様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　３０　年　　４　月　１９　日現在

機関番号：　１３９０１
研究種目：　挑戦的萌芽研究
研究期間：　2014～2017
課題番号：　２６５４００２５

研究課題名（和文）ソフトウェアセキュリティのための量を扱う計算モデルの提案

研究課題名（英文）Formal models for quantitative analysis of software security

研究代表者
関　浩之（Seki, Hiroyuki）

名古屋大学・情報学研究科・教授

研究者番号：８０１９６９４８

交付決定額（研究期間全体）：（直接経費）　　2,800,000円

研究成果の概要（和文）：ソフトウェアのセキュリティやプライバシーの尺度として，量的情報流，差分プライバシー等が提案され種々議論されている．本研究課題では主に量的情報流を中心にいくつかの定量的尺度について，それを計算するアルゴリズムや，そのような尺度に基づくソフトウェアの安全性検証手法の開発に取り組んだ．具体的に，復号アルゴリズムの時間攻撃による漏洩量の計算法，XMLデータベースにおけるk-安全性の検査法，#SMTソルバーを用いた量的情報流の計算手法，および文字列データの量的情報流解析の基礎となる認識可能級数および代数的級数に対する計数アルゴリズムの開発と，シミュレーションまたは実装ツールに基づく評価実験を行った．

研究成果の概要（英文）：A few quantitative notions for security and privacy of software such as quantitative information flow (QIF) and differential privacy have been proposed. In this research, we developed methods that analyze given programs or systems based on such notions. Specifically, we proposed an approximation algorithm that computes leakage by timing attack against an RSA decoder, a verification algorithm of k-secrecy of XML databases. Furthermore, as a theoretical basis for QIF analysis of programs that dynamically generate strings, we propose algorithms that counts, for a given recognizable or algebraic series S and a natural number d, the summation of the coefficients (or weights) of words of length d in S efficiently. The proposed methods were shown to be effective either by computer simulation or by experiments based on the implemented tools.

研究分野：　ソフトウェア基礎理論

キーワード：　セキュリティ　量的情報流　k-安全性　XMLデータベース　時間攻撃　差分プライバシー　SMT　SAT

１．研究開始当初の背景

ネットワークやセンシング，組込み技術の進歩により計算機システムは現代社会を支える基盤としてその安全性と信頼性が求められている．特に計算機システムにおけるセキュリティとプライバシーの問題は喫緊に解決すべき重要課題である．この問題の解決へ向け，現在，システムが外部の観測者に対してどの程度の情報を漏えいするかを定量的に定義し，そのような尺度に基づいてシステムの安全性を解析し保証することを目指す研究が盛んとなっている．定量的な尺度として，量的情報流[Sm09]，差分プライバシー[Dw06, MT07]，*k*-匿名性[MKGV07 他]がよく知られているが，それらの相互関係については限られた考察しかなされていない．研究代表者らは，*k*-匿名性を拡張した概念を用いて XML に代表される構造化文書データベースに対するセキュリティ安全性に関する研究を行ってきた経験から，形式言語理論を中心とする計算理論的アプローチと情報理論的アプローチを組み合わせることにより既存の量的セキュリティの概念を拡張することの着想を得た．

[Dw06] C. Dwork, Differential privacy, ICALP 2006, LNCS 4052, 1-12.
[MT07] F. McSherry and K. Talwar, Mechanism design via differential privacy, IEEE FOCS 2007, 94-103.
[Sm09] G. Smith, On the foundations of quantitative information flow, FOSSACS 2009, LNCS 5504, 288-302.
[MKGV07] A. Machanavajjhala, D. Kifer, J. Gehrke and M. Venkitasubramaniam, l-diversity: privacy beyond k-anonymity, ACM Trans. on Knowledge Discovery from Data, 1(1), 2007.

