# 믹스 시스템에서의 키 변경

이주형*, 홍만표

*아주대학교 정보통신전문대학원

# Key Change in the Mix System

Ju-Hyeong Yi*, Man-Pyo Hong

*Graduate School of information and communication, Ajou University.

**요 약**

익명통신을 위한 믹스 시스템에서 공격의 위험에 대응하기 위하여 효과적인 키 관리가 중요하다. 공개키 암호 시스템을 사용할 경우, 믹스 노드의 비밀 키가 노출된 것으로 의심됐을 때, 또는 노출의 예방을 위해서 주기적으로 키를 변경할 필요가 있다. 믹스 시스템의 특성상 메시지의 송수신 사이에는 전송딜레이 이상의 딜레이 시간을 갖게 되고, 키 변경 이후에 일정 시간 동안에는 예전의 키를 적용한 메시지와 새로운 키를 적용한 메시지가 동시에 나타날 수 있다. 이러한 키 교차 시간을 충분히 가질 경우 공격에 노출될 위험이 커지고, 반대로 짧은 교차 시간을 갖는 것은 정상적인 메시지의 드랍을 증가시킨다. 믹스 시스템의 환경을 고려하여, 메시지의 평균적인 딜레이 시간을 알아내고 적당한 교차 시간을 적용해 볼 수 있겠다.

## I. 서론

인터넷은 개방된 네트워크로서 디자인 되었으므로 같은 랜 안에 있는 컴퓨터는 다른 사람의 트래픽을 볼 수가 있고, 네트워크의 라우터들 또한 자신을 경유하는 모든 트래픽을 볼 수 있다. 패킷 내에 있는 라우팅 정보 또한 개방적이므로 IP 패킷의 헤더를 읽음으로써 송신자와 수신자를 인식할 수 있다. 따라서 수동적인 관찰자의 입장만으로도 누가 누구와 통신하고 있는지 쉽게 알아낼 수 있는 것이다.

메시지에 대한 암호화는 메시지의 내용은 숨길 수 있으나, 통신 주체에 대한 신원(identity) 정보는 숨길 수 없다. 패킷의 페이로드(payload)만이 암호화 될 뿐, 라우팅 정보는 여전히 공개적이기 때문이다.

따라서 통신 중에 사용자의 신원 정보를 보호하는 것은 암호화가 아닌 다른 방법으로 해결될 수 있으며, 이러한 프라이버시 보장 문제를 해결하기 위한 대안으로서 익명통신이 연구되고 있다. 익명통신이 적용될 수 있는 분야로는 전자메일에서 발신자를 추적하지 못하게 하거나, 신분을 위장한 범죄수사, 구매자의 신분을 숨길 수 있는 전자 화폐, 익명성이 보장되는 전자 투표, 검열을 피할 수 있는 출판등이 있다.

익명성(Anonymity)이란 통신 주체들의 집합 내에서 구별되지 않을 수 있는 상태를 말한다. 행위와 행위의 주체, 예를 들어 이메일과 이메일을 보내는 사람이 서로 연결되지 않는 것은 연결불가능성(Unlinkability) 이며 익명성보다 한 단계 위의 상태이고, 행위가 일어나는 것조차 알아채지 못하는 것은 관찰불가능성(Unobservability) 이며 실현하기 가장 어려운 상태이다.

이러한 익명통신을 위한 시스템 중 대표적인 것으로 믹스넷(Mix-net), 크라우드(Crowds), 토르(Tor), 어니언라우팅(Onion Routing), DC넷(DC-net), 셔플넷(Shuffle-net)등이 있다. 여기

에서는 믹스넷을 실제 활용할 경우 고려해볼 만한 문제로서, 공개키 암호화 시스템의 키 분배 문제를 다뤄보려고 한다.

