様式 C-19

科学研究費補助金研究成果報告書

平成２３年６月１７日現在

機関番号：２５４０３
研究種目：基盤研究（C）
研究期間：2008～2010
課題番号：２０５０００５４
研究課題名（和文）ネットワーク侵入検知のための高速ストリングマッチングハードウェアに関する研究
研究課題名（英文）A Study on High Performance String Matching Hardware for Network Intrusion Detection
研究代表者
若林　真一（WAKABAYASHI SHINICHI）
広島市立大学・情報科学研究科・教授
研究者番号：５０２１０８６０

研究成果の概要（和文）：ネットワーク侵入検知のための、ウィルスパターンを実行時に設定可能なシストリックアルゴリズムに基づく正規表現マッチングハードウェアを提案した。また、正規表現マッチングと近似文字列照合を組み合わせた近似正規表現マッチングに対しても新しいマッチングハードウェアを提案した。これらの研究成果により、未知のネットワークウィルス等に対しても柔軟かつ迅速に対処可能なネットワーク侵入検知システムの構築が可能になった。
研究成果の概要（英文）：　We have newly proposed a regular expression matching hardware engine based on a systolic algorithm, in which virus patterns can be set during the execution, for network intrusion detection.　We have also presented a matching engine for approximate regular expression matching, in which regular expression matching and approximate string matching were combined.　From those results, it has become possible to construct a network intrusion detection system, which can handle unknown network viruses in a flexible and efficient manner.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２００８年度	1,000,000	300,000	1,300,000
２００９年度	1,300,000	390,000	1,690,000
２０１０年度	1,000,000	300,000	1,300,000
年度			
年度			
総　計	3,300,000	990,000	4,290,000

研究分野：論理設計、組合せ最適化
科研費の分科・細目：情報学　・　計算機システム・ネットワーク
キーワード：リコンフィギャラブルシステム、ストリングマッチング、FPGA、ネットワーク侵入検知

１．研究開始当初の背景

インターネットが情報化社会を支える最も重要なインフラストラクチャとなった現在において、インターネットに侵入するウィルスメールやワームから情報システムを防護する技術の重要性は非常に高くなっている。ウィルスメールやワームの検知はネットワーク侵入検知（Network Intrusion Detection, NID）と総称され、多くの研究が行われている。NID は、従来はソフトウェアにより実現されていた。しかしながら、伝送速度がギガビットを超える高速ネットワークが一般的になるにつれて、NID の処理の中核であるパターンマッチングをソフトウェアで実現することは事実上不可能になってきており、このため、NID をハードウェアで実現する研究が盛んになってきた。

ネットワークへの侵入や攻撃を行うウィ

ルスメールやワーム（以下では総称してウィルスという）はパケットに分割されて伝送されるため、ルータやネットワークスイッチにおいてリアルタイムにそれらを検知し、異常を見つければそのパケットを破棄することが最も効率的で、かつネットワークシステムに対しても安全である。ウィルスの検知はネットワーク上を伝送されるデータがあらかじめ登録されたパターンに一致するかどうかで判断される。このパターンは一般には正規表現で記述される。このため、NID への応用を前提とした正規表現に対する高速パターンマッチングを実現するハードウェアの研究が 1990 年代後半から盛んに行われている。

NID のための正規表現に対する専用マッチングハードウェアに関してはすでに多くの研究成果が得られており、実用化を前提とした実機システムにおける検証も実施されている。これらの従来手法の大部分は、パターンの集合に対する専用マッチングハードウェアを FPGA 上に実現することで NID システムを実現している。各パターンは専用回路として FPGA に実現されるので、ギガビットネットワークの伝送速度にも対応した高速ストリングマッチングが可能になっている。

しかしながら、NID に対する従来手法にはいくつかの問題点がある。その一つは、パターンがハードウェアとして FPGA 上に実現されるため、パターンの変更は回路構成の変更につながることである。FPGA 上に実現する回路を変更するためには、論理合成、配置・配線、コンフィグレーションデータの生成等の設計工程を必要とし、これらの工程は現状では少なくとも 30 分以上の時間を要する。一方、ネットワークウィルスは時として短時間に爆発的にインターネットに拡大することがあり、NID のパターン更新は可能な限り短時間であることが望ましい。

