트래픽 수집 지점의 차이로 인한 통계 기반 트래픽 분석 방법론의 한계와 해결방안

김민석, 안현민, 김명섭

고려대학교 컴퓨터정보학과

{vmpd, queen26, tmskim}@korea.ac.kr

A Method to resolve the Limit of Traffic Classification caused by Traffic collection point

요 약

응용 레벨 트래픽 분석은 네트워크의 효율적인 운영과 안정적인 서비스를 제공하기 위한 필수적인 요소이다. 최근에는 응용 레벨 트래픽 분석을 위해서 플로우 통계 정보에 기반한 트래픽 분류 방법론이 제안되고 있다. 통계 정보로는 패킷 전송 순서, 패킷 전송 방향, 패킷 크기 등을 이용한다. 하지만 트래픽 수집 지점의 차이로 인해서 트래픽은 통계 정보의 일관성을 잃게 되며 이는 신뢰할 수 없는 분석결과로 나타나게 된다. 따라서 본 논문에서는 트래픽 수집 지점의 차이로 인한 문제를 분석하고, 이를 해결하는 방법론을 제안한다. 제안하는 방법론은 학내 망에서의 실험을 통해 그 성능을 검증한다.

1. 서 론[1]

오늘날의 네트워크는 다양한 응용의 등장으로 인해 트래픽이 복잡해짐에 따라 응용 레벨 트래픽 분석은 네트워크의 효과적인 운용과 관리를 위한 필수적인 요소이다. 응용 레벨 트래픽 분류를 위해 플로우의 통계정보에 기반한 방법론들이 최근 많이 사용되고 있다. 패킷의 크기와 전송 방향, 전송순서, 캡쳐 시간 등의 Feature 를 사용하는 플로우 통계 정보 기반 트래픽 분류 방법론은 네트워크 환경에서의 한계가 있는데 바로 TCP 세션의 이상동작인 Retransmission 과 Out-of-order[1], 트래픽 수집 지점의 차이로 인한 패킷 순서의 변동이 그것이다. 전 논문에서 TCP 세션의 이상동작인 Retransmission 과 Out-of-order 문제의 한계를 해결하여 개선하였다[2].

본 논문에서는 트래픽 수집 지점의 차이로 인해 발생하는 패킷의 순서 이상 문제의 한계를 해결하기위하여 Seq 넘버와 Ack 넘버를 비교하여 Reordering 하는 방법을 제안한다.

학내 망에서 수집한 트래픽으로 확인한 결과 무시할 수 없는 양의 이상 동작이 감지되었고 실험을 통해 패킷의 크기와 전송 순서 등을 Feature 로 사용하는 트래픽 분석 방법론에 영향을 끼침을 알 수 있었다. 또한 제안하는 탐지 방법을 사용하여 이를 분석한 결과 높은 탐지량을 보임으로써 이를 해결하기 위한 방법론을 제안한다.

* 이 논문은 정부(교육과학기술부)의 재원으로 2010 년도 한국연구재단-차세대정보컴퓨팅기술개발사업(20100020728) 및 2012 년도 한국연구재단(2012R1A1A2007483)의 지원을 받아 수행된 연구임.

본 논문은 2 장에서 관련 연구에 대해서 기술하고, 3 장에서는 이러한 트래픽 수집 문제를 발생하는 이유와 탐지, 해결방안에 대해 자세히 설명한다. 4 장에서 실험 환경과 결과에 대해서 기술한다. 마지막으로 5 장에서 결론과 향후 연구 방향에 대해서 기술한다.

2. 관련연구

본 장에서는 패킷 전송 순서, 패킷 전송 방향, 패킷 크기 등을 이용하는 트래픽 분석 시스템과 트래픽 분석 시스템에서의 Feature 이용 방법에 대해 설명 한다.

다양한 방법들이 인터넷 트래픽의 응용 프로그램별 분류를 위해 제시되고 있다[3-6]. 방법론들은 크게 3 가지로 구분할 수 있는데, 이들은 시그니쳐 기반 분석[3], 트래픽 상관관계 기반 분석[4], 머신러닝 기반의 분석[5,6]이다. 본 논문의 대상은 첫 번째 시그니쳐 기반 분석 방법 중 통계 시그니쳐 기반 분석 방법과 세 번째 머신러닝 기반의 분석방법이다. 본 논문에서 관련연구로 실험을 위해 통계 시그니쳐 기반 트래픽 분석 방법론[3]을 간략히 설명한다.

