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

令和　３ 年　 ６ 月　 ７ 日現在

機関番号：１３９０１
研究種目：基盤研究(C)（一般）
研究期間：2017～2020
課題番号：１７Ｋ００００９８

研究課題名（和文）量的情報流解析のための投射モデル計数ソルバの開発

研究課題名（英文）Projected Model Counters for Quantitative Information Flow Analysis

研究代表者
橋本　健二（Hashimoto, Kenji）

名古屋大学・情報学研究科・助教

研究者番号：９０５４８４４７

交付決定額（研究期間全体）：（直接経費）　　3,500,000円

研究成果の概要（和文）：ソフトウェアのセキュリティ・プライバシー解析における定量的尺度としてどの程度情報が漏洩したかを表現する量的情報流という概念が知られている。量的情報流解析の一手法として、命題論理式の解の個数を計算するモデル計数ソルバを利用した方法がある。本研究課題では、投射モデル計数の様々な計数技術を複合的に利用できる実行基盤を目指して、複数の計数手法を同じSATソルバ上に実装した上で、各々計算手法の実装の高速化や省メモリ化を実現した。また、量的情報流解析の一種である動的情報流解析に対して、実装したBDD/d-DNNF構築ツールを利用することで解析が高速化できることを確認しその有効性を示した。

研究成果の学術的意義や社会的意義
モデル計数・列挙技術を共通のSATソルバ上に実装することでそれらを複合的に利用できる環境を構築した。投射モデル計数・列挙を行うツールは一般のモデル計数・列挙のためのツールに比べて公開されているものはまだ少ない。今回、投射モデル計数を行うGPMCや投射モデル集合を表現するROBDDの構築を直接的に行うツールPC2BDDを開発し一般公開した。投射モデル計数・列挙は、量的情報流解析に限らず、規模の多い解析対象の部分的な性質を解析するのに有用であり今回のツールはそのような解析にも役立つ。

研究成果の概要（英文）：Quantitative Information Flow (QIF) was introduced for a quantitative security criterion to capture how much information is leaked though software systems. One way to analyzing QIF is based on projected model counting. In this work, we implemented on the SAT solver GlueMiniSat ROBDD and d-DNNF construction techniques for projected models. Besides, we used our BDD/d-DNNF compilers to analyze dynamic information flow analysis, a variant of QIF analysis, and confirmed its effectiveness.

研究分野：モデル計数

キーワード：投射モデル計数　量的情報流解析

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

情報システムにおける機密情報やプライバシーの保護の問題には、それらに対する社会通念やシステムサービス要件など様々な社会的および工学的要件が関係している。一般にどのようなサービスシステムにも可用性とプライバシーとのトレードオフが存在するので、要件に応じた適切なセキュリティレベルで運用可能なソフトウェアが求められる。そのため、トレードオフに関する客観的かつ定量的尺度とそれに基づいた設計技法の開発が望まれる。そうした尺度の１つとしてソフトウェア解析分野では量的情報流と呼ばれる概念が導入されている。プログラムの量的情報流とは、データに機密レベルを設定したとき、機密レベルの低い出力を観測することで機密レベルの高い入力についてどの程度知ることができるかを表す定量的な尺度である。近年、量的情報流が適切に定めた閾値を超えないことをセキュリティ要件とする考え方が提案され、このような新しいセキュリティ要件に関する正確かつ高速な解析技術の開発が重要となっている。プログラムの量的情報流解析の手法の 1 つに、命題論理式の解（モデル）の個数を計算するモデル計数（#SAT）ソルバを利用した方法がある。この手法では、解析対象のプログラムの入力関係を表現する命題論理式を構成し、それをモデル計数ソルバに与えてモデル数を計算することで量的情報流を求める。従来の型推論を利用する手法と比べて、この手法はある条件下のプログラムについては量的情報流を正確に計算できるという精度に関する長所をもつ。一方で、一般にはスケーラビリティが低く既存の実装では比較的小さいプログラムでも計算に数十分～数時間程度かかる場合もある。実行時間のほぼすべてはモデル計数ソルバによる求解時間であるため、いかにその求解時間を短縮するかが実用上の課題となっている。

