C-016

IDS におけるハードウェアでのパターンマッチング性能検証

永松　優児 †　　小柳　滋 ‡
† 立命館大学大学院理工学研究科　　‡ 立命館大学

1　はじめに

不正侵入に対するセキュリティシステムの一つに侵入検知システム (IDS：Intrusion Detection System) があり，特にネットワーク型のものをネットワーク型侵入検知システム (NIDS：Network IDS) と呼ぶ．

NIDS の処理性能は向上し続けているが，攻撃パターン数の増加や回線速度の向上により，更なる NIDS の性能向上が必要とされている．NIDS においてパターンマッチング処理は一般的にソフトウェアで構成されているが，現在の処理性能では対応しきれない場合がある．その原因を解消するために，[1][3][5] のようなパターンマッチング部をハードウェアで実装する研究が行われている．

本研究では複数アルゴリズムをそれぞれハードウェア上で実装して，回路規模，処理性能の検証により，ハードウェアでパターンマッチングを実装する上での有効性について比較検討する．

2　アルゴリズム

本研究で利用した NFA，BM 法，Karp-Rabin，Bloom Filter の各アルゴリズムについて概要の説明を行う．

2.1　NFA　(Nondeterministic Finite Automaton)

[1] による NFA パターンマッチング回路の研究では，パターンの集合を一つの正規表現で表し，それに対応する NFA 回路を構築する方法を利用している．

その方法を図 1 に表す．図 1(1) はシングルキャラクタ”c”の NFA 回路構成を示しており，flip-flop の状態が”1”を保持している．その際，入力された Text character があらかじめ用意されている比較器と一致する場合，このシングルキャラクタを受理したとして信号”1”を出力する．(2) は N1，N2 と順に一致した場合受理する．N1，N2 は図 1(1) を簡略化したものを表す．(3) は N1 または N2 が一致する場合受理する．(4) は N1 の 0 回以上の繰り返しを受理する．NFA ではこれらの回路構成を用いてパターンを生成する．

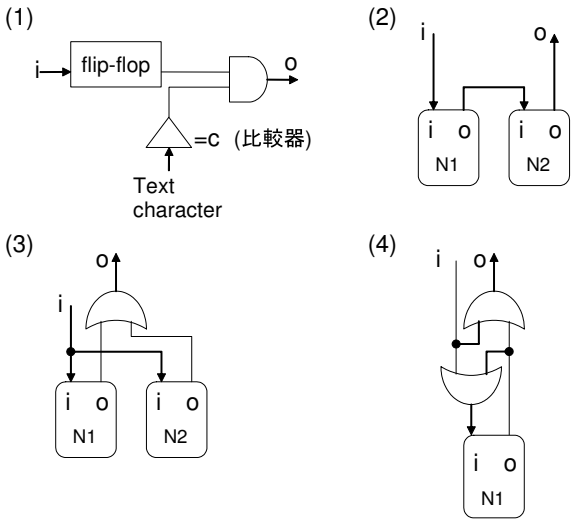

図 1: NFA ステートマシンに対応した回路

2.2　BM 法 (Boyer-Moore 法)

BM 法は検索パターン文字列の末尾よりパターンマッチングを行うアルゴリズムである．BM 法の最大の特徴はテキスト上の全ての文字を比較することなく，ある法則に従いテキストをスキップしてマッチングすることが可能なため，高速なパターンマッチングが可能な点である．

その法則とはパターンとテキストでマッチングを行った際に不一致が発生すると，その不一致が発生した際のパターンとテキストの位置情報により無駄な検査をすることなく次に検査すべきテキストの位置情報を求めることである．この法則に [4] では，skip table と next table の 2 つの表が用いられている．

2.3　Karp-Rabin

Karp-Rabin でのパターンマッチングは初めにパターンのハッシュ値を記憶する．次に対象となるテキストの同数のすべての文字列に対してハッシュ値の計算を行い，記憶しているシグネチャとのマッチングを行う．ハッシュ値の計算例を次に示す．

$$\begin{aligned} 14152 &\equiv (31415 - 3 \times 10000) \times 10 + 2 \pmod{13} \\ &\equiv (7 - 3 \times 3) \times 10 + 2 \\ &\equiv 8 \pmod{13} \end{aligned}$$

この数式は 13 を法とする剰余をハッシュ関数として，テキスト中にある 31415 という文字列のハッシュ値が 7 だと求められている場合，テキストを 1 文字進め文字列 31415 から文字列 14152 のハッシュ値を求める例である．これを繰り返すことによりテキストのハッシュを行う．このように Karp-Rabin の最大の特徴はテキストを 1 文字シフトでずらした場合のハッシュ値の計算が容易な点である．

2.4　Bloom Filter

Bloom filter は Karp-Rabin 同様ハッシュを用いる．Bloom Filter の構造を図 2 に示す．対象となる要素 (パターン) を k 個のハッシュ関数に入力して k 個のハッシュ値を m ビットの配列のインデックスとして，その位置のビットを 1 とする．

