Vol.2010-AL-128 No.12
2010/1/26

非公開データベース間における機密性の高い情報共有の検討

隅　　崇　佳[†1]　上土井　陽子[†1]　若　林　真　一[†1]

情報共有を行うことは新たな知識を獲得できる可能性があるが，非公開データベース間の情報共有の場合，データベース内の独自情報まで公開することは望ましくない．そのため，R.Agrawal らは機密性を保持したプロトコルを提案している．ここで，機密性を保持しているとは，情報共有を行った後に，相手機関独自の情報を推測できないことを意味している．本研究では，独自情報の推測に利用可能な情報流出を抑制することで，従来プロトコルの高機密化を目指す．

Secure Information Sharing Across Private Databases

Takayoshi Sumi,[†1] Yoko Kamidoi[†1]
and Shin'ichi Wakabayashi[†1]

There is a new increasing need for information sharing driven by several trends, i.e., end-to-end integration, simultaneously competing and cooperating and privacy preserving. The need is to share necessary information across multiple entites without revealing any additional information exclusive of information admitted to share. To solve this issue, Agrawal et al. proposed a protocol for minimal information sharing across private databases. Minimal information sharing means that it is allowed to be revealed some minimal additional information during the communication followed the protocol. In this paper, we consider secure information sharing across private databases by not releasing information related additional information as much as possible.

1. はじめに

複数の非公開データベース間で情報共有を行う際，機密性の観点から質問と無関係な情報が明らかにならないような方法で必要な情報のみを共有する要求が高まってきた．例えば，図 1 のように二つの機関それぞれが持つデータベースの積集合に属している要素の情報を共有したい場合，相手機関と共有できる部分がどこであるかを相手に応じて求める必要がある．その方法として，R.Agrawal らは可換性のある暗号化関数を用いることで積集合に属さない要素を明らかにしないという機密性を保持することを目的としてプロトコルを提案している．ここでは，機密性を保持したプロトコルとは情報共有を行った後にプロトコルを通して得た情報を利用して相手機関独自の情報を推測できないプロトコルとする．R.Agrawal らのプロトコル (従来プロトコル) はプロトコルに用いている暗号が部分解読できなければ，機密性を保証できると文献[1] で証明されている．しかし，ある限定された状況での暗号の解読に関する研究は日々進んでおり，プロトコルの適用環境によっては保証された機密性が十分でない可能性もある．

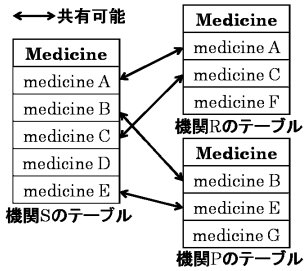

図 1　情報共有の例

本研究では，プロトコルを通して得られる相手機関独自の情報を推測するために利用できる情報に関して，極力，公開しないことで従来プロトコルの機密性が破られたとしても機密性を保持した情報共有を行う手法について検討する．

2. 準　　備

まず，プロトコルの機密性と動作モデルについて説明する．

2.1 プロトコルの機密性

本研究でのプロトコルの機密性とは，参加機関すべてがプロトコルに従うとしたとき，プロトコルを通して相手から得た情報を利用することで相手が独自に持つ情報を推測することがプロトコルを行う前と比較して容易にならないことを保証することである．プロトコルの機密性は，以下の二つの事項により決定される．

†1 広島市立大学大学院 情報科学研究科
Graduate School of Infomation Sciences, Hiroshima City University

- プロトコル中で用いられる暗号の強度
- プロトコルを通して渡される暗号解析に利用可能な情報

暗号の強度とは，プロトコルに用いられる暗号を解析する問題の困難さにより決まり，暗号の強度が高ければ，より解析が困難であることからプロトコル全体の機密性が高くなる．一方，暗号解析に利用可能な情報とは，相手機関独自の情報を推測するための解析に利用可能な情報により決まる．利用可能な情報が少ない，もしくは，利用方法が限られるなら，相手機関独自の情報を推測する際の解析難度が高いためプロトコル全体の機密性も高くなる．ここで，相手機関独自の情報の推測とは，プロトコルを介して得たメッセージの中で独自情報の解析に利用できる情報を用いて暗号解析を行い相手機関独自の情報の一部を求めることとする (図 2).

