제 19 회 한국정보처리학회 춘계학술발표대회 논문집 제 10 권 제 1 호 (2003. 5)

패킷 마킹을 이용한 IP 역추적 기술

강동호, 한승완, 장종수
한국전자통신연구원 네트워크보안연구부

IP based Traceback technology Using Packet Marking

Dong-ho Kang, Se0ung-wan Han, Jong-soo Jang
Network Security Department, ETRI

요 약

최근 인터넷의 급속한 보급 확대 및 대중화에 따라 네트워크상의 서비스 속도와 정보보안은 인터넷 사용자에게 중요한 문제로 대두되고 있고, 이로 인해 네트워크의 트래픽 증가와 공격자로부터의 공격이 점점 증가하고 있는 현실이다. 인터넷 사용자의 급속한 증가로 인한 복잡한 TCP 반응과 연관되어 네트워크 서비스에 많은 패킷 손실을 야기하게 되었다. 이러한 문제를 해결하기 위하여 대학 연구소와 기업체에서는 패킷에 임의의 정보를 마킹하고, 마킹된 정보를 추출하는 기법들을 연구하는 중이다. 그러한 기법은 네트워크 트래픽 제어를 위하여 패킷에 우선순위를 두어 차등 서비스 제공을 목적으로 패킷마킹 기법을 사용하기도 한다.

1. 서론

현재 인터넷의 급속한 보급 및 대중화에 따라 DDoS 공격이 손쉽게 이루어지고 있음으로 인해 DDoS 공격대응을 위한 연구가 활발하게 진행되고 있다. 공격자의 공격에 대응하는 근본적인 공격차단을 연구하는 과정에서 기존의 침입탐지시스템이 소극적인 대응이었다면 이제는 좀더 적극적인 대응 방안으로 해커의 실제 위치를 찾아서 공격에 대한 대응하는 방법으로 발전하고 있다. 여러 시스템을 경유하여 특정 시스템의 접근권한을 획득하고자 하는 전통적인 해킹기법을 위한 역추적은 TCP 연결 역추적 기술이 필요하며, 시스템 또는 네트워크를 마비시키기 위한 Bandwidth Consumption 공격은 IP 기반 역추적 기술이 필요하다.

본 논문에서는 Bandwidth Consumption 공격의 대응을 위한 IP 역추적 기술에 대해서 논하였고, 해당 기술들이 가지는 문제점에 대해서 언급하였다. 이러한 문제점들을 개선시킴으로써 현 인터넷 망에 적용하지 못하고 있는 역추적 기술에 대한 적용 가능성을 살펴보고자 한다.

2. IP 역추적 기술

대부분의 DDoS 공격은 해커의 위치를 숨기기 위해서 IP 주소를 변경하여 공격을 시도한다. 이러한 공격에 대응하기 위해서는 우선적으로 해커의 실제 위치를 찾아 대응하는 방법이 필요하며, 이를 위해서 해커의 공격 패킷으로부터 별도의 부가적인 정보를 수집하여 공격 패킷의 실제적인 주소를 찾는 연구가 진행중 이다. 이러한 접근 방법을 IP 역추적이라 불리며 다음과 같이 4 가지 기법이 존재한다.

- Traceback Marking 기법

역추적 마킹 기법은 패킷이 라우터를 거쳐갈 때, 라우터에서 자신의 특정 정보를 덧붙이고, 피해 시스템은 라우터 정보가 포함된 수신 패킷들을 통해 역추적하는 기법이다.

- Traceback Logging 기법

역추적 로깅 기법은 키 라우터로 하여금 일정기간 동안 키 라우터를 거쳐가는 모든 패킷 정보를 기록하여, 데이터 마이닝 기법을 이용하여 패킷을 역추적하는 기법이다.

