Application-Specific Power Behaviors	5．発行年 2020年
3．雑誌名 2020 IEEE 26th International Symposium on On-Line Testing and Robust System Design (IOLTS)	6．最初と最後の頁 NA
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IOLTS50870.2020.9159732	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Ryota Ishikawa, Masashi Tawada, Masao Yanagisawa, and Nozomu Togawa	4．巻 13
2．論文標題 Scalable Stochastic Number Duplicators for Accuracy-flexible Arithmetic Circuit Design	5．発行年 2020年
3．雑誌名 IPSJ Transactions on System LSI Design Methodology	6．最初と最後の頁 10～20
掲載論文のDOI（デジタルオブジェクト識別子） 10.2197/ipsjtsldm.13.10	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

1．著者名 Ryota Ishikawa, Masashi Tawada, Masao Yanagisawa, and Nozomu Togawa	4．巻 NA
2．論文標題 Error Correction System using Stochastic Numbers in Symmetric Channels and Z Channels	5．発行年 2019年
3．雑誌名 2019 26th IEEE International Conference on Electronics, Circuits and Systems	6．最初と最後の頁 578-581
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ICECS46596.2019.8965039	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

1．著者名 Ryota Ishikawa, Masashi Tawada, Masao Yanagisawa, and Nozomu Togawa	4．巻 NA
2．論文標題 Error Correction Coding of Stochastic Numbers Using BER Measurement	5．発行年 2019年
3．雑誌名 2019 IEEE 25th International Symposium on On-Line Testing and Robust System Design (IOLTS)	6．最初と最後の頁 243-246
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/IOLTS.2019.8854450	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計5件（うち招待講演　0件／うち国際学会　0件）

１．発表者名
高崎和成，木田良一，戸川望

２．発表標題
LSTMによる定常状態電力波形生成を利用した消費電力解析にもとづくデバイスの異常動作検知手法

３．学会等名
電子情報通信学会暗号と情報セキュリティシンポジウム2021

４．発表年
2020年

１．発表者名
[17]栗原樹，戸川望

２．発表標題
トリガ回路の性質にもとづく特徴量を利用したランダムフォレストによるハードウェアトロイ識別

３．学会等名
電子情報通信学会技術報告

４．発表年
2020年

１．発表者名
石川遼太，多和田雅師，柳澤政生，戸川望

２．発表標題
ストカスティック数を用いた非対称通信路の誤り訂正

３．学会等名
2020年電子情報通信学会総合大会

４．発表年
2020年

１．発表者名
石川遼太，多和田雅師，柳澤政生，戸川望

２．発表標題
ストカスティック計算におけるステップ関数の実装と評価

３．学会等名
電子情報通信学会技術報告

４．発表年
2019年

１．発表者名 石川遼太，多和田雅師，柳澤政生，戸川望
２．発表標題 ストカスティック数を用いた再帰的分割による解像度解釈可変な画像形式
３．学会等名 電子情報通信学会技術報告
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	木村　晋二 (Kimura Shinji) (20183303)	早稲田大学・理工学術院（情報生産システム研究科・センター）・教授 (32689)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関