様式 C-19

科学研究費補助金研究成果報告書

平成 23 年 4 月 12 日現在

機関番号：１２１０２
研究種目：若手研究（B）
研究期間：2009～2010
課題番号：21700093
研究課題名（和文）　クラウド環境におけるセキュアでトレーサブルな XML データ流通機構

研究課題名（英文）　A secure and traceable XML distribution mechanism in a cloud computing environment

研究代表者
天笠　俊之（AMAGASA　TOSHIYUKI）
筑波大学・大学院システム情報工学研究科・准教授
研究者番号：70314531

研究成果の概要（和文）：
クラウドコンピューティング環境（以下，クラウド環境と呼ぶ）が新たなコンピューティング環境として注目されている．また，XML（Extensible Markup Language）はデータフォーマットとして普及し，XML 形式で蓄積されるデータ量は増加しつつある．このため，クラウド環境においてプライバシーを保護したまま XML データを流通する基盤が必要になる．このような背景から，本研究ではクラウド環境を対象に，セキュリティとトレーサビリティを両立した XML データ流通機構の研究･開発を行った.プライバシー保護 XML データ流通機構の基礎技術である，プライバシー保護データベース問合せ，プライバシー保護オントロジー統合，XML データへの注釈付けについて研究を行った.

研究成果の概要（英文）：
Could computing environment has been an emerging as a new platform. Meanwhile, XML (Extensible Markup Language) has been adopted in many applications as a de facto standard data format, and the volume of data being described using XML is increasing. For this reason, a novel mechanism, that enables data exchange in the cloud while preserving privacy, is needed. In this research work, we investigated a secure and traceable XML distribution mechanism for the cloud. Specifically, we have developed a privacy-preserving database query method, a privacy-preserving ontology integration scheme, and a scheme of XML data annotation.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2009 年度	1,700,000	510,000	2,210,000
2010 年度	1,500,000	450,000	1,950,000
総　計	3,200,000	960,000	4,160,000

研究分野：データ工学
科研費の分科・細目：情報学　・　メディア情報学・データベース
キーワード：XML，クラウド，トレーサビリティ

１．研究開始当初の背景

(1) 新しいコンピューティング環境として，クラウドコンピューティング（cloud computing）が注目されている．クラウドとはインターネットそのものやインターネットに接続された膨大な計算機資源のことを指しており，利用者は，サービスとして提供されたそれらの計算機資源を利用することで，コンピュータによる計算処理を享受できる．利用者にとっては，手元に処理プログラムや大規模な計算機を持つ必要がなく，また計算機の維持管理や知識も必要ないため，新しいコンピ

ュータの利用形態として脚光を浴びている．Google，Amazon，Yahoo，Microsoft などが環境を提供し出しており，今後さらなる普及が予想されている．

(2) 他方，XML (Extensible Markup Language) は，誕生から 10 年を迎え，データフォーマットのデファクトスタンダード（事実上の標準）として広く利用されている．その用途は Web 文書，ビジネス文書，科学データ，ビジネスデータ，ネットワークプロトコルなど多岐に渡り，現在のコンピューティング環境において欠かすことができない重要な位置を占めている．クラウド環境においても，さまざまな情報の表現手段や交換フォーマットとして XML データが利用されることが予想される．

(3) ところで，クラウド環境においては，主たる計算処理を行う計算機やデータベースが利用者の環境にないことから，これまでには存在しなかった次のような問題が生じる：

①データや処理プログラムの管理運用が保有者以外の第三者（サービスプロバイダ）の手に委ねられる．サービスプロバイダは，業務遂行上セキュリティの保持に最大限の努力をするものの，内外からの攻撃を完全に排除することは不可能である．このため，外部に置かれたデータベースに対応した新たなセキュリティ対策が必要である．

②クラウド環境は，インターネットを基板とする分散環境であり，処理の過程で，さまざまな情報源から取得された XML データが複合的に組み合わされる．このとき，導出された XML データの系統管理が極めて重要となる．