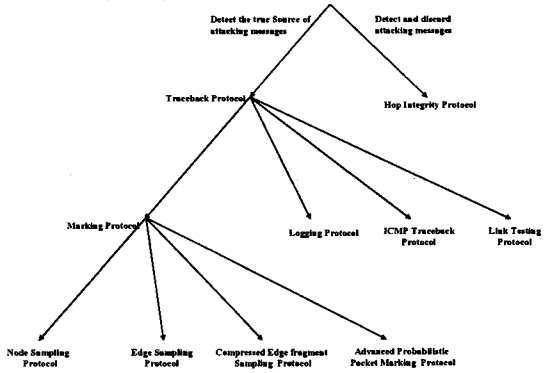
- Link Testing 기법

공격이 이루어지고 있는 동안 피해 시스템에 가장 가까운 라우터에서 시작하여 전달되는 패킷의 위치를 거슬러 올라가는 기법이다. 이 기법은 공격이 진행되는 중일 때만 역추적이 가능하며, 공격이 중단되며 역추적도 불가능하다는 단점이 있다.

- ICMP 역추적(ICMP Traceback) 기법

ICMP 역추적 기법은 라우터에 거쳐가는 패킷에 대해서 패킷의 일부가 포함된 ICMP 역추적 패킷을 생성하여 목적지 주소로 전송하고, 전송 받은 시스템은 해당 정보를 수집하여, 공격이 검출되면 수집된 정보를 이용하여 해커를 역추적하는 기법이다. 이 기법은 해커가 임의의 ICMP 역추적 패킷을 생성하여 피해 시스템에 전송할 수 있다는 단점이 존재한다.

위와 같은 방법은 DDoS 와 같은 공격을 대응하기 위한 하나의 과정이며, 이 자체가 DDoS 공격을 막을 수는 없다. 따라서, 해커의 실제 위치를 추적 후, 해커의 공격 패킷을 차단하는 방법이 요구되며, 대표적인 방법으로는 Hop Integrity 기법이 있다. [그림 1]은 DDoS 공격에 대응하기 위한 기법을 설명하고 있다.



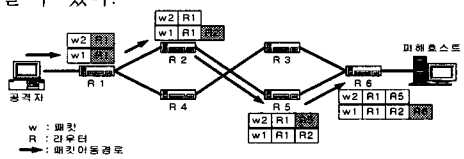
[그림 1] DDoS 공격 대응을 위한 프로토콜

3. IP 패킷 마킹 기법

패킷 마킹 기법은 네트워크에서 패킷이 전송되는 동안 부분적인 경로 정보를 가지고 라우터에서 확률적으로 패킷에 마킹을 하게 된다. 이렇게 마킹된 패킷을 받은 피해 호스트는 마킹된 패킷을 이용하여 공격자의 근원지를 네트워크상에서 찾아가게 되는 기법이다.

1) Node Sampling 기법

Node Sampling 기법의 구조는 [그림 2]와 같이 라우터에서는 확률 p 를 이용하여 지나가는 패킷에 대하여 마킹 하게 된다. 공격자에 의해 보내진 패킷 w1 이 R1, R2, R5, R6 를 경유하여 피해호스트로 전송될 경우, 피해호스트에서는 패킷 w1 의 마킹 정보는 R1, R2, R6 가 될 것이다. 이렇게 패킷 w2 의 마킹 정보는 R1, R5 가 될 것이다. 이러한 방법으로 공격경로를 재설정할 수 있을 정도의 패킷을 받는다면 공격자의 근원지를 찾아낼 수 있다.