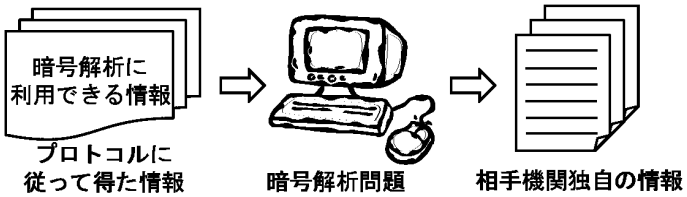

図 2 独自情報の解析

本研究で考察するプロトコルの機密性の高さは相手機関独自の情報を推測する暗号解析の困難さに等しい．従来プロトコルでは，この困難さが Decisional Diffie-Hellman 問題の困難さと等価であると文献[1] で主張されている．

2.2 情報共有を行う機関のモデル

情報共有を行う各機関は以下の好奇心に豊む (Honest-but-Curious) 振る舞い[1] に従うとする．

Honest-but-Curious behavior

プロトコルに参加する機関は忠実にプロトコルに従う．しかし，プロトコル実行中に受け取ったメッセージや行った計算を全て記録してのちに付加情報を得ることを目的として記録を解析するかもしれない．

上記モデルから，本研究では，情報共有を行う機関が故意に偽データを含ませたデータベースを情報共有して相手機関の情報を得ようとすることはないとする．また，第三者によるデータの改ざんなどでの情報漏えいも機密性を破ることの対象としない．

2.3 離散対数問題と Decisional Diffie-Hellman 問題

離散対数問題とは，p，g，y が与えられたとき $y = g^a \bmod p$ となる a を求める問題である．ここで，p を十分に注意深く選ばれた大きな素数，g を素数位数 p の巡回群 G の原始元，y を G の元とする．離散対数問題は素数 p が十分に大きいとき現在の計算機環境では現実的な時間内で解くことが難しいとされている問題である[2].

Diffie-Hellman(DH) 問題は離散対数問題の困難性に基づいて計算が困難とされている問題で Diffie-Hellman 鍵配送方式の安全性に関係する問題である．DH 問題には CDH(Computational Diffie-Hellman) 問題と DDH(Decisional Diffie-Hellman) 問題と呼ばれる問題がある．以下に 2 つの問題について説明する．

[CDH 問題]

CDH 問題は G を素数位数 p の巡回群，g を G の原始元，$a, b(0 \leq a,b \leq p-1)$ をランダム値とする．また，p は十分に大きい値とする．このとき，(G, p, g, g^a, g^b) が与えられたときに，g^{ab} を求める問題のことである．そして，CDH 問題を解く効率的なアルゴリズムが存在しないという仮定のことを，CDH 仮定と呼ぶ． □

暗号に使われているほとんどの群においては，CDH 問題と離散対数問題は多項式等価である．つまり，離散対数問題と同様に CDH 問題は解くことが難しいとされる問題である．

[DDH 問題]

G を素数位数 p の巡回群，g を G の原始元，a, b を，$0 \leq a,b \leq p-1$ を満たす，ランダム整数値とする．また，CDH 問題と同様に p は十分に大きい値とする．このとき，(G, p, g, g^a, g^b) が与えられたときに g^{ab} の部分情報を求める問題を DDH 問題と呼ぶ．この DDH 問題は識別可能という表現を使って，言い換えると次のようになる．G を素数位数 p の巡回群，g を G の原始元，a, b, c をランダム値とする．このとき，$(G, p, g, g^a, g^b, g^{ab})$ と (G, p, g, g^a, g^b, g^c) を識別する問題である．また，この 2 つのタプルが計算的に識別不可能であるという仮定を DDH 仮定と呼ぶ． □