２．研究の目的

(1) 本研究では，クラウド環境におけるセキュアな XML 問合せ方式およびクラウド環境における XML データのトレーサビリティ機構に関する研究を行なう．

３．研究の方法

(1) セキュアな XML 問合せ方式については，関係データベースで過去数年に渡り研究されたデータベースアウトソーシングの技術の枠組みを利用し，これを XML データに適用した新たな XML データベースアウトソーシング技術を開発する．基本的なアイデアは，問合せしたい XML データ事前に暗号化し，（プロバイダが提供する）サーバにアップロードする．問合せ処理を行う際は，利用者から発行された問合せを変換し，第三者が分からない形のままサーバで処理を行い，暗号化された結果に後処理を行う．サーバ側でデータの復号化を行わないため，セキュリティを保ったままクラウド環境での XML データ問合せが実現できる．

(2) XML データのトレーサビリティ機構に関する研究については，先行研究課題「e サイエンスのためのオーバレイネットワークに基づく XML データ流通機構の研究」で提案した，アノテーションに基づく XML データトレーサビリティの手法を拡張し，クラウド環境に対応したトレーサビリティ機構を開発する．

４．研究成果

(1) セキュリティに関しては，クラウド環境においてアウトソースされたデータベースに対して機密を保護したまま問合せを行うための，プライバシー保護検索手法を開発した．従来の Order-Preserving Encryption Scheme (OPES) を発展させ，データベースに対する問合せを可能にしつつセキュリティを向上した Multi-Valued OPES (MV-OPES)を提案した．MV-OPES により，暗号化されたデータベースを利用して各種関係代数演算が処理可能となる．実験により，その提案手法の有効性を評価した．

(2) さらに，MV-OPES を発展させた，MV-POPES (Multi-Valued Partial Order Preserving Encryption Scheme) を提案した．これは，MV-OPES に区間のランダム化を導入した手法であり，より強いセキュリティを実現するとともに，各種問合せ最適化を組み合わせることで，実用的な問合せ処理性能を実現している．

(3) トレーサビリティについては，XML データに対してトレーサビリティ情報の付与を目的としたアノテーションのモデルを提案した．これは，任意の XML データに対して，XML データで記述したアノテーションを付与することを可能とするモデルである．アノテーションが付与された XML データに対しては，拡張した XPath 式によってアノテーションの特徴を利用した問合せ行うことができる．これらを一般の XML データベースを利用して実現するための格納構造と，問合せ変換について議論し，その有効性を実験によって評価した．

(4) XML データ処理に関しては，近年普及しつつあるマルチコア CPU を利用したスケーラブルな問合せ処理アルゴリズムを開発した．Holistic Twig Join (HTJ) と呼ばれる問合せ処理アルゴリズムに着目し，

これをオンラインで分割，各 CPU コアに割り当てを行うことによって並列処理が可能となる．

(5) クラウド環境におけるデータ流通のためには，プライバシーを保護したオントロジーの統合が必要である．このため，プライバシーを保護したオントロジー統合手法を提案した．これは，信頼できる第三者の存在を仮定した手法であり，まずオントロジーを FastMap と呼ばれる手法で多次元空間に写像し，それを信頼できる第三者に送信することによって，オントロジーの詳細を明かすことなくオントロジーの近接度を評価できる．

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 7 件）

① Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″MV-OPES: Multivalued-Order Preserving Encryption Scheme: A Novel Scheme for Encrypting Integer Value to Many Different Values,″ IEICE Trans. Info. & Syst., Vol. E93-D, No. 9, pp. 2520-2533, Sept. 2010. 査読有

② Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Parallel Holistic Twig Joins on a Multi-core System,″ International Journal of Web Information Systems (IJWIS), Vol. 6, No. 2, pp. 149-177, 2010. 査読有

③ Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Mixed Encryption over Semi-Trusted Datebase″, MASAUM Journal of Basic and Applied Science (MJBAS), Vol. 1, Issue. 2, pp. 302-312, 2009. 査読有

④ 渡辺知恵美，新井裕子，天笠俊之，“ブルームフィルタを用いたプライバシ保護検索における攻撃モデルとデータ撹乱法の一検討”，日本データベース学会論文誌，Vol. 8, No. 1, pp. 113-118. 2009 年 6 月．査読有

