鍵付きレベル 2 準同型暗号の構成および機械学習による予測への秘密計算の応用

数理情報学専攻 48-196230　前田 勇作
指導教員　縫田 光司 准教授

1　はじめに

暗号とは重要な情報を秘匿することでその安全性を保証する技術であり，現代社会において必要不可欠な技術である．しかし，暗号技術は情報の保護のために重要な一方で，一般に暗号化された情報に対しては計算などの処理をすることができず，利便性が失われてしまうという欠点がある．そこで本研究ではデータを秘匿したまま安全に処理する方法およびその応用について研究した．

第 I 部 鍵付きレベル 2 準同型暗号の構成

2　研究背景

準同型暗号とは，秘密鍵を用いることなく暗号文の間の計算をすることが可能な暗号技術である．しかしその便利な機能の反面，準同型暗号は公開鍵暗号方式で一般に望ましい安全性とされている IND-CCA2 安全性を達成できないことが知られている．

これに対し，Emura らは誰でも演算が可能であるという点が強い安全性を達成するための問題点であったことを指摘し，秘密鍵 sk・公開鍵 pk に加えて演算鍵 ek を設けることで演算の権限を管理する，鍵付き準同型暗号 (KH-PKE) と呼ばれる方式を提案した [5]．また，Emura らは KH-PKE に対し KH-CCA 安全性を定義した．これは，ek を持たない攻撃者に対しては IND-CCA2 安全性を，ek を持つ攻撃者に対しては IND-CCA1 安全性を達成するとともに，ek の漏洩も考慮した安全性である．更に，Emura らは Hash Proof System (HPS) と呼ばれる暗号プリミティブを用いた KH-PKE の一般的な構成も与えている．

KH-PKE の提案以降，これまでにもいくつかの仮定に基づく具体的な方式が提案されてきたが，これらは加算・乗算のうちの一方しかできない，または加算と乗算が両方とも可能だが実用的ではない仮定 ($i\mathcal{O}$) を元に構成している，などといった問題点があった．

3　本研究の成果

本研究では，無制限の加算と 1 回の乗算が可能な KH-PKE である，鍵付きレベル 2 準同型暗号 (Keyed-2LHE) を構成し，これが以下で定義される KH-CCA 安全性を SXDH 仮定の下で満たすことを示した．提案方式は実用的な仮定の下で加算と乗算の双方を行うことのできる初の方式である．

定義 1 (KH-CCA 安全性). KH-PKE 方式が KH-CCA 安全であるとは，任意の PPT 攻撃者 $\mathcal{A}$ に対してその優位度

$$\left|\Pr[\mathsf{pp} \leftarrow \mathsf{ParamGen}(1^\ell); (\mathsf{pk}, \mathsf{sk}, \mathsf{ek}) \leftarrow \mathsf{KeyGen}(\mathsf{pp}); (m_0^*, m_1^*, \mathsf{st}) \leftarrow \mathcal{A}^{\mathcal{O}}(\mathsf{find}, \mathsf{pk}); b \leftarrow \{0,1\}; C^* \leftarrow \mathsf{Enc}(\mathsf{pk}, m_b^*); b' \leftarrow \mathcal{A}^{\mathcal{O}}(\mathsf{guess}, \mathsf{st}, C^*) : b = b'] - \frac{1}{2}\right|$$

が ℓ について negligible であることをいう．

ここで，$\mathcal{O}$ は攻撃者が利用できるオラクルで，演算鍵を返すオラクル RevEK，復号オラクル Dec，演算オラクル Eval の 3 つからなる．

Keyed-2LHE を構成するために，本研究では Attrapadung らによるレベル 2 準同型暗号 [4] に着目した．この方式は現在提案されているレベル 2 準同型暗号の中では最も効率の良い方式である．本研究では [4] の方式を HPS の枠組みに拡張し，これを [5] の一般的な構成に適用することで Keyed-2LHE を構成した．また提案方式は従来の方式にない特徴として，ek の一部のみを渡すことで特定の演算のみを行う権限を与えることが出来る，という特徴を持っている．

4　性能評価

提案方式の各アルゴリズムの演算回数をまとめたのが次の表 1 である．ただし，Op, Exp における 3 つの数字はそれぞれ群 $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T$ における群演算および定数倍の計算回数を，また P はペアリングの計算回数を，DL は離散対数の計算回数を表す．なお，$\mathbb{Z}_p$ 上の演算はこれらよりも高速に計算できるため，省いている．

表 1. 提案方式の各アルゴリズムにおける演算回数

		Op	Exp	P	DL
KeyGen		4, 4, 15	8, 8, 24	0	0
Enc	レベル 1	2, 2, 0	7, 7, 0	0	0
	レベル 2	0, 0, 6	0, 0, 21	0	0
Dec	レベル 1	4, 2, 0	6, 4, 0	0	1
	レベル 2	0, 0, 18	0, 0, 24	0	1
Eval	$\mathsf{Add}^{(1)}$	11, 11, 0	10, 10, 0	0	0
	$\mathsf{Add}^{(2)}$	0, 0, 41	0, 0, 38	0	0
	Mult	2, 2, 19	4, 4, 22	12	0

