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　 ６ 月　 ４ 日現在

機関番号：３２６４４
研究種目：基盤研究(C)（一般）
研究期間：2019～2021
課題番号：１９Ｋ１１９７１

研究課題名（和文）Deep Learningを利用した共通鍵暗号の脆弱性発見手法の研究開発

研究課題名（英文）A Study on Evaluation Method for Symmetric Key Encryption Scheme using Deep Learning

研究代表者
大東　俊博（Ohigashi, Toshihiro）

東海大学・情報通信学部・准教授

研究者番号：８０５０８１２７

交付決定額（研究期間全体）：（直接経費）　　3,300,000円

研究成果の概要（和文）：多層型のニューラルネットワーク（Deep Learning）を利用することで共通鍵暗号の脆弱性を自動的に発見する手法について検討した．まず初めに共通鍵ストリーム暗号に対して暗号解読につながる特徴的な偏りを検出する汎用的な手法を提案している．次に，共通鍵ブロック暗号に対して秘密鍵を使わずに平文から暗号文の作成，または暗号文から平文の推測が可能になる方法について検討した．これらの提案手法は複数の暗号アルゴリズムに対して実行し，有効性を検証している．

研究成果の学術的意義や社会的意義
共通鍵暗号の安全性は既存の全ての暗号解読法への耐性を網羅的に調べることで評価されるが，暗号方式の提案後に設計者の想定していない解読法が発見されて安全性が低下するケースが数多くある．本研究課題では解読法で用いられる暗号関数内の特徴量をDeep Learningにより検出する方法を開発することで，人手では発見が困難であった未知の攻撃手法等を自動的に発見できるようになることを目指している．本研究の成果および更なる発展により暗号設計者の安全性評価の負担の軽減や未知の攻撃に対するリスクの低減が実現できると考えている．

研究成果の概要（英文）：We studied methods for automatically evaluation of vulnerabilities in symmetric key ciphers by using DNN (Deep Neural Networks). First, we proposed a method to detect statistical biases related to the cryptographic attacks in the output of stream ciphers. Second, we propose deep learning-based output prediction attacks for block ciphers, which can realize ciphertext prediction and plaintext recovery in a high probability. In addition, we demonstrate our methods work against several cipher by experimentations.

研究分野：情報セキュリティ

キーワード：共通鍵暗号　ニューラルネットワーク　暗号解読

１．研究開始当初の背景

共通鍵暗号は情報セキュリティの基盤技術であり，データの秘匿や改ざん検出等を提供する．処理速度が高速であることからサーバでのリアルタイム処理やビッグデータの暗号化に適している．また，シンプルな演算で実現可能であることから RFID やセンサーノード等の小型なデバイスでも利用可能であり，IoT に必須の技術である．暗号方式を数学的な問題に帰着させることで安全性を証明する公開鍵暗号と異なり，共通鍵暗号は既存の暗号解読法（以降，単に解読法と呼ぶ）への安全性を証明するアプローチをとっている．

共通鍵暗号に対する解読法では，秘密情報（秘密鍵）および平文と暗号文が暗号関数の中で特徴的な関係を持つことに注目し，その特徴量を利用して秘密鍵や平文を復元する．例えば暗号関数の中で内部メモリのビット情報が有意な確率で線形の関係を持つことを利用する攻撃（線形攻撃），暗号関数の変換を方程式で表現したときに変数の代数次数が高くならないことを利用する攻撃（高階差分攻撃，補間攻撃等），暗号関数内での鍵情報の拡散が不十分という性質を使った攻撃（中間一致攻撃）など特徴量は多岐に渡って検討されている．

研究代表者や研究分担者は共通鍵ストリーム暗号および共通鍵ブロック暗号について暗号評価を目的とする解読法の研究に取り組んできており，特にロシアの前の標準ブロック暗号 GOST への解読法や SSL/TLS におけるストリーム暗号 RC4 への解読法では複数の手法の特性を利用した攻撃手法を採用することで解読能力を引き上げ，対象の暗号方式が安全でないことを初めて示した．このように解読法は高度化してきており，個別の解読方法のみではなく複数の特徴量の組み合わせを検討する必要がある．暗号設計者の負担が年々大きくなり，全ての解読法に対する評価は困難である．また，特徴量は研究者の経験や試行錯誤によって発見されることから，設計者の予期しない未知の特徴量を用いた解読法が残されているリスクも抱えている．もし，暗号方式の入出力から解読につながる特徴量が自動的に抽出できれば，暗号設計の際の安全性評価の負担は減り，より網羅的な安全性評価が可能である．なおかつ設計者の予期しない未知の解読法に対しても事前に発見することが可能であり，暗号方式の安全性を大幅に高めることが可能である．

