プライバシーポリシー執行を保証する関数評価

佐久間 淳 †, 陸 文傑 †, 西出 隆志 †, 國廣昇 ‡

† 筑波大学 大学院 システム情報工学研究科
305-8577 茨城県つくば市天王台 1-1-1
jun@cs.tsukuba.ac.jp, riku@mdl.cs.tsukuba.ac.jp, nishide@risk.tsukuba.ac.jp

‡ 東京大学大学院新領域創成科学研究科
277-8561 柏市柏の葉 5-1-5
kunihiro@k.u-tokyo.ac.jp

プライバシーポリシー執行を保証する関数評価（FE-PPE）の新たな枠組みを提案する．評価者は個人によって提供された個人データを用いて関数評価を行い，クライアントは評価結果を取得するとする．FE-PPE は，関数評価において，2 種類のプライバシーポリシー：評価者ポリシーとクライアントポリシーの執行を保証する．評価者ポリシーは，データを用いて関数評価を行うことができる entity を，クライアントポリシーは，評価結果を得ることができる entity を制限する．個別化医療への応用の実装実験を行い，プライバシーポリシー執行により生じるオーバーヘッドは，執行しない場合と比較して 10%未満であることを示した．

Function Evaluation with Privacy Policy Enforcement

Jun Sakuma†, Wenjie Lu†, Takashi Nishide†, Noboru Kunihiro‡

†Tsukuba University
1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, JAPAN
jun@cs.tsukuba.ac.jp, riku@mdl.cs.tsukuba.ac.jp, nishide@risk.tsukuba.ac.jp

‡The University of Tokyo
5-1-5 Kashiwanoha, Kashiwa, 277-8561, JAPAN
kunihiro@k.u-tokyo.ac.jp

Abstract We propose a novel framework for secure function evaluation with privacy policy enforcement (FE-PPE). Suppose an evaluator evaluates a function with private data contributed by an individual and a client obtains the result of the evaluation. FE-PPE enforces two different kinds of privacy policies to the process of function evaluation: evaluator policy and client policy. An evaluator policy restricts entities that can conduct function evaluation with the data. A client policy restricts entities that can obtain the result of function evaluation. We demonstrate our construction with personalized medication. Experimental results show that the overhead caused by enforcing the two privacy policies is less than 10% compared to function evaluation by homomorphic encryption without any privacy policy enforcement.

1 研究背景

オンラインサービスは，我々の日常生活に不可欠なツールである．近い将来，機密性の高い情報（個人の医療記録や金融取引など）で動作するオンラインサービスが広く利用されることになる．オンラインサービスの高度な個別化を実現するためには，利用者が，プライバシー保護と利便性のバランスをコントロールできるような仕組みを構築することが重要である．本研究では，安全な関数評価を行う上で，プライバシーポリシーの「執行」に着目する．

各個人が，自分の個人情報の暗号文をクラウドストレージに保管し，サービス提供者が，安全な関数評価に基づくサービスのために，その暗号文を用いる状況を考える．我々のフレームワークでは，誰が関数評価を実行してよいか，誰が関数評価の結果を得ることができるかを，各個人がコントロールできる仕組みを実現する．プライバシーポリシー執行を保証する関数評価の典型的な例として，個別化医療の例を述べる．

次の状況を考える．患者 X は，自分の治療のために，診療所 A で医療記録が保管されている．患者 X はまた，疫学研究のために gene bank B に，遺伝子データを提供している．患者 X が，地域病院 C を訪れたときに，医師は，なんらかの薬を投薬することを決めたとする．薬は重篤な副作用を引き起こす可能性があり，薬剤の適切な投与量は，臨床的および遺伝的因子が異なる患者の間では異なる可能性がある．薬を生産する製薬会社 D は，臨床検査の結果や患者のゲノムデータに応じて適切な投薬量を評価するオンラインサービスを提供している．このため，診療所 A と gene bank B は，臨床検査の結果とゲノムデータを製薬会社と共有する必要がある．

この例では，診療所により収集された臨床データと gene bank によって収集された遺伝子データを適切に組み合わせることができれば，遺伝子検査を実現することが可能である．しかし，診療所と gene bank で行われたデータ使用のデータ提供者の同意が，製薬会社のサービスに適合しない場合には，うまく機能しない．具体的には，gene bank B が，研究目的のためだけに，遺伝的データを収集する場合には，gene bank は，患者自身の治療のためであっても，遺伝子データを提供することはできない．診療所 A と製薬会社 D は，患者のデータ交換との合意がない場合は，診療所は，どちらにも，患者の臨床データを提供することはできない．

このようなミスマッチは，使用条件がデータ提供者に対して独立に設定されていることにより生じる．一度，データ提供者が，データ提供の際に，使用条件に合意してしまうと，後で，プライバシーポリシーをコントロールすることは困難である．個別化医療の例のように，複数のデータ提供者からの個人情報の組み合わせが必要な場合には，ますます複雑になる．データ提供者が，自分自身でより柔軟に，プライバシーポリシーを設定でき，いつでも，更新することができれば，個人の機密情報に基づくオンラインサービスの利便性は劇的に向上する．

本稿では，プライバシーポリシー執行を保証する安全な関数評価（FE-PPE）の新しい枠組みを導入する．我々のフレームワークでは，高機能暗号を使用し，プライバシーポリシーの自己コントロール可能性を提供することにより，プライバシー保護と利便性のバランスを取ることが可能となる．FE-PPE のコンセプトは，以下で与えられる．

