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　２９　年　　６　月　　３　日現在

機関番号：　２５４０３
研究種目：　基盤研究(C)（一般）
研究期間：　2014～2016
課題番号：　２６３３００６９

研究課題名（和文）次世代超高速イーサネットのためのネットワーク侵入検知ハードウェアに関する研究

研究課題名（英文）A Study on Network Intrusion Detection Hardware for Next-Generation High-Speed Ethernets

研究代表者

若林　真一（Wakabayashi, Shin'ichi）

広島市立大学・情報科学研究科・教授

研究者番号：５０２１０８６０

交付決定額（研究期間全体）：（直接経費）　　3,600,000円

研究成果の概要（和文）：　本研究では、伝送速度が毎秒40ギガビット以上の超高速インターネットに対するネットワーク侵入検知（NID）システムの新しい構成を提案した。提案システムにおいては、システムに入力されるパケットをまず専用回路で高速にスクリーニングし、ウィルス感染が疑われる「怪しい」パケットだけに対して完全なマッチングを行うことで、スループットを確保しながら、回路規模の削減を可能とした。計算機実験により提案システム構成法の有効性を確認した。

研究成果の概要（英文）：　This research proposes a new structure of network intrusion detection systems for very high-speed internet, whose transmission rate is more than 40Ggps. In the proposed system, a dedicated circuit performs “screening” of input packets, and for suspicious packets which might contain “computer virus”, complete matching will be performed to determine they are in fact infected by a virus. The proposed system can achieve a high throughput while the circuit size can be reduced. Experimental results show the effectiveness of the proposed system.

研究分野：　情報工学

キーワード：　ネットワーク侵入検知　FPGA　正規表現マッチング　有限オートマトン　スクリーニング

１．研究開始当初の背景

ネットワークに対する攻撃や悪意を持ったプログラムの侵入検知（Network Intrusion Detection、以下 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、1000BASE-T などの伝送速度が毎秒 1 ギガビットあるいはそれ以上の高速 LAN が普及するにつれて、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ネットワークパフォーマンス確保の観点から事実上不可能になり、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、それらを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専用ハードウェアの研究が 1990 年代後半から盛んに行われてきた。

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知られており、製品化されて市場にも出ている。しかしながら、NID に対する従来手法の大半は、ネットワーク伝送速度が高々10Gbps 以下にしか対応できていないという大きな課題がある。一方、現在、伝送速度 1Gbps のギガビットイーサネットワークが幅広く普及し、10Gbps イーサネットワークも普及が進みつつある。さらに、40Gbps および 100Gbps の超高速イーサネットワークについても標準規格が制定され、今後、普及が急速に進むことが予測される。このような状況において、40Gbps 以上の次世代超高速イーサネットに対する NID 技術を確立することが急務となっている。

ネットワークの高速化に対応した NID の従来研究としては、正規表現を受理する有限オートマトン（Finite Automaton、FA）の構成を工夫するものが多く知られている。例えば多文字遷移オートマトンやビット並列マッチングアルゴリズム、制限された正規集合に対する Shift-AND 法などが提案されているが、いずれも 10Gbps 以下の伝送速度のネットワークにしか対応できず、 伝送速度 40Gbps、100Gbps、あるいはそれ以上の超高速イーサネットに対する NID を実現するためにはブレークスルーが必要である。また、既存手法の大半は与えられたウィルスパターンに特化したハードウェア回路を FPGA 上に実現するインスタンス依存回路として実現されているが、インスタンス依存回路はパターン更新における瞬時の回路更新ができないことから、セキュリティ確保の点で問題がある。

２．研究の目的

本研究では、伝送速度が毎秒 100 ギガビット（Gbps）以上の超高速イーサネットに対するネットワーク侵入検知（NID）システム構築を目的とした正規表現マッチングハードウェアの新しい構成法を提案する。既存 NID 技術は伝送速度 10Gbps の高速イーサネットへの適用が限度だったが、本研究では、100 ギガビット（Gbps）以上の次世代超高速イーサネットに適用可能な NID 技術を確立することを目的とし、スクリーニング回路の導入により、NID の高速化とハードウェア資源の削減の両立を可能とする新しい NID システムの構成方法を提案する。本研究の成果により、今後、急速に普及することが予想されている伝送速度 40Gbps 以上の超高速イーサネットにおけるネットワーク侵入検知が可能とな

り、将来の高度 ICT 社会のインフラストラクチャである次世代超高速イーサネットで構築された LAN（Local Area Network）のセキュリティ確保に大きな貢献をすることが期待される。