２．研究の目的

本研究では，暗号解読につながる特徴量を自動で抽出できる手法として，機械学習を用いた計算機探索により暗号解読手法自体を発見する方法を確立することを目的とする．具体的には Deep Learning として知られている Deep Neural Networks（以降，DNN）を用いて暗号関数内の特徴的な振る舞いを発見し，その情報から解読法が成立する特徴量を抽出する．DNN を用いる理由は，暗号関数内には安全性を向上させるために非線形な部品が多く組み込まれており，それらによって作られた暗号文と平文および鍵の関係等を分析するには非線形分離が可能なニューラルネットワークが適しているからと考えたからである．

従来の解読法のための計算機探索では既知の特徴量が暗号関数の中に存在するかどうかを探索しており，現実的に実行可能な探索範囲においてのみの特徴量の有無を保証するものでしかなかった．そのため，暗号方式が提案されたのち，設計者が評価できなかった特徴量による解読法が提案され，暗号の安全性が低下するケースは数多くある．本研究では特定の暗号解読手法に依らず理想的な振る舞いから外れた特徴量を得ることができるため，既存手法では探索不可能であった特徴量の有無を判定できると予想される．また，本研究で開発される技術により暗号設計者が暗号方式を設計する際の安全性評価の負担が大きく軽減される．特に未知の攻撃に対するリスクを下げることができることは，暗号アルゴリズムが情報システムに実装されて広まったあとに重大な欠陥が発見されるリスクを低減できることが期待される．

３．研究の方法

本研究で開発する DNN を用いる安全性解析手法は，未知の脆弱性をも発見できるようにするために，暗号方式の内部構造や探索する特徴量の性質に依らず実行できるような方法を検討する．DNN の学習の際には具体的なアルゴリズムの演算は入力とせず，暗号方式をブラックボックスとした上で得られる入出力で学習し，それが有効であるかの検討の際には解読法の効果によりランダムな探索より効果的であるかという基準で暗号方式の脆弱性を検出することを試みる．本研究課題では，共通鍵暗号方式をストリーム暗号とブロック暗号に分類し，それぞれに関して自動的に脆弱性を検出する方法を提案する．

(1) ストリーム暗号に対するアプローチ

ストリーム暗号は秘密鍵および初期化ベクトルを入力とし，暗号関数により疑似乱数系列を出力させ，その疑似乱数系列と平文を排他的論理和することで暗号文を得る．もし，パケットヘッダなどの情報から平文の一部が類推可能である場合，既知となった平文の部分情報と暗号文の排他的論理和をとることで疑似乱数系列の一部を得ることができる．この疑似乱数系列の一部から疑似乱数系列の他の情報を推測したり，秘密鍵やそれと等価な内部情報を推測することができることは脆弱性であるため，疑似乱数系列の安全性はストリーム暗号の安全性の本質と

なる．さらに，疑似乱数系列自体が偏っている場合も，同じ平文を異なるユーザ秘密鍵で暗号化して送信するようなモデル（broadcast setting）において平文が復元されることにつながるため，そのような性質は望ましくない．

本研究ではストリーム暗号への解読法は疑似乱数系列に何らかの特徴量が発生することに起因して得られることに注目し，疑似乱数系列が真にランダムな系列と区別ができるかという観点で DNN を用いて評価することで脆弱性の有無を調べる方法を提案する．具体的には 2 値分類の分類問題として DNN を用い，評価対象の疑似乱数系列のラベルを 1，真正乱数列のラベルを 0 として学習する．疑似乱数系列が真正乱数列と区別できるような特徴量が存在する場合，DNN は自動的にそれを検出し，暗号関数の不備がある可能性を設計者に提示できる．この際，暗号関数から生じる特徴量は微量であることが多いため，評価フェーズにおけるラベルの推測値（0〜1 の数値によってどのラベルに近いかを示す指標）については，画像などの DNN の分類問題で用いられるような(0.9，0.1)のようなそれぞれのラベルに近い閾値ではなく，微妙な違いを拾うことができる 0.5 前後のものを選ぶことにする．また，ニューラルネットワークの種類は疑似乱数系列の離れた位置のバイト間の関係の特徴も得られることを期待して長期的な時系列データに適している LSTM(Long Short-Term Memory)を採用した．