FE-PPE のステークホルダー． FE-PPE は，5 種類のステークホルダー：データ提供者，評価者（サービスプロバイダ），クライアント，認証局，クラウドストレージにより構成される．簡潔に言うと，FE-PPE は，データ提供者と評価者の間の二者間計算である．データ提供者は，関数評価のために自分自身の個人情報を入力する．評価者は，サービス提供のために用意した機密情報を入力する．上記の例では，患者がデータ提供者に対応し，病院や gene bank を通じて，秘密情報をクラウドストレージに提供する．製薬会社は評価者に対応しており，個別化医療のサービスを提供する．診療所はクライアントとして機能し，患者の治療のための個別化医療を提供する．

評価者ポリシーの執行． データ提供者は，自らのデータを準同型暗号を用いて暗号化し，評価者が，分析のためにデータをダウンロードできるように，暗号文をクラウドストレージに保管する．各暗号文に対して，2 種類のプライバシーポリシー（評価者ポリシーとクライアントポリシー）が付加される．データ提供者は，任意のタイミングで，クラウドストレージで保管された評価者ポリシーとクライアントポリシーを更新することができる．評価者は，クラウドストレージから，どの暗号文もダウンロードすることができる．評価者が暗号文に付加された評価者ポリシーを満たしている場合には，評価者は，暗号文を用いて関数評価を行うことが可能となる．満たしていない場合には，評価者は暗号文に対して何もすることはできない．例えば，データ提供者は，「政府機関のみが私のデータを用いて関数評価を行うことができる」,「誰もが，2015 年 4 月 1 日から十年間のみ，私のデータで分析を行うことができる」というような評価者ポリシーを設定することができる．評価者が，評価を行う権限がある場合には，関数評価をおこない，クライアントが結果を利用できるように，評価結果は暗号化されて，再び，クラウドストレージに保管される．関数評価は，準同型暗号で暗号化されて処理されているため，評価者は，

個人データについては何も得ることができない．

クライアントポリシーの執行．評価者により提供された解析結果を得たいクライアントは，クラウドストレージから，その暗号文をダウンロードすることができる．クライアントが，結果の暗号文に付加されるクライアントポリシーを満たしている場合，そのクライアントは，暗号文を復号し，評価結果を得ることができる．満たさない場合には，クライアントは何も得ることはできない．例えば，データ提供者は,「医師免許を持つ医師のみが結果を得ることができる」,「誰も私の承認後に結果を得ることができる」というようなクライアントポリシーを設定することができる．

ポリシーの秘匿性．FE-PPE では，入力だけでなく，評価者ポリシーとクライアントポリシーも，個人情報として扱う．例えば，クライアントポリシーが,「マンハッタンの乳癌の専門医師だけが評価結果を見ることができる」であったとする．このポリシーは，データ提供者が，マンハッタンに住んでいて，乳がんにかかりやすい女性であることを示唆している．クラウドストレージから，誰もが，自由にデータをダウンロード可能であることを考慮すると，ポリシーの秘匿性を保持する必要がある．

1.1 成果

問題および安全性の定式化．プライバシーポリシー執行を保証する安全な関数評価のための新たな枠組みを定式化する．ここでは，２種類のプラバシーポリシー（評価者ポリシーとクライアントポリシー）を導入する．前者は，誰がデータの解析を行う権限があるかを規定し，後者は，誰が，計算結果を得ることができる権限があるかを規定する．2 つのポリシーは，データ提供者により独立に付加される．このことは，ポリシーは，それぞれのデータに対して，独立に設定されることを意味している．我々の知る限り，このような定式化は，安全な関数評価の文脈では行われていない．提案プロトコルの安全性は，基となる属性ベース暗号および準同型暗号の安全性から直接示される．

高機能暗号を用いた構成．プライバシーポリシー執行を保証する関数評価手法の構成例をいくつか示した．評価できる関数のクラスは，用いる準同型暗号化の能力によって決まる．目標関数が，加法準同型暗号を用いて計算できるほど十分に簡単であるときには，ElGamal 暗号や Paillier 暗号などの「枯れた」暗号方式を使用することができる．より複雑な関数を評価するためには，完全（もしくはレベル付）準同型暗号を使用することにする．数値実験では，加法準同型暗号と完全準同型暗号の両方の実装を行っている．我々の構成では，ポリシー秘匿性を持つ暗号文ポリシー属性ベース暗号を用いている．これにより，プライバシーポリシーを秘密情報として取り扱うことが可能である．

実問題に対する実装実験．3 種類の関数評価（個別化医療：遺伝子/臨床因子を用いたロジスティック回帰による疾患リスク予測，遺伝疫学：χ^2 検定による疾患に鋭敏な遺伝子同定，機械学習：サポートベクターマシンによる分類）に対して実装実験を行った．最初の例では，modified Paillier 暗号を用い，後者の 2 つの例は，ring-LWE 仮定に基づくレベル付き準同型暗号を用いた．プライバシー執行を行う場合と行わない場合で，処理時間の比較を行った．実験結果では，プライバシーポリシー執行により生じるオーバーヘッドは，最初の例では，1.4%，二つ目の例では，2.7%，三番目の例では，6.6%である．この結果は，我々の方式の有効性を示している．

2 定式化

FE-PPE は，5 種類のステークホルダー：データ提供者，評価者，クライアント，クラウドストレージ，認証局，により構成される．データ提供者，評価者，クライアントが，主要な entity であり，クラウドストレージと認証局は，プライバシーとセキュリティ要件を実現するために導入された補助的な entity である．

データ提供者は，独立に個人データを保持し，他の entity に対して，自身のデータは秘密にしておきたい．評価者は，データ評価に用いる秘密関数を保持している．データ提供者は，自身が定めた評価者ポリシーを，評価者が満たす時に，評価者による関数評価を許可する．同時に，データ提供者は，自身が定めたクライアントポリシーを，クライアントが満たす時に，クライアントが評価結果を得ること許可する．評価者ポリシーとクライアントポリシーは，データ提供者の秘密情報である．FE-PPE で考慮する秘密情報については，表 1 を参照されたい．