⑤ 寺島慎太郎，天笠俊之，北川博之，“木直列化に基づく XML データの類似結合における木構造の統合”，日本データベース学会論文誌，Vol. 8, No. 1, pp. 47-52. 2009 年 6 月．査読有

⑥ Chantola Kit, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Algorithms for Efficient Structure-Based Grouping in XML-OLAP″, International Journal of Web Information Systems (IJWIS), Volume 5, Issue 2, pp. 122-150, 2009. 査読有

⑦ Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″XML Data Partitioning Schemes for Parallel Holistic Twig Joins″, International Journal of Web Information Systems (IJWIS), Volume 5, Issue 2, pp. 151-194, 2009. 査読有

〔学会発表〕（計 25 件）

1) 張帆，天笠俊之，北川博之，“プライバシーを保護したオントロジーマッピング手法”，情報処理学会全国大会講演論文集，5Y-3，東京工業大学，2010 年 3 月 6 日～8 日．

2) 添野隆之，天笠俊之，北川博之，“確率的スキーママッチングによる複数 XML データに対する問合せ”，情報処理学会全国大会講演論文集，4N-5，東京工業大学，2010 年 3 月 6 日～8 日．

3) 駒水孝裕，天笠俊之，北川博之，“XML データに対するファセット検索のユーザビリティ評価”，情報処理学会全国大会講演論文集，4N-3，東京工業大学，2010 年 3 月 6 日～8 日．（学生奨励賞受賞）

4) 駒水孝裕，天笠俊之，北川博之，“キーワード検索が可能な XML データに対するファセット探索”，第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM2011)，E6-1，ラフォーレ修善寺，2011 年 2 月 27 日～3 月 1 日．

5) Hasan Kadhem, Toshiyuki Amagasa, Hiroyuki Kitagawa, ″Optimization Techniques for Range Queries in the Multivalued-Partial Order Preserving Encryption Scheme″, 第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM2011)，C5-6，ラフォーレ修善寺，2011 年 2 月 27 日～3 月 1 日．

6) 金子静花，渡辺知恵美，天笠俊之，“Semi-ShuffledBF:ブルームフィルタを用いた安全かつより高速なプライバシ保護検索手法の提案”，第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM2011)，C5-3，ラフォーレ修善寺，2011 年 2 月 27 日～3 月 1 日．

7) 寺島慎太郎，天笠俊之，北川博之，“XML データにおける類似極大部分木の効率的な探索”，第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM2011)，E1-2，ラフォーレ修善寺，2011 年 2 月 27 日～3 月 1 日．

8) Ngo Sy Viet Phu, Toshiyuki Amagasa, Hiroyuki Kitagawa, ″Extracting XML Data from the Web,″ The 12th International Conference on