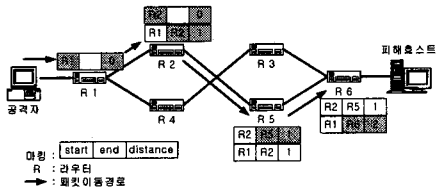


[그림 2] Node Sampling 기법

2) Edge Sampling 기법

Edge sampling 기법은 앞에 제시된 기법의 문제들을 해결하기 위해 개별적인 노드들 보다는 공격경로의 edge 들을 암호화하는 것이다. 이를 위해 피해호스트에서는 edge 샘플의 거리를 표현하는 필드와 링크의 각 끝에 라우터의 IP 주소를 표현하기 위하여 각 패킷에 2 개의 고정된 크기의 start 주소 필드와 end 주소 필드를 예약하게 된다. 라우터가 패킷에 마킹을 결정하게 되면, start 필드에 라우터 자신의 주소를 기록하고 거리 필드에 0 을 기록한다. 만약 그렇지 않고, 이미 거리 필드에 0 이 기록되어 있다면, 이전의 라우터에서 마킹된 것이므로 라우터는 자신의 주소를 end 필드에 기록하고 거리필드를 1 증가시킨다. 이것이 이전 라우터와의 edge 를 나타내는 것이다. 피해호스트에서 거리가 d 홉 떨어진 라우터로부터 마킹된 패킷을 받을 확률은 1 / p(1-p)d-1 이다. 그리고 라우터에서 마킹된 패킷이 도착할 확률은 적어도 dp(1-p)d-1 이다. 예를 들면, 만약 확률 p=1/10 이고 공격경로의 길이가 10 이면, 피해호스트는 공격자로부터 75 패킷을 받은 후에 공격경로를 재설정할 수 있다.

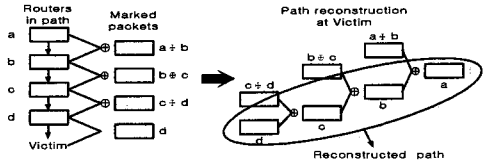
Edge sampling 기법의 구조는 [그림 3]과 같다. 그림에서 나타내듯이 라우터에서는 지나가는 패킷에 대하여 확률 p 를 적용하여 마킹하는 것이다. 라우터 R1 에서 마킹이 결정되면 패킷의 start 필드에 자신의 주소를 마킹하고 distance 필드에 0 을 마킹한다. 다음 라우터 R2 에서는 패킷에 이미 마킹된 것을 판단하고 end 필드에 자신의 주소를 마킹하고 distance 필드를 1 증가시킨다. 이렇게 마킹된 start 필드와 end 필드의 정보를 이용하여 라우터 사이의 edge 를 나타낼 수 있다.



[그림 3] Edge Sampling 기법

3) Compressed Edge Fragment Sampling 기법

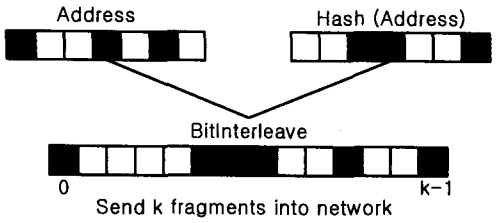
Compressed Edge Fragment sampling 기법은 마킹할 공간을 줄이기 위해 두 가지 기법을 사용한다. 첫째로, edge 를 구성하는 두 개의 주소(start 필드, end 필드)를 XOR 연산을 통하여 저장 공간을 반으로 줄이는 것이다. 즉, 라우터 a 에서 패킷에 마킹을 결정하게 되면, 그 패킷에 a 의 주소를 마킹하고, 다음 라우터 b 에서는 edge 필드에 a, b XOR 연산 값을 패킷에 저장하게 된다. 이 값을 edge-id 라 한다. 패킷이 경유한 라우터들의 정보 값인 피해호스트에서 받은 패킷의 edge 필드를 a XOR b XOR a = b 를 이용하여 공격경로를 재설정할 수 있다. [그림 4]는 XOR 연산을 이용하여 경로 정보를 마크하고 마크된 정보를 이용하여 경로를 재설정하는 것을 나타낸다.



[그림 4] XOR 연산을 이용한 인코딩과 디코딩

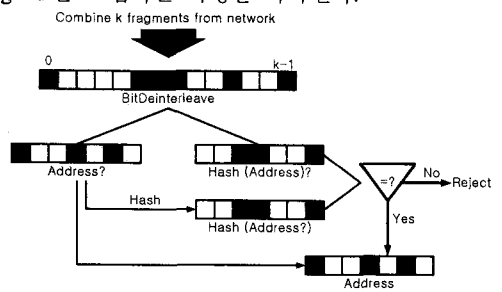
둘째로, edge-id 를 k 개로 분할하여 전송함으로써 저장 공간을 줄일 수 있다. [그림 5]는 k 개로 분할된 edge-id 의 구조를 나타낸다. 라우터가 패킷에 마킹을

결정하게 되면 분할된 k 개 중 하나를 랜덤 하게 선택하여 패킷에 저장하고 k 에 대한 $\log_2 k$ 의 저장 공간 추가하게 되는데 이것이 offset 이다. 공격자에 의해 충분한 패킷이 보내진다면 결국, 피해호스트는 k 개의 모든 fragment 를 받을 것이다.