命題論理式のモデル計数のためのソルバはいくつか存在するが、量的情報流解析に応用するためには特定の変数集合（投射変数集合）への真理値割り当ての違いにのみ注目してモデルの数を数える（投射モデル計数）機能が必要とされる。そのような機能をもつソルバとして sharpCDCL や #clasp などが知られている。研究代表者のグループでも、sharpSAT とよばれる既存のモデル計数ソルバを投射モデル計数に対応させたソルバ GPMC を開発している。得意とする問題のタイプはソルバによって大きく異なる。たとえば、sharpCDCL や#clasp などは複雑だがモデル数が比較的少ない問題を高速に解くことができ、sharpSAT などは単純だがモデル数が多い問題を得意とする。しかし、いずれのソルバを用いても、規模が大きい論理式だけでなく比較的な単純なプログラムから生成される論理式でさえも数時間かけても解けない例が少なくない。そのため、研究代表者は、量的情報流解析においてはプログラムを論理式に変換した後に既存の汎用のソルバでただ解くだけでは求解時間の短縮に限界があると考えている。加えて、その例の中には、プログラムを解析することでわかる事実から、結果が保存される範囲で論理式を簡単化、あるいはソルバ中での決定変数の順序を強制的に変更することで計算が数分もかからず完了するものがあることがわかっている。このことから、解析対象のプログラムを単に論理式に変換するだけでなく、そのプログラムの解析を別途行うことによってソルバによる計算戦略を決定する上で有効な情報が得られるのではないかと考えている。しかしながら、研究代表者が知る限り、既存のモデル計数ソルバにはそのような情報を直接受け付けて十分に活用できるだけの自由度は与えられていない。

２．研究の目的

(1) モデル計数・列挙の既存技術の量的情報流解析における有効性の調査と考察

命題論理式のモデル列挙や計数のための技術はこれまでに様々提案されている。本研究では、量的情報流解析問題から生成された論理式群に対して、それらの技術がどのような場合にどの程度有効かを明らかにすることを目的の１つとする。まず Minisat ベースの代表的な SAT ソルバである Glucose や GlueMiniSat を元にしてテストソルバを開発しその上に既存の計数技術を再実装する。研究代表者が知る限りでは既存のソルバは基本的には CDCL（Conflict-Driven Clause Learning）型の SAT 判定法をベースにしているものが多いが、実装レベルではデータ構造等に差異が存在する。加えて、特定のソルバにしか実装されていない技術もある。それらの技術を組み合わせて利用する場合についても検討・評価を行う。

(2) プログラム解析結果に基づく計算戦略の提案と実装

プログラムの事前解析結果を外部入力として受け付けてそれに応じてモデル計数の計算戦略を変更できるように、(1)で開発するソルバを拡張する。論理式以外の入力として、変数の選択順と、最初に優先して探索する空間を指定するための（一部の）変数への真理値割り当てを想定する。モデル計数ではどの順番で変数に真理値を割り当てながらモデルを探すかは実行時間に大きく影響を与える。たとえば、部分式の計数結果をキャッシュして重複計算を避ける機能を実装している場合、再利用される機会が多くかつなるべく大きい部分式を優先して計算しておくことが望ましい。本研究では、既存の汎用ヒューリスティクスだけに頼るのではなく、プログラム中に現れる演算の性質なども利用し、なるべく一度に多くのモデルが数えられる変数選択順や再利用できる可能性の高い部分式を導出できるような計算戦略を提案・実装する。

(3) 異なる計数処理技術の複合的な利用の検討

モデル計数問題は#P クラスに属する計算量が大きい問題であるため、問題の規模が大きくなれば必要な計算コストも急激に増加することは避けらない。そこで、マルチスレッド/マルチプロセスを使用した計数処理の並列化を検討する。並列化の方法として、異なる特徴をもった複数のソルバで同じ問題を解く方法（ポートフォリオ型）や、モデル計数の計算コストが均等になるように探索空間を分割する方法が考えられる。本研究では、(1)での結果を参考にしたポートフォリオ型ソルバの開発や、プログラムの事前解析結果を利用して探索空間を分割する手法の検討を行う。

３．研究の方法

(1) モデル計数・列挙の既存技術の調査と動的量的情報流解析への応用

モデル計数・列挙の既存技術として、成分分割およびキャッシングを利用した方法、2 分決定図（Binary Decision Diagram、 BDD）などのデータ表現を構築していく方法、阻止節を追加するソルバ方法について調査を行う。また、評価実験を通して手法の長所や短所を調査し、量的情報流解析への応用について検討する。

(2) GlueMinisat ベースでのモデル計数技術の再実装・改良と評価

異なる計数処理技術の複合的な利用に向けて、異なるタイプの計数技術を MiniSat 2.2.0 ベースの SAT ソルバをもとに再実装を行う。具体的には、まず投射モデル BDD を構築してモデル計数・列挙を行う方法の実装を行う。投射モデル集合ではなく全変数を対象とした通常のモデル集合に対する BDD を構成するツールとして cnf2obdd が知られているため、その実装を参考に GlueMinisat をベースとした投射モデル集合 BDD を構成するツールの開発を行う。