計算的に識別不可能とは次のように定義できる．

[識別不可能性]

$\Omega_k \subseteq \{0,1\}^k$ を k ビットの有限領域とする．また D1 と D2 を Ω_k 上の分布とする．$A_k(x)$ を $x \in \Omega$ が与えられ真か偽を返すアルゴリズムとする．計算的に区別がつかないことを式を用いて定義する．以下の式において $A_k(x)$ は多項式時間内に真か偽を出力するアルゴリズムとし，どんな多項式 $p(k)$ に対しても以下の式は成り立つと定義する．

$$|Pr[A_k(x)|x\sim D1] - Pr[A_k(x)|x\sim D2]| < 1/p(k)$$

$Pr[Ak(x)]$ はアルゴリズムが入力に対して真を返す確率であり，$x\sim D$ は分布 D 上のある要素 x ということを意味している．□

以降では，計算的に識別できないことを単に識別不可能と表現する．

3. 従来プロトコル

Agrawal らが文献[1] で提案した情報共有プロトコルを以下に示す．

1 機関 S と R はそれぞれが持つ集合 Vs，Vr をハッシュ関数 h より変換する．変換後の集合をそれぞれ Xs=h(Vs)，Xr=h(Vr) とする．機関 S と R は定義域 $KeyF$ からランダムに鍵 e_s，e_r を選ぶ．

2 機関 S と R はハッシュ化された集合をそれぞれが選んだ鍵で暗号化する．暗号化後の集合をそれぞれ Ys=f_{es}(Xs)，Yr=f_{er}(Xr) とする．

3 機関 S は集合 Ys の要素を辞書式順に並び換えた列を機関 R に送る．

4 (a) 機関 S は集合 Ys の要素を辞書式順に並び換えた列を機関 R に送る．
(b) 機関 S は集合 Yr に含まれるすべての要素 y を鍵 e_s で暗号化する．暗号化後の集合を Zr=f_{es}(Yr) とする．それから機関 R に集合 Yr に含まれる全ての要素 y に関しペア $< y, f_{es}(y) >$ を送り返す．

5 機関 R はステップ 4 の (a) で機関 S から得た集合 Ys を鍵 e_r で暗号化し集合 Zs=f_{er}(Ys) を作成する．また集合 V_r に含まれる要素 v に対してステップ 4(b) で得たペア $< y, f_{es}(y) >=< f_{er}(h(v)), f_{es}(f_{er}(h(v))) >$ からペア $< v, f_{es}(f_{er}(h(v))) >$ を得る．

6 機関 R はステップ 4 で送られてきた暗号化集合 Zr とステップ 5 で作成した暗号化集合 Zs から集合 Zs ∩ Zr を求める．求めた集合 Zs ∩ Zr=$f_{es}(f_{er}(h(\text{Vs} \cap \text{Vr})))$ に属する暗号文とペアになっている集合 Vr に属する要素の集合が積集合 Vs ∩ Vr となる．

上記のプロトコルでは，共有する属性の取り得る定義域を U としたとき，定義域 U が小さい場合においてもプロトコルの機密性を保持することを目的としている．そのため，積集合を求め，かつ，プロトコルに機密性を保持させるために用いる暗号化関数は次節にて述べる特性を満たす必要がある．

3.1 暗号化関数

従来プロトコルにおいて用いられる暗号化関数 f(鍵 e での暗号化を f_e とする) は以下の 4 つの暗号化特性を満たす必要がある．

(1) 可換性：全ての $e, e' \in KeyF$ に対して $f_e.f_{e'} = f_{e'}.f_e$ が成り立つ．

(2) 各関数 f_e は全単射．

(3) 鍵 e を用い暗号化された情報は鍵 e が与えられると鍵長 k に関する多項式時間内に復号が可能．

(4) 識別不可能性：$< x, f_e(x), y, f_e(y) >$ の分布は $< x, f_e(x), y, z >$ の分布と区別がつかない．$(x, y, z \in {}_rDomF,\ e \in {}_rKeyF)$

