망침입검출체계에서 흐름분할에 기초한 병렬패턴정합에 대한 연구

박성호, 황철진

경애하는 최고령도자 **김정은**동지께서는 다음과 같이 말씀하시였다.

《통신망의 보안능력을 높이는것은 당과 국가, 군사비밀을 철저히 보장하고 나라의 존엄과 안전을 수호하기 위한 중요하고도 책임적인 사업입니다.》

망침입검출체계(NIDS)에서 FPGA나 전용TCAM소편을 비롯한 장치기술을 리용하여 패턴정합연산을 처리하고 병렬화를 도입하여 수Gbps의 높은 속도를 달성하고있으나 입력자료흐름을 순차로 검사해야 하는것으로 하여 성능을 보다 높이지 못하는 문제[1, 2]가 있다.

론문에서는 FPGA에 기초한 NIDS에서 병렬화방식을 보다 개선함으로써 높은 성능을 달성하여 고속망환경에서 NIDS의 실시간성을 높이기 위한 방법을 제안하였다. 패턴정합의 입력자료에 대하여 흐름분할을 적용하고 분할된 흐름토막을 단위로 하여 독자적인 패턴정합을 실행하여 정확성을 보장하면서도 병렬화를 보다 개선하여 높은 패턴정합속도를 달성하였다.

1. 문 제 설 정

현재 많이 쓰이는 오용검출방식의 NIDS의 전체 연산에서 패턴정합연산은 70～80%를 차지하며 이 부분에서 성능개선을 달성하는것은 전체 체계의 성능에 큰 영향을 주게 된다.

패턴정합속도를 개선하기 위하여 FPGA기술을 리용하는 방안이 제안되여 Gbps급의 높은 속도를 달성하고있다. 여기서 병렬화에 의한 속도제고를 취급하였으나 관흐름방식과 규칙모임분할에 머물러있고 여전히 입력자료를 선형으로, 순차적으로 처리함으로 하여 보다 높은 정합속도를 달성하지 못하고있다.

론문에서는 FPGA에서 입력자료흐름을 분할하고 분할된 매 토막들에 대하여 독립적으로, 병렬로 패턴정합을 실행함으로써 전체적인 정합속도를 높이기 위한 방법을 제안하였다.

입력자료흐름의 분할에서는 흐름이 분할되는 경우 하나의 악성패턴이 여러개의 토막들에 분산되는것으로 하여 패턴정합에서 정확한 패턴을 검출하지 못할 가능성이 존재할 수 있는 문제점이 제기된다.

패턴정합의 정확성을 유지하면서 흐름을 분할하는데서 제기되는 이러한 문제를 해결하기 위하여 론문에서는 선행연구[2]에서 제안된 TCAM에 기초하여 침입검출에서 부의 패턴(NP)의 개념을 리용하였다. 흐름에서 나타나는 부의 패턴안의 위치들에서 흐름을 분할하면 흐름안의 악성패턴이 여러 토막에 갈라지지 않으며 그것을 놓치지 않는다는것이 담보된다.

2. 흐름분할에 의한 병렬화개선

1) 흐름분할문제

패턴정합의 속도를 높이기 위하여 전체 입력파케트흐름을 순차적으로 검사하는것이 아니라 그것을 분할하고 분할된 매 토막에 대하여 동시에 독립적으로 패턴정합을 진행한다면 정합속도를 크게 높일수 있다.

입력흐름의 길이가 LB이고 정합처리단위가 WB라고 하자.

입력흐름을 순차적으로 1B씩 옮기면서 하나의 패턴에 대하여 정합을 진행한다면 정합조작의 회수는 $L-W+1$이다.

일반적으로 입력흐름을 n개의 토막으로 분할할 때 정합조작의 회수 T[2]는 다음의 식으로 표시할수 있다.

$$T=\sum_{i=1}^{n}\max(L_i-W+1,\,1)$$

웃식으로부터 입력흐름을 될수록 FPGA의 정합처리단위의 크기에 가까운 크기를 가지는 토막들로 분할할수 있다면 즉 $L_i\approx W$ 이면 정합조작의 회수를 크게 줄일수 있으며 병렬조작에 의하여 효률을 높일수 있다는것을 알수 있다.

