クラウドストレージにおける実用的検索可能暗号の実現

代表研究者	西出 隆志	筑波大学 システム情報系 准教授
共同研究者	櫻井 幸一	九州大学 大学院システム情報科学研究院 教授
共同研究者	菅 孝徳	九州大学 システム情報科学府 修士 2 年(現 NEC 勤務)

1 はじめに

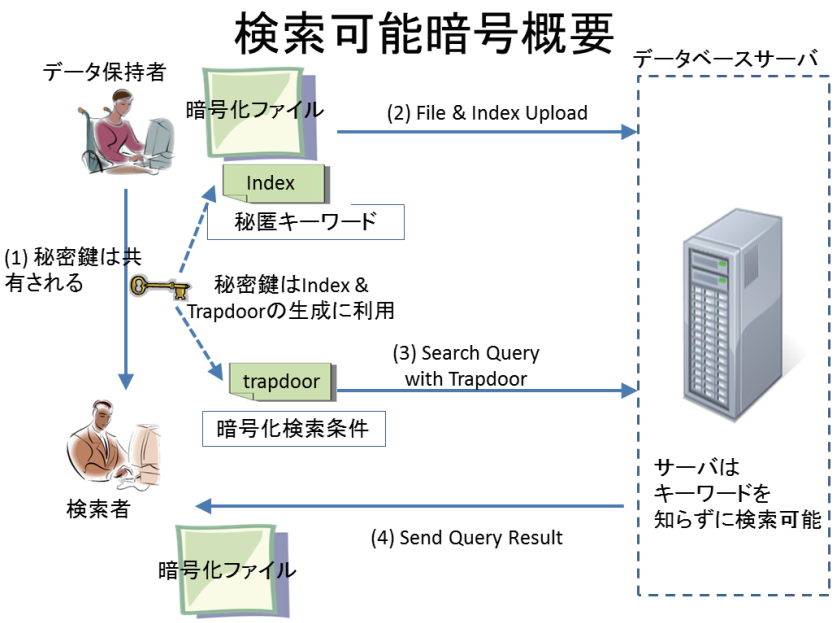
本課題で取り組んだ研究の目的とその成果について述べる.

2 研究内容

2-1 研究目的と概要

電気通信技術の発達により，どこでも計算資源へのアクセスを可能とする「クラウドコンピューティング」が普及しつつある．これに伴い，データをクラウドサーバ上に置き，どこからでもデータアクセス可能となることが期待されている．この技術は一般にクラウドストレージと呼ばれている．大量のデータをどの端末からでもアクセスできる，また容易にデータを複数のデータ操作者の間で共有できるなどの利点を有している．しかしながら，データを外部のサーバ上に置くというアウトソーシングの形態を取るということから，データの秘匿性に関する問題が生じうる．特にビジネスや公共事業での利用を考えた場合，企業秘密や個人情報に関するデータを素のままでクラウドストレージに保存しておくのはリスクが高く，また一度データが管理者のミスやソフトウェアの脆弱性に起因して漏洩してしまった場合，データを回収するのはほぼ不可能な現状がある.

単純な対応策としてデータの暗号化が考えられる．データを暗号化し，暗号鍵を適切に管理することで万が一，クラウドサーバが攻撃者に侵入されたり，クラウドサーバの管理者に悪意があったとしても被害を最小限に抑えることが期待できる.



しかしながらデータを暗号化することで安全性を得ることができるが，それと同時に弊害も生じうる．その代表的な例としてキーワード検索機能が挙げられる．クラウドコンピューティングの利点としてデータのアウトソースのみならず，計算処理のアウトソースも同時に考えられる. しかしながら，データを暗号化してしまうことによって，利用者は指定したキーワードの検索処理をクラウドサーバにアウトソースすることができなくなってしまい，指定したキーワードに関するデータのみをダウンロードするということが困難となる．またデータ利用者が毎回全ての暗号データをクラウドストレージよりダウンロードし，復号を行ってから情報を利用することも考えられるが，データサイズが巨大なとき，クラウドストレージを利用するメリットが失われてしまうということも考えられる.

