익명성을 제공하는 P2P 시스템 비교 분석[1)]

*김현철 ** 김범한 ***이동훈
고려대학교, 정보경영공학전문대학원
{*hckim, **anewholic}@cist.korea.ac.kr ***donghlee@korea.ac.kr

Analysis of Anonymous P2P System

*Kim, Hyun-Cheol **Kim, Bum-Han ***Lee, Dong-Hoon
Graduate school of Information Management & Security, Korea Univ.

요약

Peer-to-Peer 네트워크는 사용자들이 컨텐츠를 공유하고 배포하기 위한 매우 유용하고 널리 알려진 클라이언트-클라이언트 네트워크이다. 근래에 다양한 방송 컨텐츠들이 생겨나고 광고 수익을 얻기 위한 기업들의 전략과 컨텐츠를 필요로 하는 사용자들의 요구사항이 맞아 떨어지면서 새로운 수익 구조가 발생하였다. 이러한 컨텐츠를 사용자들이 공유하고 배포하기 위한 수단으로 P2P 네트워크가 가장 효과적인 방법이다. P2P 네트워크에서는 사용자들의 정보가 드러나지 않도록 익명성을 제공하는 것이 관건이다. 본 논문에서는 서로 다른 접근을 통해 익명성을 제공하는 P2P 시스템을 비교 분석하고 P2P 네트워크에서 가능한 공격들을 다루고 이를 막기 위한 대응 방법을 알아본다.

1. 서론

과거에는 컨텐츠 제작자와 제공자들이 제작된 컨텐츠를 보호하고 일정 권한을 가진 사람에게만 컨텐츠를 제공하고자 했다. 제작된 컨텐츠를 사용자들에게 팔아서 수익을 얻는 단순한 구조를 취하고 있었던 것이다. 하지만 점차 제작된 컨텐츠를 보호하는데 비용이 많이 들게 되고 기존의 수익 구조로는 큰 이익을 얻기가 힘들게 되었다. 또한 기업들은 새로운 제품이나 기술을 홍보하기 위한 광고를 하는데 비용 부담이 있었다. 사용자들은 돈을 지불해야하는 컨텐츠에는 접근을 꺼려 결과적으로 컨텐츠 배포자가 수익을 얻는데 어려움을 격게 되었다. 이러한 문제를 해결하기 위해 기업의 광고와 컨텐츠가 합쳐져 무료로 사용자들에게 제공되어 질수 있게 되었다. 이러한 컨텐츠를 사용자들이 쉽고 빠르게 공유 할 수 있도록 하는 방법으로 P2P 시스템이 가장 효과적이다.

P2P는 Peer들 사이에 데이터를 공유하고 배포하기 위한 가장 유용한 방법이다. 각 Peer들은 원하는 데이터를 찾아서 공유할 수 있고 그 과정이 복잡하지 않아서 널리 사용할 수 있었다. 하지만 점차 데이터 제공자의 프라이버시와 데이터 요청자의 프라이버시가 필요해지고 프라이버시가 제공되지 않는 곳에서는 Peer들 사이에 데이터 공유가 점차 줄어들 가능성이 있다. 따라서 효과적인 P2P 서비스를 하기 위해서는 데이터 제공자의 프라이버시, 데이터 요청자의 프라이버시, 데이터 기밀성이 필요하다. 본 논문에서는 이러한 요구조건을 만족하는 효과적인 P2P 시스템을 알아본다.

2. P2P 시스템 요구조건

P2P 시스템을 위한 요구조건으로 크게 데이터 제공자 익명성, 데이터 요청자 익명성, 데이터 기밀성을 들 수 있다. 다음 요구조건을 만족하게 위해서 P2P 시스템에서 가능한 방법에 대해서 알아보도록 하겠다.