３．研究の方法

（１） スクリーニング回路を用いたNIDS

本研究では、NIDの構成要素としてスクリーニング回路を新たに提案する。スクリーニング回路とは、ネットワーク上を流れるパケットが正常か否かを、簡易的な検査によって高速にふるい分けするための回路である。これにより、ウィルス感染の恐れがあるパケットだけを詳細な検査に回し、正常なパケットはすぐに検査完了となるため、全体としてのパケットの伝送効率(スループット）を向上することができる。ネットワーク上を流れるの99.9%以上が正常なパケットという事実から、すべてのパケットに対して完全な検査を行う必要はなく、危険性の高いパケットだけ精密な検査を行えば十分であり、この事実がスクリーニング回路の有用性を保証している。

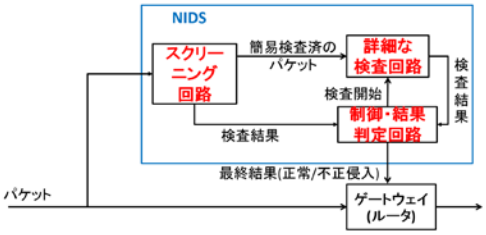

図１　提案NIDS

提案スクリーニング回路を用いたNIDS の構成を図１に示す。このNIDS では、まずスクリーニング回路でパケットを簡易的に検査し、結果を制御・結果判定回路に送信する。スクリーニング回路の検査で明確に正常なパケット、あるいは不正なパケットと判断できた場合は、詳細な検査は行わずに、結果を出力する。正常でも不正でもない不審なパケットに対しては、詳細な検査回路で従来通りの詳細な検査(正規表現マッチング)を行い、最終判断を下す。最終結果は、ネットワーク管理者またはパケットフィルタリング機能を持ったゲートウェイやルータに送られ、不正パケットの破棄やネットワークの切断などの対応が取られる。

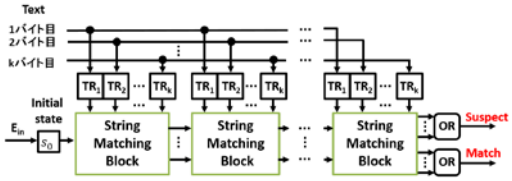

図２　スクリーニング回路

（２） スクリーニング回路の構成

本研究で提案したスクリーニング回路の概略図を図２に示す。スクリーニング回路は、kバイト遷移NFA（非決定性有限オートマトン）に基づいており、単純なkバイト文字の文字比較を行う部分回路(ストリングマッチング回路）を1次元配列状にカスケード接続することで構成される。ストリングマッチング回路をそれぞれString Matching Block（SMB)と呼ぶ。図中のE_{in} はマッチング開始信号を意味し、Textとはネットワークから入力されたパケットのkバイト分であり、このkバイトが1つのkバイト文字として回路に同時に入力される。Suspectは正常か不正かの明確な判断がつかない不審なパケットに対する出力信号を意味しており、Matchは正常もしくは不正なパケットであることが確定した場合の出力信号を意味している。

（３）スクリーニングパターンの生成

スクリーニング回路で正規表現マッチングを行うと、ハードウェアリソースが膨大になり、それに伴い処理速度が十分に上がらず、高速ネットワークへの対応が困難になる。そのため、ストリングマッチングを採用する。しかし、ストリングマッチングのパターン(スクリーニングパターン）として正規表現マッチングと等価な長い文字列でマッチングを実行しても、与えられる正規表現パターンによってはかえって遅くなる場合がある。そのた

め、スクリーニング回路では短い文字列をパターンとして用いることで、ハードウェアリソースの削減を図りつつ、ストリングマッチングを高速に実行する。一方、スクリーニングパターンを短くした場合、パケットの分類精度が低下し、不正パケットの見逃し(False Negative)や誤検知(False Positive) が増加する可能性があり、スクリーニングパターンの長さによって、スクリーニング回路の回路サイズと有効性にトレードオフが生じる。また、同じ長さのパターンであっても、誤検知を減らすために不正パケットのみに共通するパターンで、なおかつ回路サイズが小さくなり処理速度が向上するパターンを用いることが望ましい。そのため、これらを考慮した最適なスクリーニングパターンの生成が必要となる。

