

## クラウドサービスにおける盗聴防止及び改ざん検知 の一方式

柯, 陳毓強  
九州大学大学院システム情報科学府情報学専攻

穴田, 啓晃  
九州先端科学技術研究所

川本, 淳平  
九州大学大学院システム情報科学研究院 | 九州先端科学技術研究所

櫻井, 幸一  
九州大学大学院システム情報科学研究院 | 九州先端科学技術研究所

<https://hdl.handle.net/2324/1662103>

---

出版情報：電気・情報関係学会九州支部連合大会。平成26年度(第67回)(1), 2014-09-18  
バージョン：  
権利関係：



# クラウドサービスにおける盗聴防止及び改ざん検知の一方式

柯陳毓トウ\* 穴田啓晃\*\* 川本淳平\*\*\* 櫻井幸一\*\*\*  
(\*九州大学院システム情報科学情報学専攻 \*\*公益財団法人九州先端科学技術研究所)

## 1 はじめに

クラウド環境を利用したサービスの発展に伴い、個別の機能に特化したサービスとそれらのサービスを連携させ付加価値を加えるサービスが出現してきた。例えば、日々の食事情報すなわち摂取カロリー情報の管理に特化したサービスに運動記録の管理に特化したサービス、そしてそれらを連携させ情報を加工し、健康管理支援という付加価値を提供するサービスなどである。

一方で、サービスに預ける個人データ的安全性や攻撃者による改ざんが問題となる。実際、情報管理サービスからの個人情報漏洩がしばしば発生し問題となっている。また、健康管理支援サービスがユーザに提示する情報が攻撃者によって改ざんされ、健康管理支援サービスが誤った情報を提示してしまうことも、場合によっては重大な問題に発展してしまう。

本稿で我々は上記のサービスをクラウドサービスと呼んで考察し、暗号化及び改ざん検知を実現する一方式を提案する。

既存研究[1]においては、著者らは多項式で表された計算をサーバに移譲するサービスを、秘密計算として実現している。技術的には、準同型暗号の内積準同型性が本質的に利用され、秘密計算が可能となっている。本稿では既存研究[1]に倣い、しかし上記のクラウドサービスについて考察する。

本稿の対象とするクラウドサービスは、データベース、クラウドサーバ、ユーザの三者からなるものとする。データベースは、データを格納し、ユーザからのクエリに応じ検索し抽出する。クラウドサーバは、抽出されたデータを受け取り、付加価値を付けるための係数を乗ずるという加工をする。そしてユーザは、加工されたデータを受け取り、サービスを楽しむ。

データの盗聴防止及び改ざん検知の観点では、サーバ-ユーザ間についてはSSLサーバ証明書の確認等により改ざんの恐れのない安全な通信が期待できる。一方、データベース-サーバ間については改ざんの恐れがある。

そこで本稿では、データベース-サーバ間については暗号化に加え電子署名を適用する。またサーバ-ユーザ間については暗号化のみを適用する。ここで鍵管理のコストの観点から、公開鍵インフラの利用を想定し、暗号化にはユーザの公開鍵を用いるものとする。暗号アルゴリズムには準同型性を持つものを用いる。準同型性により、データベースで生成された暗号文に対し、サーバで生成された暗号文を乗じ、積の結果をユーザが秘密鍵で復号できるようになる。具体的には、暗号化及び署名に署名付きエルガマル暗号[2]を用い、サービスを実現する方式を組み立てた。

なお、本稿ではデータベースやサーバへの侵入による盗聴や改ざんについては検討しない。

本研究の結果として、本稿では、データベース-サーバ間においては盗聴防止及び改ざん検知を、またサーバとユーザ間においては盗聴防止を可能にするクラウドサービ

スを提案することができた。

## 2 準備

### 2.1 署名付きエルガマル暗号

#### ・鍵生成

2つの素数  $p, q$  と整数  $g$  を  $g^q \equiv 1 \pmod{p}$  となるように選ぶ。0以上  $(q-1)$ 以下の範囲の乱数  $x$  を選び、 $g$  の法  $p$  での  $x$  乗である  $y = g^x \pmod{p}$  を計算する。組  $(p, q, g, x)$  を秘密鍵  $sk$  とし、組  $(p, q, g, y)$  を公開鍵  $pk$  とする。