[그림 5] K 개로 분할된 edge-id

하지만, 다중 공격자들이 존재한다면 피해호스트는 같은 거리에 존재하는 여러 개의 edge fragment 들을 받을 수 있다. 다른 경로로부터 들어온 fragment 를 조합함으로써 잘못된 경로의 재설정을 방지하기 위해 각 라우터의 주소와 랜덤 한 hash 함수 값을 interleaving 한 단순한 오류검출코드를 추가한다. 피해호스트에서는 이렇게 interleaving 한 값을 받아 deinterleaving 하여 라우터의 주소 값과 hash 함수의 값으로 나눈 후 hash 함수 값을 검사하여 정확한 라우터 주소를 알아냄으로써 공격경로를 재설정할 수 있다. [그림 6]은 k 개로 분할되어 패킷에 의해 전송된 edge-id 를 조립하는 과정을 나타낸다.



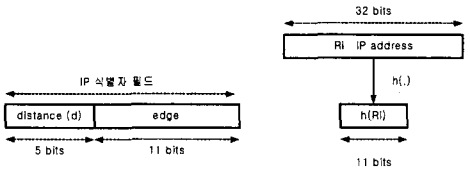
[그림 6] K 개로 분할된 edge-id 조립과정

4) Advanced Probabilistic Packet Marking 기법

이 기법의 전제 조건은 upstream 라우터들의 맵으로 한정 지으며, 피해호스트는 라우터들의 맵을 가질 수 있다고 가정한다. 이 기법은 공격자의 근원지를 역추적하기 위해 패킷헤더의 IP 식별자 필드 16 비트를 이용하여 패킷이 경유하는 라우터의 정보를 마킹한다.

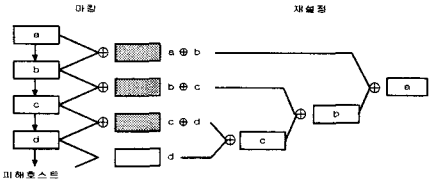
가) Advanced Marking Scheme I

이 기법은 [그림 7]에서 나타나듯이 IP 식별자 필드 16 비트를 마킹 필드로 사용하기 위해 피해 호스트로부터 라우터의 거리를 나타내는 5 비트의 거리필드와 11 비트의 edge 필드로 구분한다. 5 비트의 거리필드는 32 홉을 표시할 수 있으므로 인터넷 경로들을 충분히 나타낼 수 있다. 패킷의 IP 식별자 필드에 32 비트의 라우터 주소를 hash 함수를 이용하여 11 비트로 암호화한 후 마킹하게 된다. 그 후 각 라우터를 경유할 때마다 XOR 연산을 통해 라우터의 정보를 암호화하여 마킹하게 되고, 공격경로를 재설정하게 된다.



