

플로우 내의 패킷 별 가중치를 이용한 통계 정보 기반 트래픽 분류 방법론의 성능 향상

안현민^o, 함재현, 김명섭

고려대학교 컴퓨터정보학과

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

요 약

네트워크의 고속화와 다양한 서비스의 등장으로 오늘날의 네트워크 트래픽은 복잡 다양해지고 있다. 효율적인 네트워크 관리를 위해서는 네트워크에서 발생하는 트래픽에 대한 다양한 분석이 필요하다. QoS, SLA 와 같은 정책을 적용하기 위해서는 트래픽 분석 중 트래픽 분류의 중요성이 크다. 최근에는 플로우의 통계 정보를 이용한 트래픽 분류 방법론이 많이 연구되고 있다. 본 논문에서는 응용에서 발생한 플로우 내의 패킷 전송 순서 별 가중치를 이용하여 통계 정보 기반 트래픽 분류 방법론의 성능을 향상시키는 방법을 제안한다. 학내 망에서 수집한 트래픽을 분석하여 응용 트래픽 분류에 패킷 별 가중치의 적용 타당성을 검증하고, 각 응용의 플로우 내 패킷 전송 순서 별 페이로드 크기 분포를 이용하여 응용의 플로우 내의 패킷 전송 순서 별 가중치를 산출한다. 선정된 통계 시그니처 기반 트래픽 분류 시스템을 이용한 학내 망에서의 실험을 통해 응용 특성에 따른 패킷 별 가중치 적용 유무의 성능을 검증한다.

1. 서론

네트워크의 고속화와 더불어 다양한 서비스와 응용프로그램이 개발됨에 따라 트래픽은 복잡 다양해지고 있다. 이와 같은 현실 속에서 네트워크의 효율적 운용과 관리를 위한 응용 레벨의 트래픽의 모니터링과 분석의 필요성이 커져가고 있다. QoS(Quality of Service), SLA(Service Level Agreement), CRM(Customer Relationship Management)과 같은 정책을 적용하기 위해서는 다양한 종류의 응용 레벨 트래픽을 정확하게 분류할 수 있는 방법과 고속 링크에서 발생하는 대용량의 트래픽을 실시간으로 처리하는 방법이 요구된다. 최근에는 높은 정확도를 가지며 분석 속도도 빠른, 플로우의 통계 정보를 이용한 트래픽 분류 방법론[1,2,3]이 많이 연구되고 있다.

플로우의 통계 정보를 이용한 분류 방법은 패킷 크기, 패킷 간의 시간 간격, 윈도우 크기 등 플로우를 구성하는 패킷들로부터 얻어지는 다양한 통계적 특징을 이용하여 머신 러닝의 특정 알고리즘들을 사용하여 트래픽을 분류하는 방법이 주로 제안되어 왔다[4]. 또한, 특정 통계적 정보를 이용하여 자체적인 알고리즘을 개발한 연구들도 진행되었는데, 그 중 패킷 또는 페이로드 크기 분포를 이용한 분류 방법들[3, 5, 6, 7, 8]이 많이 제안되고 높은 정확도를 나타내었다.

본 논문에서는 기존 연구의 통계 정보 기반 트래

픽 분류 방법론에서 모든 패킷에 동일하게 가중치를 적용하는 것에 대한 한계점을 파악하고, 그 한계점을 극복하기 위해 응용의 특성을 조사한 뒤 각 응용의 특성에 따라 패킷 별 가중치를 다르게 적용하여 트래픽을 분류하는 방법을 제안한다.

성능 검증을 위해 통계 시그니처 기반 트래픽 분류 시스템을 선정하고 해당 시스템에 제안하는 방법의 적용 전후 트래픽 분류 실험을 통해 분석률 및 정확도의 향상을 분석한다. 실험은 학내망에서 수집한 트래픽을 이용한다.