不正パケットの見逃しを避け、不正パケットのみに共通するパターンを効率良く見出すために、本研究ではウィルスパターンを表現する正規表現集合から、与えられたパターン長(文字数) を満たした最適なスクリーニングパターン集合を抽出する問題を定義する。また、不正パケットのみに共通するパターンを効率良く発見するために、すべてのパケットに(正常にも不正にも) 共通するパターンの集合は、頻出語の集合として与えられるものとする。以上を踏まえたパターン生成問題を以下のように定式化した。

最適パターン生成問題
入力：正規表現パターンの集合RP
　　文字数k、
　　頻出語の集合FW
出力：制約1、2、3 を満たし、目的関数を最小化するスクリーニングパターンの集合SP
目的関数：SPを受理するNFAの状態数
制約1:RPの各要素が表す正規言語のすべての要素のk文字のプレフィックスはSPに含まれる。
制約2：SPの各要素の長さはk
制約3：SPの各要素はFWには含まれない。

スクリーニング回路の大きさはNFAの状態数に依存するため、最適パターン生成問題は状態数の最小化を目的とした。制約1は、元の正規表現パターンにマッチするパケットは、それから抽出したより短いスクリーニングパターンには必ずマッチすることを保証するものであり、不正パケットの見逃し(False Negative) を防ぐ制約である。制約2はスクリーニングパターン長kに対する制限である。kの値を変化させることでパケットの分類精度とNFA の状態数のトレードオフをユーザが考慮できる。制約3は頻出語をパターンとすることを回避するためのものであり、これにより誤検知(False Positive) を抑制している。

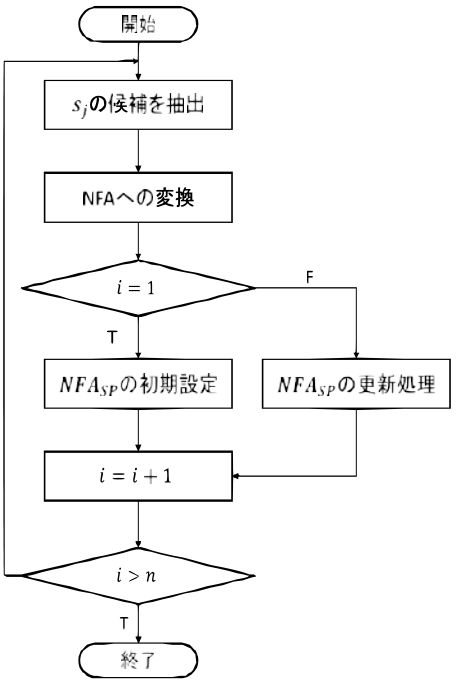

図３　提案解法のフローチャート

（4）最適化問題の発見的解法

上記の最適化問題に対する解を厳密に求めるのは、解空間が広いため困難である。そこで本研究では、与えられた正規表現パターンに共通する文字列を抜き出し、NFAの状態を共

有することで、状態数の最小化を目指す発見的手法を提案した。本提案手法では、与えられた正規表現パターンはk文字以上の単純文字列(正規表現の演算子を含まない文字列)を含むものと仮定する。これは用いるk の値は高々10 程度であり、ほとんどの正規表現パターンが10 文字以上の文字列を含んでいる事実に基づく。入力で与えられた正規表現パターンの集合RP の各要素r_i に対して、図3に示す処理を順に適用する。ただし、NFA_{SP} を目的のNFA とし、NFA_{SP}の初期値は$\varnothing$(空集合)とする。

４．研究成果

（１）計算機実験

提案手法の有効性を評価するために、先行研究で用いたスクリーニングパターンと提案手法で得られたスクリーニングパターンの両方をNFAで表現した際の状態数を比較した。本実験では、スクリーニングパターン長k を8とし、 10種類の正規表現パターン（正規表現1～10と表示）からスクリーニングパターンを抽出した。先行研究と提案手法で抽出したスクリーニングパターンをNFA に変換したときの状態数を表１に示す。

表１　実験結果

対象正規表現	先行研究	提案手法
1～5	58	46
6～10	195	24
1～10	253	64

（2）評価と考察

表１より、正規表現1～5 では、提案手法がわずかだがNFAの状態数を削減するスクリーニングパターンを生成できたことがわかる。状態数があまり減らなかった理由としては、正規表現1～3、5 は単純文字列で始まっているため、先行研究の方法でも、提案手法同様、単純文字列が抽出され、更に正規表現1～5 には共通する部分文字列がほとんどないため、Prefix tree やSuffix tree で状態の共有があまり起こらなかったためだと推測される。