[그림 7] Advanced Marking Scheme I 인코딩 구조

거리필드는 각 라우터가 확률 p 로 패킷을 마킹할 때 마킹된 라우터와의 거리가 된다. 즉, 자신이 마킹되면 그 라우터의 거리 값은 0 이 된다. edge 필드에는 32 비트의 IP 주소가 hash 함수를 이용하여 나온 11 bit 의 결과값으로 저장되는데, 먼저 확률 p 로 마킹이 결정된 한 라우터 *Ri* 가 있으면 edge 필드에 IP 주소를 hash 함수를 이용하여 나온 결과값 *h(Ri)*를 넣고 거리필드에 0 을 기록한다. 거리필드의 값이 0 이면 이전 라우터가 마킹을 한 것으로 간주하여 edge 필드에 upstream 라우터상에서 다음 라우터의 IP 주소 *h'(Ri)* 와 XOR 한 값을 덮어씌운다. 지나온 상위 경로를 알아내는 재설정절차는 $z = x\ h'(y)$로서 y 는 upstream 상의 한 라우터이며 x 는 *h(Ri)*의 hash 값을 나타낸다. 재설정절차는 $z = x\ XOR\ h'(y)$로서 y 는 upstream 상의 한 라우터이며 x 는 *h(Ri)*의 hash 값을 나타낸다. 재설정은 a XOR b XOR a = b 의 공식에 의해 마킹된 edge 필드를 피해호스트로부터 경로를 따라 XOR 연산을 이용하면 최종적으로 공격자 위치를 역추적 할 수 있음을 보여준다. [그림 8]은 XOR 연산을 이용한 마킹절차와 경로 재설정절차를 나타낸다.

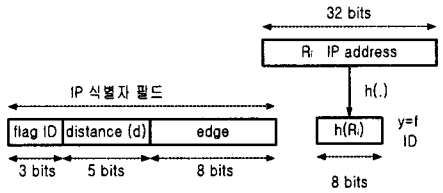


[그림 8] XOR 연산을 이용한 마킹과 재설정

나) Advanced Marking Scheme II

Advanced Marking Scheme I 이 기존의 FMS(Fragment Marking Scheme)에 비해 DDoS 공격에 많이 강해지기는 했지만 60 개 이상의 분산 공격이 있고, 공격 그래프상에서 거리 값이 같은 많은 라우터가 있으면 11 비트의 hash 값에서 충돌이 일어나는 문제점을 가지고 있다. 그래서 Advanced Marking Scheme I 를 확장한 것이 Advanced Marking Scheme II 이다. 이 기법은 2 개의 hash 함수를 사용하는 대신 독립적인 hash 함수 세트를 이용하는 것이다. Advanced Marking Scheme II 는 [그림 9]에서 보여주듯이 IP 헤더의 식별자 필드의 16 비트를 마킹할 필드로 사용하는 것은 동일하지만 flag ID 라는 필드를 추가하였다. 거리 필드는 32 홉을 표현할 수 있는 5 비트를 그대로 사용하며 edge 필드는 flag ID 필드를 뺀 8 비트가 사용된다. [그림 9]는

Advanced Marking Scheme II 인코딩 구조를 나타낸다.



[그림 9] Advanced Marking Scheme II 인코딩 구조

주어진 flag ID 를 가지고 라우터 *Ri* 를 인코딩 하는 hash 함수는 h(<*f*ID,R*i*>)로 구할 수 있으며, 8 비트 edge 필드에 저장된다. 한 라우터 *Ri* 에 패킷마킹이 결정되면 라우터는 랜덤하게 선택되는 숫자와 그 라우터의 IP 주소 값으로 hash 값을 구하게 되는데, 그때 사용된 숫자를 알리기 위해 flag 필드를 사용한다. Advanced Marking Scheme II 는 flag ID 가 추가되어 1500 개의 분산공격에서도 견딜 수 있게 강화되었으며, 그 외의 것은 Advanced Marking Scheme I 과 비슷하다. 인코딩 부분은 *edge = edge h(<fID, Ri>)*로 edge 필드에 덮어쓰기를 하며 경로상의 라우터들을 디코딩하여 공격자 위치를 역추적 할 수 있다. 하지만 Advanced Marking Scheme 에서의 마킹은 인증을 거치지 않아 마킹 위조에 대한 위험은 Advanced Marking Scheme I, II 모두 가지고 있는 것이 단점으로 지적되고 있다.