본 논문은 다음과 같이 구성된다. 서론에 이어 2 장에서는 관련 연구로써 선정된 통계 시그니처 기반 분류 방법에 대해 간략히 설명한다. 3 장에서는 정답지 트래픽을 이용해 응용 별 패킷의 크기 분포를 조사하고, 4 장에서는 3 장에서 조사한 정보를 이용하여 가중치를 결정한다. 5 장에서는 제안한 방법론의 우수성을 검증하기 위해 각 응용 특성에 따른 패킷 별 가중치 적용 유무를 소규모 네트워크에서 수집한 트래픽에 대해 실험한 내용과 분류 결과를 기술한다. 마지막으로 6 장에서는 결론 및 향후 연구에 대해 기술한다.

2. 관련 연구

본 논문에서 제안하는 방법론의 대상은 통계 시그니처 기반 분석 방법과 머신 러닝 기반의 분석 방법이다. 통계 시그니처 기반 분석 방법은 플로우의 여러 통계정보를 이용하여 응용 별 시그니처를 추출하고 이를 이용하여 트래픽을 응용 별로 분류

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

하는 것이다. 머신러닝 기반의 분석 방법은 응용 별 인터넷 트래픽의 특징이 될 수 있는 항목들을 머신러닝의 classification clustering 기법을 이용하여 트래픽을 분류하는 방법이다. 두 방법론은 모두 패킷 크기, 전송 순서 등을 사용 하기 때문에 제안하는 패킷 별 가중치를 적용할 수 있다.

본 장에서는 관련연구로써 제안하는 방법의 성능 검증 실험을 위해 선정된 통계 시그니처 기반 트래픽 분류 시스템을 간략히 설명한다. 본 논문에서 사용할 통계 정보 기반 트래픽 분류 방법론은 통계 시그니처 기반 트래픽 분류 방법론[9]이다.

[9]에서는 5-tuple 정보(source IP, source port, destination IP, destination port, protocol)를 기반으로 양방향 패킷의 순서 집합을 플로우라고 정의하고, 플로우의 첫 N 개 패킷의 전송 순서, 방향, 페이로드 크기를 이용해 플로우를 N 차원 벡터화 한다. 그 뒤 Feature 값이 일정 기준치 내인 플로우들을 그룹핑하고 각 그룹별로 하나의 시그니처를 추출한다. 이렇게 추출된 시그니처를 이용하여 트래픽을 분류하는 방법이다. 페이로드를 포함하지 않는 TCP 컨트롤 패킷(SYN, RST, FIN, ...) 등은 N 개 패킷에 포함되지 않는다. 즉, 페이로드 크기는 데이터 패킷(페이로드가 존재하는 패킷)의 페이로드 크기만을 의미한다. 방향은 양수와 음수로 표현하는데 TCP 의 경우 양수는 클라이언트에서 서버로 향하는 패킷, 음수는 서버에서 클라이언트로 향하는 패킷을 의미한다. UDP 는 서버/클라이언트의 구분이 명확하지 않기 때문에, 양수/음수의 의미는 단지 방향이 서로 반대라는 것만 표현할 수 있다. 따라서 UDP 의 경우에는 발생하는 첫 패킷을 양수로 표현하고 뒤에 이어지는 패킷은 첫 패킷을 기준으로 방향이 같으면 양수, 다르면 음수로 표현한다. 양방향 플로우는 최대 N 차원의 벡터로 표현되며, [9]에서는 이를 PSD(Payload Size Distribution) 벡터 또는 PSD 패턴이라 지칭한다. 그림 1 과 같은 플로우를 PSD 벡터로 표현하면 {+20, -30, +20,+25, -15}이다.

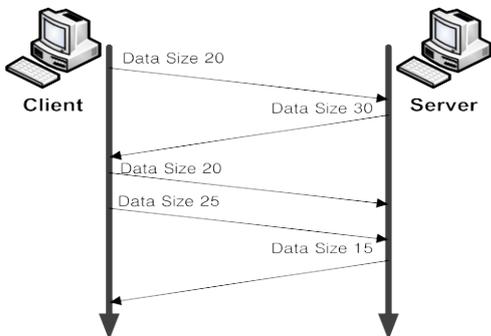


Figure 1. The bi-directional flow

모든 플로우를 PSD 벡터화 한 후 City-block distance 를 이용하여 PSD 벡터 사이의 거리를 계산한다. 그 거리가 일정 기준치 이내인 플로우들을 그룹핑한다. 이때 그룹은 하나의 벡터 값이 아닌, 일정 범위를 지니게 된다.