２．研究の目的

近年，ソフトウェアのセキュリティやプライバシーの尺度として，量的情報流，差分プライバシー，*k*-匿名性が提案され種々議論されているが，これらセキュリティ強度を表す概念間の関係は十分には解明されていない．これらの独立に提案された概念を包摂するモデルを提案し，それに基づいてセキュリティやプライバシーを保証する系統的方法論を開発することは挑戦的課題である．本研究ではまず，既存のものを含め定量的尺度をいくつか具体的に設定する．次に，コスト（重み）を表現可能な計算モデルを導入し，ソフトウェアに対する上記の定量的尺度をこれらのモデルに基づいて表現する．そして，効率のよいコスト計数アルゴリズムの提案ならびに，与えられたソフトウェアが，定量表現されたセキュリティ要求を満たすかどうかの静的解析法や動的監視法を提案することを目的とする．これらの研究課題の解決過程を通じて，一般的なソフトウェアセキュリティの概念を構築するための指針を見出す．

３．研究の方法

研究期間の前半では，現実のソフトウェアにおけるセキュリティの定量化を行うのに十分な能力と，精密な解析を許容する数学的に良好な性質を合せもつモデルを構築するための理論的考察に集中する．後半ではこれらの理論的成果を学会で公表する一方で，現実のプログラムの解析を行う手法を提案し，実装と評価を行うフェーズに移る．解析アルゴリズムは形式言語理論における言語演算に対する閉包性や，空判定問題等の決定可能性に帰着させる予定である．さらに最近注目されている別のアプローチとして，SAT/SMT ソルバの援用を予定している．想定する応用分野として，暗号アルゴリズムのサイドチャネル攻撃（特に時間攻撃）耐性，データマイニングアルゴリズムのプライバシー保全性，XML データベースにおける推論攻撃耐性，web システムにおけるクライアント側プログラムの脆弱性解析等を取り上げる．

４．研究成果

（１）導入的ケーススタディ

①差分プライバシー

一つ目は，隠れマルコフモデルにおけるパラメータ推定アルゴリズムの差分プライバシー化である．k-mean 法や決定木構成アルゴリズム等，機械学習やデータマイニングのためのいくつかのアルゴリズムがε-差分プライバシーを満たすように改良されている．しかし，状態空間をもつモデルの学習アルゴリズムの差分プライバシー化はあまり検討されていなかった．そこで本研究では，隠れマルコフモデル(HMM)におけるパラメータ推定アルゴリズムの差分プライバシー化を行った．

②量的情報流

二つ目は，暗号システムへのタイミング攻撃に対する耐性の定量化である．この攻撃では，攻撃者がシステムの実行時間を観測し分析することで，内部で使用されている鍵の情報を推測する攻撃手法である．タイミング攻撃に対するシステムの耐性を定量的に評価するため，情報理論的手法を活用することが検討されている．情報理論的な見地からは，暗号システムを，未知の鍵を入力とし実行時間を出力とする通信路であると解釈することができる．通信路の出力(実行時間）から入力(鍵)を推測する行為がタイミング攻撃に相当するが，攻撃者が得ることのできる情報量は，通信路入出力間の相互情報量（量的情報流）を超えることはない．本研究では，暗号システムとして RSA 暗号の復号アルゴリズムを想定し，上記のように定義される通信路の量的情報流を精密に算出した．

（２）#SAT ツールを用いた k-安全性検査

データベース内の機密情報の漏洩を防ぐ一般的な手法としてアクセス制御がある．アクセス制御とは，データベースに対し，機密情報が得られるような問合せの実行をアクセス権

のあるユーザのみに許可するというものである．以下では，問合せを許可問合せと禁止問合せに分類することによりアクセス制御を行うと仮定する．このようにユーザごとに許可問合せ，禁止問合せが適切に指定されていれば機密情報の漏えいは直接的には生じない．しかし，ユーザがデータベースに対して許可問合せ，データベーススキーマ，問合せのコード（意味）等の公開情報を組み合わせることで，禁止問合せの結果を推論できることがある．このようなユーザの行為を推論攻撃という．推論攻撃によって禁止問合せの結果がk個未満に絞りこめないとき，データベースインスタンスは（これらの問合せとスキーマに対して）k-安全であるという．本研究ではXMLデータベーススキーマ，問合せ，インスタンスが与えられたとき，これらを命題論理式のモデル計数問題に帰着してk-安全性を検査する手法の提案を行った．この手法では，スキーマの適合性や問合せの入出力関係を充足可能性問題に帰着し，命題論理式のモデル（命題論理式を充足するような，論理変数への真偽値割り当て）を数えることによりk-安全性を判定する．

