整数上完全準同型暗号の研究

完全準同型暗号は暗号化したままデータの処理を可能にする暗号分野における最先端の注目技術です．この技術はクラウドセキュリティの分野で幅広く使用されることが期待されており，実用的な性能を満たすためにどのような方式で実装していくのかが１つの課題になっています．NTTセキュアプラットフォーム研究所では完全準同型暗号の具体的な構成方式である「整数上完全準同型暗号」を対象に議論的改良と効率向上を行い，世界最速の実装方式を実現しました．

Tibouchi Mehdi

NTTセキュアプラットフォーム研究所

完全準同型暗号とは

電子メール，ビジネス文書，個人情報など，増える一方の電子データがクラウドに託されつつあります．これらの電子データの秘密やプライバシをどのように保護できるかは大きな課題の１つとなっています．Cloud Security Allianceの調べによると，過失，マルウェア，内部攻撃などによるクラウドデータの漏洩事件がここ数年に激増しています[(1)]．このような問題を回避するために，暗号分野ではさまざまな手法が提案されています．その中で現在もっとも注目されている技術の１つは完全準同型暗号（Fully Homomorphic Encryption）と呼ばれるものです．

クラウドに送信する前に電子データを暗号化すると，悪質なハッカーやクラウドサーバの管理者を含めて復号化の鍵を持たない利用者は電子データの中身を参照することはできません．そのため，万が一電子データの漏洩が起きても秘密は守れます．しかしながら，蓄積されている電子データに対してクラウドサーバ上で処理をして何らかのサービス（検索サービス等）を提供しようと思っても，復号化の鍵がないために暗号化された電子データに対する処理ができず，サービスが提供できないという問題があります．完全準同型暗号はこの問題を克服することのできる技術です．完全準同型暗号で暗号化された電子データは，通常の暗号による秘密の保護に加え，暗号化された電子データに対して，任意の処理ができるようになります．処理結果についても暗号化されているため，復号化の鍵の持ち主にしか読み取れません．この特徴を基にして，安全性が高く多種多様なクラウドサービスの実現が期待されています．

潜在用途の具体例

⑴　委託計算

完全準同型暗号のもっとも直接的な応用は，いわゆる安全な委託計算です．１つのクライアントが慎重に取り扱うべきデータに対して特定の処理をしたいが計算力が足りない，必要な情報やノウハウを持たない，などの理由からクラウドサービスを利用して処理をしたいという状況において，完全準同型暗号の利用が効果的です（**図１**）．クライアントが暗号化したデータをクラウドサーバに送り，サーバはデータを暗号化したままの状態で処理（特定の計算の「準同型評価」）をし暗号化したままの出力を返送します．最後に，クライアントが出力データを手元で復号化することで，データを漏らさず処理結果を得ることができます．例えば，株式市場のトレーディング戦略のバッ