第 II 部 3 者秘密計算による異種混合モデルを用いた予測の実装評価

5 研究背景

近年，ビッグデータに基づいて高精度の予測が可能な技術である機械学習が様々な分野で利用されている．機械学習の実用における課題の一つとして，解釈性と予測精度の両立がある．この精度と解釈性の両立を実現するモデルの一つとして，異種混合予測モデルの一種である，Hierarchical Mixture of Experts (HME) モデルが挙げられる．これは使用する線形モデルを二分木を用いて選択するモデルで，実用化もされているモデルの一つである [6].

機械学習に関するもう一つの課題としてセキュリティの問題がある．例えば，機械学習を実用する場面で訓練の実行者と予測の実行者が異なる場合は，訓練済みのモデルと予測の際の入力データを互いに明かしたくない状況が考えられる．この他にも，訓練済みのモデルパラメータから学習に用いたデータの情報が洩れる危険性なども指摘されている．

この問題に対する解決策の一つとして，秘密計算技術の応用が考えられる．秘密計算とは，データを秘匿した状態から復元することなくその処理を行うことができる技術である．汎用的かつ効率的な秘密計算の方法としては，秘密分散ベース，準同型暗号ベース，Garbled Circuit ベースの大きく 3 つの方法がある．秘密分散ベースの秘密計算は他の方式と比べて必要な通信回数は多いが，計算量と通信量が小さいため，一定時間当たりの処理量 (スループット) の観点からは秘密分散ベースの秘密計算の方が優位である．

6 研究成果

本研究では秘密分散ベースの秘密計算により，HME モデルでの予測を固定小数点数・浮動小数点数の両方の場合について実装し，実行時間および精度を評価・比較した．その結果，データの次元 3，木の高さ 6 の HME モデルで 10 個の入力データに対し，固定小数点数で 1 秒程度、浮動小数点数で 12 秒程度で予測が実行可能であった．また [6] で扱っているものと同じ 14 次元のデータに対する高さ 6 の HME モデルによる予測は，固定小数点数の場合約 1.5〜3.6 秒，浮動小数点数の場合はおよそ 21.6〜48.7 秒で計算できることが分かった．秘密分散方式は，整数・固定小数点数の計算に関しては Araki ら [2]，浮動小数点数の計算に関しては Aliasgari ら [1] の方式に準じた．また，実装は NEC-SPDZ [3] を用いて行った．なお，本研究は SCIS 2020 での発表内容 [7] を元にしている．

参考文献

[1] Mehrdad Aliasgari, Marina Blanton, Yihua Zhang, and Aaron Steele. Secure Computation on Floating Point Numbers. In: *NDSS 2013*, 2013.

[2] Toshinori Araki, Jun Furukawa, Yehuda Lindell, Ariel Nof, and Kazuma Ohara. High-Throughput Semi-Honest Secure Three-Party Computation with an Honest Majority. In: *Proceedings of ACM CCS 2016*, pp. 805–817, 2016.

[3] Toshinori Araki, Assi Barak, Jun Furukawa, Marcel Keller, Yehuda Lindell, Kazuma Ohara, and Hikaru Tsuchida. Generalizing the SPDZ Compiler For Other Protocols. In: *Proceedings of ACM CCS 2018*, pp. 880–895, 2018.

[4] Nuttapong Attrapadung, Goichiro Hanaoka, Shigeo Mitsunari, Yusuke Sakai, Kana Shimizu, and Tadanori Teruya. Efficient Two-level Homomorphic Encryption in Prime-order Bilinear Groups and a Fast Implementation in WebAssembly. In: *Proceedings of ASIACCS 2018*, pp. 685–697, 2018

[5] Keita Emura, Goichiro Hanaoka, Koji Nuida, Go Ohtake, Takahiro Matsuda, and Shota Yamada. Chosen ciphertext secure keyed-homomorphic public-key cryptosystem. *Designs, Codes and Cryptography*, Vol. 86, No. 8, pp. 1623–1683, 2018.

[6] Riki Eto, Ryohei Fujimaki, Satoshi Morinaga, and Hiroshi Tamano. Fully-Automatic Bayesian Piecewise Sparse Linear Models. In: *Proceedings of AISTATS 2014*, pp. 238–246, 2014.

[7] 前田勇作, 土田光, 大原一真, 古川諒, 寺西勇, 縫田光司. 3 者秘密計算による異種混合モデルを用いた予測の実装評価. *2020* 年暗号と情報セキュリティシンポジウム（*SCIS2020*）予稿集, 2020.