従来手法の第二の問題点としては、従来手法の大部分は、NID の代表的なソフトウェアプログラムである Snort のパターンを用いることを前提として開発されているが、従来手法では Snort のパターンの一部については扱うことができないことである。このため、Snort のすべてのパターンを効率よく実現できる専用マッチングハードウェアの研究開発が望まれる。

２．研究の目的

NID は大量の入力系列データから与えられたパターンにマッチする部分系列を検索・抽出するストリングパターンマッチング(以下ではストリングマッチングという)として定式化される。本研究は、研究代表者がこれまでに行ってきたストリングマッチング専用ハードウェアの研究を基礎として、これまでに得られた研究成果をさらに拡張し、再構成可能論理デバイスである FPGA を利用してネットワーク侵入検知システム（NIDS）を専用ハードウェアとして構成し、ギガビットネットワークに対してリアルタイムで NID を実現する手法を開発することで、ネットワークシステムのセキュリティを飛躍的に高めることを目的とする。

また、従来の NID におけるパターンマッチングでは、パターンに正確にマッチするウィルスのみを検出することとしていたが、既知のウィルスの亜種、変種が日々、発見されている現状を考慮すると、パターンに正確に一致するウィルスだけでなく、入力データがパターンにある程度一致すればウィルス感染の疑いのあるデータとして検知することが望ましい。このため、正規表現マッチングに近似文字列マッチングの概念を取り入れた近似正規表現マッチングに対しても専用ハードウェアを開発することを目的とする。

３．研究の方法

(1) ストリングマッチングのハードウェア解法の開発

NID のための正規表現をパターンとするストリングマッチングのハードウェア解法を開発する。正規表現に対するストリングマッチングハードウェアの開発にあたっては研究代表者がすでに得ている研究成果を基礎として、これの拡張を行う。

研究代表者らが以前に提案したストリングマッチングハードウェアはクリーネ演算子のネスト回数を１に制約した正規表現に対する並列マッチングアルゴリズムを提案している。このアルゴリズムを基礎として、クリーネ演算子のネスト回数に関する制約をはずした一般の正規表現に対するマッチングアルゴリズムを開発する。また、FPGA の内部メモリを用いたテーブル検索など、できるだけ FPGA のハードウェア資源を有効利用するような回路構成を工夫する。

(2) 近似正規表現マッチングハードウェアの開発

ネットワーク侵入検知(NID)ではパターンの定義は正規表現を用いることが一般的であるが、パターンの効率のよい定義には文字列間の距離（編集距離という）に基づく定義も有用である。このため、本研究においては、正規表現マッチングと編集距離に基づくマッチングを組み合わせたストリングマッチングを近似正規表現マッチングと定義し、近似正規表現マッチングを効率よく実行するマッチングハードウェアを開発する。

(3) ハードウェア解法の FPGA 上での実現と評価

(1), (2)で開発したハードウェア解法を FPGA 上に回路として実現するため、回路の HDL 記述を行って論理シミュレーションにより動作を検証する。次に、FPGA 設計ツールを用いて論理合成、配置配線を行い、合成された回路が正しく動作するかどうかをシミュレーションにより検証する。さらに、FPGA 評価ボード上に提案ハードウェアを実現して動作を確認するとともに、回路の規模や動作周波数について評価する。

(4) 研究成果の公表

本研究で得られた成果については、関連する国内研究会等で発表すると共に、学術論文にまとめて学術雑誌論文に投稿する。

４．研究成果

(1) 文字列の繰返しを指定する量指定子をパターンとして使用できるシストリックアルゴリズムに基づく正規表現マッチングハードウェアを提案した。さらに、この研究成果を拡張し、シストリックアルゴリズムと非決定性有限オートマトンを組み合わせた新しいマッチングハードウェアを提案し、以前のアルゴリズムよりさらに広い正規言語のサブクラスをパターンとして使用することを可能にした。