正規表現6～10 では、提案手法が状態数を大きく削減しながらスクリーニングパターンを生成できているのがわかる。その理由としては、正規表現6～10 の先頭8 文字には繰り返しを表す正規表現の演算子‘+’、‘*’が含まれており、正規集合が大きくなったためだと考えられる。また、これらの正規表現には共通する部分文字列があり、それらをPrefix tree とSuffix treeを用いて状態を共有し、削減できたことも理由になっている。

実験結果から、提案手法は、接頭文字や接尾文字の共有により、先行研究の手法よりもNFAの状態数をさらに削減することが可能であることが分かった。提案手法では、正規表現パターンの数が多くなるほどスクリーニングパターン間で共有できる状態が増えるため、正規表現の数が増加しても状態数の増加を抑えることができ、スクリーニングパターン回路の回路規模の大幅な削減が可能となる。

（３）本研究のまとめと今後の課題

本研究ではNIDシステムの新しい構成手法としてスクリーニング回路を提案した。スクリーニング回路が実現するNFA の状態数の最小化を目的に最適なパターン生成のための発見的手法を提案し、その有効性を実験的に検証した。本研究で得られた成果に基づいて、NIDシステムを構成すれば、 40Gbps以上の高速インターネットに対する侵入検知に対しても、従来手法より少ないハードウェア量で効率よく侵入検知を行えることが分かった。

今後の課題としては、スクリーニングにおける誤検知(False Positive) の最小化の検討等がある。

また、本報告書では詳細を述べなかったが、ネットワーク侵入検知への応用が可能なストリームデータ内の外れ値を検出する問題についても本研究では考察し、マハラノビス距離に基づく新しい外れ値検出手法を提案した。マハラノビス距離の計算では、行列演算が多

用されているため計算量が大きい。そこで、行列演算の効率のよい実現が容易なFPGA を用いることで処理の高速化を図った。マハラノビス距離の計算に必要な平均値ベクトルと共分散行列の差分計算を提案し、回路はパイプライン回路として実現した。FPGA 実装の結果、データの属性数が4の場合、提案回路はソフトウェアによる外れ値検出と比較して、37倍の高速化を達成した。また、属性数が5以上への回路の拡張手法についても提案した。

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）

〔学会発表〕（計６件）

① 橋本智明, 永山忍, 稲木雅人, 若林真一:“ネットワーク侵入検知のためのスクリーニング回路に対する最適スクリーニングパターン生成について”, 信学技報, vol.116, VLD2016-108, pp.37-42, 2017年3月2日, 沖縄県那覇市.

② 荒井悠人, 若林真一, 永山忍, 稲木雅人: “ストリームデータに対するマハラノビス距離に基づく外れ値検出手法のFPGA実装”, 進学技報告, vol.116, RECONF2017-72, pp.141-146, 2017年１月24日, 神奈川県横浜市.

③ Yuto Arai, Shin’ichi Wakabayashi, Shinobu Nagayama, Masato Inagi: “An Efficient FPGA Implementation of Mahalanobis Distance-Based Outlier Detection for Streaming Data”, Proc. 2016 International Conference on Field-Programmable Technology, pp.253-256, 2016年12月3日, 中国・西安市.

④ Tomoaki Hashimoto, Shin’ichi Wakabayashi、Shinobu Nagayama, Masato Inagi, Hiroki Takaguchi:“A High-Speed Programmable Network Intrusion Detection System Based on a Multi-Byte Transition NFA”, Proc. 9th International Conference on Advances in Circuits, Electronics and Microelecronics, pp.45-51, 2016年7月28日, フランス・ニース市.

⑤ 高口裕貴, 若林真一, 永山忍, 稲木雅人: “高速ネットワークにおける侵入検知に対するスクリーニング回路とFPGA実装”, 信学技報, vol.115, VLD2015-119, pp.49-54, 2016年3月1日, 沖縄県那覇市.

⑥ 橋本智明, 永山忍, 稲木雅人, 若林真一: “ネットワーク侵入検知のための高速正規表現マッチングハードウェア専用コンパイラの開発”, 第17回IEEE広島支部学生シンポジウム論文集, B-64, pp.480-483, 2015年11月21,22日, 岡山県岡山市.

〔その他〕
所属研究室ホームページ
http://www.lcs.info.hiroshima-cu.ac.jp/

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

(2)研究分担者

(3)連携研究者
永山　忍（NAGAYAMA, Shinobu）
広島市立大学・大学院情報科学研究科・教授
研究者番号：10405491

(4)研究協力者