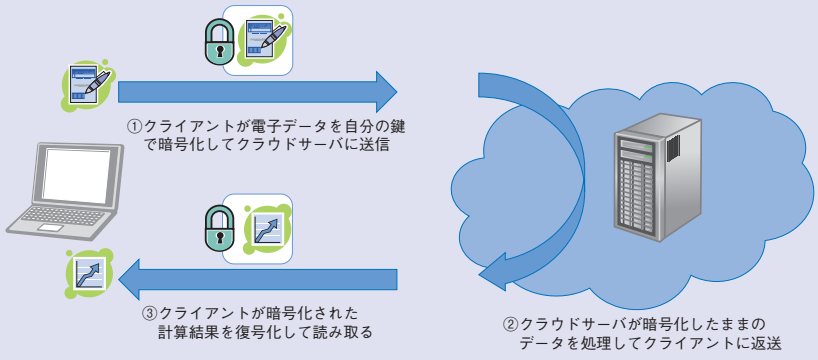

図１　完全準同型暗号に基づいた委託計算のイメージ

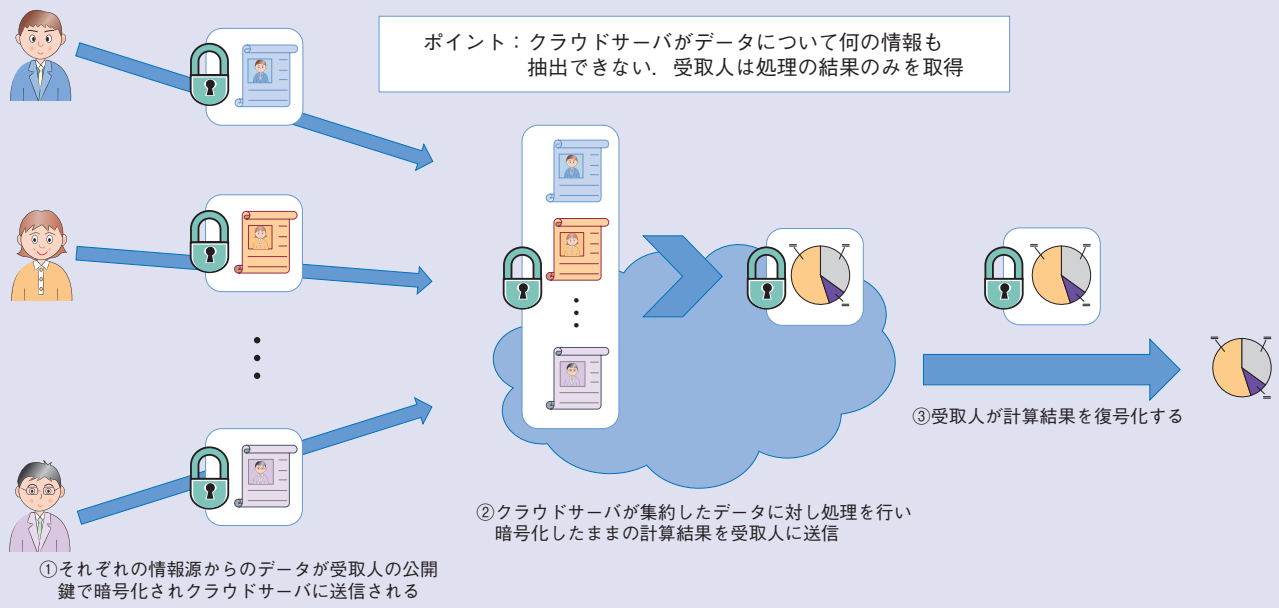

図２　完全準同型暗号に基づいた匿名データ処理のイメージ

クテストをしてくれるクラウドサービスがすでに存在していますが，効率的な戦略は価値が高いので第三者のサービスに明かしたくないとすれば，完全準同型暗号を使うことで解決できます．同様に，医療機関や捜査当局向けにプライバシを守りつつDNA分析をできるサービスの提供や，航空工業や建設業向けにデザインの秘密を守ったままで構造分析を行うサービスの提供などが考えられます．

⑵　匿名処理

完全準同型暗号の応用として，データの匿名処理も挙げられます．さまざまな情報源からデータがクラウドに集約され，集計や統計処理を行い受取人に送信されます．このとき，クラウドサーバの管理者がデータの内容を読み取れないこと，受取人が集計のみを獲得し各々の情報源のデータを得られないことを保証する必要があります．このような状況でも完全準同型暗号が効果的になります（**図２**）．それぞれの情報源が自分のデータを受取人の公開鍵を用いて暗号化しクラウドサーバに送信します．その後，クラウドに集約されたデータに対する集計処理などの準同型評価が実施され，その出力が受取人に送信され，復号化されます．この仕組みにより安全な電子投票を行うことができます．投票者個々人が一票を開票担当者の公開鍵で暗号化して集計サーバに送ります．集計サーバは準同型評価を使って暗号化された集計と有効性のチェックを計算して開票担当者に送ります．最後に開票担当者が集計などを復号化して開票結果を公開します．同じようなプロトコルでセキュアオークションや医療データに対する統計分析なども実現できます．