ここで，鍵の定義域を $KeyF$，暗号化前の値，暗号化した値の定義域を $DomF$ とする．$a \in {}_rb$ は“集合 b からランダムに選ばれた a”ということを意味する．

条件 (4) の識別不可能性はある暗号化前後のペアが与えられたとき，次に与えられるペアが同じ鍵で暗号化されたものかどうかを鍵長 k に関する多項式時間内に判定できるアルゴリズムがないということを意味している[1]．

上記の 4 つの暗号化特性を満たす暗号として以下の Diffie-Hellman 暗号化関数 f_e を従来プロトコルでは挙げている．

$$f_e(x) \equiv x^e \bmod p$$

以降，暗号化の計算において mod p を省略する．暗号化関数 f を用いることによる機密性がどの程度かを文献[1] では考察している．この識別不可能性にある二つのタプルを識別する問題の困難さは，DDH 問題を解くことの困難さと等価であることが多項式時間帰着により示されている．暗号化関数 f を適用した識別不可能性は $< x, f_e(x), y, f_e(y) >$ の分布は $< x, f_e(x), y, z >$ の分布と識別不可能という特性である．この分布 $< x, f_e(x), y, z >$ を鍵の定義域 $KeyF$ から適当な値 d を用いて変換する．

$$< x, x^e, y, z >$$
$$=< g^d, g^{ad}, g^b, g^c >$$

x は g^d，y は g^b，z は g^c，鍵 e は a に対応している．上記のように変換すると，分布 $< g^d, g^{ad}, g^b, g^{ab} >$ と分布 $< g^d, g^{ad}, g^b, g^c >$ の識別が不可能と言い換えることができる．この二つの分布 $< g^d, g^{ad}, g^b, g^{ab} >$ と分布 $< g^d, g^{ad}, g^b, g^c >$ を区別する問題を distinguish 問題と定義する．一方，DDH 問題は $< g^a, g^b, g^{ab} >$ と $< g^a, g^b, g^c >$ を区別する問題である (巡回群 G，原始元 g，素数 p は省略)．distinguish 問題と DDH 問題を比較すると図 3 のようになる． 識別不可能性にある二つのタプルを識別する問題の困難さは，DDH 問題を解くことの困難さと等価であることを distinguish 問題と DDH 問題の多項式時間帰着によ

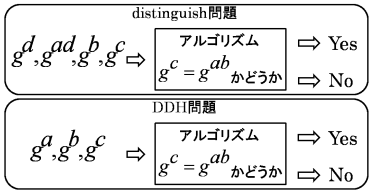

図 3 distinguish 問題と DDH 問題

り示す.

DDH 問題を distinguish 問題に多項式時間帰着することについて考える．図 3 の DDH 問題の入力の g^a と原始元 g を鍵の定義域 $KeyF$ から適当な値 d を用いてそれぞれ g^{ad}，g^d に変換する．すると，変換した 2 つの値 g^{ad}，g^d と DDH 問題の入力である g^b，g^c の 4 つのタプル $< g^d, g^{ad}, g^b, g^c >$ を入力とする distinguish 問題の出力より，DDH 問題の出力を得ることが可能である．このようにして，文献[1] で DDH 問題が distinguish 問題に多項式時間帰着可能と示されている．

また，distinguish 問題から DDH 問題への多項式時間帰着は文献[6] にて示している．双方向の多項式時間帰着が可能なことから distinguish 問題と DDH 問題が多項式時間同値であることが分かる．つまり，暗号化関数 f を適応した従来プロトコルの機密性を破る問題は DDH 問題と同じ困難さを持つ.

3.2 従来プロトコルの機密性を破る問題

この節では，識別不可能性で対象とした distinguish 問題に対する解法を利用して従来プロトコルの機密性を破る問題を解くことにかかる具体的な計算量を考察する.