NIDS의 패턴정합을 위한 흐름분할에서는 하나의 악성패턴이 흐름분할에 의하여 여러 토막에 나뉘여 존재함으로 하여 패턴정합에서 그 패턴을 검출하지 못하는 문제를 어떻게 해결할것인가 하는 문제가 제기된다.

2) 부의 패턴을 리용한 흐름분할

망침입검출에서 패턴정합은 입력파케트흐름에 주어진 패턴모임의 특정한 패턴이 존재하는가를 검사하는 행위이다. 선행연구[2]에서는 부의 패턴의 개념을 제기하였는데 이것을 리용하면 패턴정합의 정확도를 보장하면서 입력흐름을 분할하는 문제를 해결할수 있다.

부의 패턴을 다음과 같이 정의한다.

패턴모임 $S=\{P_1, P_2, \cdots, P_n\}$ 이 주어졌을 때 문자렬 np를 그것의 임의의 k바이트(k>1)뒤붙이가 임의의 $P_i\in S,\ i=1, 2, \cdots, n$ 의 부분문자렬이 아니라면 S의 부의 패턴이라고 부른다.

부의 패턴의 개념을 리용하여 분할후에 그 어떤 패턴도 2개이상의 토막에 걸쳐서 나타나지 않도록 담보하면서 흐름을 분할할수 있다. 검사하려는 흐름에서 부의 패턴 np를 찾고 그것의 마지막 2B사이에서 흐름을 자른다면 S의 그 어떤 패턴도 두 토막에 나뉘여 나타나지 않는다는것이 담보된다.

결국 부의 패턴의 마지막 두 바이트사이에서 흐름을 분할하면 매 토막은 순차적으로가 아니라 동시에 검사될수 있고 패턴이 갈라지는 현상은 나타나지 않게 된다.

3) 흐름분할알고리듬

먼저 흐름분할에 사용할 부의 패턴들을 구한다. 여기서 부의 패턴의 뒤붙이(suffix)들이 중요하며 특히 마지막 두 바이트사이에서 흐름을 분할하여야 하므로 주어진 NIDS의 패턴모임에 대하여 두 바이트로 이루어지는 부의 패턴들을 모두 구한다. 이 공정은 전처

리공정으로서 침입검출체계가 기동하기 전에 비직결, 비실시간으로 진행되므로 계산량은 비교적 크나 체계성능에 영향을 주지 않는다.

흐름분할과정은 실지 입력파케트흐름에 대하여 실시간적인 과정으로 수행된다.

처음에 흐름의 W바이트위치에서 2B 부의 패턴의 존재를 검사하며 그것이 검출되면 그 중간위치에서 분할이 진행된다. 부의 패턴이 검출되지 않으면 1B 옮겨 조사한다.

이와 같은 방법으로 부의 패턴이 발견될 때까지 전진하며 발견되면 분할을 진행하고 다음에 그 위치로부터 W바이트만큼 옮겨 새로운 부의 패턴조사를 진행한다.

그림 1에 흐름분할알고리듬의 한가지 실례 즉 부의 패턴을 리용한 흐름분할을 보여주었다.

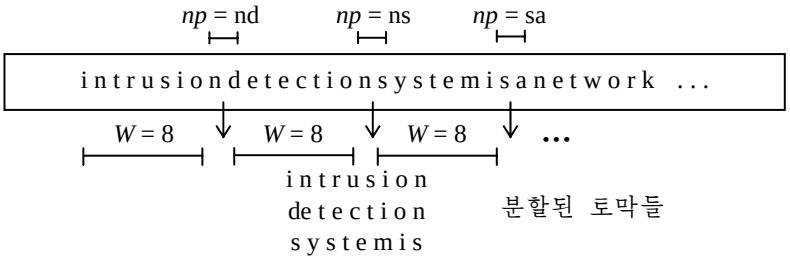

그림 1. 부의 패턴을 리용한 흐름분할

이 실례에서 패턴모임은

$$S = \{\text{intrusion, detection, system}\}$$

이고 흐름내용은《intrusiondetectionsystemisanetwork…》이며 $W = 8$이라고 가정한다.