## II. 관련 연구

믹스넷은 연결불가능성을 통해 익명성을 실현하는 것이다. 송신자 S 는 수신자 R 에게 익명의 메시지를 보내기 위해 한 개 이상의 믹스 노드를 거치도록 한다. 그림 1 과 같이 메시지가 M1 과 M2 의 노드를 지나게 될 경우, S 는 R 의 공개키로 메시지를 암호화하고, 이것을 다시 M2 의 공개키로 암호화하며, 다시 M1 의 공개키로 암호화한다. 이렇게 여러 겹으로 암호화된 메시지를 M1 이 받으면 자신의 비밀키로 복호화하고 라우팅 정보를 읽어 다음 노드인 M2 로 메시지를 보낸다. M2 또한 복호화 과정을 거친 후 최종 수신자인 R 에게 메시지를 보내게 되고, R 은 안전하게 메시지를 읽을 수 있다. 이때 M1 과 M2 는 메시지의 내용을 알 수 없을 뿐만 아니라, 송신자와 수신자를 연결해 낼 수 없는 것으로 메시지에 대한 기밀성과 사용자에 대한 익명성을 보장할 수 있다.

## III. 문제 제기

### 3.1 키의 지속 시간(duration time)

대부분의 믹스 시스템은 암호화를 지원하고 있고, 여러 가지 공격에 대응하기 위한 방법으로 이러한 키 셋을 주기적으로 바꾸어 사용하도록 한다. 예를 들어 믹스 시스템에 복사된 메시지를 다시 보내봄으로써 메시지의 불연결성을 깨뜨리는 replay attack 에 대하여, 일정 시간마다 믹스의 비밀키를 바꾸게 되면 똑같은 메시지가 도착한다 하더라도 믹스는 메시지를 처리할 수 없으므로 공격당하지 않을 수 있다.

믹스 시스템의 특성 상 메시지는 전송 중에 여러 믹스 노드를 거치게 되고, 배칭 방법(batching strategy) 에 따라 각 노드에서 일정 시간의 딜레이 시간을 갖게 된다. 그러므로 키를 바꾼 믹스 노드가 새로운 키 정보를 사용자들에게 알려주어도, 네트워크 내에는 아직 예전의 키 정보를 가지고 노드를 이동하는 메시지가 존재하게 된다. 따라서 키를 변경한 믹스 노드는 일정 시간 동안 예전 키와 변경된 키를 동시에 가지고 복호화할 필요가 있다.

두 개의 키가 적용되는 지속시간 Dt 을 충분히 주는 것은 공격에 대한 약점이 된다. 예전 키로 암호화된 모든 메시지를 놓치지 않고 처리할 수는 있겠으나, 한편으로는 예전 키가 사용된 공격자의 메시지를 통과시킴으로써 공격에 노출될 위험이 커지는 것이다. 또한 지속시간 Dt 을 짧게 두는 것 공격에 관련된 메시지뿐만 아니라 많은 정상적인 메시지까지 잃어버리게 된다. 믹스 시스템은 메시지 수신에 대한 확인절차(ACK)가 없으므로 전달을 실패하는 것에 대해 많은 부담을 가진다.

### 3.2 적절한 지속시간 Dt 의 필요성

지속시간 Dt 를 적절하게 적용하는 것은 시스템의 안정성(과 신뢰성에 있어 꼭 필요한 문제이다.

기존의 믹스 타입 중 하나인 믹스미니언(Mixminion)에서는 일정 시간 간격으로 키를 바꾸어 사용하는 것을 권고하였으나, 변경된 이후에 키를 적용하는 문제에 대해서는 언급이 없었다. 또한 지금까지 익명 시스템에서 키 사용에 관련한 이슈는 안전하게 키 셋을 전달하는 방법에 관한 것이 대부분이었다. 그러나 실제로 익명통신에 적용하기 위해 믹스 시스템을

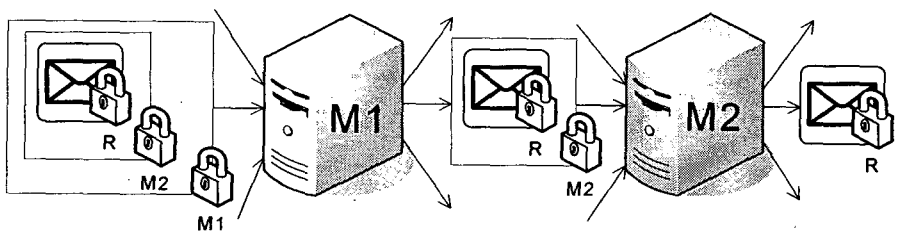


그림 1 믹스 노드를 통과하는 메시지

구현하는데 있어, 키 변경 시 두 가지 키를 어떻게 사용하는지는 전체 시스템의 성능에 큰 영향을 미치는 문제가 될 수 있다.