2.1 概要

i 番目のデータ提供者は，個人データ $x_i \in \{0,1\}^*, (i = 1,\ldots,N)$ をクラウドストレージへ委託する．j 番目

の評価者は，秘密関数を保持する．一般性を失うことなく，j 番目の評価者は常に関数 f_j を用いると仮定する．関数 f_j は，入力として，クラウドストレージから得たデータの部分集合を受け取る．ここで，S は，部分集合のサイズとして，$f_j : \{0,1\}^{*\times S} \to \{0,1\}^*$ で定義される．原理的には，FE-PPE は，データ提供者と評価者との間の安全な関数評価として定義され，評価結果はクライアントだけに明らかにされる．

評価者とクライアントは，クラウドストレージに委託された任意のデータをダウンロードすることが許可されている．一方，データ提供者は，誰がデータ提供者のデータを用いて関数評価を行う権限があるのか，誰がデータ提供者のデータに関連付けられた関数評価の結果を得る権限があるのか，個別にをコントロールすることができる．このようなポリシー，権限の執行は，述語によって記述される．i 番目のデータ提供者の評価者ポリシーは $\mathsf{pol}_i^{eval} \in \mathcal{P}$ であるとする．j 番目の評価者の資格は，$\mathsf{eli}_j^{eval} \in \mathcal{E}$ であるとする．$\gamma : \mathcal{P} \times \mathcal{E} \to \{\mathsf{true}, \mathsf{false}\}$ を，権限が評価者ポリシーを充足させるかを判定する述語関数とする．$\gamma(\mathsf{pol}_i^{eval}, \mathsf{eli}_j^{eval}) = \mathsf{true}$ であるならば，j 番目の評価者は，データ x_i と関数 f_j により，$f_j(x_i)$ の評価をすることができる．それ以外の場合では，評価者は，x_i を用いて関数評価を行うことはできない．評価者とクライアントの権限は，認証局によって正当化される．本稿では，認証局は信頼されていると仮定する．pol_i^{eval} は**評価者ポリシー**と呼ぶ．

同様に，i 番目のデータ提供者が定めるクライアントポリシーを，$\mathsf{pol}_i^{client} \in \mathcal{P}$ で記述する．k 番目のクライアントの権限は，$\mathsf{eli}_k^{client} \in \mathcal{E}$ とする．$\gamma(\mathsf{pol}_i^{client}, \mathsf{eli}_k^{client}) = \mathsf{true}$ が成り立つならば，k 番目のクライアントは，$f_j(x_i)$ を得ることが許される．成り立たない場合には，クライアントは，何も得ることができない．pol_i^{client} を，**クライアントポリシー**と呼ぶことにする．評価者ポリシーとクライアントポリシーは、独立して設定される．

2.2 攻撃者モデルと安全性要件

データ提供者，評価者，クライアント，クラウドストレージの行動について議論する．

データ提供者（患者や銀行口座保有者など）は，クラウドストレージへ秘密データを預ける entity である．データ提供者は，他のデータ提供者の秘密データや評価者/クライアントポリシーを得るために，能動的な攻撃者として振る舞う．

表 1: 攻撃者モデル

ステークホルダー	秘匿情報	攻撃者	結託
i 番目のデータ提供者	x_i, $\mathsf{pol}_i^{\{eval,client\}}$	能動的	anyone
クラウドストレージ	—	受動的	no one
j 番目の評価者	f_j	受動的	no one
クライアント	—	能動的	anyone
認証局	—	—	no one

クラウドストレージは，データ提供者が，秘密データを格納するデータ保存機関である．評価者とクライアントは，クラウドストレージ上のデータをダウンロードすることが許可されている．クラウドストレージは，サービスプロバイダとして良い評判を維持するために，指定されたプロトコルから逸脱しないようなインセンティブを持っている．しかし，クラウドストレージは，サービス改善のための保管情報（アクセスログ解析を含む）を悪用する可能性がある．したがって，クラウドストレージが受動的攻撃者であると仮定する．

評価者は，データ提供者のデータ（遺伝子検査サービスプロバイダや信用調査サービス）を用いて，特定の関数評価を行うため，価値のある（時には機密）の知識を持っている entity である．指定された評価者ポリシーを満たす評価者だけが，関数評価を行うことが許可されている．評価者は，サービス・プロバイダーとしての良い評価を維持するために，指定されたプロトコルから逸脱しないようにインセンティブを持っている．例えば，遺伝子検査サービスプロバイダや信用調査サービスでは，法律によって信頼性の高い分析結果を提供することが規定されている．評価者が，個人情報を得るために，指定されたプロトコルを逸脱しないと仮定することができる．このため，評価者は，受動的攻撃者である．安全性評価において，評価者は，データ提供者の個人データを得るために，他のステークホルダーと結託しないことを仮定する．

クライアントは，データ提供者のデータを用いた関数評価を要求する entity である．指定されたクライアントポリシーを満たすクライアントだけが，評価結果を得ることができる．クライアントは，能動的攻撃者であるとみなし，誰とも結託することができる．クライアントは，他のデータ提供者や評価者の秘密のデータを得るために，悪意を持って不正な要求をすることを許す．

最後に，FE-PPE によって達成すべき安全性要件についてまとめる．

- データ秘匿性: i 番目のデータ提供者以外は，x_i を知ることはできない.
- 関数秘匿性: j 番目の解析者以外は，f_j を知ることはできない.
- 評価者ポリシー秘匿性: i 番目のデータ提供者以外は，pol_i^{eval} を知ることはできない.
- クライアントポリシー秘匿性: i 番目のデータ提供者以外は，pol_i^{client} を知ることはできない.
- 評価整合性: pol_i^{eval} を満たす評価者だけが，x_i の準同型暗号文を得ることができ，x_i の暗号文を用いて，$f(x_i)$ の暗号文を得ることができる.
- 強解析結果秘匿性: pol_i^{client} を満たすクライアントのみが，$f(x_i)$ を得ることができる.
- 弱解析結果秘匿性: i 番目のデータ提供者と pol_i^{client} を満たすクライアントのみが，$f(x_i)$ を得ることができる.
- データフロー秘匿性: クラウドストレージは，データ提供者から評価者，評価者からクライアントへのデータの流れを特定することができない.