처음에 $W = 8$인 위치에서의 두 바이트 np를 찾으면 on으로서 np가 아니다. 한바이트 전진하여 다시 두 바이트 np를 찾으면 nd로서 np이다. 그러므로 nd의 중심점을 분할점으로 한다. 다음에 이 위치로부터 $W = 8$인 위치를 다시 검사한다. on으로서 np가 아니며 한 바이트 전진하여 $np = \text{ns}$를 얻는다. 다음번 $W = 8$의 위치에서는 $np = \text{sa}$를 얻는다.

이와 같은 방법으로 분할하여 토막들을 얻는다.

4) 병렬패턴정합

패턴정합은 흐름준위의 병렬화뿐아니라 패턴준위의 병렬화도 리용하여 진행된다. 주어진 NIDS의 패턴모임을 일정한 부분모임들로 분할하여 매 부분모임이 서로 다른 흐름토막들에 대하여 동시에 패턴정합조작을 실행하도록 한다.

병렬패턴정합알고리듬의 한 단계를 그림 2에 보여주었다.

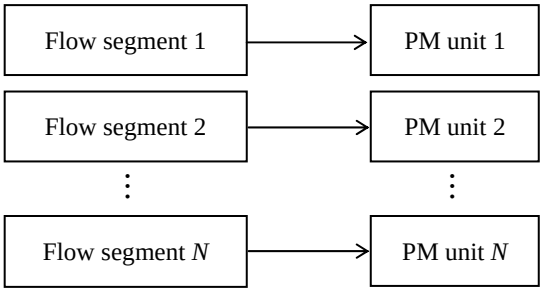

그림 2. 병렬패턴정합알고리듬의 한 단계

패턴정합은 *N*개의 PM장치에서 동시에 진행된다. 여기서 *N*은 부분패턴모임의 수이며 매 PM장치에는 하나의 부분패턴모임이 대응된다.

먼저 FPGA의 내부기억기의 흐름토막들로부터 *N*개의 토막이 추출되여 길이가 *W*B인 토막등록기들에 적재되며 각각 *N*개의 PM장치들에 입력된다. 매 PM장치들에서 패턴정합이 진행되고 끝나면 한 단계가 완료된다.

다음단계에서 흐름토막들은 라운드－로빈방식으로 다음번 PM장치에 입력된다.

이와 같은 방식으로 *N*개의 단계가 진행되면 *N*개의 흐름토막들에 대한 PM과정이 완료되며 다시 새로운 *N*개의 흐름토막들을 불러들여 우와 같은 단계들을 실행한다.

정합이 발견되면 패턴정합은 끝나게 되며 이 과정은 정합이 발견될 때까지 또는 모든 흐름토막들에 대한 PM과정이 완료될 때까지 계속된다.

3. 체 계 실 현

우에서 서술된 병렬패턴정합체계의 실현을 그림 3에 보여주었다.