가. 데이터 제공자 익명성 및 데이터 요청자 익명성

데이터 제공자 익명성은 데이터 요청자가 데이터를 찾을 때 데이터 제공자의 개인정보가 노출 되지 않아야 하고 특히 IP 주소가 드러나지 않아야 한다. 데이터 제공자의 IP주소를 드러내지 않으면서 데이터를 제공받는 것이 핵심이다. 데이터 요청자 익명성은 데이터를 요청자의 개인정보가 노출되지 않아야 한다는 요구조건이다. 데이터 제공자와 데이터 요청자의 익명성을 제공하면서 데이터를 공유하는 방식으로 크게 세 가지 방법이 있다.

먼저 Mix-net[1]을 이용하는 방식으로 가장 고전적인 방법이다. Mix-net은 여러 Peer들 중에서 데이터를 전송하기를 원하는 Peer를 숨기기 위해서 데이터와 관계없는 여러 Peer들을 이용한다. 여러 Peer들중에 하나의 Peer가 데이터를 전송하고자 하면 전체 Peer들이 참가하게 되므로 결과적으로 누가 데이터를 보내왔는지 알 수 없게 된다. 또한 데이터 제공자도 똑같은 방식으로 데이터를 제공하기 때문에 데이터 제공자도 익명성이 제공한다. 이 방법은 네트워크에 참가하고 있는 Peer들에게 큰 부담을 주므로 바람직한 방법은 아니다. 하지만 Mix-net에서 알 수 있듯이 데이터가 여러 Peer를 거치게 되면 과연 누가 최초의 데이터 제공자인지 알 수 없게 되고 결국 데이터 제공자

1) "본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음" (IITA-2008-(C1090-0801-0025))

의 익명성을 제공할 수 있다는 것을 알 수 있다. 특히 Peer들이 많고 여러번 거칠수록 확실한 익명성이 제공될 수 있다. 하지만 그렇게 되면 부담이 많고 데이터 요청자가 데이터를 받는데 시간이 오래 걸리게 된다.

두 번째로 데이터 요청자가 데이터를 전달받을 Path를 선택해서 요청 데이터와 함께 쿼리를 던지는 방법이다. 데이터를 전달받을 Path는 Onion 구조로 되어 있다. 데이터 제공자는 Path를 보고 데이터를 전달한다. 데이터를 받은 Peer는 자신의 개인키로 Onion을 복호화 하고 나머지를 Path에 있는 다음 Peer에게 데이터를 전달한다. 이 방법은 데이터 제공자가 누구인지 모르지만 자신이 설정한 Path를 따라서 데이터를 전달하도록 함으로서 데이터 제공자의 익명성을 제공하는 방법이다. 하지만 데이터 요청자가 정한 리턴Path에서 가장 첫 Peer는 데이터 제공자가 누구인지 알 수 있고 첫 Peer가 mailicious Peer면 데이터 제공자에 대한 정보를 노출할 수 있다. 또 애초에 데이터 요청자가 자신과 사전에 공모한 Peer를 리턴Path에 첫 번째 Peer로 정할 수도 있다. 이렇게 되면 데이터 제공자에 대한 익명성이 제공되지 못하게 된다. 리턴 Path에 있는 Peer들중에서 리턴이 끝나기 전에 Peer가 로그아웃하면 데이터가 정상적으로 리턴되지 않는 문제점도 있다.

데이터 요청자는 여러 peer들 중에 하나를 랜덤하게 voluntary peer로 정해서 해당 peer가 데이터를 요청하도록 한다. 데이터 요청자는 중간 voluntary peer까지 랜덤하게 정해진 Path로 요청 데이터를 전달한다. 그 후에 중간 voluntary peer가 데이터를 요청하는 패킷을 보내게 되고 외부에서 봤을 때 누가 실제 데이터 요청자인지 모르게 된다. 이 방법은 중간 voluntary peer에게 큰 부담을 준다. 또한 랜덤하게 voluntary peer가 설정되므로 데이터 요청자는 누가 voluntary peer가 될지 예상할 수 없다.

