様　式　Ｃ－１９、Ｆ－１９、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

平成　２７　年　　４　月　２２　日現在

機関番号：　１３１０１
研究種目：　基盤研究(C)
研究期間：　2012～2014
課題番号：　２４５０００７６

研究課題名（和文）悪質なサイトからの攻撃的な通信を遮断するコンテンツフィルタの実現

研究課題名（英文）On the Study of Content Filter Blocking Malicious Communication from Bad Websites

研究代表者

三河　賢治（Mikawa, Kenji）

新潟大学・学術情報基盤機構・准教授

研究者番号：００３４４８３８

交付決定額（研究期間全体）：（直接経費）　　3,900,000円

研究成果の概要（和文）：本研究は，近年増加する悪質なWebサイトからの攻撃的な通信からパソコンやモバイル端末などのネットワーク機器を守るためのフィルタリング技術の開発を目的としている。既存のフィルタリング技術では，自由度の高いルール記述ができず，多様化する脅威に対抗できなかった。本研究では，高速性能と省メモリ性能の両方を兼ね備えた，自由度の高いルール記述できる新しいフィルタリング技術を提案した。

研究成果の概要（英文）：In this study, we aim at developing a content filter to block malicious communication from bad websites, which attempt to intrude into our computers, mobile phone, and other network devices. Because they have difficulty in dealing with arbitrary bitmask rules, almost existing filtering algorithms recently have a trouble blocking such various malicious communication. We propose a novel content filter dealing with arbitrary bitmask rules which has good performance of high-speed filtering and consumed amount of memory.

研究分野：　計算機科学

キーワード：　ネットワーク　セキュリティ　パケット分類　フィルタリング　データ構造

１．研究開始当初の背景

近年ブロードバンド回線やスマートフォンが普及して「携帯アプリ」「インターネット通販」など Web サイトを利用するユーザが急増している。インターネット上には多数の悪質な Web サイトが存在し，偽の Web サイトに誘導して個人情報（ユーザ ID とパスワード，クレジットカードの番号等）を盗み取ったり，コンピュータウイルスに感染させたり，大きな被害を与えている。被害を食い止めるためには，ファイアウォールで HTTP 通信を完全にブロックしてしまえばよいが，ほとんどの情報検索やインターネットサービスが Web 化しているため，HTTP 通信を制限できない状況にある。また，最近の P2P ソフトウェアは，通信に使用するポート番号を自由に変更して通信してしまう。既存のファイアウォールは IP アドレスとポート番号を指定して通信をフィルタリングしているため，このような最近の脅威に対抗できない。このため，既存の方式に代わり，通信に含まれる固有のパタンを指定できるフィルタリング技術が必要不可欠であった。

既存のファイアウォールでは，探索木を利用した方式や，探索空間の分割を利用した方式が提案されている。これらの方式は，IP アドレスとポート番号で通信をフィルタリングする専用アルゴリズムであり，フィルタリング処理が高速である反面，通信に含まれる固有のパタンを自由に指定できない制約を抱えている。一方，固有のパタンを自由に指定できる方式の開発は発展途上にあり，最近の報告でも，フィルタリングの性能を犠牲にして省メモリ性能を上げるか，フィルタリング性能を上げて莫大なメモリ空間を浪費するか，決定的な方式が提案されていない。どちらのアルゴリズムも，固有のパタンを指定できる方式としては良い性能を示しているが，IP アドレスとポート番号を指定した場合には，既存の専用アルゴリズムの性能に遠く及ばない。固有のパタンを自由に指定できる方式，かつ，IP アドレスとポート番号を指定した場合にも専用アルゴリズムに匹敵するようなフィルタリング技術の実現が待ち望まれていた。

２．研究の目的

本研究の目的は，固有のパタンをフィルタリングルールに自由に指定できるフィルタリング技術の新しい構成方法を提案し，実際の運用に耐えうるフィルタリング技術を開発することである。要求されるフィルタリング技術の性能は，フィルタリング時の処理速度と使用メモリ量である。特に，最悪の処理速度，使用メモリ量の評価が重要視されている（最悪時の性能を明らかにすることで，使用環境に依存しない性能保証が可能となるためである）。具体的な目標は，本研究で提案するフィルタリング技術について，これらの性能を明らかにして既存のフィルタリング技術よりも優位であることを示す。