통계 시그니쳐 기반 트래픽 분석 방법론[3]은 TCP, UDP 를 전송 프로토콜로 사용하는 플로우들을 입력으로 받는다. 플로우 내의 컨트롤 패킷을 제외한 페이로드가 있는 패킷만을 Feature 로 사용하는데 먼저 플로우의 첫 N 개 패킷의 페이로드 크기와 전송방향, 순서를 이용하여 N 차원의 플로우 벡터로 표현한다. 플로우 벡터에서 패킷의 페이로드 크기는 정수로 표현되고 전송방향은 TCP 의 경우 '+'는

클라이언트에서 서버로 향하는 패킷, '-'는 서버에서 클라이언트로 향하는 패킷을 의미한다. UDP 는 클라이언트/서버의 구분이 명확하지 않기 때문에 발생되는 첫 패킷을 '+'로 표현하고 뒤에 이어지는 패킷은 첫 패킷을 기준으로 방향이 같으면 '+', 다르면 '–'로 표현한다. 다음의 식 1 은 플로우 벡터를 표현한 것으로 $V(f)$ 는 플로우 f 의 벡터, d_i 는 플로우 f 내 i 번째 패킷의 전송 방향(+/-), s_i 는 i 번째 패킷의 페이로드 크기를 의미하며 각 요소의 순서는 플로우 내 패킷의 전송 순서이다.

$$V(f) = \{d_1 \times s_1, d_2 \times s_2, ..., d_n \times s_n\} \quad \text{(식 1)}$$

예를 들어 두 종단호스트 A 와 B 가 통신을 하는데 A 가 먼저 B 에게 데이터 크기가 30 인 패킷을 두 개 연속으로 보내고 B 에서 A 로 데이터 크기 40 인 패킷을 하나 전송한다면 두 종단호스트를 연결하는 플로우 f 의 벡터 $V(f)$ 는 식 2 와 같이 표현할 수 있다.

$$V(f) = \{+30, +30, -40\} \quad \text{(식 2)}$$

트래픽에서 플로우들을 앞에서 정의한 N 차원의 플로우 벡터로 표현 한 뒤 추출된 시그니쳐와 비교, 분석한다.

3. 트래픽 수집 지점의 따른 문제와 해결

본 장에서는 패킷 전송 순서, 패킷 전송 방향, 패킷 크기 등을 Feature 로 이용하는 통계 시그니쳐 기반 분석의 한계와 해결방안을 설명한다.

3.1 트래픽 수집지점에 따른 Feature 변화 문제

TCP 에서 두 호스트에서 통신 중일 때 한 종점 호스트에서 전송 되는 패킷을 기다리지 않고 상대 호스트에서 동시에 패킷을 발생시키는 경우가 있는데 이때 트래픽 수집 지점에 따라 플로우의 통계 정보가 변하게 된다. 즉, 통계 정보를 Feature 로 사용하는 트래픽 분석 방법론은 수집 지점의 차이로 인해 일정한 Feature 를 보장받지 못하며 분석 결과를 신뢰할 수 없다.

예를 들어 그림 1 에서 보듯 두 종단 호스트 Client 와 Server 가 통신을 할 때 트래픽 수집 지점인 C1, C2, C3, C4 에 따라 패킷의 순서가 달라지고 또한 통계 정보가 달라진다. 이로 인해 같은 플로우에서 Feature 의 값이 달라져 일관성을 잃은 잘못 된 분석을 하게 되는 문제점이 발생한다. 본 논문에서는 이러한 문제를 탐지 및 해결 하는 방안을 제안한다.