세 번째로 브로드 캐스트 방법이 있다. 이 방법은 데이터 요청자가 요청 데이터와 리턴 받을 자신의 정보를 pseudonym으로 정하여 하여 브로드 캐스팅 하고 데이터 제공자도 자신의 정보를 pseudonym으로 정하고 데이터와 함께 브로드 캐스트 하는 방법이다. 모든 데이터 패킷은 브로드 캐스트 되기 때문에 데이터 제공자가 누구이고 어디로 가는지, 데이터 요청자가 누구인지 알 수 없다. 네트워크에 참가한 Peer가 적다면 약한 익명성이 제공되겠지만 어느 정도 숫자가 되는 상황에서는 강한 익명성을 제공한다. 이 방법은 Onion 구조처럼 암호화와 복호화를 반복해서 하지 않기 때문에 각 peer들에 부담이 덜하고 단순한 구조로 익명성을 제공할 수 있다. 하지만 브로드 캐스팅은 전체 네트워크에 부하를 일으킬 수 있고 브로드 캐스트 스톰(Broadcast Storm)과 같은 문제를 일으킬 수 있다. 브로드 캐스트 방법을 이용하기 위해서 브로드 캐스트 횟수를 줄이면서 강한 익명성을 제공하는 해결책이 제시 되어야 한다.

나. 데이터 기밀성

데이터 기밀성을 제공하기 위해서는 기본적으로 암호화를 사용한다. P2P 시스템에서는 어떠한 암호 시스템을 사용하고 어떠한 방법으로 암호 시스템을 사용하는지 알아본다.

기밀성을 제공하기 위한 방법으로 데이터 요청자와 데이터 제공자사이에 세션키를 공유하여 데이터를 전송할 때 데이터를 세션키로 암호화하는 것이 가장 일반적인 방법이다. 이 방법은 가장 단순하고 효과적인 방법이나 세션키를 안전하게 전달하는 것이 핵심이다. 대부분 P2P 시스템들이 이 방법을 사용하고 있고 안전하게 세션키를 전달하기 위해서 여러 방법을 사용하고 있다. 데이터 요청자가 요청하는 데이터 쿼리에 대해서는 암호화를 하기도 하고 안하기도 한다. 익명성이 제공되면 데이터 요청자와 데이터 쿼리가 매칭되지 않기 때문에 데이터 쿼리에 대한 암호화는 필요하지 않기 때문이다. 하지만 경우에 따라서는 데이터 쿼리도 암호화가 필요하므로 시스템에 따라 적절한 결정이 필요하다.

3. 익명성을 제공하는 P2P 시스템 비교

익명성을 제공하는 P2P 시스템을 크게 암호화/ 복호화 기반의 시스템과 브로드 캐스팅 방법을 이용한 시스템으로 나누어 비교해본다. 암호화/복호화 기반의 시스템은 Han[2]의 스킴을 선택하였고 브로드 캐스팅 방법은 Chou[3]의 스킴을 선택하였다.

가. Han[2]의 스킴

본 스킴에서는 데이터 요청자의 익명성을 제공하기 위해서 중간 voluntary peer를 선택하는 방법을 사용하였다. 데이터 요청자는 암호화된 요청 쿼리와 리턴Path를 한쪽 peer로 보내고 암호화된 요청 쿼리를 복호화 할 수 있는 키를 다른 peer로 보낸다. 데이터를 받은 Peer는 데이터를 로컬 스토리지에 저장하고 다음 peer에게 데이터를 전달한다. 데이터를 가지고 있는 Peer가 키를 받으면 데이터를 복호화하고 복호화된 요청 쿼리가 의미있는 데이터이면 그 peer는 voluntary peer가 된다. voluntary peer는 요청 쿼리를 복호화 하고 요청쿼리를 flooding 한다. 랜덤하게 voluntary peer가 선택되었으므로 누가 데이터 요청자인지 알 수 없게 된다.