Advanced Marking Scheme II 마킹 알고리즘은 hash 함수를 이용하여 암호화한 결과값들이 동일한 거리상에서 중복됨을 피하기 위해 P.fID 필드가 추가되었다. P.fID 필드는 3 비트를 사용하기 때문에 0 에서 7 까지의 8 가지의 hash 함수가 표시될 수 있다. 기본적인 마킹절차는 Advanced Marking Scheme I 과 유사하다.

```
Advanced Marking Scheme II at router Ri:
  for each packet P
    let v be a random number from [0,1)
    if  v ≤ p then
        let x be a random integer from [0,7)
        P.fID ← x
        P.distance ← 0
        P.edge ← g(<x, Ri>)
    else
        if(P.distance == 0) then
          P.edge ←  P.edge g'(<P.fID, Ri>)
  P.distance ←  P.distance + 1
```

Advanced Marking Scheme II 재설정 알고리즘은 마킹된 패킷을 받은 피해호스트에서는 라우터 맵과 z=x g'(<l,y>)의 연산으로 g(<l,y>)와 z 값의 비교를 통해 패킷이 마킹되어진 경로를 역추적하는 기법을 나타내고 있다. 여기서 y 는 거리가 d 홉 떨어진 라우터를 의미하며, l 은 hash 함수의 종류를 나타낸다.

```
Reconstruction procedure at victim v:
  let Sd be empty for 0 ≤ ←  d ≤ maxd
  for each child R of v in Gm
  let count = 0
    for l = 0 to 2^w - 1
        if g(<l, R>) ∈  £0,1 then
          count = count +1
  if count > m then
    insert R into S0
for d:= 0 to maxd - 1
  for each y in Sd
    for each child u of y in Gm
      let count = 0
      for l:= o to 2^w - 1
        for each x in  £d+1,l
          z = x XOR g'(<l, y>)
          if g(<l, u>) = z then
            count = count + 1; break
    if count > m then
      insert u into Sd+1
  output Sd  for  0 ≤ d ≤ maxd
```

다) Authenticated Marking Scheme

Advanced Marking Scheme 은 인증되지 않은 마킹이므로 신뢰할 수 없는 라우터가 그것을 위조할 위험이 높다. 이러한 문제를 해결하기 위해 디지털 서명이 사용되지만 비용과 오버헤드의 문제로 MAC(Message Authentication Code)을 사용한다. MAC 의 종류 중 하나인 HMAC-MD5 는 디지털 서명에서 사용하는 1024-bit RSA 보다 1,000~10,000 배정도 빠르며 기존의 암호화에서는 128 비트가 필요한 반면에 16 비트만 있으면 되기 때문에 오버헤드를 줄일 수 있는 장점이 있다.

4. 결론

본 논문에서는 해커의 위치를 속이고 공격하는 DDoS 공격에 대해서 실제 위치를 찾을 수 있도록 라우터에서 패킷에 특정 정보를 삽입하는 IP 패킷 마킹 역추적 기술에 대해서 살펴 보았다. 현재 해커의 실제 위치를 찾을 수 있는 기법들이 실험 환경에서 많이 연구되고 있으나, 네트워크 노드 변경등과 같은 많은 변수가 존재하기 때문에 현 인터넷 망에 적용하기에는 매우 어려움이 따른다. 또한, 위에서 언급한 역추적 기술들이 현재 망에 적용되더라도 DDoS 공격 시, 해커의 실제 위치를 찾기가 불가능하다. 따라서, 현재 망에 손쉽게 적용할 수 있는 IP 역추적 기술 개발이 필요하며, 또한 IP 역추적 기술과 TCP 역추적 기술에 대한 통합된 형태의 새로운 역추적 기술에 대한 연구가 필요하다고 하겠다.

참고문헌

[1] Chun He, "Formal Specifications of Traceback Marking Protocols", 2002.

[2] S. Savage, D. Wetherall, A. karlin, and T. Anderson, "Network Support for IP Traceback", IEEE/ACM transactions on networking, vol. 9, No. 3, June 2001.

[3] R. Stone, "CenterTrack: An IP overlay network for tracing DoS floods", in Proc, 2000 USENIX Security Symp., July 2000, pp. 199-212.

[4] D. Song and A. Perrig, "Advanced and authenticated marking schemes for IP Traceback", in Proc. IEEE INFOCOM, vol. 2, April 2001, pp. 878-886.

[5] H. Burch and B. Cheswick, "Tracing Anonymous Packets to Their Approximate Source", in Proc. 2000 USENIX LISA Conf. December 2000, pp319-327