そこで本研究ではデータを暗号化してもなお，キーワードすら秘匿したままでキーワード検索処理をクラウドサーバにアウトソース可能で，キーワード検索にヒットしたファイルのみをクラウドサーバが利用者に返す機能を持つ検索可能暗号について研究調査を行った．実用的な検索可能暗号が利用可能となればクラウ

ドストレージの利便性とデータの秘匿性を両立させることが可能となる意義がある.

2-2 関連研究と本研究の経緯

本研究では既存方式と比較して，より柔軟な検索条件の指定が可能な検索可能暗号の構成に取り組んだ．ここでいう”柔軟”の意味することはキーワードが完全に一致した場合のみ検索がヒットするだけでなく(完全一致検索などと呼ばれる),利用者が指定したキーワードに類似する他の暗号化キーワードも検索ヒット可能であることを意味している.

ここでまず検索可能暗号の既存の研究動向について簡単に述べる.

Song らが最初の検索可能暗号技術を提案した[6]．Song らの提案は共通鍵暗号方式を利用したもので検索方法しては完全一致のみをサポートしている．また Goh はブルームフィルタ(Bloom filter)と呼ばれるデータ構造を用いた共通鍵暗号方式に基づく構成を提案している[7]．Goh の方式でも検索方法は完全一致のみがサポートされている状況である.

公開鍵暗号方式に基づく検索可能暗号としては ID ベース暗号を用いた Boneh らの提案が最初である[8]．また匿名 ID ベース暗号と検索可能暗号の関係が Abdalla らによって形式化されている[9]．これらはやはり完全一致検索のみをサポートしている状況である.

より進んだ検索機能として Golle らは複数の検索条件を AND で結合して指定することが可能な方式を提案した[10]．しかし個々の検索条件では完全一致のみが可能である.

通常検索可能暗号ではクラウドサーバ側で保存された全ての暗号ファイルをチェックして検索にヒットしているか否か調べる必要がある．そのため大規模なクラウドストレージの環境では効率性の問題が生じる可能性がある．このようなサーバ側での全ファイル探索処理を回避し，効率を向上するための技術として Bellare らは確定的公開鍵暗号による検索可能暗号を提案している[11].

また Li らは完全一致だけでなく類似検索も可能とする共通鍵暗号に基づいた方式を提案した[12].しかしながらサーバ側で保存すべきデータ量が大幅に増加するという欠点を持っていた.

Sedghi[13]らは完全一致だけでなく，キーワード検索条件としてワイルドカードも指定可能な公開鍵暗号方式に基づく構成を提案している(より具体的にはペアリングと呼ばれる数学的な演算を用いている)．ここでいうワイルドカードとはどのような文字にも相当することが可能な検索条件の指定の中で用いる特別な文字である.Sedghi らの方式ではクラウドサーバに検索条件のどこでワイルドカードを利用しているかは通知しなければならない制限がある.

また近年，マイクロソフトリサーチでもクラウド暗号というプロジェクト名でクラウドコンピューティング環境に適した暗号要素技術の重点テーマとして検索可能暗号の開発に取り組んでいる[14].

さらに検索可能暗号を利用した製品も少なからず販売されている[15]．しかしながら，これら製品の技術的詳細は明らかにされておらず，安全性が厳密に研究された技術が望ましいと言える.

我々はこれらの関連研究の調査を進める中で，検索可能暗号として既存研究では指定できる検索条件が限られたもの，特に完全一致での指定のみをサポートしているものが大半であるという結果を得た．このような完全一致のみがサポートされている状況では,例えば 2011 年の日付を持つファイルを探すための検索条件として”2011/01/01”から”2011/12/31”の複数のキーワードを指定することが必要となり，クラウドサーバ側での処理時間が長くなる，またクライアントサーバ間のデータ通信量が多くなるという欠点があることに着目した.

これに対し，例えば類似検索機能として”2011/??/??”を検索条件として指定すると”2011/06/05”や”2011/12/10”のようなキーワードがヒットすることが可能となれば処理時間の短縮が期待できる.

このような完全一致以外の検索条件の指定を可能とすることで，より使い勝手がよく実用レベルで利用可能な検索可能暗号が提案できないかと考えた.