研究代表者らのこれまでの研究成果により，理論的に効率的なフィルタリングルールの探索アルゴリズムの開発は完了している（引用文献②）。この探索アルゴリズムは，固有のパタンをフィルタリングルールに自由に指定できるフィルタリング技術であるが，まだまだ荒削りな部分も多い。しかしながら，この探索アルゴリズムは，フィルタリング処理速度の高速化の可能性を秘めている。本研究では，この探索アルゴリズムをチューニングして，高速化を実現しつつ省メモリ化に配慮した探索アルゴリズムの開発を目標とする。

３．研究の方法

本研究で提案するフィルタリング技術の性能について，理論的な性能と実践的な性能の両方を評価するため，次のように研究をすすめる。

(1) 実証実験のためのネットワーク環境を構築する。研究期間の初期の段階では，実際の通信を利用して，提案フィルタリング技術の性能を評価するための環境を構築する。そこで，実際の通信を保存するための仕組みの構築と提案フィルタリング技術の性能評価のための実証実験環境を構築する。

① 実際の通信を保存するための機構の構築については，外部接続用ルータのミラーポート（すべての通信のコピーを転送するポート）に実験用ルータを接続して，実際の通信を保存用ストレージに記録する。保存された通信記録は，今後の実証実験で利用するためのものである。

② 提案フィルタリング技術の性能評価のための実証実験環境については，評価用ファイアウォールに対して，パケット送信用コンピュータＡとパケット受信用パソコンＢを接続する環境を構築する。フィルタリング時の処理速度は，評価用ファイアウォールに探索アルゴリズムを置き，ファイアウォールを通過する同一パケットの送信時刻と受信時刻の差分を計測する。

(2) フィルタリングルール探索アルゴリズムをチューニングする。引用文献②のフィルタリングルール探索アルゴリズム（以下，ベースアルゴリズム）を基礎に，高速化／省メモリ化をそれぞれ検討する。

① 高速化の検討については，ベースアルゴリズムに決定木のプロセスを取り入れ，ベースアルゴリズムの高速化の可能性を探る。ベースアルゴリズムは，決定木のプロセスを取り入れることが容易なデータ構造を持つ。このため，決定木のプロセスを取り入れることでフィルタリング処理の高速化が期待できる。理論的な評価で言えば，ベースアルゴリズムは，ルール数がある程度増えると，ルール数に比例してルール探索の処理時間が長くなってしまう。提案アルゴリズムは，ベースアルゴリズムに対して，ルール数の増加に

関係なく，ルール探索の処理時間が一定になることが期待できる。
② 高速化と省メモリ化はトレードオフの関係であることが知られており，省メモリ化の検討は前項の高速化の検討とは正反対の検討となる。一方，引用文献②で研究代表者らは，フィルタリング処理に関しては実際の運用に影響がない程度に省メモリ化が可能であることを示した。実際の観点から，ベースアルゴリズムの改良を検討する。また，前項で検討する決定木のプロセスを取り入れた高速化アルゴリズムに対して，省メモリ化を検討する。決定木のプロセスを単純に取り入れるだけでは，フィルタリング処理時の使用メモリ量が指数関数的に爆発してしまう。ルールの探索に不必要な探索経路を決定木から削除（枝刈り）することにより，高速化を犠牲にすることなく省メモリ化の達成が期待できる。

４．研究成果

(1) 実証実験のためのネットワーク環境を構築に関する研究成果は次のとおりである。実証実験で利用するため，実際の通信を保存用ストレージに保存し，蓄積された通信データが実証実験に役立つものであるか検討を行った。実際の通信の挙動に関して，正常の通信がほとんどを占めており，フィルタリングすべき情報が少ないことが確認できた。蓄積された実際の通信を実証実験に利用しても，フィルタリング性能を正確に評価できないため，フィルタに高負荷を与える疑似的な通信データを生成するアルゴリズムの検討を行った。
① フィルタに高負荷を与える疑似的な通信データを検討した結果，フィルタの構造から得られる通信データが最も高負荷を与えることが分かった。フィルタを自動生成するツール，パケット転送先をランダムに決定するツールの開発を行った。この成果は，雑誌論文②，③，⑥，⑦，学会発表④，⑤，⑥，⑦，⑨で発表した。

(2) フィルタリングルール探索アルゴリズムのチューニングに関する研究成果は次のとおりである。
① ベースアルゴリズムに決定木のプロセスを取り入れ，ベースアルゴリズムの高速化を検討した。ベースアルゴリズムは，ビット展開したフィルタリングルール（ワイルドカードを含む）を連に区切りトライを構成する。

表 1：フィルタリングルール

番号	ルール	番号	ルール
R_1	*0*1	R_7	**10
R_2	0000	R_8	01**
R_3	0*00	R_9	*11*
R_4	0*1*	R_{10}	*000
R_5	1100	R_{11}	*1*1
R_6	*01*	R_{12}	***1