Information Integration and Web-based Applications & Services (iiWAS2010), pp. 117-114, Paris, France, November 8-10, 2010.
9) Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″A Secure and Efficient Order Preserving Encryption Scheme for Relational Databases,″ International Conference on Knowledge Management and Information Sharing (KMIS 2010), Valencia, Spain, October 25-28, 2010.
10) Toshiyuki Amagasa, Noriyoshi Ishii, Tomoteru Yoshie, Osamu Tatebe, Mitsuhisa Sato, Hiroyuki Kitagawa, ″A Faceted-Navigation System for QCDml Ensemble XML Data,″ Fifth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2010) in conjunction with BWCCA-2010, pp. 132-139, Fukuoka, Japan, November 4-6, 2010.
11) Hang Shi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Fast Detection of Functional Dependencies in XML Data,″ 7th International XML Database Symposium (XSym 2010) in conjunction with VLDB2010, pp. 113-127, Singapore, September 17, 2010.
12) SY VIET PHU NGO, 天笠俊之, 北川博之, “属性の共起関係に着目した WWW からの効率的な XML データ抽出”, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会講演論文集, 東京大学, 2010 年 3 月 8 日～12 日.
13) 駒水孝裕, 天笠俊之, 北川博之, “異種 XML データに対するファセット検索システムの性能評価”, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会講演論文集, 東京大学, 2010 年 3 月 8 日～12 日.
14) 清野真奈, 天笠俊之, 北川博之, “注釈によるトレーサビリティ機能を持つ XQuery 処理系の実装”, 情報処理学会全国大会講演論文集 第 72 回平成 22 年(1), ″1-907″-″1-908″, 東京大学, 2010 年 3 月 8 日.
15) 史航, 天笠俊之, 北川博之, ″On Finding Functional Dependencies in XML Data″, 情報処理学会全国大会講演論文集 第 72 回平成 22 年(1), ″1-903″-″1-904″, 東京大学, 2010 年 3 月 8 日.
16) 李重頡, 天笠俊之, 北川博之, ″Efficient Privacy Preserving Query Processing using GPGPU″, 情報処理学会全国大会講演論文集 第 72 回平成 22 年(1), ″1-811″-″1-812″, 東京大学, 2010 年 3 月 8 日.
17) 山口卓郎, 天笠俊之, 北川博之, “センサネットに対する XML ビューの提案”, 情報処理学会全国大会講演論文集 第 72 回平成 22 年(1), ″1-671″-″1-672″, 東京大学, 2010 年 3 月 8 日.
18) 駒水孝裕, 天笠俊之, 北川博之, “異種 XML データに対するファセット検索における多様な検索”, 第 2 回データ工学と情報マネジメントに関するフォーラム (DEIM 2010), C7-5, 淡路夢舞台国際会議場, 2010 年 2 月 28 日～3 月 2 日. (学生奨励賞受賞)
19) Hasan Kadhem, Toshiyuki Amagasa, Hiroyuki Kitagawa, ″An Encryption Scheme to Prevent Statistical Attacks in the DAS Model″, 第 2 回データ工学と情報マネジメントに関するフォーラム (DEIM 2010), B5-5, 淡路夢舞台国際会議場, 2010 年 2 月 28 日～3 月 2 日.
20) Imam Machdi, Toshiyuki Amagasa, Hiroyuki Kitagawa, ″Task Parallelism for TwigStack Algorithm on a Multi-core System″, 第 2 回データ工学と情報マネジメントに関するフォーラム (DEIM 2010), C8-2, 淡路夢舞台国際会議場, 2010 年 2 月 28 日～3 月 2 日. (学生奨励賞受賞) (最優秀論文賞受賞)
21) 駒水孝裕, 天笠俊之, 北川博之, 「異種 XML データに対するファセット検索手法の提案」, 情報処理学会研究報告 (DD), 2009-DD-073(7), pp. 1-8, 東京大学, 2009 年 9 月 25 日.
22) 天笠俊之, ゴーシーヴィエトフー, 北川博之, 「Web ページを対象とした XML データ抽出手法の検討」, 情報処理学会研究報告 (DD), 2009-DD-073(4), pp. 1-7, 東京大学, 2009 年 9 月 25 日.
23) Imam Machdi, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Executing Parallel TwigStack Algorithm on a Multi-core System″, The 11th International Conference on Information Integration and Web-based Applications & Services (iiWAS2009), pp. 176-184, Kuala Lumpur, Malaysia, December 14-16, 2009.
24) Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″A Novel Framework for Database Security based on Mixed Cryptography″, The Fourth International Conference on Internet and Web Applications and Services (ICIW 2009), pp. 163-170, Venice, Italy, May 24-28, 2009.
25) Hasan Kadhem, Toshiyuki Amagasa, and Hiroyuki Kitagawa, ″Encryption over

Semi-trusted Database", DASFAA 2009 Ph. D Workshop (DASFAA Workshops 2009), pp. 358-362, Brisbane, Australia, April 21-23, 2009.

〔その他〕
ホームページ等
http://www.kde.cs.tsukuba.ac.jp/~amagasa/

６．研究組織
(1)研究代表者
天笠　俊之（AMAGASA　TOSHIYUKI）
筑波大学・大学院システム情報工学研究科・准教授
研究者番号：70314531