3 構成要素

FE-PPE の構成において，modular approach を取る．提案方式は，4 章で示すように共通鍵暗号 (SKE)，準同型暗号 (HE)，属性ベース暗号 (ABE) により構成されている．

最初に，HE の文法を与える．4 つのアルゴリズム：KeyGen, Enc, Eval, Dec により構成される．

KeyGen$(1^\kappa) \to (\mathbf{pk}, \mathbf{dk})$**:** 入力としてセキュリティパラメタ κ を受け取り，公開鍵 $\mathbf{pk}$ と復号鍵 $\mathbf{dk}$ を出力する．

Enc$(m, \mathbf{pk}) \to \mathsf{CT}$**:** 入力としてデータ m，公開鍵 $\mathbf{pk}$ を受け取り，暗号文 CT を出力する．

Eval$(\mathsf{CT}, \mathbf{pk}, f) \to y$**:** 入力として（複数の）暗号文 CT，公開鍵 $\mathbf{pk}$，関数 f を受け取り，y を出力する．

Dec$(\mathsf{CT}, \mathbf{dk}) \to z$**:** 入力として暗号文 CT，復号鍵 $\mathbf{dk}$ を受け取り，z を出力する．

復号の正当性は，確率 1 で，$z = f(m)$ が成り立つことで与えられる．準同型暗号の KeyGen アルゴリズムを，HE.KeyGen で記述する．他も同様である．いくつかのアプリケーションに対して，加法準同型暗号 [7] では，不十分であり，完全（もしくは，somewhat）準同型暗号 [4] が必要となる．

我々の FE-PPE の構成では，暗号文ポリシー ABE (CP-ABE) を用いる．ABE は，4 つのアルゴリズム Setup, KeyGen, Enc, Dec により構成される．また，属性空間を，Σ であらわす．

Setup$(1^\kappa) \to (\mathsf{MPK}, \mathsf{MSK})$**:** 入力としてセキュリティパラメタ κ を受け取り，属性ベース暗号用のマスター公開鍵 MPK，マスター秘密鍵 MSK を出力する．

KeyGen$(\mathsf{MPK}, \mathsf{MSK}, A) \to \mathbf{sk}^{(A)}$**:** 入力としてマスター秘密鍵 MSK，マスター公開鍵 MPK，属性 $A \subseteq \Sigma$ を受け取り，復号鍵 $\mathbf{sk}^{(A)}$ を出力する．

Enc$(m, \mathsf{MPK}, \Psi) \to \mathsf{CT}$**:** 入力としてメッセージ m，マスター公開鍵 MPK，ポリシー Ψ を受け取り，暗号文 CT を出力する．

Dec$(\mathsf{CT}, \mathbf{sk}^{(A)}) \to z$**:** 入力として暗号文 CT と復号鍵 $\mathbf{sk}^{(A)}$ 受け取り，z を出力する．

復号の正当性条件は，“$z = m$ if $\Psi(A) = \mathtt{true}$ and $z = \perp$ otherwise” で与えられる．

この研究では，暗号文ポリシー ABE (CP-ABE) の中でも，特に，[5, 6] のように，暗号文ポリシーを秘匿することができる方式を用いる．

最後に，共通鍵暗号 SKE の文法を示す．SKE は，3 つのアルゴリズム KeyGen, Enc, Dec で構成される．

KeyGen$(1^\kappa) \to K$**:** 入力としてセキュリティパラメタ κ を受け取り，秘密鍵 K を出力する．

Enc$(m, K) \to \mathsf{CT}$**:** 入力としてデータ m，秘密鍵 K を受け取り，暗号文 CT を出力する．

Dec$(\mathsf{CT}, K) \to z$**:** 入力として CT と秘密鍵 K を受け取り，z を出力する．

4 提案プロトコル：FE-PPE

この章では，我々の提案プロトコルを示す．まず，提案プロトコルで用いるアルゴリズムの文法を示し，各ステークホルダーが，どのアルゴリズムをどのように用いるかを詳細に示す．さらに，各アルゴリズムの具体的な構成法を示す．ここで，構成要素として，共通鍵暗号，準同型暗号，属性ベース暗号などの「枯れた」暗号技術を採用している．最後に，利便性を高めた方式を示す．

4.1 FE-PPE で用いるアルゴリズム

Σ を属性空間とする．Σ は，評価者用の属性 Σ_E とクライアント用の属性 Σ_C に分けることができる．我々のプロトコルは，8 つのアルゴリズム Setup, Hom-KeyGen, SKE-KeyGen, ABE-KeyGen, DataEnc, KeyEncap, Evaluate, Dec により構成される．

Setup$(1^\kappa) \to (\mathsf{MPK}, \mathsf{MSK})$**:** 入力としてセキュリティパラメタ κ を受け取り，属性ベース暗号用マスター公開鍵 MPK とマスター秘密鍵 MSK を出力する．