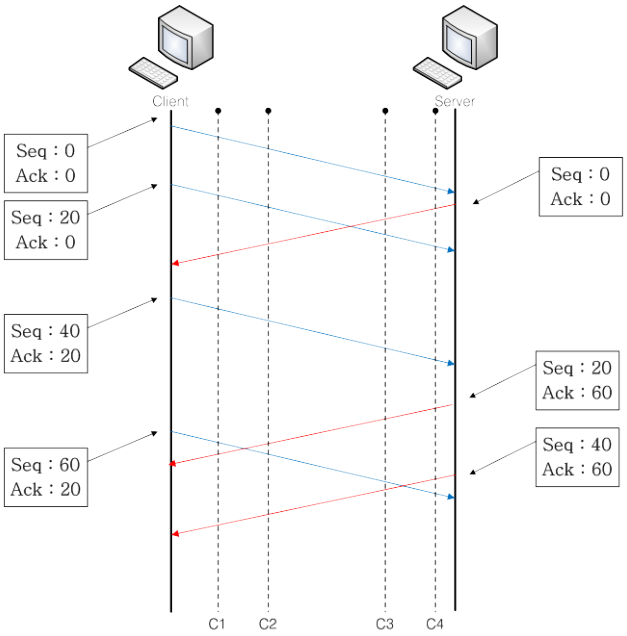

그림 1. 트래픽 수집 지점에 따른 문제

3.2 탐지 알고리즘

본 절에서는 트래픽 수집 지점의 차이로 인해 수집하는 패킷의 순서 이상이 발생하는 경우를 탐지하는 알고리즘에 대해 제안한다. 그림 1 에서 C1, C2, C3, C4 인 트래픽 수집 지점에 따라 패킷의 순서가 각각이 다르게 된다. 그러나 명확한 기준점이 없으면 어느 것이 옳은 순서인지 잘못된 순서인지 알 수가 없다. 본 논문에서 제안하는 올바른 순서는 Client 를 기준으로 한다. Client 기준이라는 것은 Client 가 패킷을 보내고 받는 순서를 일컫는 것이다. Client 기준인 이유는 대부분의 트래픽이 Client 에서 Server 로 요청 함으로써 통신이 발생하며 Server 에서 이에 응답하는 형식으로 이어지기 때문이다. 즉, Client 의 요청에 따라 트래픽이 변화하기 때문이다. 따라서 본 논문에서는 트래픽 수집 지점에 상관없이 Client 기준으로 재정렬 한다.

그림 2 는 탐지 알고리즘의 의사코드이다. 패킷 Retransmission, Out-of-order 가 제거된 TCP 트래픽 중 페이로드가 있는 패킷만을 입력으로 받는다. P(n)은 n 번째 패킷을 가리키며 P(n)(Seq), P(n)(Ack)는 각각 n 번째 패킷의 Sequence 넘버와 Acknowledge 넘버를 가리킨다. Forward packet 은 Client 에서 Server 로 전송하는 패킷이며 Backward packet 은 Server 에서 Client 로 전송하는 패킷이다. 탐지 방법은 P(n)의 Seq 넘버, Ack 넘버와 P(n+1)의 Seq 넘버, Ack 넘버를 비교하여 각각의 4 가지 판단 조건으로 normal 과 abnormal 을 탐지한다.

모든 case 들 각각의 case 에 대해 예시그림과 설명을 함에 있어 Normal case 는 n.1~n.4, Abnormal case 는 a.1~a.4 로 나타낸다. C1, C2, C3 는 본 논문에서 문제가 되는 트래픽 수집 지점을 나타내고 각 패킷 간의 교차라는 것은 Client 와 Server 간의 Forward 패킷과 Backward 패킷의 전송 도중 수집 지

점의 시점에 보았을 때 두 패킷이 만나는 것을 나타낸다. 즉, Forward 패킷이든 Backward 패킷 중 어느 것이든 먼저 전송을 시작하였고 그 패킷이 상대 호스트에게 도착 전에 반대 방향의 패킷이 전송을 하는 도중 겹치는 현상을 나타낸다.

Resolve retransmission and out-of-order problem
Remove all non-payload packets from the packet sequence

1: **procedure** Detect Normal Packet Sequence
2: Input : packet sequence of a TCP flow
3: **if** (P(n) (direction) == P(n+1) (direction)) **then** normal
4: **if** (P(n) == Fp && P(n+1) == Bp) **then** normal
5: **if** (P(n) == Bp && P(n+1) ==Fp) {
6: **if** (P(n)(Ack) <= P(n+1)(Seq) && P(n)(Seq) < P(n+1)(Ack))
7: **then** normal
8: **else** abnormal
9: }
10: **end procedure**

1: **procedure** Detect abnormal Packet Sequence
2: Input : packet sequence of a TCP flow
3: **if** (P(n) ==Bp && P(n+1) ==Fp) {
4: **if** (P(n)(Ack) <= P(n+1)(Seq) && P(n)(Seq) >= P(n+1)(Ack)) }
5: **then** abnormal
6: **else** normal
7: }
8: **end procedure**

Payload packet : a packet with payload data
Non-Payload packet : a packet without payload data
P(n): n-th payload packet in a TCP flow
Forward packet (Fp) : a packet going from client to server
Backward packet (Bp) : a packet going server to client