(2) ブロック暗号に対するアプローチ

ブロック暗号に対しては秘密鍵無しで暗号化や復号が可能な関数を DNN で作成可能かという観点から評価方法を考える．DNN で模倣ができると，任意の平文に対応する暗号文を第 3 者が作成できる暗号文推測攻撃や暗号文から平文を復元する平文回復攻撃が可能となる．特に暗号化関数の鍵無しでの逆関数を実現する補間攻撃という解読法の枠組みがブロック暗号では知られており，本提案手法における平文回復攻撃はアルゴリズムの構造の情報無しで補間攻撃と似たようなことを DNN を使って自動的に行うという位置づけとなる．なお，この方法は回帰問題を DNN を使って解く形に対応する．

提案手法では平文ブロックや暗号文ブロックのバイト間の関係も特徴量の探索対象にするためストリーム暗号に対するアプローチと同様に LSTM を採用している．まず学習フェーズでは既知の情報として平文と暗号文のペアを学習させ，そこで得られたモデルを使って推測フェーズにて（学習フェーズに使っていない）平文/暗号文から暗号文/平文を推測する．推測に成功した確率がランダムな探索をした場合より向上していれば暗号化関数に何らかの特徴量が存在すると判定する．ここで，ブロック長が n ビットのブロック暗号の場合，学習フェーズで 2^x の平文・暗号文対を学習させたとき，推測フェーズでは推測に成功する確率が $1/(2^n - 2^x)$ より高いかで判定していることに注意する．ブロック暗号では平文と暗号文は 1 対 1 対応の置換をしていることから，学習フェーズで得られた情報を使うと，それ以外の候補からランダムに探索するという戦略が自然である．例えば 16 ビットのブロックの場合 2^{16} の平文・暗号文対があるが，学習フェーズで 2^{15} の平文・暗号文対を与えられたとすると，残りの平文や暗号文を攻撃するときにはランダムに残された候補から推測しても $1/2^{15}$ で正しいものを推測できる．これは完全にランダムに 16 ビットを推測する場合の確率の 2 倍である．したがって，DNN による推測が有効であるとみなすためには $1/2^{16}$ ではなく $1/2^{15}$ 以上の確率で推測が成功することを示す必要がある．

提案手法をフルサイズのブロック暗号に適用する場合，平文・暗号文対の全ての空間を利用した実験は現実的でなく，ある特定の条件での結果が得られてしまう可能性がある．そこで，我々はブロックサイズや暗号化部品を縮小した Toy model の暗号を対象に実験により性質を調査し，その傾向がフルバージョンの暗号にも適用されると見なして解析を行うこととした．具体的には 16 ビットブロック程度であれば全平文・暗号文対を発生させることは可能であるため，SPN 構造、Feistel 構造，軽量暗号用の構造などブロック暗号の構成法別に実験によって効果を検証する．なお，DNN による推測の攻撃能力と既存の解読手法である線形解読法や差分解読法の攻撃能力との比較も行う．

４．研究成果

(1) ストリーム暗号に対するアプローチ

3-(1)で説明した提案手法の有効性を検証するため，既知の脆弱性がある RC4 暗号（疑似乱数系列に微量の偏りがある）および線形合同法（疑似乱数系列の各バイトに線形の関係がある）について適用して実験を行った．具体的には，DNN に入力する疑似乱数系列のデータ量を増やした場合に，特徴量の事前情報なしに分類ができるかを確かめた．ここで，真にランダムな乱数を大量に用意することは大きなコストがかかり現実的でないため，比較的良質な疑似乱数が出力されると信じられている HMAC DRBG の出力を真にランダムな乱数とみなして実験をしている．