以下では紙面の都合上、拡張したアルゴリズムの結果のみを示す。NIDSのハードウェア実装として、非決定性有限オートマトン(NFA)を用いた手法がよく知られている。正規表現で記述されたウィルスパターンからNFAを生成し、生成したNFAをFPGAのような再構成可能デバイスに実装する事でストリングマッチングを実現している。この手法ではパターンに依存した回路構成となるため、パターンの更新が行われるたびに回路の再合成・再構成が必要となる。そのため、パターン更新が頻繁に起こるNIDSにおいては新たなウィルスパターンへの対応が遅れるという問題が生じる。

本研究では、定数パターン文字数内で任意の正規表現のマッチングが可能なNFA回路(パターン非依存NFA回路)と我々が提案したシストリックアルゴリズムに基づくマッチングハードウェアを組み合わせることで、従来では扱えなかったクリーネ演算のネスト（KK項と呼ぶ）を扱うための手法を提案した。具体的には、KK項以外の部分パターンについては、シストリックアルゴリズムに基づくマッチングハードウェアでマッチングを実現し、KK項については、NFAの状態遷移と文字列マッチングの機能を分離し、状態遷移についてはパターン非依存NFA回路で実現し、文字列マッチングについては、文字列マッチング回路（セルとよぶ）で実現した。

上記のようなハードウェア構成とすることで、提案手法ではパターンをマッチング実行時に設定可能となり、従来手法のようなパターン更新時の回路の論理合成・配置配線・コンフィギュレーションが不要となるため、ウィルスパターンの更新に対しても瞬時の対応が可能となり、ネットワークのセキュリティ向上に貢献すると考えられる。

本研究で提案したマッチングハードウェアについて面積と速度の評価を行った。統合開発環境はXilinx社のISE12.1、FPGAはVertex-4（XC4VLX100-11F1513）を用いた。扱うパターンはSnort ルールv2.7とした。 面積評価は実験で用いたFPGA に実装できるセル数で評価した。提案ハードウェアの合成の結果、動作周波数271.7MHzであり、上記のFPGA上に1000セルを実装可能であることが分かった。この結果から、提案手法はKK項を効率よく実現できることがわかった。

(2) 正規表現マッチングと近似文字列照合を組み合わせた近似正規表現マッチングに対し、シストリックアルゴリズムに基づく新しいマッチングハードウェアを提案し、実際にFPGA上に実現して評価した。

具体的には、柔軟なテキスト検索が可能となるように、正規表現マッチングと近似文字列照合を組み合わせ、パターンに正規表現を使用し、そのパターンに類似する部分文字列を入力系列から探し出すというテキスト検索問題を解くための効率的なハードウェアアルゴリズムを提案した。提案アルゴリズムは１次元シストリックアーキテクチャで実現される。

提案１次元シストリックアーキテクチャでは文字を比較する回路（セル）を１次元配列状に相互接続することで実現される。近似正規表現マッチングは、与えられた文字列とパターンの編集距離を計算することで実現される。本研究では、近似文字列照合問題のパターン記述にユニオンやクリーネ演算子等の正規表現演算子を導入し、正規表現演算子を含むパターンに対する編集距離を新たに定義した。入力系列とパターンの編集距離を１次元シストリックアルゴリズムで並列計算することで、入力系列の長さと同じクロック数でマッチングを実行できる。

本研究で提案したシストリックアルゴリズムに基づく専用マッチングハードウェアをVerilog HDLで記述し、FPGA上に実装した。FPGAはAltera社のStratix EP1S60F1020C7を使用した。比較対象として近似正規表現マッチングのアルゴリズムをソフトウェアで実現し、提案したハードウェアアルゴリズムとの計算時間の比較を行った。ソフトウェアはCPUがPentium D 3.0GHz、メインメモリが2GBのPC