まず，distinguish 問題を解くことが可能になると，識別不可能性にある二つのタプル $< x, f_e(x), y, f_e(y) >$ と $< x, f_e(x), y, z >$ の区別が可能となる．この二つのタプルの区別が行えるということは，$x = g^d$，$y = g^b$ とすると，$< x, f_e(x), y, z >$ を $< g^d, g^{ad}, g^b, g^{ab} >$ と変換でき，図 3 のような判定アルゴリズムが存在することを意味している.

図 3 の判定アルゴリズムは，ある暗号化前後のペアともう一つ別の暗号化前後のペアが与えられたとき，それらが同じ鍵を用いた暗号化前後のペアかどうかを判定するアルゴリズムである．図 3 の判定アルゴリズムを利用して，相手機関独自の情報を推測する方法について考える.

まず，機関 R はプロトコルを通して得た相手機関 S の鍵 e_s で暗号化された機関 S のテーブル (f_{es}(Xs)) に注目する．このテーブルはプロトコルのステップ 4(a) で得られる．(4.1 節参照) 機関 R は機関 S との積集合が分かるとテーブル f_{es}(Xs) 内の各データが積集合に属しているか，そうでないかを求めることができる．また，その積集合に属している各暗号文の平文も求めることが可能である.

ここで，テーブル f_{es}(Xs) において，積集合に属する値を $f_{es}(h(a))$，その平文を $h(a)$，積集合に属さない値を 1101 (機関 R にとって 1101 の平文は未知) とする．以上の値を distinguish 問題の入力 g^d，g^{ad}，g^z として与える．また，入力 g^b としてハッシュ化した定義域 $h(U)$(U:属性の取り得る値の定義域) の要素から積集合に属さない値 1101 の平文の候補 $h(x) \in h(U)$ を選択し，distinguish 問題の入力 (g^d，g^{ad}，g^b，g^z) を ($h(a)$，$f_{es}(h(a))$，$h(x)$，1101) とする．ここで，全体集合 U の取り得る要素数は有限であり，かつ，全数探索が可能な大きさとすると，従来プロトコルにおいて distinguish 問題を利用して情報共有を行った機関が相手機関独自の情報を解析するとき，タプル $< x, f_e(x), y, z >$ の各要素の取り得る定義域は $DomF$ と異なる．暗号値 $f_e(x)$，z の定義域は $DomF$ であるが，x，y の定義域に関しては定義域 U のハッシュ関数適用後の値の集合 $h(U)$ となる．しかし，暗号化特性 4 より，distinguish 問題はタプル $< x, f_e(x), y, z >$ の要素 x，y，z の取り得る値の定義域が暗号の取り得る値の定義域 $DomF$ 程度大きければ，二つのタプル $< x, f_e(x), y, f_e(y) >$ と $< x, f_e(x), y, z >$ が計算的に識別ができないことを意味している．よって，相手機関独自の情報を得るための解析において，タプルの取り得る要素 x，y の定義域が定義域 $DomF$ より小さくなる (定義域が $h(U)$ になる) ため，従来プロトコルにおける distinguish 問題の困難さは定義域の大きな distinguish 問題の困難さより容易になる可能性がある．さらに，従来プロトコルにおける distinguish 問題は厳密には DDH 問題と等価でない．従来プロトコルにおける distinguish 問題はタプル $< x, f_e(x), y, z >$ の要素 x，y の定義域が小さいことから，DDH 問題の入力における原始元 g の値域が小さいとき等価といえる．例えば，以下の図 4 のように全体集合 U が a〜j の 10 個の要素を持つならば，$h(x)$ の候補は $h(a)$〜$h(j)$ となる．このとき，$h(x)$ として $h(j)$ を選択し入力 ($h(a)$，$f_{es}(h(a))$，$h(j)$，1101) から yes と判定されたならば，1101 は $h(j)$ の暗号文と推測できるため相手機関 S が独自に持つ情報が j であると機関 R は知ることができる.