그림 2. 탐지 알고리즘 의사코드

그림 3 은 Normal case 의 그림이다. Normal case 1 인 n.1 은 n 번째 패킷과 n+1 번째 패킷이 같은방향 일 때, Nomal case 2 인 n.2 는 n 번째 패킷이 Forward 패킷이고 n+1 번째 패킷이 Backward 패킷 일때의 경우이다. 탐지 알고리즘에서 명시되어 있듯이 n 번 n+1 번째 패킷이 같은 방향 또는 n 번째 패킷이 Forward 패킷인 경우는 항상 normal 이다. 그림에서 보듯이 트래픽 수집지점이 다르지만 Client 에서의 기준과 모두 동일한 순서로 잡히게 된다. 이는 트래픽 수집지점에 따라 패킷의 순서가 변화가 되기 위해서는 Forward 패킷과 Backward 패킷이 두 패킷 중 하나 이상의 패킷이 반대 방향의 패킷이 도착하기전보다 먼저 전송을 시작하게 됨으로써 서로 교차

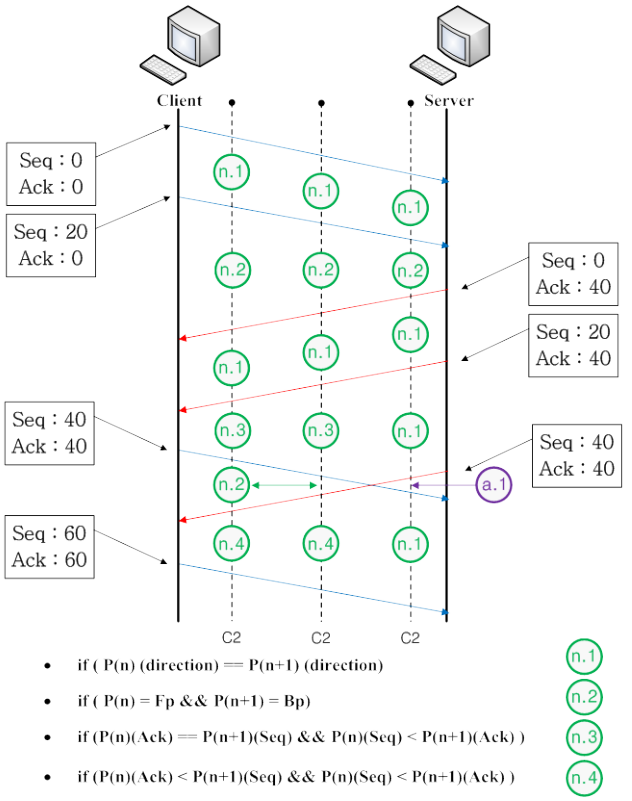

그림 3. Normal case

되어 수집지점에 따라 변화가 될 수 있는 원인이 된다. Forward 패킷이 n 번째 패킷으로 수집이 되면 기준점으로 정한 Client 에서 보내고 받는 순서로 모든 트래픽 수집지점에서의 패킷 순서를 맞춰주는 것이 목표이기 때문에 교차가 일어나려면 Backward 패킷이 Forward 패킷이 도착하기 전에 전송을 시작하여 교차가 일어나는 것이기 때문에 Forward 패킷이 n 번째로 수집된다면 항상 normal 이 된다. Normal case 1, 2 를 제외한 나머지 모든 case 는 P(n) 번째 패킷이 Backward 패킷이며 P(n+1)번째 패킷이 Forward 패킷에서의 비교 간에 탐지 한다.

Normal case 3 인 n.3 은 n 번째 패킷이 Backward 패킷이지만 Forward 패킷과 Backward 패킷간의 교차가 없어 어느 지점에서 트래픽을 수집하든 패킷 순서에 영향을 미치지 않는 경우일 때의 경우의 예시를 보여 준다. 하지만 다음 Normal case 4 인 n.4 와 같은 경우에는 n.4 가 탐지되기 전 패킷 비교에서 Forward 패킷과 Backward 패킷이 교차하여 Abnormal case 가 발생한다. 즉, n.4 가 일어나기 전에는 항상 Abnormal case 의 하나인 a.1 이 먼저 발생 한다는 것이다.

Normal case 4 의 경우 Abnormal case 1 이 일어난 후에 발생되는 Normal case 로써 n 번째 패킷이 Backward 패킷이며 n+1 번째 패킷이 Forward 패킷이 되는 경우 인데 Normal case 4 와 Abnormal case 1 은 각 패킷의 seq 넘버, ack 넘버의 비교를 통해 normal 과 abnormal 로 판단 한다.