⑶　データベース検索

一方，暗号化されたデータベース検索は完全準同型暗号があまり適していないクラウドサービスの例です．なぜならば，サーバが検索式の内容について何の情報も得られないので，検索の準同型評価を行うためにデータベースの一部だけではなく最初から最後までの確認が必要となり，計算量が非常に高いものになります．結果として，完全準同型暗号を基にした安全なWeb検索サービスについてはあまり現実的ではありません．また，準同型評価の出力も暗号文なので，暗号化された電子メールに対するスパムフィルタなど，クラウドサーバ自体が結果を読み取ってフィルタリングしてほしい処理でも，クライアントへの連絡が必要となり効果が発揮できません．

このような制限があるにもかかわらず，完全準同型暗号はクラウドセキュリティのためにとても有意義な技術であり，実用的な実装方法の実現

が期待されています．NTTセキュアプラットフォーム研究所ではこれらの期待にこたえられるように取り組んでいます．

整数上完全準同型暗号

完全準同型暗号の概念は約35年前に提案されましたが，長い間実現方法が見つからず，暗号学者の間では不能問題ではないかとも考えられました．しかし2009年にスタンフォード大学のGentry博士が完全準同型暗号を実現できることを初めて証明しました．その翌年にGentryら4名の研究者が，概念的によりシンプルな構成である「整数上完全準同型暗号（Fully Homomorphic Encryption over the Integers）」を紹介しました．どちらもとても重要な理論的ブレイクスルーとなりましたが，この段階では極めて非効率的なため実用には適していませんでした．

まず，整数上完全準同型暗号について簡単に説明します．完全準同型暗号を構成するには，0か1かである1ビットのメッセージの安全な暗号化と，この1ビットのメッセージに対するXOR演算*1とAND演算*2の準同型評価を定義することがポイントになります．そして，任意の長さのデータをビット列として表現し暗号化することによって，データに対する任意の処理をXORゲートとANDゲートから成るブール回路として表現し準同型評価を行います．

整数上完全準同型暗号では，秘密鍵を比較的に大きいサイズ（600桁ぐらい）の整数p（奇数）とします．この整数pの倍数$q_i \times p$をいくつか公開すると，pを逆算することは最大公約数の計算で容易にできますが，pの倍数の近傍の整数$q_i \times p + e_i$（e_iは比較的に小さい「ノイズ」であり，20～30桁の整数とします）からはpを逆算するのは難しいと考えられます．実はq_iが十分に大きい（1000万桁ぐらい）であれば，pを知らない攻撃者にとって$q_i \times p + e_i$は同じ長さの乱数とは識別不可能になります．したがって，1ビットのメッセージmを次のように暗号化できます．小さくてランダムな偶数$2r$（20～30桁のノイズ）と非常に大きいpのランダムな倍数$q \times p$（1000万桁ぐらい）の和$c = q \times p + 2r + m$を暗号文として出力します．mが0であっても1であってもcは乱数とは識別不可能なので，pを知らない攻撃者がcを取得してもmを復元できません．一方，秘密鍵pを知っている正規の利用者はcをpで割って，剰余$2r+m$が偶数か奇数かによってmが0か1かが分かり，復号化できます．結果として，安全な暗号方式が定義されました．

準同型評価

次に，XOR演算とAND演算による準同型評価を説明します．1ビットのメッセージm_1, m_2の暗号文$c_1 = q_1 \times p + 2r_1 + m_1$，$c_2 = q_2 \times p + 2r_2 + m_2$を足せば，$m_1$ XOR m_2の暗号文が求められます．次に$c_1 + c_2 = (q_1 + q_2) \times p + 2(r_1 + r_2) + (m_1 + m_2)$を復号化すると，$m_1 = m_2$の場合は$2(r_1 + r_2) + (m_1 + m_2)$が偶数なので0が，$m_1 \neq m_2$の場合は$2(r_1 + r_2) + (m_1 + m_2)$が奇数なので1が求められます．つまり$c_1 + c_2$は$m_1$ XOR m_2の暗号文となります．同様に，$c_1 \times c_2 = (q_1 q_2 p + 2q_1 r_2 + q_1 m_2 + 2q_2 r_1 + q_2 r_1) \times p + 2(2r_1 r_2 + r_1 m_2 + r_2 m_1) + m_1 m_2$は$m_1$ AND m_2の暗号文になることもチェックできます．