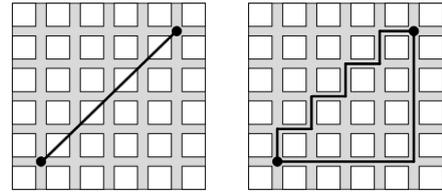


Figure 2. The methods to measure the distance between two points

시그니처는 응용 프로그램 이름, 전송 계층 프로토콜, 각 패킷 별 크기 및 패킷 별 크기 변위값을 가진다. 패킷 별 크기와 변위값을 통해 범위를 지니게 된다. 트래픽 분류 시에는, 분류하고자 하는 플로우의 PSD 벡터가 이러한 시그니처의 범위 안에 존재할 때 해당 플로우는 시그니처를 포함하는 응용으로 분류된다.

[9]에서 플로우를 그룹핑하는 조건으로 플로우와 그룹 사이의 City-block distance 를 이용한다. 대다수의 경우 N 차원의 공간에서 두 점 사이의 거리는 Euclidean distance 를 사용해 측정하는데 실시간 트래픽 분석을 위해 Euclidean distance 와 크기는 비례하나 계산속도가 빠른 City-block distance 를 이용하는 것이다.

하지만 [9]에서 실험한 결과에 따르면 유사한 Feature 를 가진 플로우들을 그룹핑하여 살펴보았을 때 패킷 별로 고정 크기를 가진 그룹의 양이 전체 그룹의 양의 81.8%를 차지한다. 패킷 별로 고정 크기를 가진 경우, 해당 패킷의 변위는 0 이 되고 이를 Feature 로 이용하는 것은 더욱 정확한 트래픽의 분류를 가능하게 한다. 시그니처의 변위가 0 인 패킷과 작은 차이가 나는 플로우도 City-Block distance 로는 같은 값을 가질 수 있기 때문이다. 따라서 통계 정보 기반 트래픽 분류 방법론에서 N 차원의 두 점 사이의 거리를 측정할 때에는 각 차원 별 거리 계산법을 사용해야 한다.

[9]에서 시그니처의 대표값을 산출할 때 그룹 내 각 패킷들 크기의 평균을 이용한다. 하지만 이도 최소값 혹은 최대값을 대표값으로 이용하는 것으로 바꾸어야 한다. 대량의 데이터를 이용하는 해당 방법론의 특성 상 모든 추출된 Feature 가 같은 값을 가리킬 수는 없기 때문에 변위가 존재하게 된다. 대표값으로 평균이나 중앙값을 이용할 경우 그림 3 과 같이 변위는 대표값에서 +/- 형태로 존재하게 되며 이는 그림 4 와 같이 최소/최대값을 이용하여 한 방향의 변위만을 가진 형태보다 계산이 복잡하며 속도도 느리다. 한, 두 번의 계산에선 차이가 나지 않으나 요즘과 같이 인터넷이 발달하여 분 당 트래픽이 대량인 경우 실시간 분석에서 그 성능 차이가 드러날 수밖에 없다. Feature 의 대표값으로 최소/최대값을 사용해서 변위를 이용한 포함 범위 계산을 한번만 해야 한다.

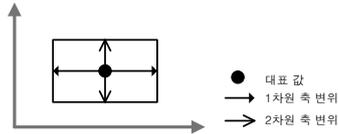


Figure 3. The representative and variation using median value

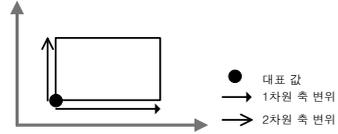


Figure 4. The representative and variation using minimum value

그림 5 는 [9]의 시스템에 위 두 방법을 적용한 트래픽 분류 방법의 예시로서 세 개의 플로우 a, b, c 에 대한 분류 결과를 보여준다. 플로우 a 는 응용 A 의 시그니처 중 하나에 의해 분류되는 경우로서, PSD 벡터가 응용 A 의 시그니처의 포함 범위에 속하므로 응용 A 로 분류한다. 플로우 b 는 어떠한 시그니처에도 속하지 않는 경우로서, 미확인(unknown) 트래픽으로 분류한다. 플로우 c 는 같은 응용 내에서 두 개 이상의 시그니처에 속하는 경우이며, 예시에서는 응용 B 로 분류된다.