이처럼 normal 과 abnormal 은 같이 일어 날 수 있으며 트래픽 수집 지점에 따라 Client 기준의 순서와 같을 수 있고 다를 수 있다는 것이 본 논문에서 탐지하고 해결하여야 하는 문제이다.

그림 3 에서 보듯 각각의 트래픽 수집 지점인 C1, C2, C3 를 보면 Client 기준의 순서와 C1 과 C2 는 같지만 C3 의 경우에는 패킷 순서가 달라져서 일관성을 잃게 되는 문제가 된다. 이러한 트래픽 수집 지점에 따른 Feature 의 변화 문제를 해결하기 위해 Abnormal case 의 4 가지 판단 조건이 필요하다.

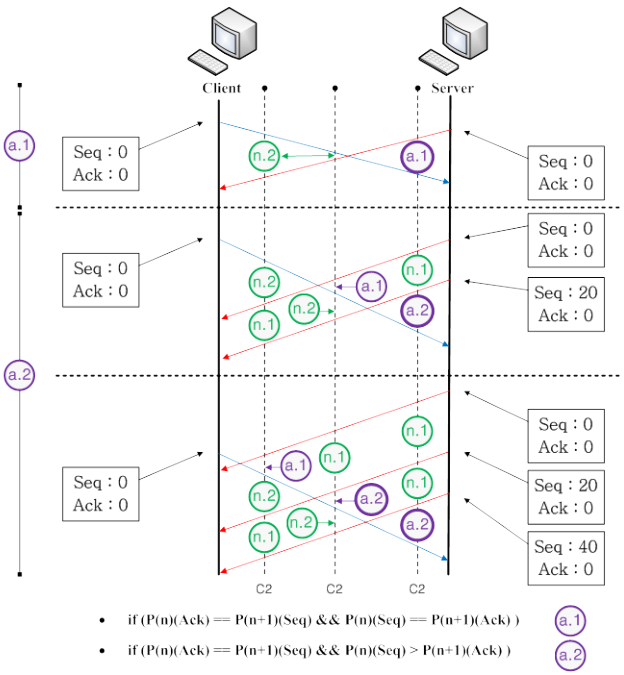

그림 4. Abnormal case 1, 2

그림 4 는 Abnormal case 1 과 2 인 a.1 과 a.2 의 그림이다. 그림에서 보는 것과 같이 트래픽 수집 지점 C1, C2, C3 는 확연히 다른 패킷 순서를 갖게 된다. a.1 의 경우에는 C1 과 C2 에서 트래픽 수집을 하였을 경우는 문제가 되지 않는다. 그러나 C3 에서 트래픽 수집을 하였을 경우 Client 기준의 순서와 동일 하지 않게 된다. 이와 같이 a.1 은 n 번째 패킷이 Backward 패킷이고, n+1 번째 패킷이 Forward 패킷인 경우에서 Backward 패킷과 Forward 패킷이 각각 하나의 패킷이 교차가 일어났을 경우 이러한 문제가 발생한다. 이처럼 한번의 교차가 일어났을 경우 트래픽 수집지점이 두 패킷 간의 교차점을 기준으로 두 분류로 나뉘게 되어 간단하게 탐지와 처리를 할 수 있지만 Forward 패킷 또는 Backward 패킷이 여러 개의 패킷이 교차가 되는 경우 매우 복잡하게 된다.

a.1 을 제외한 a.2, a.3, a.4 는 모두 Forward 패킷 또는 Backward 패킷이 여러 번 교차하여 발생하게 된다. Abnormal case 2 인 a.2 는 하나의 forward 패킷이 도착하기 전에 Backward 패킷이 2 개 이상의 패킷을 전송하여 교차가 발생하였을 때 일어나는 경우이다.