上記で説明した暗号方式はまだ完全準同型暗号ではなく，正確には制限付き準同型暗号（Somewhat Homomorphic Encryption）と言います．さらなる問題は，特にAND演算の準同型評価を行うたびに，暗号文ノイズのサイズが大きくなってしまいます．ノイズの桁数はおおよそ倍増します．結果的には，AND演算数個を含めた処理を行うと，ノイズがpより大きくなり復号化が不正確になってしまう場合があります（**図3**）．この問題を克服するために，ノイズをある程度小さくする方法が必要となり，Gentryの革新的なブートストラップ手法を使います．これにより完全準同型暗号へ変換を行います．

ただし，整数上の制限付き準同型暗号方式の性能は低く，1ビットのメッセージに相当する暗号文の長さで分かりますが，暗号文サイズは平文サイズの数千万倍になってしまいます．そのため，XOR演算・AND演算という単純な処理の準同型評価は巨大整数に対する算術操作になるので，メモリ消費も時間計算量もかなり大きくなります．そしてブートストラップ手法による完全準同型暗号への変換をすると，効率性がさらに下がってしまいます．

NTTからの提案

NTTセキュアプラットフォーム研

*1　XOR演算：入力される1と0の組み合わせのうち，その値が一致しないときに限り1（真）を出力する方式．XORはeXclusive ORの略で，排他的論理和とも呼ばれます．