Hom-KeyGen$(1^\kappa) \to (\mathbf{pk}, \mathbf{dk})$**:** 入力としてセキュリティパラメタ κ を受け取り，準同型暗号用公開鍵 $\mathbf{pk}$ と復号鍵 $\mathbf{dk}$ を出力する．

SKE-KeyGen$(1^\kappa) \to K$**:** 入力としてセキュリティパラメタ κ を受け取り，共通鍵暗号用秘密鍵 K を出力する．

ABE-KeyGen$(\mathsf{MPK}, \mathsf{MSK}, A) \to \mathbf{sk}^{(A)}$**:** 入力として属性ベース暗号用のマスター公開鍵 MPK，マスター秘密鍵 MSK，属性 $A \subseteq \Sigma = \Sigma_E \cup \Sigma_C$ を受け取り，復号鍵 $\mathbf{sk}^{(A)}$ を出力する．

DataEnc$(x_i, \mathbf{pk}, K) \to \mathsf{CT}_i$**:** 入力としてデータ x_i，準同型暗号用の暗号化鍵 $\mathbf{pk}$ 共通鍵暗号用の秘密鍵 K を入力として受け取り，暗号文 CT_i を出力する．

KeyEncap$(K^E, \Psi^E, \mathbf{dk}, K^C, \Psi^C) \to (eK^E, eK^C, eK^{\mathbf{dk}})$**:** 入力として秘密鍵 K^E, K^C, 準同型暗号用の復号鍵 $\mathbf{dk}$，評価者ポリシー Ψ^E，クライアントポリシー Ψ^C を入力として受け取り，カプセル化された鍵 $eK^E, eK^C, eK^{\mathbf{dk}}$ を出力する．

Evaluate$(\{\mathsf{CT}\}_{i=1}^n, \mathbf{sk}^{(E)}, f, \mathbf{pk}, eK^E) \to \perp$ **or** CTwA**:** 入力として暗号文（の集合）$\{\mathsf{CT}\}$，属性 $E \subseteq \Sigma_E$ に対応する属性ベース暗号用の秘密鍵 $\mathbf{sk}^{(E)}$，関数 f，公開鍵 $\mathbf{pk}$，カプセル化された鍵 eK^E を受け取り，解析結果の暗号文 CTwA か $\perp$ を出力する．

Dec$(\mathsf{CTwA}, \mathbf{sk}^{(C)}, eK^C, eK^{\mathbf{dk}}) \to z$**:** 入力として暗号文 CTwA，属性 $C \subseteq \Sigma_C$ に対応する属性ベース暗号用の秘密鍵 $\mathbf{sk}^{(C)}$，カプセル化された鍵 $eK^C, eK^{\mathbf{dk}}$ を受け取り，メッセージ z を出力する．

4.2 FE-PPE の具体的な記述

各ステークホルダーが，どのようにアルゴリズムを用いるかを述べる．図 1 に，データの流れを示す．

データ提供者の一人

最初に，Hom-KeyGen アルゴリズムと SKE-KeyGen を動作させ，$1^\kappa \to (\mathbf{dk}, K^E, K^C)$ を得る．次に，評価者ポリシーとクライアントポリシー Ψ^E, Ψ^C を定める．ついで，KeyEncap アルゴリズムを動作させ，復号鍵のカプセル化を行う：

$$(K^E, \Psi^E, \mathbf{dk}, K^C, \Psi^C) \to (eK^E, eK^C, eK^{\mathbf{dk}}).$$

最後に，$(eK^E, eK^C, eK^{\mathbf{dk}})$ をクラウドストレージにアップロードする．

i 番目のデータ提供者 $(1 \le i \le n)$

データ x_i を持っているとする．DataEnc アルゴリズムを動作させ，暗号文を得る：$(x_i, \mathbf{pk}, K^E) \to \mathsf{CT}_i$. CT_i を，クラウドストレージにアップロードする．

属性 $\mathtt{Att}^E$ を持つ評価者

準備として，認証局と通信を行う．属性 $\mathtt{Att}^E$ に関連した復号鍵 $\mathbf{sk}^{(E)}$ を受け取る．鍵は，ABE-KeyGen アルゴリズムにより生成される．暗号化されたデータ $\{\mathsf{CT}_i\}_{i=1}^n$ とカプセル化された鍵 eK^E をクラウドストレージからダウンロードする．次に，Evaluate アルゴリズムを動作させる．

$$(\{\mathsf{CT}\}, \mathbf{sk}^{(E)}, f, \mathbf{pk}, eK^E) \to \begin{cases} \mathsf{CTwA} & \text{if } \Psi^E(\mathtt{Att}^E) = \mathtt{true} \\ \perp & \text{otherwise.} \end{cases}$$

出力が $\perp$ でなければ，CTwA をアップロードする．

属性 $\mathtt{Att}^C$ を持つクライアント

準備として，認証局と通信を行う，属性 $\mathtt{Att}^C$ に関連した復号鍵 $\mathbf{sk}^{(C)}$ を受け取る．鍵は，ABE-KeyGen アルゴリズムにより生成される．暗号化された結果 CTwA とカプセル化された鍵 eK^C と $eK^{\mathbf{dk}}$ をクラウドストレージからダウンロードする．次に，Dec アルゴリズムを動作させる．

$$(\mathsf{CTwA}, \mathbf{sk}^{(C)}, eK^C, eK^{\mathbf{dk}}) \to \begin{cases} f(x_1, \ldots, x_n) & \text{if } \Psi^C(\mathtt{Att}^C) = \mathtt{true} \\ \perp & \text{otherwise.} \end{cases}$$