また、阻止節追加型投射モデル計数法についても GlueMinisat をベースに試作する。阻止節追加型のモデル計数は、SAT 判定を繰り返して異なるモデルを発見していく方法である。この方法では、モデルが 1 つ見つかるごとに、見つかったモデルの否定を表す節（阻止節）を追加して同じモデルが重複して発見されないようにする。阻止節を追加する際にその節から冗長なリテラルを取り除くことで高速化を図る手法が提案されている。それらの手法についても、投射モデル計数でも利用可能であるかを検討・改良を行う。

４．研究成果

(1) 投射モデル集合を表現する d-DNNF 及び BDD の構築と動的情報流解析への応用

研究代表者が所属する研究グループでは量的情報流解析の一種として動的情報流解析を提案している。動的情報流解析では、これまでのように解析対象のプログラムの入力関係を表現する命題論理式を構成してそのモデル数を一度計算するだけはなく、一部の変数に対する条件を変更しながら繰り返し計数を行う必要がある。生成された命題論理式に条件を追加した式に対して毎回計数処理を一から行うと時間がかかりすぎるため、論理式と投射変数集合からモデル計数・列挙が容易になる表現をあらかじめ生成しておくことで、追加条件が与えられてから計数結果が得られるまでの実行時間の短縮を図ることを考えた。BDD と d-DNNF（deterministic decomposable negation normal form）などのデータ表現は構築さえできてしまえばモデル計数・列挙は効率的にできることは知られている。そこで、投射モデル集合を表現する BDD や d-DNNF を構築するツールと、BDD や d-DNNF に対して追加条件のもとでの計数・列挙処理を行うツールの開発を行った。そして、GPMC を利用して毎回一から計数する方法と、BDD や d-DNNF を構築してから計数する方法の実行時間を比較した。その結果、最初に BDD、 d-DNNF を構築する時間は計数よりも時間がかかってしまうものの、構築された BDD、 d-DNNF に対して追加条件を与えて計数する時間は最大で約 100 分の 1 に抑えることができた。

(2) 異なるタイプの計数・列挙技術の GlueMinisat ベースでの再実装・改良

(1) で開発した BDD 構築ツールは cnf2obdd と同様に Minisat 1.14 であったが、GlueMiniSat をベースに再実装を行った。また、生成途中の BDD は簡約化された状態で維持されるようにした．これによりメモリ使用量を抑えることで、メモリ不足で BDD が構築できなかった問題に対しても構築できる問題が増加した。さらに、変数順序を入力として指定できるようにし、BDD サイズが小さくなるような変数順を求める Force などの既存ツールを投射モデル BDD 構築に適用することを試みたが、実行時間に大きな影響は観測できず改善には至らなかった。阻止節追加型ソルバの実装も GlueMiniSat をベースに行い，阻止節の既存の簡約化手法はそのまま投射モデル計数には利用できなかったため、発見したモデルを用いて再度投射変数優先の再割り当てを導入することで既存の簡約化手法を利用できるようにした。

(3) モデル計数競技会への参加とソルバの公開

国際会議 SAT Conference 2020 に併設された Model Counting Competition 2020 に参加し、投射モデル計数トラックにおいて GPMC は 3 位となった。なお、上位 1、2 位はいずれも近似ソルバが大きな役割を果たしているソルバであったのに対して、GPMC は厳密解のみを出力ソルバである。また、この分野で活躍する海外グループが開発した厳密解を計算する ganak や d4 などのソルバと比べて求解数で大きく上回った。ソルバ公開のための Web ページを用意し、ソルバのソースコードは gitlab にて公開している。

５．主な発表論文等

〔雑誌論文〕　計1件（うち査読付論文　1件／うち国際共著　0件／うちオープンアクセス　0件）

１．著者名 Bao Trung Chu, Kenji Hashimoto, Hiroyuki Seki	４．巻 E102-D
２．論文標題 Quantifying Dynamic Leakage - Complexity Analysis and Model Counting-based Calculation -	５．発行年 2019年
３．雑誌名 IEICE Transaction on Information and Systems	６．最初と最後の頁 1952-1965
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2019EDP7132	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計2件（うち招待講演　0件／うち国際学会　1件）

１．発表者名 磯貝 孝明, 橋本 健二, 酒井 正彦
２．発表標題 命題論理式の全ての投射モデルを表現するBDDの構成法
３．学会等名 人工知能学会 第116回人工知能基本問題研究会
４．発表年 2021年

１．発表者名 Bao Trung Chu, Kenji Hashimoto, Hiroyuki Seki
２．発表標題 On the Compositionality of Dynamic Leakage and Its Application to the Quantification Problem
３．学会等名 The 30th International Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2019)（国際学会）
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

Projected Model Counting and Enumeration: GPMC/PC2BDD
https://www.trs.cm.is.nagoya-u.ac.jp/~k-hasimt/tools/gpmc.html

6．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考

7．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

8．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関