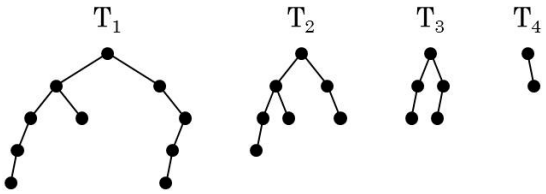

図 1：連分割トライ

例えば，ベースアルゴリズムは，表 1 に示すフィルタリングルールに対して図 1 のトライを構成する。ベースアルゴリズムに対して，決定木のプロセスを取り入れることによって，フィルタリング処理の計算量をトライの高さの 2 乗まで短縮し，ルール数の増加に関係なくルール探索の処理時間が一定となるフィルタリング技術を確立した。この成果は雑誌論文①で発表した。
② 省メモリ化に関しては，決定木プロセスを取り入れた探索アルゴリズムに対する不要な探索経路の枝刈りを行った。決定木プロセスを取り入れた提案アルゴリズムの使用メモリ量は指数関数的に増大してしまったが，枝刈りを行うことによって，連分割トライの格納に必要な総メモリ量とほとんど変わらないことが実証実験で確かめられた。この実証実験の過程で，ベースアルゴリズムの連分割トライをシンプルに再構成した新しい探索アルゴリズムを開発した。連分割トラ

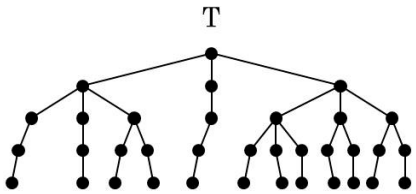

図 2：提案トライ

イと同様に，表 1 に示すフィルタリングルールに対して図 2 のトライを構成する。連分割トライと比べて，使用メモリ量は増加してしまうが，既存の線形探索によるフィルタリング技術よりも格段の高速化を達成することができた。

提案手法の性能を評価するため，代表的な既存手法である線形探索（固有パタンを自由に指定できるフィルタリング技術）と階層トライ（探索木を利用した高速であるが固有パタンを指定できない専用アルゴリズム）との比較実験を行った。実験の内容は次のとおりである。負荷の異なる 3 種類の疑似フィルタ（ACL，FW，IPC）を各アルゴリズムに搭載し，処理時間（処理時間で比較すると計算機の性能や環境に依存するので，フィルタリングルールとパケット（通信データ）の比較回数とした）を計測した。また，ルール数による影響も考慮して，疑似フィルタに搭載するフィルタリングルールの個数を 1,000 個と 5,000 個に変えて実験を行った。はじめに，

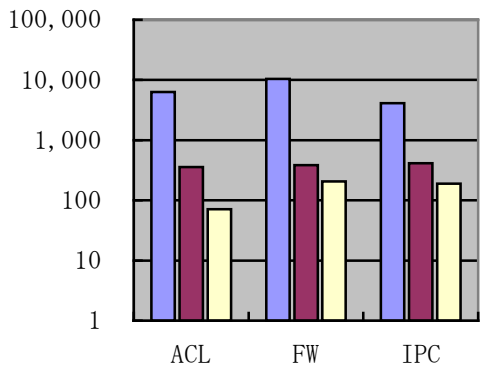

図 3：ルール数が 1,000 個の場合

図 3 は，フィルタリングルール数 1,000 個のときの実験結果である。グラフは，左から順に線形探索，階層トライ，提案手法である。提案手法は，固有パタンを自由に指定できるフィルタリング技術の代表格である線形探索と比較して圧倒的に高速である（大きな性能差となったので対数軸とした)。次に，

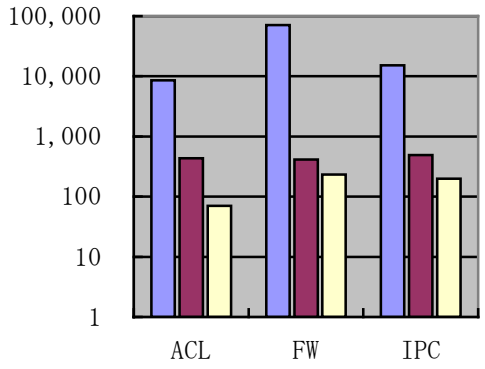

図 4：ルール数が 5,000 個の場合

図 4 は，フィルタリングルール数 5,000 個のときの実験結果である。グラフは，左から順に線形探索，階層トライ，提案手法である。線形探索は（疑似フィルタによって振る舞いは異なるが）ルール数に比例して比較回数が増加している。一方，探索木を利用した階層トライと提案手法は，ルール数に関係なく，処理速度が安定していることが分かる。