上で実行した。また、コンパイラにはBlorland C++ 5.5.1を使用し、最適化オプションとして-O2を使用した。

ハードウェアの設計結果は、250個の文字比較回路（セル）を実装した場合、論理ブロック数は41,491個となり、論理ブロック数の使用率は72.6%となった。また、セル１個当たりの論理ブロック数は166個となった。最高動作周波数は79.92MHz となった。また、FPGA上で回路を構成し、正しく回路が動作するかを確認した。ソフトウェアとハードウェアの実行時間の比較を行った結果、提案ハードウェアはソフトウェアに対する速度向上比が364倍となっており、十分な高速化が達成されていることが確認できた。

(3) NIDS におけるパターンは一般に数100 種類と数が多く、頻繁に更新されるという特徴がある。さらに、更新にあたっては瞬時の対応が要求される。本研究ではネットワークのセキュリティを高度に保つことを目的として、パターンを回路に組み込む従来のパターン依存マッチングハードウェアと、我々が提案したパターンを実行時に設定可能なパターン非依存マッチングハードウェアを組み合わせた新しい NIDS のシステム構成を提案した。提案システムでは、既知のパターンに対してはコンパクトな回路構成で高速にストリングマッチングを実行できるパターン依存ハードウェアでストリングマッチングを行い、更新されたパターンに対してはパターン非依存ハードウェアでストリングマッチングを行う。そのため、提案システムは、パターンの更新に対して瞬時に対応可能でありながら、かつ、実現するためのハードウェアコストが低い、という特徴がある。

(4) 本研究の成果により、未知のネットワークウィルス等に対しても柔軟に、かつ迅速に対処可能なネットワーク侵入検知システムの構築が可能になった。

今後の課題としては、さらなる回路の高速化によるマッチング時間の短縮、回路規模の削減による面積効率の向上、近似正規表現マッチングにおいて利用可能な文字列演算子の拡大などがある。

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９件）

[1] Yoichi Wakaba, Masato Inagi, Shinichi Wakabayashi, Shinobu Nagayama: An Efficient Hardware Matching Engine for Regular Expression With Nested Kleene Operators, Proc. 21st International Conference on Field Programmable Logic and Applications (FPL 2011), to appear.

[2] 宇丹裕一朗, 若林真一, 永山忍: 近似正規表現マッチングのためのシストリックアルゴリズムとそのFPGA実装, 電子情報通信学会論文誌D, Vol.J94-D, No.6, pp.934-944 (2011).

[3] Yuichiro Utan, Shinichi Wakabayashi, Shinobu Nagayama: An FPGA-Based Text Search Engine for Approximate Regular Expression Matching, Proc. 2010 IEEE International Conference on Field Programmable Technology (ICFPT 2010), pp.184-191 (2010).

[4] Yoichi Wakaba, Masato Inagi, Shinichi Wakabayashi, Shinobu Nagayama: An Extension of Systolic Regular Expression Matching Hardware for Handling Iteration of Strings Using Quantifiers, Proc. 16th Workshop on Synthesis And System Integration of Mixed Information technologies, R4-10, pp.412-417 (2010).

[5] 川中洋祐, 若林真一, 永山忍: パターン非依存型正規表現ストリングマッチングマシンとそのFPGA 実装, 電子情報通信学会論文誌D, Vol.J92-D, No.12, pp.2159-2167 (2009).

[6] Yosuke Kawanaka, Shinichi Wakabayashi, Shinobu Nagayama: A Systolic String Matching Algorithm for High-Speed Recognition of a Restricted Regular Set, Proc. 2009 International Conference on Engineering of Reconfigurable Systems and Algorithms (ERSA 2009), pp.151-157 (2009).

[7] Yosuke Kawanaka, Shinichi Wakabayashi, Shinobu Nagayama: A Fast Regular Expression Matching Engine for an FPGA-based Network Intrusion Detection System, Proc. 15th Workshop on Synthesis And System Integration of Mixed Information technologies, R1-14, pp.88-93 (2009).