*2　AND演算：入力された値がすべて1（真）であった場合に限り，演算結果を1（真）とする方式．論理回路が行うもっとも基本的な論理演算の1つです．

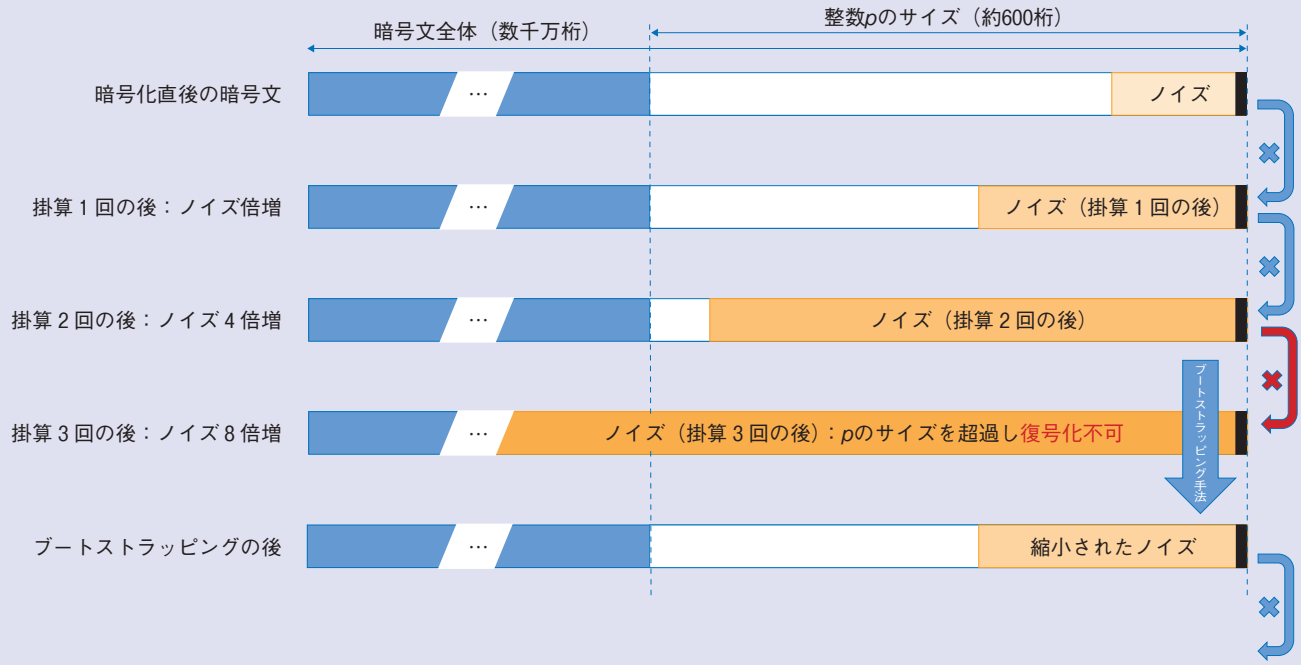

図３　整数上完全準同型暗号におけるブートストラッピング

究所では，整数上完全準同型暗号をより実用的な実現方式にするべく取り組んでいます．効率性に対するボトルネックは主に３つあります．

① 暗号文拡張率：メッセージの１ビット当りの暗号文サイズが大きすぎる点です．

② 準同型評価のオーバヘッド：AND演算のような，メッセージに対する単純な処理は準同型評価を行うと任意精度算術演算などかなり重い計算になってしまいます．ブートストラップ手法でこの問題はさらに悪化します．

③ 公開鍵と公開パラメータのサイズ：今まで説明した方式は共通鍵暗号になっていますが，匿名データ処理など一部の用途では公開鍵暗号が必要となっています．公開鍵暗号への簡単な変換手法は存在していますが，これを利用すると公開鍵は非常に大きくなり，使用に適さなくなります．また，ブートストラップ手法などを導入すると，共通鍵方式の場合でも大量の準同型評価用パラメータを公開しなければならなくなる問題があります．

NTTでは2012年までに，①と③の問題を検討し，公開鍵圧縮手法や暗号文圧縮手法，非線形暗号化手法などを開発しました．これらの成果によって，原型の整数上完全準同型暗号では公開鍵ストレージだけでも大規模データセンタを必要とする規模であったものを，公開鍵のサイズを数メガバイト程度まで減らすことができ，普通のPCで実行できるほどの改良ができました[(2)]．さらに，公開鍵や暗号文は準同型評価時に処理せざるを得ないデータなので，そのサイズを減らすことによって準同型評価の性能が向上しました．

2013年には，暗号文拡張と準同型評価のオーバヘッドを同時に向上させるような，さらなる改良を提案しました[(3)]．これは，１つの暗号文に複数のメッセージのビットm_1，…，m_nを詰め込んで，準同型評価のときにそのすべてのビットに対する並列処理を可能するものです（**図４**）．根本的なアイデアは複数の秘密奇数p_1，…，p_nを生産して暗号文cは$p_1 \times \cdots \times p_n$の倍数に近い素数とします．ただし，$c$を$p_i$で割れば$2r_i + m_i$が求められます．暗号文の足し算・掛け算をするとそれぞれの平文ビットm_iに対してXOR演算・AND演算の並列準同型評価となるのです．また最近では，整数上完全準同型暗号でのブートストラップ手法の使用を完全に避けるための新技術も提案しまし