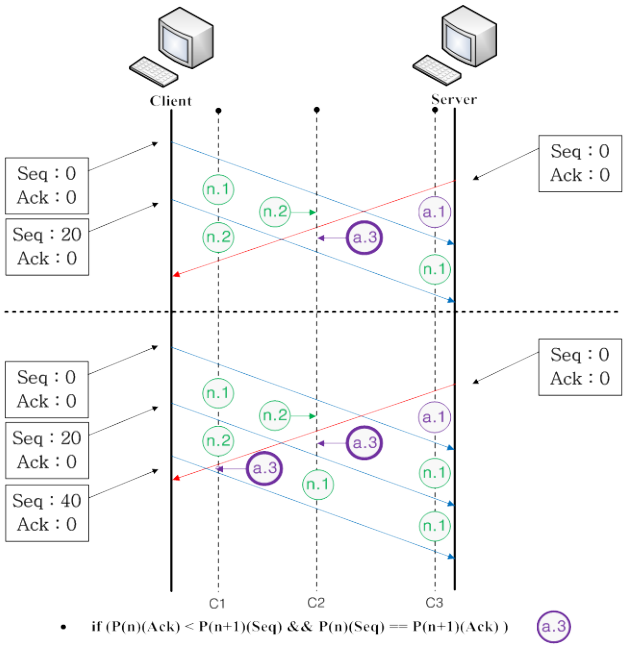

그림 5. Abnormal case 3

Forward 패킷이 전송되고 도착하기 전에 첫 번째 교차가 일어나는 Backward 패킷 이후로 전송되고 Forward 패킷과 교차가 있는 모든 Backward 패킷에서 a.2 는 탐지가 된다. 즉, a.2 은 a.1 이 먼저 탐지가 되고 난 후에 항상 발생할 수 있다.

그림 5 는 Abnormal case 3 인 a.3 의 그림이다. a.2 와 반대로 a.3 은 하나의 Backward 패킷이 도착하기 전에 여러 개의 Forward 패킷이 2 개이상의 패킷을 전송하여 교차가 일어날 경우 탐지 된다. Backward 패킷이 전송되고 도착하기 전에 첫 번째 교차가 일

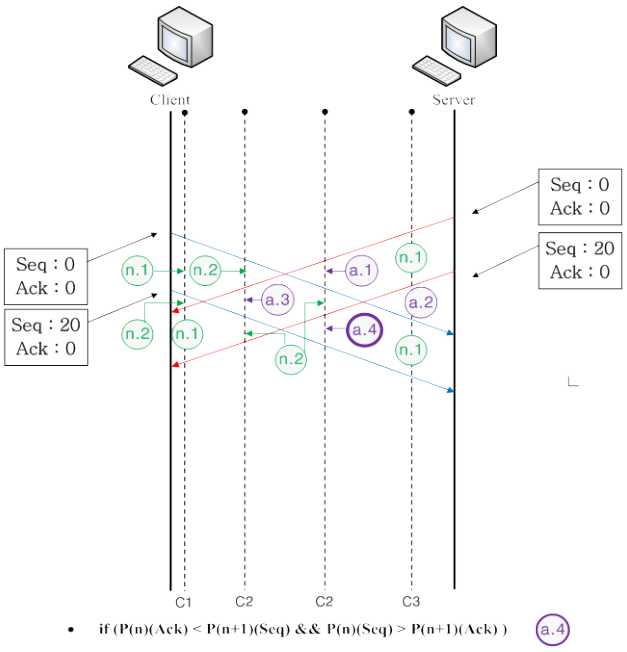

그림 6. Abnormal case 4

어 나는 Forward 패킷 이후로 전송되고 Backward 패킷과 교차가 있는 모든 Forward 패킷에서 a.3 은 탐지가 된다. 즉, a.3 또한 a.1 이 먼저 탐지가 되고 난 후에 항상 발생할 수 있다고 할 수 있다.

그림 6 은 abnormal case 4 인 a.4 의 그림이다. a.4 는 Forward 패킷과 Backward 패킷 모두 두 개 이상의 패킷이 교차될 때 탐지되는 현상이다. 즉, a.4 는 a.1, a.2, a.3 이 모두 발생한 후에 발생할 수 있고 교차된 마지막 Forward 패킷과 Backward 패킷에서 탐지 된다. 이처럼 Abnormal case 가 탐지 되었을 때 해당 case 의 문제뿐만 아니라 여러 case 가 같이 발생하는 문제를 해결 방안에서 고려하여야만 한다.