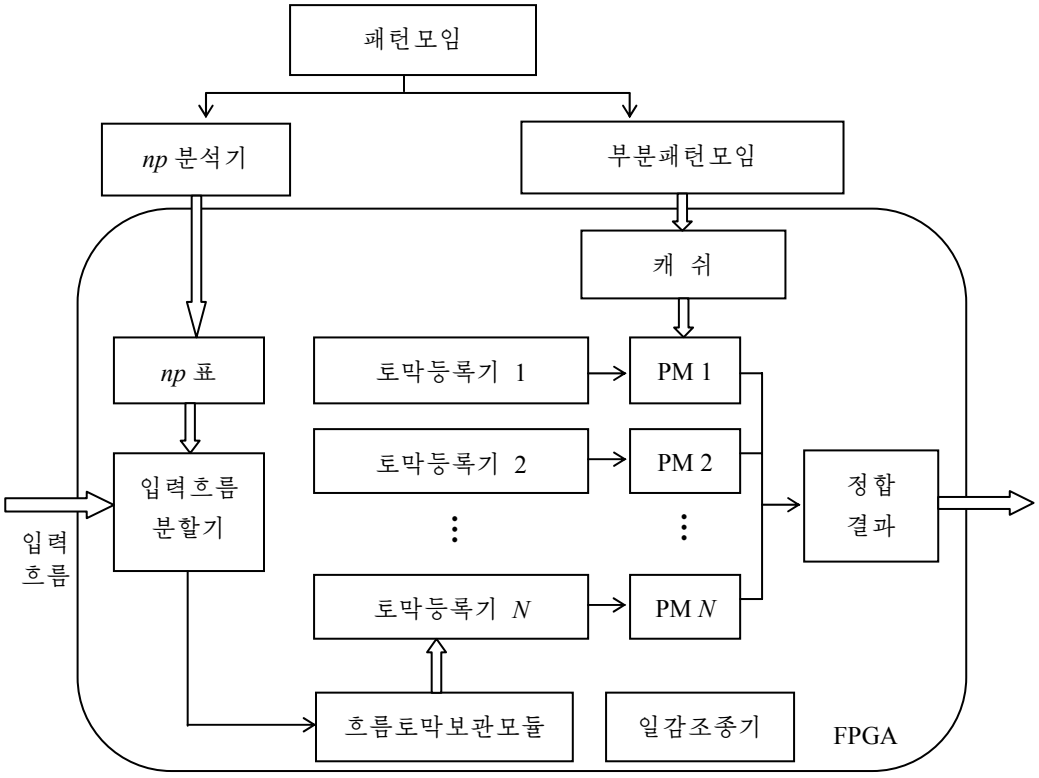

그림 3. 병렬패턴정합체계의 실현

제안된 병렬패턴정합체계는 *np*분석기와 *np*표, 입력흐름분할기, 흐름토막보관모듈, 부분패턴모임들, *N*개의 토막등록기와 PM장치들 그리고 일감조종기로 구성되였다.

비직결의 전처리단계에서 부의 패턴찾기에 의하여 부의 패턴표(*np*표)가 구축되며 패턴모임의 분할에 의하여 패턴부분모임들이 구성된다. 패턴부분모임들은 외부의 SDRAM에

적재되며 패턴정합시에는 FPGA내부의 캐쉬기억기를 통하여 PM장치들에 제공된다.

SDRAM으로부터 캐쉬기억기에로의 패턴자료전송은 패턴정합과 병렬로 진행되도록 하며 이렇게 하여 기억기접근으로 인한 시간소비를 최소화할수 있다.

실시간적인 패턴정합단계에서 입력파케트의 자료부인 입력흐름은 흐름분할기에 의하여 토막들로 분할되여 토막보관모듈에 적재되며 첫 N개의 토막이 형성되면 곧 병렬패턴정합이 시작된다. 일감조종기는 라운드－로빈방식의 일감할당을 진행한다. 정합의 결과는 다음단계에로 넘겨진다.

침입검출체계 Snort-2.8.3의 패턴모임에서 크기가 40B이하인 패턴수는 약 88%에 달하며 이것들에 대한 패턴정합을 병렬로 진행하는것으로 하여 체계의 성능을 개선할수 있다.

우리는 전체 패턴모임에서 1－40B까지의 패턴들을 N개의 부분모임으로 분할하며 이때 크기가 21－40B인 하나의 패턴은 20B이하의 2개의 부분패턴으로 분할하여 취급하기로 하였다. 또한 패턴모임은 부하균등의 원칙에서 분할하는데 패턴들의 바이트수가 커짐에 따라 부분모임의 크기는 대략 선형으로 커지도록 하였다. 병렬패턴정합체계를 실현하기 위하여 패턴모임 P를 4개의 부분패턴모임들로 분할하였으며 따라서 병렬정합장치 PM의 수는 $N=4$로 하였다.