まずは予備実験として，疑似乱数系列を分類できなかった場合に，評価フェーズのラベルの推測値がどのようになるかを確認した．その結果，単に 0.5 に値が集まるわけでなく，学習の深度によっては 0.5 から少しずれた値になることが分かった．しかしながら，どちらの系列も推測値がそのずれた値になることから，推測値の平均値の差分を見て分類の可否を判断する方法を採用した．Precision や Recall の値を使う一般的な方法をとる可能性もあったが，今回の少し不安定な数値になってしまったため，今回は採用しないこととした．

まず RC4 暗号の疑似乱数系列の実験では最も大きな偏りがある疑似乱数系列の 2 バイト目だけを抽出して，その特徴量が得られるかを確認した．RC4 の 2 バイト目の出力は 2^8 個を超えて集

めて統計的な処理をすると一定の確率で真にランダムな乱数列と識別できることが知られている．実験の結果，訓練フェーズで入力する RC4 の 2 バイトのデータが 2^{19} バイトを越えたあたりから RC4 の疑似乱数列と真正乱数とみなした HMAC DRBG を分類できる推測値になることがわかった．RC4 は疑似乱数系列の 2 バイト目以外にも先頭の 259 バイトの範囲に比較的大きな偏りを持っているため，それらにも同様に効果があるかを確認するのは今後の課題としている．

次に線形合同法と HMAC DRBG の系列の分類実験を行った．線形合同法は漸化式を用いているため，疑似乱数のバイト間に線形の関係があることから LSTM による時系列単位の特徴の抽出が有効に働くと予想する．実験結果から 2^{22} バイトを越えたあたりから HMAC DRBG と分類できることがわかった．線形合同法は明らかな特徴量がある疑似乱数系列として知られていたが，NIST SP 800-22 で示されている汎用的な乱数検定ツールでは，その特徴が検出できないことが知られていた．これは，本提案手法は NIST SP 800-22 で対応できない疑似乱数の特徴量も検出できる可能性があるということを示している．

(2) ブロック暗号に対するアプローチ

ブロック暗号に対する提案手法の評価をするために，複数の Toy Model に対する実験を行った．本実験の評価対象とした Toy model は，16 ビットに縮小したブロック暗号として，SPN 構造は AES と同じ構造で作成した Small AES, Feistel 構造は TWINE の構造から作成した Small TWINE を利用，軽量暗号として PRESENT の Toy Model 版である Small Present を採用している．これらのラウンド関数を変化させながら攻撃可能なラウンド数を評価した．

実験の結果，Small Present は暗号文予測/平文回復ともに最大 4 ラウンドで有効な攻撃ができていることが確認できた．また，線形確率及び差分確率から見積もった 16 ビットブロック版 Small Present の線形解読法や差分解読法の攻撃可能ラウンド数も 4 ラウンドであったため，提案手法は暗号解読法の専門的な知識や暗号化関数の内部構造の情報なしで既存の強力な攻撃手法と同程度の攻撃ができている場合があることがわかった．

Small AES では提案手法の攻撃可能ラウンド数は 1 ラウンドに対して，差分解読法は 2 ラウンド，線形解読法は 3 ラウンドで攻撃可能であった．また Small TWINE は提案手法では攻撃可能ラウンド数は 3 ラウンドに対して，差分解読法は 7 ラウンド，線形解読法は 9 ラウンドで攻撃可能であった．これらは必ずしも線形解読法・差分解読法と同程度の攻撃が可能とは言えないことを示している．しかしながら，線形解読法と差分解読法では特徴量を識別する識別攻撃の到達ラウンド数を議論しており，本提案手法で実現している暗号文予測攻撃/平文回復攻撃は更に強力な攻撃であることには注意が必要である．

さらに，16 ビットブロックからブロック長を 32 ビット，64 ビットに変更した場合の評価も一部実施している．結果として学習フェーズに入力する暗号文・平文対の個数は増加するものの攻撃可能ラウンド数という観点からは同様の傾向がみられることを確認した．これは Toy model の解析結果をフルバージョンの結果に適用する際の信頼性について前向きの可能性を示した結果と言える．