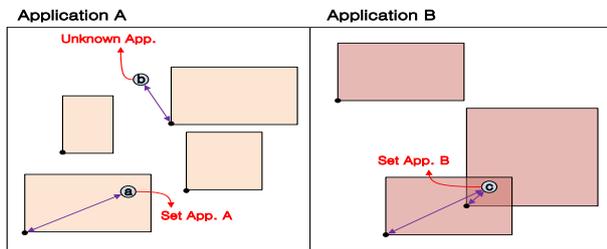


Figure 5. The example of traffic classification based on the statistic signature

PSD 기반의 통계 시그니처는 응용 별 시그니처의 충돌이 존재한다. 포함 범위가 겹치는 두 개 이상의 시그니처의 응용이 서로 다를 때 이를 충돌이라고 정의한다. 분류 시 충돌 영역에 속하는 플로우는 2 개 이상의 응용에 의해 분류 가능하므로, 이 중 하나의 응용으로 분류할 경우 잘못 분류할 가능성이 있기 때문에 이는 미확인(unknown)트래픽으로 분류한다.

본 논문에서는 해당 트래픽 분석 시스템의 플로우 그룹핑 및 트래픽 분류 단계에서 응용 특성에 따라 패킷 별 가중치를 적용하여 통계기반 트래픽 분류 시스템의 성능향상 방안을 제안한다.

3. 응용 별 패킷 크기 분포

본 장에서는 응용 별로 플로우 첫 N 개 패킷의 페이로드 크기 분포를 조사하여 패킷 별 가중치를 응용에 따라 다르게 적용하는 방법의 타당성에 대해 증명한다. 이는 정답지(ground-truth) 트래픽에 대한 분석을 통해 이루어진다.

한 응용의 플로우들 중 첫 N 개 패킷의 순서와 전송 방향이 일정한 플로우들끼리 하나의 그룹으로 그룹핑 한다. 그룹 별로 패킷 크기 분포도를 그려 다른 Feature(패킷 전송 순서, 전송 방향)가 일정할 때 패킷의 크기 변화를 분석한다.

3.1 정답지 트래픽 트레이스

트래픽 분류에 사용되는 시그니처를 생성하거나 분류 규칙을 생성하기 위해서는 정답지(ground-truth) 트래픽이 필요하다. 이러한 정답지는 매우 정확해야만 트래픽 분류 결과에 신뢰성을 보장해준다. 본 연구에서는 TMA-에이전트 기반의 정답지 생성 방법 [10]을 이용하여 정확한 정답지를 학내 망에서 10 일간 수집하였다. 표 1 은 분석에 사용된 정답지 트래픽에 관한 정보이다.

Table 1. The traffic information of collected application traces (ground-truth)

Application	Flow (x 10 ³)	Packet (x 10 ³)	Byte (x 10 ⁶)
6	109	1,197	504

수집 기간은 2012-11-30 ~ 2012-12-09 까지 10 일간이며, 응용은 총 6 개로 Dropbox, Kartrider, NateOn, Skype, Teamviewer, uTorrent 가 그것이다.

3.2 응용 별 패킷 크기 분포

본 절에서는 통계 정보 기반 트래픽 분류 시스템에서 응용에 따라 패킷 순서 별 가중치를 이용하는 것의 타당성 입증 실험 결과를 분석한다. 실험은 3.1 절의 트래픽을 이용하여 진행하였고, 대상으로 하는 방법론의 Feature 인 패킷 크기, 전송 방향, 전송 순서 중 가중치를 할당하기 적합한 패킷 크기에 대해 분석하였다. 응용에 따라 다르게 패킷 순서 별 크기에 가중치를 할당하는 것이 타당한지 검증하기 위해 한 응용 내에서 첫 N 개 패킷의 전송 방향과 순서가 일정한 플로우들끼리 그룹핑 하여 각 응용 내 그룹별로 패킷 크기 분포를 분석한다. 패킷은 해당 방법론의 Feature 인 페이로드가 있는 패킷만을 사용하였으며 첫 패킷은 TCP 의 경우 3-handshake 패킷 이후의 페이로드가 있는 패킷을 첫 패킷으로 정의한다.