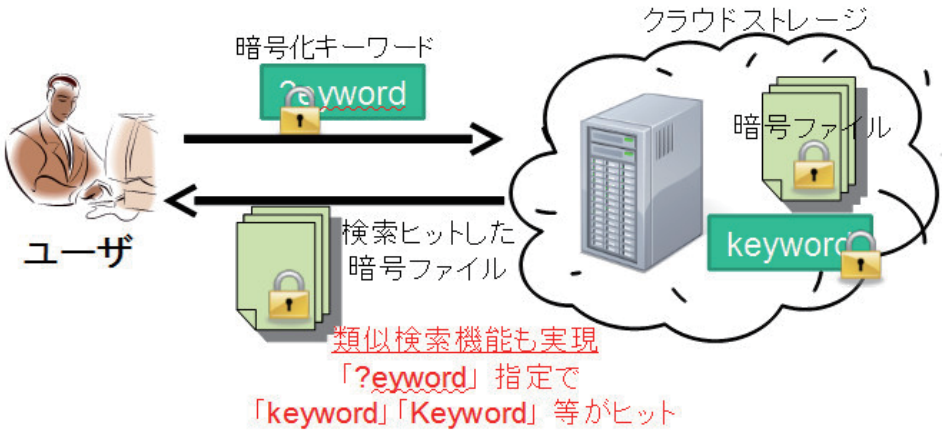
2-3 研究成果の詳細

まず既存研究の調査の結果，従来の多くの検索可能暗号方式では，例外的な[6]を除き，暗号ファイルの中身全体を検索対象とするのではなく，暗号ファイルの作成者がファイルの中身に関連する重要キーワードを最初に複数列挙し，それらを暗号ファイルの検索インデックスとして付加するアプローチを取っていた．そのため我々も，従来のアプローチに倣った検索可能暗号方式の構成に取り組んだ.

また一般に検索可能暗号は「公開鍵暗号ベースの方式」と「共通鍵暗号ベースの方式」に大別される．公開鍵暗号方式ベースでは，キーワードの暗号化は誰でも行うことができ，検索に必要なデータ(しばしばトークンと呼ばれる)はある特権を持った者のみが可能な方式である.そして一般的に公開鍵暗号ベースの方式は

ペアリングと呼ばれる楕円曲線上で定義される数学的な演算を用いることで実現される．それに対し，共通鍵暗号ベースの方式では，暗号の秘密鍵情報は暗号ファイルを共有する利用者らの間で共有することが前提となる．またペアリング演算を用いずに構成することがほとんどである．ペアリング演算は数学的に複雑な処理が必要となり，処理速度が低速であるという制限があるため，我々は共通鍵暗号ベースの検索可能暗号の構成に取り組んだ．

これまでの我々の研究の経緯に基づき「ブルームフィルタ」と呼ばれるデータ構造を用いることで完全一致の検索だけでなく，類似したキーワードの検索も実現できるのではという着想から研究を開始した．ブルームフィルタとは複数の要素をある長さの配列に効率的に一つの集合として登録し，あとからある要素がブルームフィルタに登録されている集合に属すか否かを効率的に検査可能なデータ構造のことである．



提案手法ではそれぞれのキーワードの暗号結果に相当するブルームフィルタを，クラウドサーバに登録するというアプローチを取った．ここでキーワードに対応するブルームフィルタを作成する際に，キーワードを一文字ずつに区切り，各文字とその文字位置の情報(例えばキーワードの中の 1 文字目なら 1)をペアとしてブルームフィルタに登録している．この登録の際に鍵付ハッシュ関数と呼ばれる暗号技術を用いることでキーワード自体はブルームフィルタの配列から推測できないよう設計を行った．この鍵付きハッシュ関数で用いられる鍵は，暗号ファイルを共有する利用者らの間で共有する前提となる．

ここで各文字毎にブルームフィルタに登録することにより，キーワードの完全一致のみならず，類似キーワード検索による検索操作も可能とすることに成功した．この提案手法ではデータ検索を行う検索者は例えば「2012/??/01」というような指定を行い，「2012/10/01」，「2012/03/01」といった複数の類似年月日に相当するキーワードにも検索ヒットすることが可能である．このような検索手法は一般にワイルドカード検索などと呼ばれ，検索指定方法をより柔軟なものとするため検索可能暗号の実用化において重要な機能であると考える．