その他，本研究の計画段階では予想がつかなかった結果として，Small Present に対して提案手法を適用した際に，4 ラウンドでの暗号文予測攻撃と平文回復攻撃では成功確率が大きく違うという結果が得られた．Small Present に対して線形解読法や差分解読法によって攻撃した際，暗号化関数を攻撃する場合と復号関数を攻撃する場合で性能の差は出ない．しかしながら，DNN を利用した解読では攻撃ラウンド数は変わらないものの，推測成功確率に明らかな差が出たことは興味深い結果と言える．暗号化関数と復号関数で S-box を使った換字処理とビットの位置を移動させる置換処理の適用順が異なったり，s-box による換字処理も暗号化と復号では逆関数になっていることに注目して，それぞれの部品の順番や種類の変更による検討も行っている．その変更によって成功確率が変化することは確認できたが，どのような原理でそれが起こっているかは現時点では明らかにできておらず，今後の課題とする．

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　2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 Hayato KIMURA, Keita EMURA, Takanori ISOBE, Ryoma ITO, Kazuto OGAWA, and Toshihiro OHIGASHI	４．巻 未定
２．論文標題 Output Prediction Attacks on Block Ciphers using Deep Learning	５．発行年 2022年
３．雑誌名 Proc. 4th International Workshop on Artificial Intelligence and Industrial Internet-of-Things Security (AIoTS 2022)	６．最初と最後の頁 未定
掲載論文のDOI（デジタルオブジェクト識別子） なし	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 Hayato KIMURA, Keita EMURA, Takanori ISOBE, Ryoma ITO, Kazuto OGAWA, and Toshihiro OHIGASHI	４．巻 2021/401
２．論文標題 Output Prediction Attacks on Block Ciphers using Deep Learning	５．発行年 2021年
３．雑誌名 Cryptology ePrint Archive	６．最初と最後の頁 1-27
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
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 Hayato KIMURA, Takanori ISOBE, and Toshihiro OHIGASHI	４．巻 なし
２．論文標題 Neural-Network-Based Pseudo-Random Number Generator Evaluation Tool for Stream Ciphers	５．発行年 2019年
３．雑誌名 Proc. CANDAR Workshops 2019, WCIS 2019	６．最初と最後の頁 333-338
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/CANDARW49138.2019	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7件（うち招待講演　0件／うち国際学会　0件）

１．発表者名 西平侑磨，鈴木達也，渡邉英伸，大東俊博
２．発表標題 Intel SGXを用いたアルゴリズム変換型プロキシ再暗号化システムの実装・評価
３．学会等名 2022年暗号と情報セキュリティシンポジウム(SCIS2022)
４．発表年 2022年

１．発表者名
佐藤友哉，大東俊博，近堂　徹，渡邉英伸，稲村勝樹

２．発表標題
Pub/Subプラットフォームにおける委託型秘密分散法の実装

３．学会等名
情報処理学会第84回全国大会

４．発表年
2022年

１．発表者名
木村隼人，江村恵太，五十部孝典，伊藤竜馬，小川 一人，大東俊博

２．発表標題
深層学習を用いたSPNブロック暗号への出力予測攻撃

３．学会等名
2021年暗号と情報セキュリティシンポジウム(SCIS2021)

４．発表年
2021年

１．発表者名
柳宏之，岡部大地，石橋拓哉，木村隼人，渡邉英伸，大東俊博

２．発表標題
暗号の危殆化に対応可能なオンラインストレージシステムの実装・評価

３．学会等名
2019年電子情報通信学会ソサイエティ大会

４．発表年
2019年

１．発表者名
松本大輝，鈴木達也，木村隼人，大東俊博

２．発表標題
ユーザの利用環境に応じたポリシーを設定可能なパスワードマネージャ

３．学会等名
情報処理学会CSEC研究会（2019年12月）

４．発表年
2019年

１．発表者名 岡部大地，柳宏之，木村隼人，鈴木達也，石橋拓哉，渡邉英伸，大東俊博
２．発表標題 アルゴリズム変換型プロキシ再暗号化を用いたオンラインストレージシステムの実装・評価
３．学会等名 情報処理学会CSEC研究会（2019年12月）
４．発表年 2019年

１．発表者名 鈴木達也，江村恵太，面 和成，大東俊博
２．発表標題 Intel SGXを用いた公開検証可能な関数型暗号の構成と実装評価
３．学会等名 2020年暗号と情報セキュリティシンポジウム(SCIS2020)
４．発表年 2020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	五十部　孝典 (Isobe Takanori) (30785465)	兵庫県立大学・応用情報科学研究科・准教授 (24506)	

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