入力を制約式に変換するツールを実装し，既存のSugar（順序符号化を行うツール）やsharpCDCL（射影を指定可能なモデル計数ツール）といったツールを用いて提案するk-安全性検査法に基づいた検査システムを実装し，性能評価のための実験を行った．

また，提案する手法では候補データベースのサイズに上限をおいているという点で十分条件を判定しているといえる．そこで禁止問合せがない場合について，提案する手法で判定可能なk-安全性の必要十分条件を与えた．

（３）#SMT ツールを用いた量的情報流解析

プログラムに入力された機密情報が出力に流出するかどうかを解析することは，情報漏えい対策やプライバシ保護において重要である．このような解析における定量的な尺度として，量的情報流が注目されている．プログラムの量的情報流とは，プログラム f の出力変数 Y の値を観測したときに入力変数 X の値についてどの程度の情報が得られるかを表す値であり，X と Y の相互情報量で定義される．量的情報流を解析する手法の一つとして，量的情報流の計算を背景理論付き論理式のモデル計数問題（#SMT）に帰着する手法が提案されており，背景理論付き論理式の充足可能性判定ツール（SMIT ソルバ）を用いて SMT の解を求めるツール（#SMT ツール）が実装されている．本年度は，量的情報流解析に用いられる#SMT ツールの計算効率の向上を以下の手順で行った．

はじめに，#SMT ツールにおいて充足不能コアとモデルのキャッシュを利用する改良を試みた．その改良を前提とし，出力変数値の探索順が計算効率に与える影響をベンチマークプログラムを用いて調査した．次に，SMT ツールを利用して得られるモデルから出力変数のサンプル値を取得し，そこから計算効率の向上が確率的に期待できる探索順を求める手法を提案した．また，論理式のレベルで簡単な充足可能性判定を前処理として行うことで探索空間を小さくする手法を提案した．最後に，解析対象プログラムの抽象解釈による静的解析を利用して出力変数値に関するビットレベルの解析を行い，その結果を利用して効率の良い探索順を得る手法を提案した．提案手法を実装したツールを用いて評価実験を行ったところ，モデル数が多いベンチマークについて実行時間を小さくすることができた．また，論理式のレベルで前処理や静的解析による前処理によって効率の良い探索順を得られることが確認でき，提案手法の有効性を示せた．

（４）形式級数の計数アルゴリズム

形式級数は各語に係数または重みと呼ばれる値を割当てる写像であり，形式言語の一般化となっている．特に，認識可能級数と代数的級数はそれぞれ，正則言語と文脈自由言語の一般化である．語 w の係数によって，w に対する操作のコストや w の発生確率などの量を表すことができ，形式級数はソフトウェアの定量的解析のモデルとして期待される．今年度は形式級数に関して以下の研究成果を得た．まず，形式級数に対する計数問題を２種定義し，これらに対する計数アルゴリズムを提案した．具体的に，与えられた形式級数 S と自然数 d に対し，CC(S,d)によって長さ d の語の S における係数の総和を表し，SC(S,d)によって長さ d の語で S における係数が非ゼロである個数を表す．本研究では，重み付きオートマトンで与えられた認識可能級数 S と自然数に対し，CC(S,d)を $O(\eta \log d)$時間（η は状態遷移行列の１回の演算に要する時間）で計算するアルゴリズム，および，S の状態遷移行列が乗算に関して可換であるとき，SC(S,d)を d に関する多項式時間で計算するアルゴリズムを示した．また，上記２つの計数問題を木級数に拡張しそれらを計算するアルゴリズムを与えた．さらに，代数的級数 S に対して CC(S,d)を d の 2 乗オーダの時間で計算するアルゴリズムも示した．提案アルゴリズムの有効性とソフトウェア脆弱性の定量評価への応用可能性を実証するため，代数的級数に対する提案アルゴリズムを実装したツールを Kaluza ベンチマークに対して実行したところ，既存の計数ツール ABC よりも多くの場合でより正確な計数を行えることを示した．