この提案手法ではクラウドストレージ側で暗号化されたキーワードはブルームフィルタ配列からなり，二分木のデータ構造で管理することが可能である．よってデータ検索者が指定した検索キーワード(これもブルームファイル配列で構成されている)を用いて，検索にヒットする暗号ファイルを特定するために，二分探索に類似した方法を用いて比較的高速に検索処理を行えることも利点として挙げられる．この構造により，クラウドストレージ上の全暗号ファイルの検索インデックスを毎回チェックしなければならない，という制限を取り除くことが可能である．

この我々の提案手法は論文[1]として，暗号の国際会議で採択されている．また研究成果の周知を行うため，他の国際会議や研究集会での発表も行った[3,4]．

さらに提案手法をタブレットなどのモバイルデバイス上でウェブアプリケーションとして実装し，性能測定も行い，それらの測定結果から提案手法が実用上，問題ないレベルで実行可能であることを確認できている．性能測定結果などもまとめた成果は現在ジャーナル論文として投稿中である．

今回の提案手法の検索で指定できるワイルドカードは一文字単位に相当するワイルドカードである．さらなる利便性の向上として，任意の長さの文字列に対応するワイルドカード指定も可能であることが望ましいといえる．つまり例えばデータ検索者が「2013*講義」というような検索指定を行ったとき，「2013 数学講義」や「2013 数学の講義」，「2013 情報科学講義」など「*」に相当する文字列が任意の長さの文字列でも検索ヒット可能であると，より自由度の高い検索指定ができると考えられる．現在の提案手法を拡張し，このような検索指定も可能な方式の構成についても今後検討を続けていくことを計画している．

また検索可能暗号の安全性評価をさらに確実なものとするために，検索可能暗号への攻撃手法についても検討を進めてきた．特に検索可能暗号の既存の安全性モデルを超えた攻撃手法の可能性について検討を行った．従来の検索可能暗号の複数の安全性モデル(例 IND-CKA モデル)の中では，頻度解析と呼ばれる攻撃手法

は考慮範囲外であった．頻度解析とは攻撃者が，キーワードがどのような頻度で文書に存在するかを事前に知っているという前提で，検索結果から，秘密にされているキーワードを推測する攻撃のことである．例えばあるキーワード1での検索にヒットする暗号ファイルの個数が10,000であり,あるキーワード2での検索にヒットする暗号ファイルの個数が5,000であったとする．攻撃者が「東京」というキーワードが「大阪」というキーワードより，より多く出現すると事前に知っていた場合，キーワード1は東京であると推測できてしまう可能性がある．より強力な攻撃者を想定したとき，このような頻度解析が現実の脅威となりうることは十分に考えられる．そこで既存の検索可能暗号方式が我々の想定する頻度解析攻撃に対して，どの程度耐性を持ちうるか実装実験などを通して評価を行った.またこの攻撃手法を,ブルームフィルタを用いた我々の提案方式にも適用し，どの程度攻撃が成功するのか，またどのようにすれば攻撃を回避できるのかについても検討を行った.調査の結果,ある程度の偽陽性を許容することで頻度攻撃を回避できることが分かった．ここで偽陽性とは本来，指定したキーワードの検索結果として含まれるべきでないものまでもが検索結果に(ある種のノイズとして)含まれてしまうことである．これは通常は望ましくはないが，検索結果に少量の余分なノイズが入ることは実際の利用の中では許容可能範囲であると考えられるため，安全性と性能のトレードオフとして捉えることができる．この成果については国際会議での論文[2]にまとめ，現在ジャーナル論文を作成中である．