그림 6 부터 11 은 각 응용의 그룹 중 가장 많은 플로우를 포함한 그룹 두 개를 선정하여 그룹 내의 패킷 크기 분포를 조사한 것이다. x 축은 패킷의 순서이며 y 축은 패킷의 페이로드 크기이다. 각 그룹에 포함된 모든 플로우의 패킷 순서에 맞는 페이로드 크기를 나타낸 것이다. 본 논문에서는 N 의 값으로 5 를 사용하였다.

각 응용의 그룹 내 패킷 크기 분포도를 살펴보면, Dropbox 는 5 패킷 전부 변화가 심하며 Kartrider 는 전체적으로 변화가 거의 없다. NateOn 은 1 번째, 4, 5 번째 패킷의 변화 차이가 크다. Skype 는 그룹별로 조금 다른 양상을 띈다. Teamviewer 는 그룹 내 한,

두 패킷의 변화만 보이며 uTorrent 또한 5 패킷 전부 변화가 심하다. 이렇듯 각각의 응용은 패킷의 전송 순서 별로 페이로드 크기의 변화 폭이나 빈도가 다르다. 즉, 응용 별로 크기 변화가 작은 패킷, 큰 패킷이 다르며 따라서 비교적 더 중요하게, 혹은 덜 중요하게 여겨야 하는 패킷이 달라야 한다. 패킷의 크기, 전송 순서와 전송 방향 등을 Feature 로 사용하는 모든 통계 정보 기반 트래픽 분류 방법론에서 플로우 내 패킷 별 가중치를 상이하게 적용하는 것이 타당하다는 것을 확인할 수 있다.

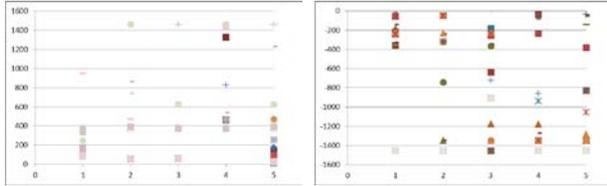


Figure 6. The packet size distribution in order of first 5 packets of each flow in each flow group of Dropbox

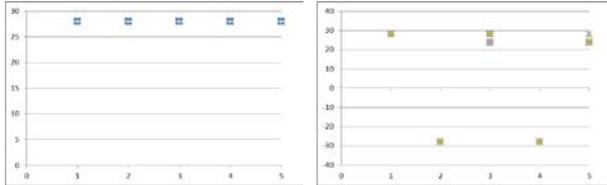


Figure 7. The packet size distribution in order of first 5 packets of each flow in each flow group of Kartrider

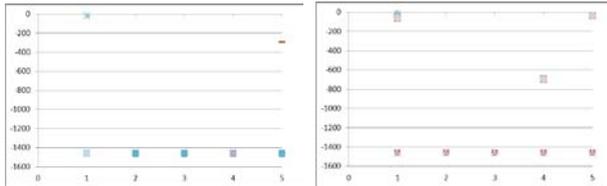


Figure 8. The packet size distribution in order of first 5 packets of each flow in each flow group of NateOn

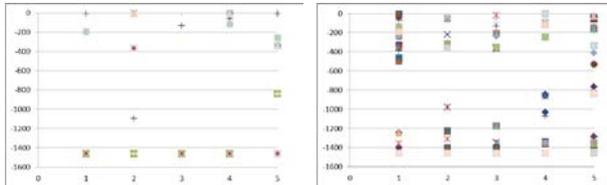


Figure 9. The packet size distribution in order of first 5 packets of each flow in each flow group of Skype

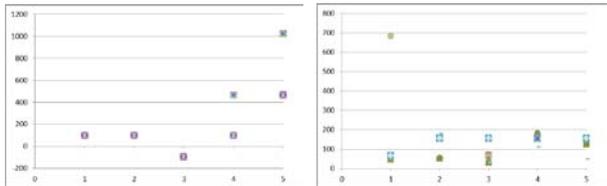


Figure 10. The packet size distribution in order of first 5 packets of each flow in each flow group of Teamviewer