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１件）

(1) Trung CHU Bao, Kenji Hashimoto and

Hiroyuki Seki, Counting Algorithms for Recognizable and Algebraic Series, IEICE Transactions on Information and Systems, E101-D(6), June 2018, to appear.
DOI: 10.1587/transinf.2017FOP0003

〔学会発表〕（計１１件）

(1) 中島聖斗, 橋本健二, 酒井正彦, 関浩之, モデル計数を用いた量的情報流解析のための論理式簡約と静的解析, 電子情報通信学会技術研究報告 SS2016-61, 116(512), 7-12, March 9, 2017.

(2) Hiroyuki Seki, Kenji Hashimoto and Trung Chu Bao, Counting for Recognizable and Algebraic Series, 情報処理学会第 113 回プログラミング研究会, 2016-5-(4), March 3, 2017.

(3) 中島聖斗, Trung Chu Bao, 橋本健二, 酒井正彦, 関浩之, #SMT ツールを用いた量的情報流解析手法の高速化, 電子情報通信学会技術研究報告 SS2016-26, 116(277), 49-54, Oct 27, 2016.

(4) 浅井孝俊, 橋本健二, 関浩之, モデル計数を用いた XML データベースの k-安全性検査システムの高速化, 電子情報通信学会技術研究報告 SS2015-52, 115(420), 47-52, Jan 25, 2016.

(5) 浅井孝俊, 上杉正紀, 橋本健二, 関浩之, モデル計数を用いた XML データベースの k-安全性検査, 電子情報通信学会技術研究報告 SS2015-15, 115(20), 71-76, May 12, 2015.

(6) 関浩之, 量的情報流と差分プライバシー, 電子情報通信学会技術研究報告 SS2015-11, 115(20), 17-22, May 11, 2015.

(7) 小林靖幸, 楫勇一, 関浩之, 伊藤実, RSA 暗号の高速化手法に対するタイミング攻撃の情報理論的安全性評価, 2015 年暗号と情報セキュリティシンポジウム (SCIS2015), 4F1-2, Jan 23, 2015.

(8) 関浩之: セキュリティやプライバシーの定量的尺度について, 電子情報通信学会技術研究報告 IT2014-52, 114(353), 13-18, Dec 9, 2014.

(9) Yasuyuki Kobayashi, Yuichi Kaji, Hiroyuki Seki, Information Theoretical Evaluation of the Bucketing Technique to Mitigate Timing Attacks, International Symposium on Information Theory and Its Applications (ISITA 2014), 574-578, Oct 29, 2014.

(10) Yasuyuki Kobayashi, Yuichi Kaji, Hiroyuki Seki and Minoru Ito, Quantitative Evaluation of the Key Information that is Learned through Timing Attack – The contribution of Bucketing Technique for RSA Cryptosystem -, 電子情報通信学会技術研究報告, ISEC2014-35, 114(115), 253-258, July 4, 2014.

(11) Nut Sornchumni, Kenji Hashimoto and Hiroyuki Seki, Towards HMM Parameter Estimation with Differential Privacy, IPSJ SIG on Mathematical Modeling and Problem Solving (情報処理学会第 98 回数理モデル化と問題解決研究会), 2014-MPS-98(25), June 27, 2014.

〔その他〕
（解説記事）

(1) 関浩之, 量的情報流 – 概要と研究動向 -, 電子情報通信学会誌, 100(9), Sept 2017.

６．研究組織

(1)研究代表者
関　浩之（HIROYUKI SEKI）
名古屋大学・大学院情報学研究科・教授
研究者番号：８０１９６９４８

(2)連携研究者
橋本　健二（KENJI HASHIMOTO）
名古屋大学・大学院情報学研究科・助教
研究者番号：９０５４８４４７