다음으로 병렬정합에서 취급되는 패턴의 최대크기가 20B이므로 PM장치들의 정합처리단위의 크기는 $W=20$B로 하였다.

4. 실험과 평가

1) 원리적평가

제안된 입력흐름분할에 기초한 병렬패턴정합에서는 입력에 패턴이 존재하는 경우 그것이 분할된 흐름토막들에 나뉘여 존재하지 않으므로 검출오유(false negative)가 나타나지 않으며 규칙분할[3]과 류사한 정도의 추가적인 병렬성이 달성된다.

입력흐름분할을 위한 부의 패턴을 얻는 과정은 사전에 비직결로 진행되므로 패턴정합성능에는 영향을 주지 않는다.

입력흐름분할과정은 첫 N개의 토막(실험에서 $N=4$)이 얻어진 이후부터는 패턴정합과 병렬로 진행되므로 그자체의 시간소비는 매우 적다. 외부의 SDRAM으로부터 규칙패턴을 불러들이기 위한 조작도 패턴정합과 병렬로 진행되며 독자적인 시간소비는 매우 적다.

결국 직접적인 패턴정합을 제외한 시간(overhead)은 매우 작으며 성능에 영향을 주지 않는다.

2) 실험적평가

우리는 이전의 실험과의 결과를 비교하기 위하여 침입검출체계 Snort-2.8.3과 그것의 기능추가에 의한 실험을 진행하였다.

Snort-2.8.3의 패턴모임의 규칙수는 15 480개, 패턴총수는 53 374개이다. 이 중에서 길이 1－40B까지의 패턴 46 802개에 대하여 부하균등의 원칙에서 4개의 부분패턴모임들을 구한다.

1－40B의 패턴모임에 대하여 2B 부의 패턴들을 구한 결과에 의하면 52 842개의 부의 패턴들이 얻어진다.

인위적으로 구성된 자료파케트모임을 리용한 실험에 의하면 부의 패턴에 의한 입력흐름분할에서 패턴이 둘로 나뉘여지는 현상이 나타나지 않았으며 100Mbps망환경에서 100%의 검출률을 보장하였다.

FPGA에서 실현할 때 기본주파수를 167MHz, 자료너비를 32bit로 한다면

$$167\text{MHz} \times 32\text{bit} \approx 5.3\text{Gbps}$$

의 처리속도가 달성될수 있다는것을 확인하였다.

전용TCAM소편을 리용한 병렬패턴정합에 기초한 NIDS에 비하면 적은 원가로서 보다 높은 성능을 보장한다는것을 알수 있다.

맺 는 말

부의 패턴의 개념을 FPGA에 기초한 침입검출체계실현에 적용하여 흐름분할에 의한 입력준위병렬화를 실현하고 패턴정합의 정확성을 보장하면서 성능을 2배이상으로 높여 선로속도에 가까운 처리능력을 달성할수 있다는것을 밝혔다.

참 고 문 헌

[1] D. Day, B. Burns; In Fifth International Conference on Digital Society, 187, 2011.

[2] Kai Zheng et al.; Computer Communications, 62, 47, 2015.

[3] Zhiping Cai et al.; IEEE Transactions on Computer, 62, 3, 417, 2013.

주체107(2018)년 11월 5일 원고접수

A Parallel Pattern Matching Algorithm Based on Flow Partition in Network Intrusion Detection System

Pak Song Ho, Hwang Chol Jin

This paper presents an algorithm for the parallel pattern matching based on the partition of the input packet flow in the network intrusion detection system and shows that this algorithm guarantee the correctness of the parallel pattern matching and achieves a high throughput close to line speed.

Key words : network security, intrusion detection system, pattern matching