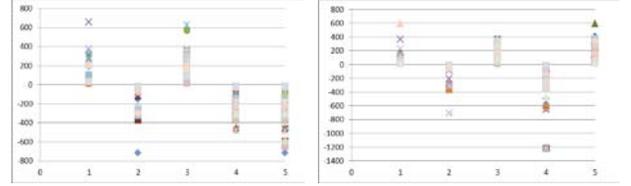


Figure 11. The packet size distribution in order of first 5 packets of each flow in each flow group of uTorrent

4. 응용 특성에 따른 패킷 가중치 결정

패킷 별 가중치 적용 실험에 앞서 가중치를 결정해야 한다. 3.2 절의 실험 결과인 응용 별 그룹들의 패킷 별 분산을 이용하여 계산한다. 먼저 각 그룹의 패킷 별 분산을 더해 (1)과 같이 패킷 별 분산의 합을 구한다. $v(p_i)$ 는 i 번째 패킷의 분산이며 $G_j(v(p_i))$ 는 그룹 j 의 i 번째 패킷의 분산이다.

$$Sum_i = \sum_{j=1}^N G_j(v(p_i)) \quad (1)$$

그 후 (2)를 이용해 한 응용 내 그룹들의 패킷 순서 별 분산의 합을 더해 전체 분산의 합을 구하고 이와 (3)을 이용해 패킷 순서 별 분산의 합의 전체 분산 합과의 비율을 계산한다.

$$TotalVariance = \sum Sum_i \quad (2)$$

$$Rate(Sum_i) = Sum_i \times \frac{100}{TotalVariance} \quad (3)$$

이를 각 응용마다 수행한다. 마지막으로 패킷 별 분산의 합의 비율을 소수 첫째 자리에서 반올림하여 $Rate(Sum)$ 을 얻었고 계산의 편의성을 위해 1 의 자리에서 반올림 한 뒤 10 으로 나누어 $WeightVector$ 를 얻었다. 그 결과는 표 2 와 같다.

Table 2. Rate(sum) and WeightVector per each Application

Application	Rate(sum)	WeightVector
Dropbox	{20, 29, 18, 13, 20}	{2, 3, 2, 1, 2}
Kartrider	{19, 19, 22, 19, 21}	{1, 1, 1, 1, 1}
NateOn	{46, 6, 7, 11, 30}	{5, 1, 1, 1, 3}
Skype	{11, 32, 9, 27, 21}	{1, 3, 1, 3, 2}
Teamviewer	{9, 7, 6, 27, 51}	{1, 1, 1, 3, 5}
uTorrent	{21, 18, 17, 23, 21}	{1, 1, 1, 1, 1}

5. 실험 및 결과 분석

본 장에서는 4 장에서 산출한 응용 특성에 따른 패킷 별 가중치에 대한 실험과 결과 분석에 대한

내용을 기술한다.

총 10 일 간 학내 망에서 발생한 트래픽을 대상으로 통계 정보 기반 트래픽 분류 방법론을 선정하여 패킷 별 가중치의 적용 유무의 분석 결과를 비교하였다. 서로 다른 둘 이상의 응용의 시그니처가 하나의 플로우를 포함할 수 있을 때 이를 시그니처의 충돌이라 하며, 두 개 이상의 어플리케이션으로 분류되는 트래픽은 분류되지 않음으로 처리하였다.

Table 3. The completeness and accuracy of traffic classification using the proposed approach

구분	Flow		Packet		Byte	
	적용 전	적용 후	적용 전	적용 후	적용 전	적용 후
분석률	75.22%	75.51%	37.03%	37.36%	35.84%	35.90%
정확도	99.48%	99.53%	99.41%	99.53%	99.89%	99.91%

표 3 은 각각 선정한 통계 기반 트래픽 분류 시스템의 기존 알고리즘과 본 논문에서 제안하는 응용에 따른 패킷 별 가중치를 적용한 알고리즘의 분석률과 정확도를 비교한 표이다. 분석률 측면과 정확도 측면에서 모두 소폭 향상된 결과를 보였다. 분석률이 증가한 이유는 패킷 별 가중치를 적용하여 시그니처를 추출하였을 때엔 패킷의 페이로드 크기에 가중치를 곱하여 Feature 로 사용함으로써 각 기능별로 뚜렷이 구분되어 가중치를 적용하지 않고 시그니처를 추출하였을 때 보다 충돌이 적게 발생하였기 때문이다. 정확도가 증가한 이유도 패킷 별 가중치로 인해 본래엔 비슷한 페이로드 크기 분포를 가진 서로 다른 응용의 시그니처와 플로우가 확연히 구분되었기 때문이다.