＜引用文献＞

① 三河賢治，田中賢，制約ルール集合上のパケット分類の領域計算量に関する考察，信学会ソサイエティ大会，2011 年 9 月 13 日，北海道大（北海道・札幌市）
② 長谷川創，三河賢治，田中賢，任意のビットマスクに対応した階層型トライの提案，信学会総合大会，2011 年 3 月 15 日，東京都市大（東京都・世田谷区）

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 Kenji Mikawa, Ken Tanaka, Run-Based Trie Involving the Structure of Arbitrary Bitmask Rules, IEICE Trans. Inf. & Syst., 査読有, Vol.E98-D, 2015（掲載決定）
② Kenji Mikawa, Ken Tanaka, Lexicographic Ranking and Unranking of Derangements in Cycle Notation, 査読有, Discrete Applied Mathematics, Vol.166, 2014, 164-169, http://dx.doi.org/10.1016/j.dam.2013.10.001
③ 三河賢治，田中賢，撹乱順列の線形時間ランキングとアンランキングについて，査読無，数理解析研究所講究録，No.1849, 2013, 12-17
④ Ken Tanaka, Kenji Mikawa, Kouhei Takeyama, Optimization of Packet Filter with Maintenance of Rule Dependencies, 査読有, IEICE Commun. Express, Vol.2, 2013, 80-85, http://dx.doi.org/10.1587/comex.2.80
⑤ Ken Tanaka, Kenji Mikawa, Manabu Hikin, A Heuristic Algorithm for Reconstructing a Packet Filter with Dependent Rules, 査読有, IEICE Trans. Commun., Vol.E96-B, 2013, 155-162, http://dx.doi.org/10.1587/transcom.E96.B.155
⑥ 三河賢治，田中賢，撹乱順列の線形時間ランダム生成について，査読無，信学技法，No.112, 2012, 53-58
⑦ 三河賢治，田中賢，巡回表記で表された撹乱順列に対する辞書順のランキングとアンランキングについて，査読無，信学技法，No.112, 2012, 93-96

〔学会発表〕（計 9 件）

① 小林由人，高橋俊彦，三河賢治，田中賢，トライを用いた高速パケット分類法の提案，信学会総合大会，2015 年 3 月 11 日，立命館大（滋賀県・草津市）
② 原田崇司，田中賢，三河賢治，決定木を用いた Run-Based Trie の探索法，信学会ソサイエティ大会，2014 年 9 月 26 日，徳島大（徳島県・徳島市）
③ 池本泰斗，田中賢，三河賢治，パケットフィルタリング最適化のためのルール集合分割法，情報科学技術フォーラム，2014 年 9 月 5 日，筑波大（茨城県・つくば市）
④ 三河賢治，田中賢，符号化を必要としない撹乱順列の線形時間ランキングとアンランキングについて，信学会ソサイエテ

ィ大会，2013 年 9 月 18 日，福岡工大（福岡県・福岡市）

⑤ 野村圭太，田中賢，三河賢治，パケットフィルタリング最適化法の有効性について，情報科学技術フォーラム，2013 年 9 月 6 日，鳥取大（鳥取県・鳥取市）

⑥ 三河賢治，田中賢，符号化を必要としない撹乱順列の線形時間ランダム生成について，情報科学技術フォーラム，2013 年 9 月 4 日，鳥取大（鳥取県・鳥取市）

⑦ 三河賢治，田中賢，撹乱順列の線形時間ランキングとアンランキングについて，LA シンポジウム，2013 年 1 月 28 日，京都大（京都府・京都市）

⑧ 秋山匠，岩本武留，田中賢，三河賢治，フィルタリングポリシー記述言語でのルール最適化について，信学会ソサイエティ大会，2012 年 9 月 12 日，富山大（富山県・富山市）

⑨ 明田川卓，三河賢治，辞書順に並ぶ順列のランク付け操作とその逆操作に対する O(nloglogn)領域を用いた線形時間アルゴリズムの簡単化，情報科学技術フォーラム，2012 年 9 月 4 日，法政大（東京都・小金井市）

６．研究組織

(1)研究代表者

三河　賢治（MIKAWA, Kenji）

新潟大学・学術情報基盤機構・准教授

研究者番号：00344838

(2)研究分担者

田中　賢（TANAKA, Ken）

神奈川大学・理学部・教授

研究者番号： 50272810