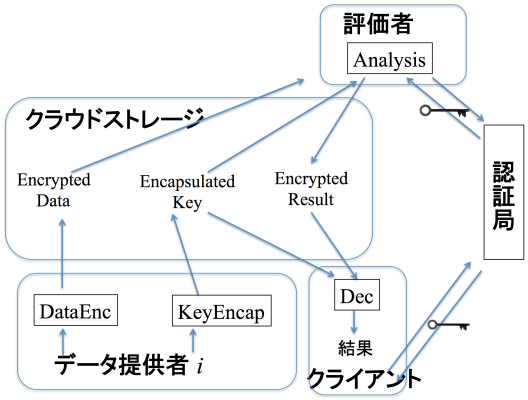

図 1: Dataflow in FE-PPE

評価者が，$\Psi^E(\mathtt{Att}) = \mathtt{true}$ をみたす属性を持つとき，準同型演算を行う権限を有する．このとき，Evaluate アルゴリズムを動作させることにより，CTwA を得ることができる．同様に，クライアントが，$\Psi^C(\mathtt{Att}) = \mathtt{true}$ をみたす属性を持つとき，解析結果を得る権限を有する．このとき，Dec アルゴリズムを動作させることにより，$f(x_1, \ldots, x_n)$ を得ることができる．これより，評価者ポリシー，クライアントポリシーの両方を満たすとき，クライアントは，$f(x_1, \ldots, x_n)$ を得ることができる．

4.3 各アルゴリズムの具体的な構成法

各アルゴリズムの具体的な構成法を与える．全てのアルゴリズムは，3 章で議論した SKE, HE, ABE 中のアルゴリズムにより構成させる．記述の簡単化のため，データ x の準同型暗号による暗号文を $[x]$ で記述する．

Setup$(1^\kappa) \to (\mathsf{MPK}, \mathsf{MSK})$**:**
1. $(\mathsf{MPK}, \mathsf{MSK}) \leftarrow \mathsf{ABE.Setup}(1^\kappa)$.

Hom-KeyGen$(1^\kappa) \to (\mathbf{pk}, \mathbf{dk})$**:**
1. $(\mathbf{pk}, \mathbf{dk}) \leftarrow \mathsf{HE.KeyGen}(1^\kappa)$.

SKE-KeyGen$(1^\kappa) \to K$**:**
1. $K \leftarrow \mathsf{SKE.KeyGen}(1^\kappa)$.

ABE-KeyGen$(\mathsf{MPK}, \mathsf{MSK}, A) \to \mathbf{sk}^{(A)}$**:**
1. $\mathbf{sk}^{(A)} \leftarrow \mathsf{ABE.KeyGen}(\mathsf{MPK}, \mathsf{MSK}, A)$.

DataEnc$(x_i, \mathbf{pk}, K^E) \to \mathsf{CT}_i$**:**
1. $[x_i] \leftarrow \mathsf{HE.Enc}(x_i, \mathbf{pk})$
2. $\mathsf{CT}_i \leftarrow \mathsf{SKE.Enc}([x_i]; K^E)$

KeyEncap$(K^E, \Psi^E, \mathbf{dk}, K^C, \Psi^C) \to (eK^E, eK^C, eK^{\mathbf{dk}})$**:**
1. $eK^E \leftarrow \mathsf{ABE.Enc}(K^E||K^C; \Psi^E)$
2. $eK^C \leftarrow \mathsf{ABE.Enc}(K^C; \Psi^C)$
3. $eK^{\mathbf{dk}} \leftarrow \mathsf{ABE.Enc}(\mathbf{dk}; \Psi^C)$

Evaluate$(\{\mathsf{CT}\}, \mathbf{sk}^{(E)}, f, \mathbf{pk}, eK^E) \to \perp$ **or** CTwA**:**
1. $K^E||K^C$ or $\perp \leftarrow \mathsf{ABE.Dec}(eK^E; \mathbf{sk}^{(E)})$
2. $[x_i] \leftarrow \mathsf{SKE.Dec}(\mathsf{CT}_i; K^E)$
3. $[f(x_1, \ldots, x_n)] \leftarrow \mathsf{HE.Eval}(\{[x_i]\}; f, \mathbf{pk})$
4. $\mathsf{CTwA} \leftarrow \mathsf{SKE.Enc}([f(x_1, \ldots, x_n)]; K^C)$

Dec$(\mathsf{CTwA}, \mathbf{sk}^{(C)}, eK^C, eK^{\mathbf{dk}}) \to z$**:**
1. $K^C \leftarrow \mathsf{ABE.Dec}(eK^C; \mathbf{sk}^{(C)})$
2. $\mathbf{dk} \leftarrow \mathsf{ABE.Dec}(eK^{\mathbf{dk}}; \mathbf{sk}^{(C)})$
3. $[f(x_1, \ldots, x_n)] \leftarrow \mathsf{SKE.Dec}(\mathsf{CTwA}; K^C)$
4. $z \leftarrow \mathsf{HE.Dec}([f(x_1, \ldots, x_n]; \mathbf{dk})$

我々の構成は，ハイブリッド暗号に基づいている．全てのデータは，Evaluate 実行中以外は，共通鍵暗号で暗号化されている．我々のプロトコルでは，権限のある評価者，クライアントへの共通鍵暗号および準同型暗号用の秘密鍵の配送に属性ベース暗号を用いるところに特徴がある．FE-PPE の正確な記述は，4.2 章と 4.3 章の記述を組み合わせることにより得られる．FE-PPE は，SKE, HE, ABE のみから構成されることに注意されたい．

提案方式は，強解析結果秘匿性とデータフロー秘匿性以外の安全性をみたしている．スペースの都合上，安全性証明および二つの安全性を満たすように安全性を高めた方式は，full version で与える．