실험 결과를 통해 응용의 특성에 따른 패킷 별 가중치를 적용하면 같은 방법론도 더 좋은 성능을 낼 수 있음을 알 수 있다.

6. 결론 및 향후 과제

본 논문에서는 통계 정보 기반 트래픽 분류 방법론에서 응용 특성에 따른 패킷 순서 별 가중치에 대해 정답지 분석을 통한 페이로드 크기 분포 분석으로 그 타당성을 증명하였고 그룹 별 패킷의 분산 비율을 이용해 가중치를 산출하였으며, 패킷 별 가중치 적용 유무 실험 결과 비교를 통하여 그 성능을 검증하였다. 응용 마다 패킷의 전송 순서 별로 페이로드 크기의 변화 폭이나 빈도가 다르다는 점은 모든 응용 모든 패킷에 같은 가중치를 두어 Feature 로 사용하는 것의 효율이 떨어진다는 점을 실험을 통해 증명하였다. 응용 특성에 맞춰 패킷 별 가중치를 적용하였을 때 분석률 및 정확도의 향상을 얻을 수 있었다.

향후 연구에서는 패킷 순서 별 가중치를 동적으로 추출할 수 있는 방법과 새로운 가중치 적용 방법, 그리고 통계 정보 기반 트래픽 분류 방법론에

영향을 끼치는 다른 요소에 관한 연구를 계획 중이다.

7. 참고 문헌

- [1] Rentao Gu, Minhuo Hong, Hongxiang Wang, and Yuefeng Ji, "Fast Traffic Classification in High Speed Networks," Proc. of the Asia-Pacific Network Operations and Management Symposium (APNOMS) 2008, LNCS 5297, Beijing, China, Oct. 22-24, 2008, pp. 429-432.
- [2] Ying-Dar Lina, Chun-Nan Lua, Yuan-Cheng Laib, Wei-Hao Penga and Po-Ching Lina, "Application classification using packet size distribution and port association" Proc. of the Journal of Network and Computer Applications, In Press, Corrected Proof, Available online, March. 20. 2009.
- [3] Huifang Feng and Yantai Shu, "Statistical Analysis of Packet Interarrival Times in Wireless LAN," Proc. of the Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference, Shanghai, China, Sept. 21-25, 2007, pp. 1888-1891.
- [4] Thuy T.T. Nguyen and Grenville Armitage, "A Survey of Techniques for Internet Traffic Classification using Machine Learning," IEEE Communications Surveys and Tutorials, to appear, 2008.
- [5] L.Bernaille, R. Teixeira, and K. Salamatian, "Early Application Identification," In: CoNext 2006. Conference on Future Networking Technologies., 2006
- [6] 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.
- [7] Gerhard Munz, Hui Dai, Lothar Braun, and Georg Carle, "TCP Traffic Classification Using Markov Models," In Proc. of Traffic Monitoring and Analysis Workshop (TMA) 2010, Zurich, Switzerland, April 2010.
- [8] Valentin Carela-Espanol, Pere Barlet-Ros, Marc Sole-Simo, Alberto Dainotti, Walter de Donato, and Antonio Pescape, "K-dimensional trees for continuous traffic classification," In Proc. of Traffic Monitoring and Analysis Workshop (TMA) 2010, Zurich, Switzerland, April 2010
- [9] Jin-Wan Park, Myung-Sup Kim, "Performance Improvement of the Statistic Signature based Traffic Identification System", KIPSTC.,18C.4., Aug. 2011, pp. 243-250.
- [10] Valentin Carela-Espanol, Pere Barlet-Ros, Marc Sole-Simo, Alberto Dainotti, Walter de Donato, and Antonio Pescape, "K-dimensional trees for continuous traffic classification," In Proc. of Traffic Monitoring and Analysis Workshop (TMA) 2010, Zurich, Switzerland, April 2010.