Performance Evaluation of Pattern Matching Hardware for IDS
†Yuji Nagamatsu　‡Shigeru OYANAGI
†Graduated School of Science and Engineering, Ritsumeikan University
‡Ritsumeikan University

要素 (テキストパターン) がその集合に含まれているかどうかを調べるには，その要素を k 個のハッシュ関数に入力してハッシュ値を得る．m ビットの配列に対して，その値の位置にひとつでも 0 が存在する場合、その要素は集合に含まれない．逆にすべてのビットが 1 ならば集合に含まれていると判定する．

Karp-Rabin を含め，ハッシュベースの手法では，集合に含まれていないが含まれていると誤検知が発生する可能性があるため，チェックが必要となる．

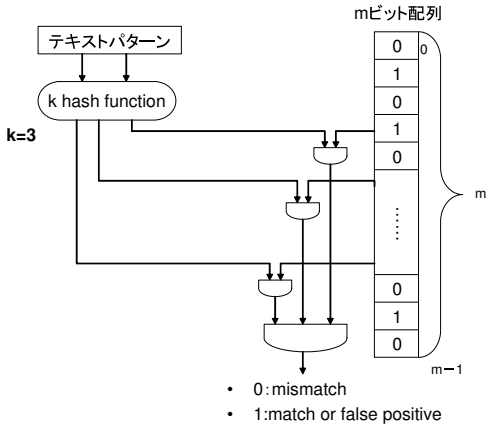

図 2: Bloom Filter の構造図

3 評価

Xilinx 社の ISE9.2i を使用し各アルゴリズムのパターンマッチング回路を実装して，シュミレーションの ModelSim XE III 6.2g を利用して回路規模，処理速度の検証を行う．回路規模は Slice，Flip-Flop，LUT の値を参照し，処理速度は周波数の数値を参照する．

表 1: NFA のパターン数による回路規模

NFA	Slice	Flip Flop	LUT	周波数
パターン数 1	24	29	43	105.485MHz
パターン数 2	35	45	62	101.999MHz
パターン数 8	121	160	213	80.080MHz
パターン数 16	208	293	334	71.731MHz

表 1 に NFA の結果を示す．パターン数が増加するにより，回路規模は増加して周波数は低下する．

表 2: BM 法のパターン数による回路規模

BM	Slice	Flip Flop	LUT	周波数
パターン数 1	66	32	128	27.928MHz
パターン数 2	118	49	226	22.472MHz

表 2 に BM 法の結果を示す．BM 法もパターン数の増加により，回路規模は増加して周波数は低下する．さらに周波数の値が他のアルゴリズムと比較すると 4 倍以上遅い結果が得られた．

表 3: Karp-Rabin のパターン数による回路規模

Karp-Rabin	Slice	Flip Flop	LUT	周波数
パターン数 1	52	37	100	175.423MHz
パターン数 2	90	65	172	175.423MHz
パターン数 8	320	209	622	175.423MHz
パターン数 16	599	369	1171	175.423MHz

表 3 に Karp-Rabin の結果を示す．Karp-Rabin はパターン数が増加することにより，回路規模である Slice，Flip-Flop，LUT は増加しているが，周波数においては変化が現れなかった．

表 3 の結果から回路規模の増加量が大きいことが分かる．この原因として誤検知のチェックを行うための回路が考えらる．その処理を行う回路部を省いた回路は現在評価中であるが，回路規模が大幅に減少すると思われる．Bloom Filter も同様に回路規模の増加が大きいため，誤検知の処理を省いた回路を評価中である．

4 まとめ

本研究では NIDS の中心的な処理であるパターンマッチングをハードウェアで実装する上での回路規模，処理性能の有効性を調べることを目的として，各アルゴリズムの比較検証を行った．

検証した結果，NFA によるパターンマッチングが回路規模に関しては一番小さい結果となった．さらに NFA は正規表現を上手く適用することにより，さらなる回路規模の縮小が考えられる．BM 法は複数パターンの検査には不向きだと考えられる．周波数はハッシュを用いた Karp-Rabin，Bloom Filter が優れていた．Karp-Rabin，Bloom Filter において誤検知の処理を行う回路部を省くと回路規模は最小となると思われる．しかし，誤検知の処理は不可欠であるため，その回路を他のアルゴリズムと組み合わせる事により，処理速度を保ち回路規模の増加量を抑えることができると考えられる．

今後の課題としてはアルゴリズムの改良を行い，パターンマッチングとしての性能を向上させることが課題となる．

参考文献

[1] Reetinder Sidhu. Viktor K. Prasanna.“ Fast Regular Expression Matching using FPGAs ”,2001

[2] 浦谷則好.“ 高速な複数文字列照合アルゴリズム：FAST“ ,1989

[3] 小林礼明. 阿部公輝.“ Karp-Rabin 法を用いたシグネチャ型 IDS の性能コスト評価 ”,2007

[4] 石畑清.“ アルゴリズムとデータ構造 ”, 1989

[5] Sarang Dharmapurikar. Praveen Krishnamurthy. Todd Sproull. John Lockwood.“ Deep Packet Inspection using Parallel Bloom Filters ”, 2004