3.3 해결 방법

메시지가 믹스넷을 통과하는데 드는 평균적인 시간은 지속시간 Dt 에 가장 큰 영향을 준다. 이러한 메시지의 평균 딜레이 시간은 믹스 노드 사이를 이동하는 전송 딜레이(propagation delay)와 믹스 노드 내에서 머무르는 리오더링 딜레이(reordering delay)로 이루어져 있다. 전송 딜레이가 통신 환경에 의존적인 것에 비해, 리오더링 딜레이는 리오더링 기법에 따라 달라지며, 확률적인 성격을 많이 가진다.

따라서,

- 두 가지 키의 지속시간을 결정하기 위하여 메시지들의 딜레이를 예측해보고, 믹스넷의 환경 변수들과 딜레이와의 관계를 알아본다.
- 딜레이의 추이를 분석하고, 시스템의 안전성과 신뢰성을 동시에 고려하여 키 지속시간 Dt 결정하는 방법을 알아본다.

## IV. 시뮬레이션

4.1 시뮬레이션 환경에 대한 설명

시스템의 사용자들은 단위 시간당 m 개의 메시지를 발생시키고, 각 메시지들은 l 의 path 길이, 즉 l 개의 믹스 노드를 지난다. 메시지들이 사용자와 믹스, 믹스와 믹스를 이동하는데 있어 전송 딜레이는 고려하지 않으며, 믹스 노드 내에서 복호화에 필요한 딜레이 또한 고려하지 않는다.

믹스에서 메시지를 내보내는 배칭 방법(리오더링)은 일반적으로 시간을 기준으로 하는 timed mix, 메시지의 개수를 기준으로 하는 threshold mix, pool 을 사용하여 일정 수의 메시지를 남기도록 하는 방법을 조합한 threshold mix 와 timed mix, 네 가지를 볼 수 있다.

| | |
|---|---|
| n | fire 기준 메시지 수 (threshold) |
| m | 단위시간 당(라운드) 발생하는 메시지 수 |
| d | 메시지를 발생시키는 라운드 수 |
| l | 메시지의 path 길이 |
| k | 믹스넷 내에 존재하는 믹스 노드의 수 |
| n | 메시지를 모아두는 threshold |
| t | 메시지를 모아두는 시간 |
| p | 믹스 노드의 pool 크기 |

표 1. 믹스넷의 파라미터

- threshold mix : 믹스에 n 개의 메시지가 모이면 n 개 메시지 모두 내보낸다.
- timed mix : 매 t 시간 마다 모인 메시지를 모두 내보낸다.
- threshold pool mix : n + p 개의 메시지가 모이면, 그중 p 개를 임의로 뽑아 남겨두고 나머지 n 개만 내보낸다.
- timed pool mix : 매 t 시간 마다 임의로 뽑은 p 개의 메시지를 남겨두고 나머지를 내보낸다.

4.2 환경 분석을 통한 예상 딜레이 시간

위의 네 가지 믹스 형식에서 동작 방법을 기준으로 메시지가 믹스넷을 통과하는데 걸리는 딜레이를 표 2 와 같이 예상해 볼 수 있다.

| 배칭 방법 | min delay | max delay |
|---|---|---|
| threshold | 1 | infinite |
| timed | 1 | 1*t |
| threshold pool | 1 | infinite |
| timed pool | 1 | infinite |