また検索可能暗号と組み合わされて用いられることの多い属性ベース暗号[18, 19, 20, 21]についても研究を行った．属性ベース暗号は暗号データの復号条件を通常の公開鍵暗号に比べ，より細かに指定できることが特長である．より具体的には復号可能な条件を各利用者の名前ではなく，利用者が所有しうる属性の組み合わせで指定できる点が既存の公開鍵暗号技術と大きく異なる点である．クラウドストレージでの検索可能暗号の利用においては，検索可能暗号でキーワードを暗号化し，ファイルの中身を属性ベース暗号で暗号化することがよく行われるため[16]，属性ベース暗号の機能向上も検索可能暗号の実用化に向けて重要な課題と考えられる.通常の属性ベース暗号では単一の鍵発行機関が存在し,利用者の所有する属性集合に基づき,鍵生成を行うことになっている．しかしながら，実際の運用を考えたとき，利用者の属性を管理する機関は複数存在することが多いため，単一の鍵生成機関では実利用に適していないという問題がある．このような状況を解決する属性ベース暗号として複数の鍵発行機関が存在可能な方式が提案されている[17, 20].これらの方式をさらに改良することでより実用的なシステムの構成が可能と考えられる．また属性ベース暗号では鍵の失効機能の実現も重要な研究課題と言える．鍵の失効とは一度，属性に対応する鍵を鍵発行機関から与えられた利用者がある時点で，その属性を所有する権限を無くしたため，それ以降の鍵の利用を無効化することである(つまり暗号ファイルを復号するために鍵を利用できなくすることである)．このような問題に対する解決策として，論文[5]において複数の鍵発行機関が存在する環境での属性ベース暗号の利用を想定し，クラウドストレージ利用者の権限の失効手法を提案した．暗号に関する鍵を所有していても失効処理を行うことで復号機能を停止する機能は企業などでの利用者の異動や一時的な権限の付与に有用であり，提案方式ではクラウドストレージでの再暗号不要な方式を実現した．提案方式では[22]で与えられた時刻情報もファイル暗号の際に用いる属性情報として利用し，さらに二分木のデータ構造で管理することで鍵の失効機能を実現できることを検証した．また時刻情報に基づく鍵の失効だけでなく，[23]での利用者の名前に基づく鍵の失効機能も取り入れて設計を行った．しかしながら，提案方式には安全性の厳密な検証や，公開鍵情報量の削減など，さらなる改善の余地があるため，このテーマに関して今後も研究を続けていく計画である．特に論文[21]で提案された公開鍵情報量の削減手法をさらに調査し，今回の提案手法に適用することを計画している．

【参考文献】

[1] Takanori Suga, Takashi Nishide, and Kouichi Sakurai, ``**Secure Keyword Search Using Bloom Filter with Specified Character Positions,**" 6th International Conference on Provable Security (ProvSec), LNCS 7496, pp.235--252, Springer-Verlag, 2012.

[2] Takanori Suga, Takashi Nishide, and Kouichi Sakurai, ``**Weakness of Provably Secure Searchable Encryption Against Frequency Analysis,**" 5th Mosharaka International Conference on Communications, Computers and Applications (MIC-CCA), ISSN: 2227-331X, pp.142--147, 2012.