つまり，相手機関が独自に持つ情報を得るには (従来プロトコルの機密性を破るには)，最悪でも，全体集合 U の要素数の回数 distinguish 問題を解けばよいことが分かる．また，従来プロトコルにおける distinguish 問題は原始元 g の値域が小さい DDH 問題と等価なので，従来プロトコルにおける distinguish 問題を解く回数分原始元 g の値域が小さい DDH 問題を解くことで従来プロトコルの機密性を破ることが可能である．本研究では，DDH 問題が

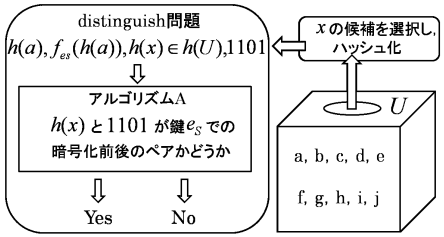

図 4 相手機関独自情報の推測方法

解けたとしても機密性を保持したまま情報を共有するプロトコル，つまり，従来プロトコルの機密性を上回るプロトコルの提案を目指す．図 5 に示すように，機密性を左右する事項である暗号解析に利用可能な情報の流出を従来より抑えることで機密性の向上を図る．

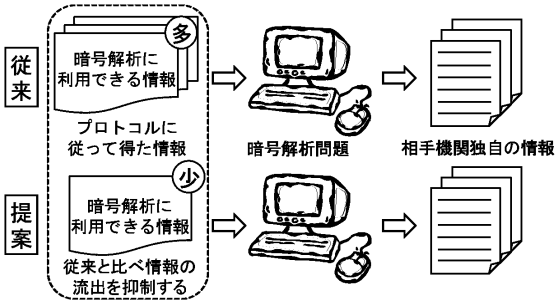

図 5 機密性向上化の方針

4. 解析に利用可能な情報流出の抑制方針

本節では，従来プロトコルの機密性を解析した結果から，機密性を向上させる方針とその方針を用いた際の機密性について述べる．

4.1 解析に利用可能な情報の流出抑制

従来プロトコルの機密性に関する議論を踏まえて，本研究で提案する機密性の向上化の方針について説明する．プロトコルの目的として，前節と同様，情報共有の一つとして，複数の集合間の積集合を求めるクエリの解決を挙げる．つまり，クエリの解決では，積集合に関しての情報（平文や暗号値など）は相手機関に必ず知られてしまう．また，プロトコルの機密性は相手機関独自の情報 (積集合に属さない部分集合) を推測することがプロトコルを行う前と比較して容易でなければ，保持される．本研究では，プロトコルを通して得た情報から相手機関独自の情報を解析困難にするため通信を行う上で，解析に利用可能な独自情報に関する暗号値などすら推測できないプロトコルを実現する方法について検討を行う．

従来プロトコルにおいて，図 3 の判定アルゴリズムの入力の中で，プロトコルを通して得られる情報は，g^d，g^{ed}，g^{eb} である．つまり，積集合の暗号化前後のペア $h(x)$，$f_e(h(x))$ と積集合に属さない要素 y の暗号値 $f_e(h(y))$ であった．そこで，解析に利用できる積集合に属さない要素の暗号値 $f_e(h(y))$ さえもプロトコルを通して流さないプロトコルについて検討する．

4.2 解析に利用可能な情報の流出を抑制した場合の機密性

解析に利用可能な情報の流出を抑制した場合の機密性がどの程度かを従来プロトコルと同様に distinguish 問題が解けたときの機密性を破る問題を解くことにかかる具体的な計算量を求めることで測る．4.1 節で述べたように解析に利用可能な情報の流出抑制を行うと，プロトコルを通して相手機関独自の情報の暗号値さえ得ることができない．また，DDH 問題と distinguish 問題の困難さが等価なことから，distinguish 問題を利用して解析に利用可能な情報の流出を抑制した場合の機密性を破る問題について考える．