표 2. 예상 딜레이 시간

threshold mix 에서 메시지 A가 믹스 노드에 들어갔다가 나오려면 n 개의 메시지 수가 채워져야 하는데, 만약 노드에 이미 n-1 개의 메시지들이 들어가 있었다면 A 는 바로 다음 라운드에 그 노드를 벗어나게 된다. 따라서 최소 딜레이는 경로 길이와 같아지는 것을 알 수 있다. 그러나 만약 노드가 메시지 n 개를 모두 내보낸 직후에 A 가 들어가게 된다면 A 는 나머지 n-1 개의 메시지가 도착할 때까지 기다려야 한다. 최악의 경우 A 가 들어있는 노드에 메시지들이 더 이상 충분히 들어오지 않을 수도 있다. 이렇게 되면 A 는 더 이상 그 노드를 벗어날 수 없으므로 최대 딜레이는 무한대가 된다.

timed mix 는 매 t 시간 마다 모여진 메시지들을 보내게 되는데 메시지 A 가 항상 t-1 시간

에 도착한다면 노드내에서 머무르는 시간이 최소화되고 최소 딜레이는 경로 길이와 같다. timed 는 메시지가 들어오는 개수에 상관없이 정해진 시간에 메시지를 내보내므로 최악의 경우라 하더라도 l*t 딜레이 시간을 넘지 않는다. threshold pool mix 는 메시지 A 가 노드에 항상 마지막으로 도착하고, 노드내의 p+n 개 메시지 중에서 내보낼 메시지 n 개를 고를 때 항상 선택된다면, 최소 딜레이는 경로길이가 된다. 그러나 만약 메시지 A 가 들어간 노드에 더 이상 메시지가 들어오지 않거나, 들어온다 하더라고 A 가 p+n 개중 n 개에 선택되지 않는다면 A 는 그 노드를 빠져나올 수가 없다.
timed pool mix 에서 또한 메시지 A 가 항상 t 의 마지막 시간에 도착하고, p+n 개중 항상 n개에 속한다면 최소 딜레이는 경로 길이가 된다. 그러나 만약 A 가 p+n 개중에서 n 개에 계속 속하지 않는다면 최대 딜레이는 무한대가 된다.

### 4.3 시뮬레이션 결과

위와 같이 딜레이에 대해 최대, 최소값을 예상해 보았으나 실제 키 변경 시 고려해야 하는 딜레이 값을 알아보기에는 부족하다. 따라서 믹스넷 환경을 가상으로 구현하고 딜레이를 측정하는 시뮬레이션을 구현하고 이를 통해 평균 딜레이 값과 환경 변수와의 관계를 알아보았다. 본 실험에서는 앞서 설명한 4가지 믹스 형태 중 첫 번째인 threshold mix를 이용하였다.

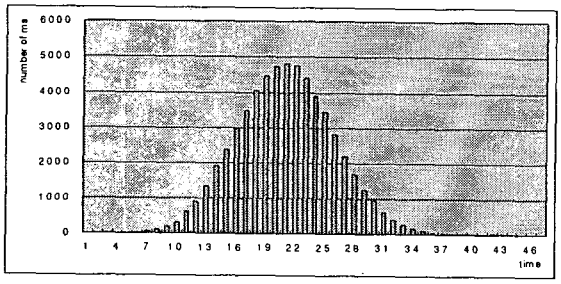


그림 2 . 시간당 도착 메시지 수

메시지들의 딜레이 시간을 재고 통계를 내보면 환경에 관계없이 공통적으로 그림 2와 그림3의 형태를 보인다. 그림 2에서 x 축은 딜레이 시간이고 y 축은 메시지의 개수를 나타낸다. 환경변수를 달리하여도 일반적으로 그래프들은 이와 같이 좌우가 비슷한 종모양의 분포를 나타낸다.