데이터 제공자는 데이터 요청 쿼리를 받으면 데이터를 랜덤Path에서 첫 peer에게 데이터를 전달한다. 데이터를 받은 peer는 데이터를 받아서 랜덤Path를 자신의 키로 복호화하고 랜덤Path에 지정된 다음 peer에게 데이터를 전달한다. 데이터는 여러 peer들을 거쳐서 전달되므로 누가 데이터 제공자인지 알 수 없게 된다.

데이터 제공자는 데이터를 제공할 때 자신이 생성한 세션키로 데이터를 암호화하고 받은 요청 쿼리에 있던 데이터 요청자 공개키를 이용해 세션키를 암호화 하여 데이터와 함께 보낸다. 즉, 데이터를 세션키로 암호화 되었고 세션키를 데이터 제공자와 데이터 요청자만 알 수 있게 되는 것이다.

나. Chou[3]의 스킴

본 스킴에서는 브로드 캐스팅 방법을 통해 데이터 요청자 익명성과 데이터 제공자 익명성을 제공한다. 데이터 요청자는 데이터 요청 쿼리를 브로드 캐스팅하고 이 데이터를 받은 peer들도 브로드 캐스팅 한다. 따라서 누가 데이터 요청자인지 알 수 없게 되고 데이터 제공자도 똑같은 방식으로 데이터를 제공하므로 익명성을 제공 받는다.

본 스킴은 두 단계로 이루어져 있다. 첫 번째 단계를 통해서 데이터 요청자와 데이터 제공자 사이에 세션키를 공유하고 두 번째 단계에서 데이터 제공자가 세션키로 데이터를 암호화하여 데이터를 보내어 데이터 기밀성을 제공한다.

또한 본 스킴에서는 브로드 캐스트를 줄일수 있는 방법을 제시하였다. 데이터 제공자가 정한 파라미터를 이용하여 중간 peer들이 자신이 브로드 캐스트를 할지 안할지 정하게 되고 브로드 캐스트 스톰을 막을수 있는 장치를 마련한 것이다. 브로드 캐스트 횟수를 줄이는 것은

익명성 제공과 trade off 관계에 있으므로 적절하게 정하여 사용해야 한다.

특히 본 스킴은 모바일 ad-hoc 네트워크 환경을 위한 P2P 시스템이므로 peer들의 참가와 탈퇴가 자유롭고 네트워크 토폴로지가 가변적인 환경에서 적합한 방법이다.

다. Hop by Hop 암호화/복호화 방식과 브로드캐스팅 방식 비교

- Hop by Hop 암호화/복호화 방식 :
  1) 암호화와 복호화를 위한 계산 오버헤드 발생
  2) peer들을 거치기 때문에 전송 실패나 peer들의 공격에 취약함
  3) peer가 패킷을 전달하지 않는 경우 대응 방법이 없음
  4) 대역폭을 효과적으로 사용할 수 있음

- 브로드 캐스팅 방식:
  1) 많은 패킷이 발생
  2) 암호화 복호화 계산이 적음
  3) 데이터 전송 실패나 peer의 공격에 강함
  4) 대역폭이 효과적이지 않음

4. P2P 시스템 공격방법

P2P 시스템 공격방법은 트래픽 분석 공격으로 이루어 지는데 분류하면 타이밍 상관관계 공격[5]과 컨텐츠 상관관계 분석 공격방법으로 나누어 볼 수 있다.

가. 타이밍 상관관계 분석 공격

타이밍 분석 방법은 공격자가 특정 지역을 모니터링 하고 테스트 메시지를 보내어 메시지의 경로를 파악하고자 메시지를 추적하는 것이다. 공격자는 자신이 보낸 메시지가 첫 peer 에게서 처리 된후 다시 전송되면 다음 peer들중 어느 peer들이 반응해서 다시 데이터를 처리하는지를 분석하여 메시지의 경로를 파악하는 방법이다. 이는 메시지가 peer에가 도달하면 peer가 일정시간후에 메시지를 다시 전송하는 것을 이용한 방법이다. 타이밍 공격을 막기 위해서는 특정 트래픽 패턴을 없애기 위해서 랜덤하게 메시지를 내보내도록 해야 한다. Min-net이 기본적인 해결방법이 될 수 있다. 공격자는 특정 지역을 전부 모니터링 하고 있어야 하므로 공격자에게 적지 않은 능력이 있어야 한다.