#### ・暗号化

0以上  $(q-1)$ 以下の範囲の乱数  $r$  を生成し、平文  $m$  の暗号文は、 $(c_1 = g^r, c_2 = my^r)$  を計算する。そして、再び0以上  $(q-1)$ 以下の乱数  $s$  を生成し、 $u = g^s \pmod{p}$  を計算し、ハッシュ関数  $H$  における数値  $e = H(c_1, c_2, u)$  を計算し、 $w = s + er$  を計算する。組  $(u, w)$  は、 $\sigma = (u, w)$  を表示する。全体として、署名付きエルガマル暗号文は、組  $w = (c_1, c_2, \sigma)$  ということである。

#### ・復号と検証

秘密鍵  $sk$  から  $x$  を取り出し、暗号文組  $c = (c_1, c_2, \sigma)$  から  $\sigma = (u, w)$  を取り出す。暗号文を復号する前に、暗号文が改ざんされたかどうかを検証する。まず、 $e$  をハッシュ関数  $H$  の組  $(c_1, c_2, u)$  におけるハッシュ値  $e = H(c_1, c_2, u)$  を計算し、検証式  $u = g^{w/c_1 e} \pmod{p}$  をチェックする。もし、検証式が成立しなければ、暗号文が改ざんされたことは分かり、復号を放棄する。一方、検証式が成立すれば、暗号文は下の式のように復号する。

$$m = (c_2 / c_1^x) \pmod{p}.$$

## 3 提案方式

本節では、第1節で述べたクラウドサービスを実現するための方式を説明する。

### 3.1 処理の流れ

本稿で提案するクラウドサービスの処理の流れを図1に示す。

データベースに保存されているデータ  $z_i$  は、摂取カロリーや運動履歴などのデータである。なお、データはデータベース内に暗号化されず保存されているが、サーバに送る前に暗号化されるものとする。

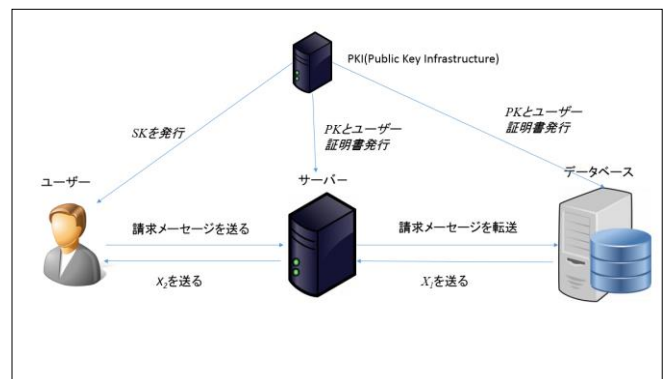


図1: クラウドサービスと盗聴防止及び改ざん検知の提案方式

クラウドサーバに保存されているデータ  $a_i$  は、データ加工サービスの(付加価値のための)係数である。係数は暗号化なしで保存されているが、データベースからもらった暗号文と乗ぜられる前に暗号化されるものとする。

ユーザが最終的に計算する計算式は下に示す形と仮定する。

$$f(z) = a_{1z_1} + a_{2z_2} + \dots + a_{nz_n}.$$

本稿では各々の単項式  $a_{iz_i}$  をユーザが自身の秘密鍵で復号した上で加算するものとする。

### 3.2 盗聴防止と改ざん検知

本稿で提案するサービスには PKI ( Public Key Infrastructure ) が必要である。秘密鍵  $sk = (p, q, g, x)$  はユーザが持っている。サーバとデータベースの公開鍵  $pk = (p, q, g, y)$  は、データを送信する際、サーバあるいはデータベースが各々PKI の認証局に申請し入手する。

サーバは、ユーザからの請求内容をデータベースにクエリする。データベースはクエリされたデータを、ユーザの公開鍵  $pk$  で暗号化する。次いで、データベースはその暗号文をサーバに送る。ここで、暗号文  $X = (X_1, X_2, X_3)$  は、2.1 節の署名付きエルガマル暗号のプロトコルに従い、次の3つベクトルで構成されている。

$$\begin{aligned} X_1 &= (c_{11}, c_{12}, c_{13}, \dots, c_{1n}), \\ X_2 &= (c_{21}, c_{22}, c_{23}, \dots, c_{2n}), \\ X_3 &= (\sigma_{11}, \sigma_{12}, \sigma_{13}, \dots, \sigma_{1n}). \end{aligned}$$

データベースはベクトルの中の元素を下に示すように計算する。

$$\begin{aligned} c_{ii} &= g^{r_i}, \quad c_{2i} = x_i y^{r_i}, \\ u_{1i} &= g^{s_i} \bmod p, \\ e_{1i} &= H(c_{1i}, c_{2i}, u_{1i}), \\ w_{1i} &= s_i + e \cdot r_{1i}, \\ \sigma_{1i} &= (u_{1i}, w_{1i}). \end{aligned}$$