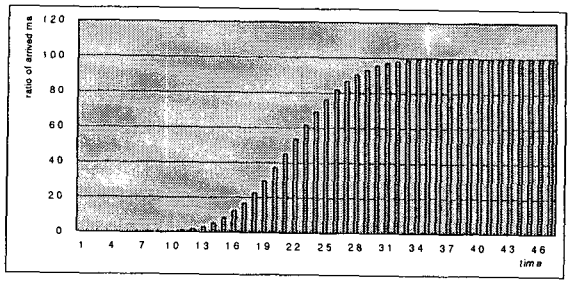


그림 3 . 시간당 메시지 도착 비율

그림 3은 각 시간까지 몇 개의 메시지가 도달했는지를 누적된 값으로 나타내는 그래프이다. x 축은 시간의 흐름을 나타내고, y 축은 전체 메시지를 1000으로 보았을 때 차지하는 비율을 나타낸 것이다. 딜레이 시간 23 이 지났을 때 1000 개중 600 개의 비율, 즉 60% 의 메시지가 도착했음을 알 수 있고, 딜레이 시간 28 정도까지가 되면 전체의 90 %가 도착했음을 알 수 있다. 그림 3의 그래프에서 x 축이 오른쪽으로 갈수록 들어오는 메시지의 개수는 증가하지만, 일정 시간 이후에는 증가율이 점차 작아지는 것을 볼 수 있다. 그래프의 변곡점이 증가율이 변하는 기준이 되며, 변곡점 전까지는 딜레이 시간을 늘림으로써 처리할 수 있는 메시지가 급격하게 증가하고 변곡점 이후에는 딜레이 시간의 증가에 비해 메시지의 증가가 아주 적다.
믹스가 딜레이 시간을 크게 허용하는 것은 공격자에게 이용당할 위험이 있으므로 시스템의 안전성을 해친다. 또한 믹스가 처리하지 못하는 메시지가 늘어나는 것은 시스템의 신뢰성을 해치게 된다. 안전성과 신뢰성의 성능이 대등하다고 보았을 때 그래프의 변곡점은 이러한 안전성과 신뢰성의 tradeoff 지점이 된다.
시스템에서 딜레이 시간을 적용할 경우, 전체 메시지 중에서 90%를 처리한다 하더라도 이에 해당하는 평균적인 딜레이 시간은 최대 딜레이 시간의 70 %를 넘지 않았다. 또한 95%를 목적으로 할 경우에도 최대 딜레이 시간의 75%를 넘지 않았음을 알 수 있었다.
환경변수와 딜레이 시간과의 관계는 다음과 같

다

- m : 시간당 발생 메시지 수가 많을수록 딜레이가 감소.
- l : 경로의 길이가 길수록 딜레이가 증가.
- n : threshold 가 클수록 딜레이가 증가.

## V. 결론

믹스 넷에서 키 변경 시 두 가지 키의 교차시간을 알아보기 위한 실험을 하고 분석해보았다. 실험 결과 메시지가 가지는 최대 딜레이 시간보다 적은 딜레이 시간을 사용하면서도 대다수의 메시지들을 처리할 수 있다는 것을 알 수 있었다. 또한 시스템을 이루는 환경변수가 변함에 따라 딜레이의 변화를 알 수가 있었고, 시스템의 신뢰성과 안전성을 동시에 고려할 수 있는 적절한 딜레이 시간을 사용할 수 있도록 근거를 제시하였다.

## VI. 향후 과제

threshold 뿐만 아니라 나머지 세 가지 모델을 이용하여 실험할 필요가 있다. 또한 각 파라미터와 딜레이 시간의 관계에 대하여 정량적으로 설명할 수 있도록 더욱 자세한 분석이 필요하다.

## [참고문헌]

[1] A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In P. Syverson and R. Dingledine, editors, Privacy Enhancing Technologies, LNCS. San Francisco, CA, 2002.

[2] A. Serjantov. R. E. Newman. On the Anonymity of Timed Pool Mixes. In the Workshop on Privacy and Anonymity Issues in Networked and Distributed Systems, Athens, Greece, pp. 427-434, May, 2003.

[3] Y. Zhu, R. Bettati , Anonymity vs. Information Leakage in Anonymity Systems, The 25th International Conference on Distributed Computing Systems (ICDCS), June 6-10, Columbus, Ohio, USA, 2005

[4] Gergely Tóth. Zoltán Hornák. Ferenc Vajda, Measuring Anonymity Revisited, Budapest University of Technology and Economics, Department of Measurement and Information Systems, Nordsec, 2004.

[5] A. Pfitzmann et al. Anonymity, Unobservability, Pseudonymity, and Identity Management - A Proposal for Terminology, Draft v0.23, Aug, 2005.

[6] George Danezis, Better Anonymous Communications, PhD thesis, University of Cambridge, December 2003.