従来プロトコルの機密性を破る問題においては，積集合に属している要素の暗号化前後のペアと積集合に属していない要素の暗号値がプロトコルに従って得られるため，積集合に属していない要素の暗号値の平文として，定義域 U のハッシュ値を distinguish 問題の入力として加えることで，最悪，定義域 U の要素数回 distinguish 問題を解けば，従来プロトコルの機密性を破ることが可能であった．この考察からもし，定義域 U が十分に大きくない場合，機密性が保証できない可能性がある．解析に利用可能な情報の流出抑制を考慮する場合では，distinguish 問題の入力として積集合に属している要素の暗号化前後のペアのみが与えられる．すなわち，distinguish 判定アルゴリズムを利用して，機密性を破るには，積集合に属していない要素の暗号化前後のペアが入力として必要となる．ここで，積集合に属していない要素の平文の候補は，従来プロトコルの機密性を破るときと同様に，定義域 U のハッシュ値から選択することが可能である．さらに，積集合に属していない要素の暗号値の候補を暗号の取り得る定義域 $DomF$ の中から選択し，判定アルゴリズムの入力することが可能である．以上から，選択した値と通信で得られる積集合に属している値を distinguish 問題への入力として，同じ鍵を用いた暗号化前後のペアかどうかを判定することが可能とな

る．これにより，判定アルゴリズムが yes と返したときに，定義域 U のハッシュ値から選択した値とそのときの暗号値が相手機関独自の情報の候補となる．

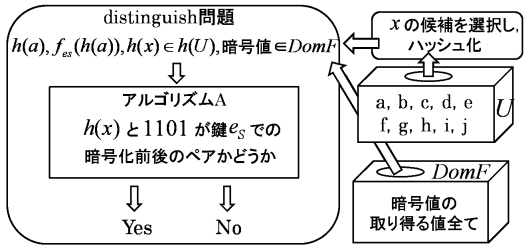

図 6 情報流出を抑制した独自情報解析

しかし，相手機関独自の情報の候補を得るには，仮に，鍵長を 2048 ビットとし，暗号値の値域も 2^{2048} とすると，相手機関独自の情報を得るには，最悪，2^{2048} 通りの入力から判定アルゴリズムの結果を得なければならない．(distinguish 問題を 2^{2048} 回解かなければならない．) このため，情報流出を抑制した場合には，従来プロトコルの独自情報解析方法では相手機関独自の情報を得ることは困難と考えられる．

独自情報の暗号文は公開せずに積集合を求めるにはデータベース内のどの情報が積集合であるかそうでないかを判断する必要がある．よって，解析に利用できる情報の流出抑制を実現したプロトコルであったとしても，積集合とそれに属さない部分を見分けるため独自情報の暗号文そのものが公開されることはないが，暗号文に関連する情報が漏えいすると考えられる．distinguish 問題を解くことが可能で，暗号文に関連する情報が漏えいしているとすると，それらを用いて独自情報の暗号文を推測することが distinguish 問題の判定アルゴリズムを利用することで可能になる．

例えば，積集合を導出するために独自情報の暗号文をさらに暗号化したものが漏えいしているとする．distinguish 問題の入力 $< x, f_e(x), y, z >$ の各要素の中で，$x, f_e(x), z$ に関する情報 (積集合に属している要素の暗号文とそれを暗号化した値，積集合に属さない要素の暗号文を暗号化した値) は得られている．また，入力の要素 x, y の定義域は独自情報の暗号文であることから $DomF$ であり，残りの要素 $f_e(x), z$ の定義域は暗号文をさらに暗号化したと考えると，暗号値の取り得る定義域 $DomF$ に等しいと考えられる．つまり，distinguish 問題の入力の中で未知の要素である，値 y を定義域 $DomF$ から選択し入力に加えることで，選択した値 y が独自情報の暗号文であるかを判定できる．独自情報の暗号文を求めるには，最悪，値 y の取り得る値域 (定義域 $DomF$) 通り判定を行わなければならない．例えば，鍵長を 2048 ビットとすると，暗号値の値域も 2^{2048} とすると 2^{2048} 通りの判定を行わなければならない．