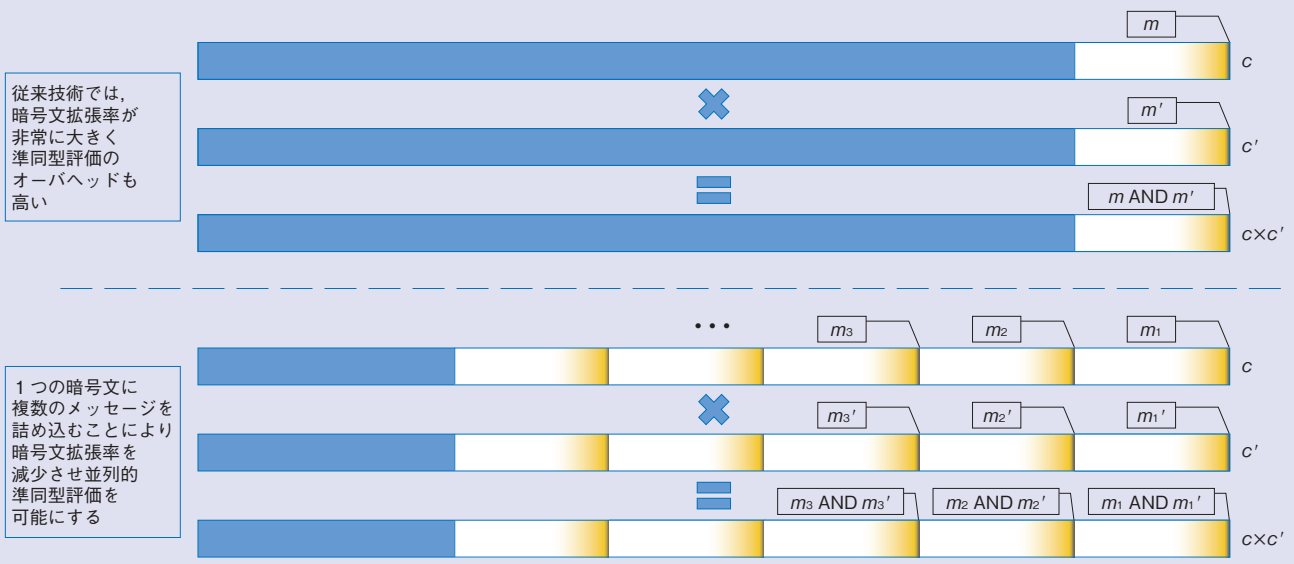

図4　整数上完全準同型暗号に対する暗号文並列化

た[4]．その技術を使うと，AND演算の準同型評価を行うたびに，ノイズが倍増しないで少しだけ大きくなります．結果として準同型処理に適応したシステムパラメータを選んだ後に，ノイズがpを超えないことを保証しブートストラップなしの完全準同型暗号をつくることができます．このような効率向上のおかげで，2013年だけで準同型暗号の100倍以上の処理高速化を達成し，世界最速の完全準同型暗号の実装を得られました．普通のPCでもブロック暗号AESの準同型評価はおよそ20秒でできます．もちろん，高度の安全性を保ったままの結果です．一部の簡単な用途には十分な性能レベルであり，完全準同型暗号の実用性がみえてきたと考えられます．

今後の展開

今後も世界トップクラスの暗号研究を維持しつつ，実用性の高い完全準同型暗号の実現に向けて革新的な改良を続けていきます．暗号文圧縮手法などの最新の整数上方式はまだ部分的にしか適応できないため，現時点での最大のボトルネックはメモリ消費だと考えられます．まずはこの問題に取り組み，実用レベルのデータ量の準同型処理を可能にする予定です．その他，完全準同型暗号とは異なりますが，構造的には類似した暗号の重要な新技術である多重線形写像やプログラム難読化の整数上の実装と効率向上などについても研究を進めていきます．

■参考文献

(1) Cloud Security Alliance: “Cloud Computing Vulnerability Incidents: A Statistical Overview,” March 2013.

(2) J.-S. Coron, D. Naccache, and M. Tibouchi: “Public Key Compression and Modulus Switching for Fully Homomorphic Encryption over the Integers,” EUROCRYPT 2012, LNCS 7237, pp.446-464, 2012.

(3) J.H. Cheon, J.-S. Coron, J. Kim, M.S. Lee, T. Lepoint, M. Tibouchi, and A. Yun: “Batch Fully Homomorphic Encryption over the Integers,” EUROCRYPT 2013, LNCS 7881, pp.315-335, 2013.

(4) J.-S. Coron, T. Lepoint, and M. Tibouchi: “Scale-Invariant Fully Homomorphic Encryption over the Integers,” PKC 2014, LNCS 8383, Buenos Aires, Argentina, March 2014.

Tibouchi Mehdi

NTTセキュアプラットフォーム研究所では，今後も暗号技術の研究を通じて，安心・安全なクラウドサービス提供の実現を目指します．

◆問い合わせ先
NTTセキュアプラットフォーム研究所
岡本特別研究室
TEL　0422-59-7743
FAX　0422-59-3285
E-mail　tibouchi.mehdi@lab.ntt.co.jp