ここで、 $r_{1i}, s_i (i = 1, 2, \dots, n)$  は、データベースが生成する乱数である。

サーバは秘密鍵を持っていないので、データベースから取り寄せた暗号文に対応する平文を知ることはできない。しかし、暗号文が改ざんされたかどうかを検証することができる。つまり、データベースから得たデータの完全性を検証することが以下の手続きで可能である。

すなわち、サーバは、データベースから得た暗号文  $X = (X_1, X_2, X_3)$  の元素を抽出し、以下の検証式に代入する。

$$u_{1i} = g^{w_{1i}} c_{1i}^{-e_i} \bmod p.$$

ここで、 $e_{1i} = H(c_{1i}, c_{2i}, u_{1i})$  である。もし、暗号文  $X$  に対し1個以上検証式が成立しなければ、その暗号文  $X$  が改ざんされたことが分かる。その場合は暗号文を放棄し、再びデータベースにデータをクエリする。

次いで、サーバは、データベースからもらった暗号文  $X = (X_1, X_2, X_3)$  に係数  $a_i$  を乗ずる。ただし、付加価値を付すための係数  $a_i$  を秘匿しておくため、 $a_i$  をまず暗号化する。サーバが係数  $a_i$  を暗号化する計算は下のようになる。

$$d_{1i} = g^{r_{2i}}, \quad d_{2i} = a_i y^{r_{2i}}, \quad \varepsilon_i = (d_{1i}, d_{2i}).$$

ここで、 $r_{2i} (i = 1, 2, \dots, n)$  はサーバが生成する乱数である。結果、係数の暗号文  $D$  は  $D = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n)$  となる。

サーバは最後に、暗号文  $D$  と暗号文  $X$  の各エレメントを乗ずる。暗号文  $D$  と暗号文  $X$  の具体的な乗算は下のようになる。

$$\begin{aligned} f_{1i} &= d_{1i} \cdot c_{1i} = g^{r_{1i} + r_{2i}}, \quad f_{2i} = d_{2i} \cdot c_{2i} = a_i m_i y^{r_{1i} + r_{2i}}, \\ X_1' &= (f_{11}, f_{12}, \dots, f_{1n})^T \\ X_2' &= (f_{21}, f_{22}, \dots, f_{2n})^T \\ X' &= (X_1', X_2'). \end{aligned}$$

ここで準同型性を利用した。サーバは、こうして得られた暗号文組  $X' = (X_1', X_2')$  をユーザに送る。

ユーザは、サーバから得た暗号文を秘密鍵  $sk$  を用いて復号する。暗号文を復号する具体的な計算を下に示す。

$$j_i = f_{2i} / (f_{1i})^x \bmod p.$$

結果、ユーザは  $J = (j_1, j_2, j_3, \dots, j_n)$  というベクトルを得る。ここで、 $j_i = a_i z_i (i = 1, 2, \dots, n)$  である。

ユーザは最終的に下に示す加算を計算し、受け取るべきデータ  $f(z)$  を得る。

$$J = j_1 + j_2 + \dots + j_n = f(z).$$

### 4 まとめ

本稿で我々は、署名付きエルガマル暗号、及びエルガマル暗号の準同型性を利用し、通信路上の盗聴と改ざんに対し安全なクラウドサービスの一方式を提案した。

今後の研究の課題は、まず、より付加価値の高いクラウドサービスのため、サーバが複数のデータベースに対しデータをクエリし取り寄せる処理を検討したい。

また、ユーザの計算負荷を軽減するため、乗算に加え加算を(つまり内積を)サーバが行うように改良することも課題と考える。

### 謝辞

第三著者は、本研究に関し部分的に JSPS 科研費 26730065 並びに 財団法人人工知能研究振興財団 の助成を受けております。

### 参考文献

- [1] 呉双, 川本淳平, 菊池浩明, 佐久間淳. 準同型性暗号に基づいたプライバシー保護オンラインロジスティック回帰. 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習 113(139), 67-74, 2013-07-11
- [2] 有田正剛, 「知識の証明と暗号技術」. 情報セキュリティ総合科学第1巻, 2009年11月, 情報セキュリティ大学院大学
- [3] W Fang, C Zhou, B Yang.: Privacy preserving linear regression modeling of distributed databases. Optimization Letters April 2013, Volume 7, Issue 4, pp 807-818, Springer-Verlag
- [4] W Du, Y S.Han, S Chen.: Privacy-Preserving Multivariate Statistical Analysis: Linear Regression and Classification, In Proceedings of the 4th SIAM International Conference on Data Mining, 2004.