[8] Yosuke Kawanaka, Shinichi Wakabayashi, Shinobu Nagayama: A Systolic Regular Expression Pattern Matching Engine and its Application to Network Intrusion Detection, Proc. 2008 IEEE International Conference on Field Programmable Technology (ICFPT 2008), pp.297-300 (2008).

[9] Sadatoshi Mikami, Yosuke Kawanaka, Shinichi Wakabayashi, Shinobu Nagayama: Efficient FPGA-based Hardware Algorithms

for Approximate String Matching, Proc. 23rd International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC 2008), pp.201-204 (2008).

〔学会発表〕(計 10 件)
[1] 宇丹裕一朗, 若林真一, 永山忍: 近似文字列照合問題の高速な解法に関する研究とそのFPGA実装, 第33回多値論理フォーラム, 多値論理研究ノート, 第33巻, No.1, pp.1-1-1-6 (2010年9月11日, 広島).
[2] 若葉陽一, 稲木雅人, 永山忍, 若林真一: クリーネ演算のネストに対応した効率的な正規表現マッチングハードウェア, 第33回多値論理フォーラム, 多値論理研究ノート, 第33巻, No.2, pp.2-1-2-7 (2010年9月11日, 広島).
[3] 宇丹裕一朗, 若林真一, 永山忍: 高速なテキスト検索のための近似正規表現マッチングアルゴリズムとそのFPGA実装, 第9回情報科学技術フォーラム(FIT2010)講演論文集, RC-003, 第1分冊, pp.69-74 (2010年9月8日, 福岡).
[4] 若葉陽一, 若林真一, 永山忍, 稲木雅人: 量指定子による文字列の繰り返しに対応した正規表現マッチング専用ハードウェア, 電子情報通信学会リコンフィギャラブルシステム研究会技術研究報告, RECONF2009-76 (2010年1月27日, 横浜).
[5] 若葉陽一, 若林真一, 永山忍: 量指定子を含む正規表現マッチングに対するハードウェアアルゴリズム, 第11回IEEE広島支部学生シンポジウム論文集, D-57 (2009年11月27日, 宇部).
[6] 若葉陽一, 若林真一, 永山忍: 量指定子を実現可能なパターン非依存型の正規表現マッチング専用ハードウェア, 平成21年度電気・情報関連学会中国支部第60回連合大会講演論文集, 20-8, pp.473-474 (2009年10月17日, 広島).
[7] 若葉陽一, 川中洋祐, 永山忍, 稲木雅人, 若林真一: Snortルールを入力とするネットワーク侵入検知ハードウェアの開発, 情報処理学会第71回全国大会講演論文集, 4K-6, pp.1-117-1-118 (2009年3月10日, 草津).
[8] 川中洋祐, 若林真一, 永山忍: NIDS専用正規表現マッチングマシンの構成とそのFPGA実装, 電子情報通信学会リコンフィギャラブルシステム研究会技術研究報告, RECONF2008-87 (2009年1月30日, 横浜).
[9] 川中洋祐, 若林真一, 永山忍: パターン非依存の回路構成を持つストリングマッチングマッチングマシンを用いたネットワーク侵入検知, 第10回IEEE広島支部学生シンポジウム論文集, D-09 (2008年11月22日, 広島).
[10] 川中洋祐, 若林真一, 永山忍: ストリングマッチングマシンのFPGAによる実現とネットワーク侵入検知システムへの応用, 情報処理学会DAシンポジウム2008, pp.91-96 (2008年8月26日, 浜松).

６．研究組織
(1)研究代表者
若林　真一 (WAKABAYASHI SHINICHI)
広島市立大学・情報科学研究科・教授
研究者番号：50210860

(2)研究分担者
なし

(3)連携研究者
永山　忍 (NAGAYAMA SHINOBU)
広島市立大学・情報科学研究科・准教授
研究者番号：10405491
稲木　雅人 (INAGI MASATO)
広島市立大学・情報科学研究科・助教
研究者番号：50468302