나. 컨텐츠 상관관계 분석 공격

컨텐츠 상관관계 분석 공격은 컨텐츠를 식별할 수 있는 특정 데이터나 변하지 않는 데이터의 길이를 이용해 메시지의 경로를 파악하는 방법이다. 공격자는 peer가 입력 받고 출력하는 메시지를 다 알아야 하므로 공격자는 강한 능력을 가지고 있어야 한다. 하지만 데이터 제공자가 보내는 메시지의 경로가 짧다면 공격자는 작은 지역을 모니터링 해도 메시지의 경로를 알아낼 수 있다. 컨텐츠 상관관계 분석 공격을 막기 위해서 메시지에 Random Nonce를 사용하거나 랜덤하게 패딩을 하여 처리할 수 있다. 하지만 이는 암호화/복호화 계산을 늘어나게 하고 전체적인 성능을 저하 시킨다.

5. 결론

P2P 시스템은 사용자간에 데이터를 공유할 수 있는 가장 유용한 방법이다. 실제 NGN2005(2005년 9월)에서는 2003년을 기점으로 인터넷에서 가장 많은 트래픽을 P2P 서비스가 차지하고 있다는 통계를 발표하였고, 같은 해 Network World 지에서도 인터넷 트래픽의 60%~89% 정도의 트래픽을 P2P 서비스가 차지하고 있다고 게재된 바 있다. 과거 P2P 서비스로는 Instant messaging, File sharing, Distributed computing 정도를 떠올렸으나, 지금은 VoIP, Video Streaming, Multicast, Web Caching, Game, Virtual office 등 다양한 분야에 응용되고 있으며, 저작권문제나 정보보호 문제와 같은 P2P 서비스의 부작용을 최소화 할 수 있는 표준화 분야 등에 관심이 집중되고 있다[4].

방송통신관련 컨텐츠는 P2P 시스템과 결합되어 새로운 수익을 창출하는 상품으로 나타날 수 있다. 사용자들이 간편하고 쉽게 컨텐츠에 접근할 수 있다는 강점을 가지고 있기 때문이다. 이러한 상품이 등장했을 때 가장 중요하게 생각되는 것은 사용자의 익명성과 프라이버시 문제일 것이다. 사용자의 익명성이 제공되지 않으면 사용자들은 접근을 꺼려 할 것이고 컨텐츠 사용이 저조하게 될 것이다. 따라서 사용자 익명성이 고려된 P2P 시스템을 이용하여 컨텐츠를 배포할 때 사용자들도 안심하고 컨텐츠를 사용할 수 있을 것이다.

6. 참고문헌

[1] D. L. Chaum, "Untraceable elcetronic main, return addresses, and digital pseudonyms", Commun. ACM, Vol. 24, No. 2, pp.84-90, 1981

[2] Jinsong Han, Yunhao Liu, Li Lu, Lei Hu, Abhishek Patil , "A Random Walk Based Anonymous Peer-to-Peer Protocol Design" , ICCNMC, LNCS, 3619, pp.143-152, 2005

[3] Chao-Chin Chou, David S. L. Wei, C.-C. Jay Kuo, Kshirasagar Kaci, "An Efficient Anonymous Communication Protocol for Peer-to-Peer Applications over Mobile Ad-hoc Networks", JSAC, IEEE, Vol.25, No.1, 2007

[4] 나재훈, "P2P 보안 표준화 동향", 한국정보통신기술협회, http://weekly.tta.or.kr/

[5] J.-F. Raymond, "Traffic analysis: protocol, attacks, design issuses, and open problems," in Proc. International workshop on Designning privacy enhancing technologies, LNCS, pp.10-29, 2001