3.3 해결 알고리즘

본 절에서는 트래픽 수집 지점의 차이로 인해 수집하는 패킷의 순서 이상이 탐지가 되었을 경우 이를 해결하는 알고리즘에 대해 기술한다.

```
Resolve retransmission and out-of-order problem
Remove all non-payload packets from the packet sequence

1: procedure Resolve Abnormal Packet Sequence
2:     Input : packet sequence of a TCP flow
3:     if ( P(n) and P(n+1) are abnormal sequence ) {
4:         change P(n) and P(n +1);
5:         n = n - 1;
6:         if ( P(n) and P(n+1) are abnormal sequence ) goto 4:
7:     }
8: end procedure
```

그림 7. 해결 알고리즘 의사코드

그림 7 은 해결 알고리즘의 의사코드이다. 수집된 트래픽을 가지고 탐지 알고리즘에 의해 탐지가 되면 해결 알고리즘에 의해 n 번째 패킷과 n+1 번째 패킷의 순서를 바꿔준다. 그러나 탐지 되었을 때 패킷의 교차가 한번 일 경우에는 n 번째 패킷과 n+1 번째의 두 패킷 간의 패킷 순서만 바꾸면 되지만 패킷의 교차가 연속적 일 경우에는 n+1 번째 패킷이 n 번째 ,n-1 번째 등 n 번째보다 이전의 패킷들 과의 교차가 있고 문제가 된다면 n 번째 패킷뿐만 아니라 이전 패킷들 과도 패킷 순서 교체를 해주어야 할 경우가 발생한다.

이러한 경우로 인해 탐지 알고리즘에서 n 번째 패킷과 n+1 번째 패킷에서 탐지가 될 경우 n 번째와 n+1 번째는 패킷 순서를 교체하고, n 번째 Sequence 가 된 Backward 패킷은 n-1 번째 패킷과 비교 후 탐지될 시 교체하고, n-1 번째 Sequence 가 된 Backward 패킷은 n-2 번째 패킷과 비교 후 탐지될 시 교체 하는 방식으로 abnormal 탐지가 되지 않을때 까지 역순으로 계속 수행한다. abnormal 탐지가 더 이상 나오지 않는다면 n+1 번째 패킷으로 돌아가 탐지알고리즘을 수행하며 abnormal 을 탐지할 시 교체하는 반복적인 방법으로 전체 플로우 끝까지 수행한다.

그림 8 은 본 논문에서 제안된 트래픽 수집 지점의 차이로 인한 문제점 탐지 및 해결 알고리즘의 플로우 차트이다.

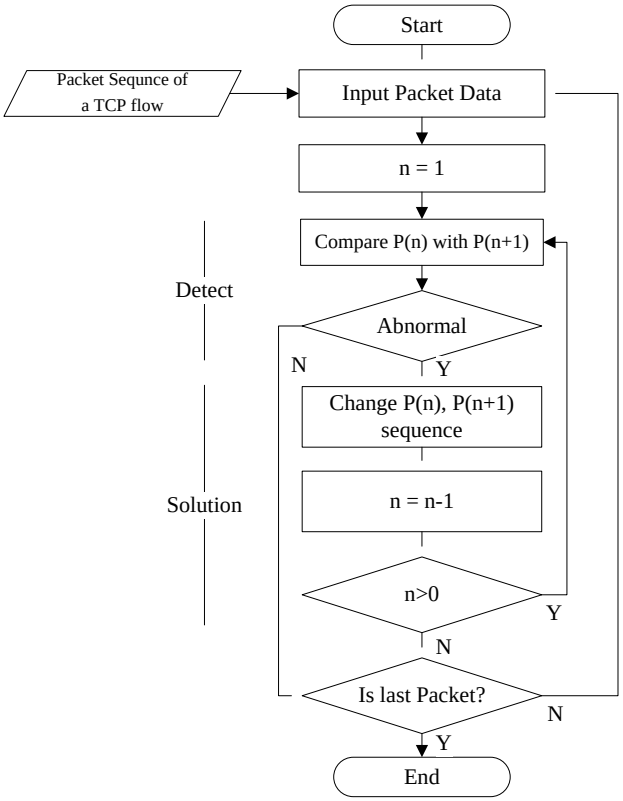

그림 8. 해결 알고리즘 순서도

4. 실험 및 결과분석

본 장에서는 3 장에서 제안된 트래픽 수집 지점의 차이로 인한 문제점 탐지 및 처리 방법의 성능을 평가하기 위해 실험한 결과를 기술한다.

탐지 실험에서는 실제 트래픽에 적용하여 실제로 얼마나 빈번하게 3 장에서 제시한 문제들이 일어나는지를 검사하였다. 제안하는 방법론의 평가를 위해 필자의 학내 망에서 수집한 트래픽을 대상으로 실험하였다. 트래픽 수집은 KU-MON[7]을 이용하여 2012-12-17 일 하루치의 데이터를 수집하여 TCP 세션만을 사용해 실험하였다.