4.4 しきい値暗号版 FE-PPE

通常の加法準同型暗号を**しきい値**加法準同型暗号に置き換えたプロトコルを考える．プロトコルの説明の前に，遺伝子解析による個別化医療への応用について説明する．次のような状況を考える．医者と患者は，両者が合意した時のみに，患者の診断結果を知りたいと考えている．これは，医者（もしくは患者）は，他者の協力なしに，単独では診断結果を見ることができないことを意味している．これにより，医者による患者の同意のない診断結果の盗み見や，診断結果の一人歩きによる患者の不理解，などを防ぐことが可能となる．

通常の FE-PPE に対して，一つのアルゴリズムを修正し，一つのアルゴリズムを追加することにより，所望のプロトコルを得ることができる．違いを明確にするために，異なる部分を中心に記す．しきい値準同型暗号中のアルゴリズム X を THE.X で記す．しきい値準同型暗号の文法に関しては，スペースの都合上省略する．$(2, 2)$ しきい値版のみを記しているが，容易に (k, n) しきい値版への拡張が可能である．n 人が復号処理に関係し，k 人以上が復号処理に協力すれば，解析結果を得ることができる．

KeyEncap$(K^E, \Psi^E, \{(\mathbf{dk}_j, K^{C_j}, \Psi^{C_j})\}_{j=1}^2)$
$\to (\{eK^{E_j}\}_{j=1}^2, \{eK^{C_j}\}_{j=1}^2, \{eK^{\mathbf{dk}_j}\}_{j=1}^2)$:
For $j = 1, 2$
1. $eK^{E_j} \leftarrow \mathsf{ABE.Enc}(K^E||K^{C_j}; \Psi^E)$
2. $eK^{C_j} \leftarrow \mathsf{ABE.Enc}(K^{C_j}; \Psi^{C_j})$
3. $eK^{\mathbf{dk}_j} \leftarrow \mathsf{ABE.Enc}(\mathbf{dk}_j; \Psi^{C_j})$

Evaluate$(\{\mathsf{CT}\}, \mathbf{sk}^{(E)}, f, \mathbf{pk}, \{eK^{E_j}\}_{j=1}^2) \to \perp$ **or** CTwA**:**
1. For $j = 1, 2$
   $K^E||K^{C_j}$ or $\perp \leftarrow \mathsf{ABE.Dec}(eK^{E_j}; \mathbf{sk}^{(E)})$
2. $[x_i] \leftarrow \mathsf{SKE.Dec}(\mathsf{CT}_i; K^E)$
3. $[f(x_1, \ldots, x_n)] \leftarrow \mathsf{THE.Eval}(\{[x_i]\}; f, \mathbf{pk})$
4. For $j = 1, 2$
   $\mathsf{CTwA}_j \leftarrow \mathsf{SKE.Enc}([f(x_1, \ldots, x_n)]; K^{C_j})$

P-Dec$(\mathsf{CTwA}_j, \mathbf{sk}^{C_j}, eK^{C_j}, eK^{\mathbf{dk}_j}) \to \tau_j$**:**
1. $K^{C_j} \leftarrow \mathsf{ABE.Dec}(eK^{C_j}; \mathbf{sk}^{(C_j)})$
2. $\mathbf{dk}_j \leftarrow \mathsf{ABE.Dec}(eK^{\mathbf{dk}_j}; \mathbf{sk}^{(C_j)})$
3. $[f(x_1, \ldots, x_n)] \leftarrow \mathsf{SKE.Dec}(\mathsf{CTwA}_j; K^{C_j})$
4. $\tau_j \leftarrow \mathsf{THE.P\text{-}Dec}([f(x_1, \ldots, x_n]; \mathbf{dk}_j)$

F-Dec$(\{\tau_j\}) \to z$**:**
1. $z \leftarrow \mathsf{THE.F\text{-}Dec}(\tau_1, \tau_2)$

5 計算機実験

3 つの異なるアプリケーション（ロジスティック回帰による疾患のリスク予測，サポートベクターマシンによる予測，χ^2 検定による統計的仮説検定）に対して，通信量および計算量の計算機実験による評価を行った．最初のアプリケーションのみ結果を示す．

それぞれのアプリケーションに対して計算時間のオーバーヘッドを調べた．実験では，3 台の計算機（Core i7 2.3 GHz CPU, 16GB RAM）を用いた．1 台はデータ提供者とクライアントの役割を，1 台はクラウドストレージの役割を，最後の 1 台は評価者の役割を行った．評価結果は，クライアントに伝達され，クライアントの計算機によって復号される．通信は，100 Mbit/s のイーサネット上で行われた．すべてのプログラムは，C/C++で実装され，完全準同型暗号として，HElib [4] を利用し，加法準同型暗号として，Paillier 暗号 [7]，しきい値加法準同型暗号として，modified Paillier 暗号 [2] を利用した．属性ベース暗号として，[6] で提案された方式を用いた．80 ビット（以上の）セキュリティを保証するために，各暗号方式のパラメタを調整した．より正確には，modified Paillier 暗号では，2048 ビットの鍵を使用した．ring-LWE に基づく準同型暗号では，格子の次元を $n = 8192$，メッセージ空間を $t = 20,011$，関数の深さの上限値を $L = 3$ と設定した．これは，

128 ビットセキュリティを実現している [3]．ABE では，Barreto-Naehrig 曲線 (ECBN254) 上で定義され最適 Ate ペアリングを使用した [1]．この方式は，126 ビットセキュリティを持つ．