以上から，情報の流出を抑制した場合，独自情報の暗号文を推測する問題の困難さは distinguish 問題の入力 $< x, f_e(x), y, z >$ の各要素 x, y, z の定義域が $DomF$ 程度の大きさになり，鍵長によっては，distinguish 問題を解けたとしても多項式時間内に独自情報の暗号文を求めることが困難な可能性があるため独自情報の暗号文を推測する問題の困難さは CDH 問題の困難さと等しいと考えられる．つまり，従来プロトコルと情報の流出を抑制した場合では，DDH 問題と CDH 問題の差だけ機密性に違いがあることが分かる．

4.3 解析に利用可能な情報の流出抑制に関する考察

前節から，解析に利用可能な情報の流出を抑制すると，独自情報の暗号文を公開しないため相手機関独自の情報を得る解析を行うには，まず，独自情報の暗号文を求める必要がある．その独自情報の暗号文を求める問題の困難さは CDH 問題の困難さと同等であることを示した．つまり，解析に利用できる情報の流出を抑制した場合，従来プロトコルと比べて機密性が高い可能性がある．具体的には，DDH 問題と CDH 問題の困難さだけ違いがあると考えられる．これは，従来と比べて積集合に属さない要素に関して，暗号値すら明らかにしない方針であるがゆえの違いである．

5. おわりに

本研究では，従来プロトコルに用いられている暗号が部分解読されたとき，従来プロトコルの機密性を破る手法を示した．また暗号が部分解読されたとしても，機密性を保持した情報共有を行うために独自情報の解析に利用可能な情報流出の抑制について方針を示した．今後，情報流出の抑制を行うことができれば，相手機関独自の情報を解析する困難さは distinguish 問題の入力として必要な要素の定義域が十分に大きい ($DomF$) ので，CDH 問題と等価な困難さを持っているため，従来プロトコルの機密性より高い機密性を保持していると考えられる．今後の課題として，情報流出の抑制を行えるプロトコルの提案，また，提案プロトコルにおいて，明らかにすることを許可した情報以外の情報が漏れていないかをゼロ知識証明で用いられるシミュレーション技法に基づいて証明する必要がある．

Vol.2010-AL-128 No.12
2010/1/26

参 考 文 献

1) R. Agrawal, A. Evfinievski and R. Srikant, “Information sharing across private databases,” Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data (SIGMOD 2003), pp.86-97, 2003.
2) D. Boneh, “The decision Diffie-Hellman problem,” Proceedings of the 3rd International Algorithmic Number Theory Symposium, volume 1423 of Lecture Notes in Computer Sciences, pp.48-63, 1998.
3) D. R. Stinson, “Cryptography Theory and Practice,” Third Edition, Chapman & Hall/CRC, 2006.
4) B. Schneier, “Applied Cryptography,” Second Edition, John Wiley and Sons, 1996.
5) 隅，村本，上土井，若林，“ 複数データベース間における機密性を保持した情報共有の一手法 ”，電子情報通信学会第 19 回データ工学ワークショップ・第 6 回日本データベース学会年次大会 (DEWS 2008)，2008.
6) 隅，村本，上土井，若林，“ 非公開データベース間の情報共有における機密性向上の検証 ”，第 1 回データ工学と情報マネージメントに関するフォーラム・第 7 回日本データベース学会年次大会 (DEIM 2009)，2009.
7) 辻井，岡本，“暗号のすべて〜ユビキタス社会の暗号技術〜 ”，電波新聞社，2002.
8) N. Zhang and W. Zhao, “Distributed privacy preserving information sharing,” Proceedings of the 31st VLDB Conference, pp.889-900, 2005.