[3] Takanori Suga, Takashi Nishide, Kouichi Sakurai, ``**Character-based Keyword Search over Encrypted Data,**" Forum "Math-for-Industry" 2012 "Information Recovery and Discovery", Fukuoka, Oct., 2012 (Poster Session).
[4] Takanori Suga, Takashi Nishide, Kouichi Sakurai, ``**Character-based Symmetric Searchable Encryption,**" 7th International Workshop on Security(IWSEC), Fukuoka, Nov., 2012 (Poster Session).
[5] Takashi Nishide, ``**Toward Revocation Mechanism for Multi-Authority CP-ABE,**" 暗号と情報セキュリティシンポジウム(SCIS), 5pages, 京都, 1 月, 2013.
[6] Dawn Song, David Wagner, and Adrian Perrig, ``**Practical Techniques for Searches on Encrypted Data,**" IEEE Symposium on Security and Privacy, pp.44--55, 2000.
[7] Eu-Jin Goh, ``**Secure indexes,**", ePrint Technical Report 2003/216, 2003.
[8] Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky, and Giuseppe Persiano, ``**Public Key Encryption with Keyword Search,**" Proc. Eurocrypt, LNCS 3027, pp.506--522, Springer-Verlag , 2004.
[9] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi, ``**Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions,**" Proc. CRYPTO, LNCS 3621, pp.205--222, Springer-Verlag, 2005.
[10] Philippe Golle, Jessica Staddon, and Brent Waters, ``**Secure Conjunctive Keyword Search over Encrypted Data,**" Proc. Applied Cryptography and Network Security(ACNS), LNCS 3089, pp.31--45, Springer-Verlag , 2004.
[11] Mihir Bellare, Alexandra Boldyreva, and Adam O'Neill, ``**Deterministic and Efficiently Searchable Encryption,**" Proc. CRYPTO, LNCS 4622, pp.535--552, Springer-Verlag , 2007.
[12] Jin Li, Qian Wang, Cong Wang, Ning Cao, Kui Ren, and Wenjing Lou, ``**Fuzzy Keyword Search over Encrypted Data in Cloud Computing,**" Proc. IEEE INFOCOM, pp.253--262, 2010.
[13] Saeed Sedghi, Peter van Liesdonk, Svetla Nikova, Pieter H. Hartel, and Willem Jonker, ``**Searching Keywords with Wildcards on Encrypted Data,**" Proc. Security and Cryptography for Networks(SCN), LNCS 6280, pp.138--153, Springer-Verlag , 2010.
[14] http://research.microsoft.com/en-us/projects/cryptocloud/
[15] http://www.air.co.jp/article.php/release74
[16] Fangming Zhao, Takashi Nishide, and Kouichi Sakurai, ``**Multi-User Keyword Search Scheme for Secure Data Sharing with Fine-Grained Access Control,**" 14th Annual International Conference on Information Security and Cryptology (ICISC'11), LNCS 7259, pp.406--418, Springer-Verlag, 2012.
[17] Tatsuaki Okamoto and Katsuyuki Takashima, ``**Decentralized Attribute-Based Signatures,**" Proc. 16th International Conference on Practice and Theory in Public-Key Cryptography, LNCS 7778, pp.125- -142, Springer-Verlag, 2013.
[18] Amit Sahai and Brent Waters, ``**Fuzzy Identity-Based Encryption,**" Proc. Eurocrypt, LNCS 3494, pp. 457--473, Springer-Verlag, 2005.
[19] Allison B. Lewko, Tatsuaki Okamoto, Amit Sahai, Katsuyuki Takashima, and Brent Waters, ``**Fully Secure Functional Encryption: Attribute-Based Encryption and (Hierarchical) Inner Product Encryption,**" Proc. Eurocrypt, LNCS 6110, pp. 62--91, Springer-Verlag, 2010.
[20] Allison B. Lewko and Brent Waters, ``**Decentralizing Attribute-Based Encryption,**" Proc. Eurocrypt, LNCS 6632, pp. 568--588, Springer-Verlag, 2011.
[21] Tatsuaki Okamoto and Katsuyuki Takashima, ``**Fully Secure Unbounded Inner-Product and Attribute-Based Encryption,**" Proc. Asiacrypt, LNCS 7658, pp. 349--366, Springer-Verlag, 2012.
[22] Amit Sahai, Hakan Seyalioglu, and Brent Waters, ``**Dynamic Credentials and Ciphertext Delegation for Attribute-Based Encryption,**" Proc. CRYPTO, LNCS 7417, pp.199--217, Springer-Verlag, 2012.

[23] Nuttapong Attrapadung and Hideki Imai, ``**Attribute-Based Encryption Supporting Direct/Indirect Revocation Modes,**" Proc. Cryptography and Coding, LNCS 5921, pp.278--300, Springer-Verlag, 2009.

〈発 表 資 料〉

題　　名	掲載誌・学会名等	発表年月
Secure Keyword Search Using Bloom Filter with Specified Character Positions	6th International Conference on Provable Security (ProvSec), LNCS 7496, pp. 235--252, Springer-Verlag	2012 年 9 月
Weakness of Provably Secure Searchable Encryption Against Frequency Analysis	5th Mosharaka International Conference on Communications, Computers and Applications (MIC-CCA), ISSN: 2227-331X, pp. 142—147	2012 年 10 月
Character-based Keyword Search over Encrypted Data	Forum "Math-for-Industry" 2012 "Information Recovery and Discovery" (Poster Session)	2012 年 10 月
Character-based Symmetric Searchable Encryption	7th International Workshop on Security(IWSEC)(Poster Session)	2012 年 11 月
Toward Revocation Mechanism for Multi-Authority CP-ABE	暗号と情報セキュリティシンポジウム(SCIS)	2013 年 1 月