표 1 과 2 는 학내 망 트래픽에서 탐지 알고리즘을 적용해 abnormal 탐지량에 대한 결과를 보여주고 있다. 전체 Flow 와 Packet 수는 하루 치 데이터를 매 분마다 통계를 낸 것을 평균으로 나타낸 것이고 Flow 탐지수는 해당 Flow 에서 여러 개의 패킷 간에 탐지가 되었지만 해당 Flow 에서는 한 번의 탐지로 의미를 두었다. Packet 패킷 탐지 수치는 예를 들어 n 번째와 n+1 번째 패킷의 비교에서 탐지가 되면 패킷 2 개가아닌 한번의 탐지로써 파악을 하였고 Bytes 양의 조사에서는 두 패킷의 Bytes 의 합으로 파악하였다.

표 1. Abnormal case by case

case	measure	Total	Detect	Bytes	Rate
a.1	Flows	16,282	349	3,345,013,424	2.141%
	Packets	50,584	95		0.187%
a.2	Flows	16,282	24	1,625,395,644	0.148%
	Packets	50,584	7		0.014%
a.3	Flows	16,282	7	332,020,732	0.042%
	Packets	50,584	2		0.004%
a.4	Flows	16,282	2,932	932,006,201,946	18.007%
	Packets	50,584	7,899		15.616%

표 2. Total Abnormal case

measure	Total	Detect	Bytes	Rate
Flows	16,282	3,120	937,308,631,746	19.162%
Packets	50,584	8,003		15.821%

Flow, Packet 에 대한 조사에 전체 Flow 중 탐지 Flow 는 19.162%이며 전체 Packet 중 탐지 Packet 은 15.821%로 높은 탐지량을 보여준다. 다시 말해서 제안 된 방법론을 적용하여 충분한 성능 개선을 이룰 수 있다고 제안한다.

5. 결론 및 향후 연구

본 논문에서는 통계 정보 기반 트래픽 분류 방법론에서 필요로 한 feature 가 트래픽 수집 지점에 따라 변화는 문제점을 분석하고 해결 방안을 제시하였다. 또한 실험을 통해 문제점 발생 원인과 발생 비율을 조사하고 탐지방법을 기술하였다. 탐지양은 플로우 기준으로 전체 트래픽의 19%가 넘는 비율을 차지하고 패킷 기준으로 전체 트래픽의 15%가 넘는 수치가 탐지 되었다. 따라서 플로우의 통계적 특징을 기반으로 트래픽을 분류하는 방법론에서 필히 트래픽 수집지점에 따른 feature 변화 문제를 개선하여야 한다.

향후 연구로는 해결 알고리즘의 프로그램 구축과 통계 기반 분석 시스템에 영향을 끼칠 수 있는 다른 문제에 대해서도 연구를 진행하고 비정상적인 세션의 다양한 분석을 제안하고자 한다.

6. 참고문헌

[1] Young-Tae Han and Hong-Shik Park, "Game Traffic Classification Using Statistical Characteristics at the Transport Layer," ETRI Journal, Vol.32, No.1, Feb., 2010, pp.22-32.

[2] 안현민, 최지혁, 김명섭, "TCP 세션의 이상동작으로 인한 트래픽 분석 방법론의 한계와 해결방안," KNOM Review, Vol. 15, No. 1, Dec. 2012, pp. 31-39.

[3] Liu, Hui Feng, Wenfeng Huang, Yongfeng Li, Xing “Accurate Traffic Classification”, Networking, Architecture, and Storage, 2007. International Conference

[4] Myung-Sup Kim, Young J. Won, and James Won-Ki Hong, “Application-Level Traffic Monitoring and an Analysis on IP Networks”, ETRI Journal, Vol.27, No.1, pp.22-42, Feb., 2005.

[5] Jeffrey Erman, Martin Arlitt, Anirban Mahanti, “Traffic Classification Using Clustering Algorithms”, Proc. of SIGCOMM Workshop on Mining network data, Pisa, Italy, pp.281-286, Sep., 2006.

[6] Andrew W. Moore and Denis Zuev, “Internet Traffic Classification Using Bayesian Analysis Techniques,” Proc. of the ACM SIGMETRICS, Banff, Canada, Jun., 2005.

[7] J. S. Park, J. W. Park, S. H. Yoon, Y. S. Oh, M. S. Kim, "Development of signature Generation system and Verification Network for Application Level Traffic Classification", in Proc. KIPS conf. Apr. 23-24, 2009, pp. 1288-1291, PuSan, Korea.