すべての実験において，プロトコルを処理するために必要な全通信時間と全計算時間を計測した．通信時間 (Comm.)，準同型演算 (Hom.：データの暗号化，$f(x)$ 評価も含む)，属性ベース暗号 (ABE: ABE.Enc and ABE.Dec)，共通鍵暗号 (AES:SKE.Enc and SKE.Dec) に要する計算時間を計測した．

プライバシーポリシー執行を適用しない場合には，関数評価に要する時間は（通信時間を除くと），Hom. のみとなる．これを，規準とする．通信時間を除いたプロトコル実行に必要な時間は，Hom.+ AES+ABE となる．したがって，FE-PPE によるオーバーヘッドは，$\mathrm{R1} = 1 - \frac{\text{Hom.}}{\text{Hom. + AES + ABE.}}$ で与えられる．実際には、通信時間が全体の処理時間の大部分を要する．通信時間を含むオーバーヘッドを，$\mathrm{R2} = 1 - \frac{\text{Comm. + Hom.}}{\text{Comm. + Hom. + AES + ABE}}$ で与える．

5.1 応用１：疾患リスク評価

$\boldsymbol{x}^G = (x_1^G, \ldots, x_{d^G}^G)$ と $\boldsymbol{x}^C = (x_1^C, \ldots, x_{d^C}^C)$ を，それぞれ，被験者の遺伝子データ及び臨床データを表すベクトルとする．1 章で示した個別化医療のシナリオと同様に，遺伝子データと臨床データは，別々に gene bank と診療所に保存されているとする．この実験では，疾患リスク評価を FE-PPE で実現した時のオーバーヘッドを調べる．ロジスティック関数 $\sigma(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}) = 1/(1+e^{-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}})$ は，広く疾患リスク予測 [8] で使用されている．ここで，$\boldsymbol{x} = (\boldsymbol{x}^C||\boldsymbol{x}^G)$ and $\boldsymbol{w} = (\boldsymbol{w}^C||\boldsymbol{w}^G)$ である．ロジスティック関数が全単射であるため，クライアントは $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ からリスク値を得ることができる．

この実験では，4.4 章で記述したプロトコルを用い，準同型暗号として，$(2,2)$ しきい値 Paillier 暗号化を使用した．データ提供者は，復号鍵の一つを保持し，クライアントは，もう一つの復号鍵を保持する．$\boldsymbol{x}^G$ と $\boldsymbol{x}^C$ は別々に暗号化され，クラウドストレージに保存される．評価者は，秘密に $\boldsymbol{w}^G$ と $\boldsymbol{w}^C$ を保持しており，$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ を加法準同型暗号を用いて評価する．$\boldsymbol{x}$ と $\boldsymbol{w}$ の次元を D とする．実験では、D は 100 から 3200 まで変化させた．

実験結果を表 2 に示す．表には，通信時間 (Comm.)，準同型計算 (Hom.)，ABE.Enc と ABE.Dec(ABE)，SKE.Enc と SKE.Dec（AES）も記載した，通信時間を考慮しないオーバーヘッド (R1)，通信時間を考慮したオーバーヘッド (R2) を記載した．通信時間を考慮しない場合のオーバーヘッドは，$3.7\% - 50.4\%$ となる．通信時間を考慮する場合は，10% 以下と小さくなる．この応用に対しては，プライバシーポリシーの執行をしても，十分効率的である．

表 2: Evaluation cost of disease risk prediction

dim.	Comm.	Hom.	ABE	AES	R1	R2
100	3.6s	0.27s	0.27s	2.24ms	50.4%	6.6%
201	11.2s	0.54s	0.28s	4.53ms	34.5%	2.4%
400	13.3s	1.05s	0.26s	8.25ms	20.2%	1.8%
800	11.9s	2.16s	0.26s	16.60ms	11.3%	1.4%
1600	25.0s	4.45s	0.27s	33.89ms	6.4%	1.0%
3200	40.1s	8.44s	0.27s	66.14ms	3.8%	0.7%

謝辞

本研究は，JST CREST「ビッグデータ統合利活用のための次世代基盤技術の創出・体系化」領域におけるプロジェクトおよび科学研究費 24680015, 26330151, 26540003, 公益財団法人倉田記念日立科学技術財団倉田奨励金の助成を受けた．

参考文献

[1] J. L. Beuchat, J. E. González-Díaz, S. Mitsunari, E. Okamoto, F. Rodríguez-Henríquez, and T. Teruya, “High-speed software implementation of the optimal ate pairing over Barreto-Naehrig curves,” in Proc. of Pairing 2010, pp. 21–39, 2010.

[2] E. Bresson, D. Catalano, and D. Pointcheval, “A simple public-key cryptosystem with a double trapdoor decryption mechanism and its applications,” in Proc. of ASIACRYPT2003, pp. 37–54, 2003.

[3] C. Gentry, S. Halevi, and N. P Smart, “Homomorphic evaluation of the AES circuit,” in Proc. of CRYPTO2012, pp. 850–867, 2012.

[4] S. Halevi and V. Shoup, “Algorithms in HElib,” in Proc. of CRYPTO2014 Part I, pp. 554–571, 2014.

[5] J. Katz, A. Sahai, and B. Waters, “Predicate encryption supporting disjunctions, polynomial equations, and inner products,” *J. Cryptology*, Vol. 26, No. 2, pp. 191–224, 2013.

[6] T. Nishide, K. Yoneyama, and K. Ohta, “Attribute-based encryption with partially hidden encryptor-specified access structures,” in Proc. of ACNS2008, pp. 111–129, 2008.

[7] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” in proc. of EUROCRYPT’99, pp. 223–238, 1999.

[8] A. Ziegler, I. R. König, and F. Pahlke, *A Statistical Approach to Genetic Epidemiology: Concepts and Applications, with an e-learning platform*, John Wiley & Sons, 2010.