Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.30, NO.1, Feb. 2020 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2020.30.1.29

IoT 환경에서 네트워크 코딩의 위장패킷 탐지와 유효한 복구의 식별 알고리즘

이 용†‡
프리랜서

Detection of Disguised Packet and Valid Reconstruction Identification Using Network Coding in IoT Environment

Yong Lee†‡
Freelancer

요 약

사물인터넷 기반의 응용서비스의 활용이 높아지고 네트워크 사용량이 급격히 증가함에 따라 네트워크 처리량을 개선하기 위하여 네트워크 코딩을 적용하는 연구가 활발하다. 네트워크 코딩에서 노드들은 주변 노드로부터 수신한 패킷을 여러 개 조합한 인코딩 패킷으로 변환하여 전송하고 목적지에서 디코딩할 수 있도록 한다. 이런 방식은 노드 간 신뢰를 기반으로 하지만 노드의 참여가 자유로운 사물인터넷 환경에서는 악의적인 노드가 구성에 참여할 경우 패킷을 조작할 수 있게 된다. 목적지에서 수신된 패킷은 하나의 소스에서 전송한 것이 아니라 여러 노드에서 생성한 여러 패킷이 조합된 것이므로 인코딩된 패킷의 진위를 식별하는 것이 어렵게 된다. 본 논문에서는 목적지에서 수신한 패킷이 전송 중에 공격을 받아 "유효하게 식별되는" 위장된 패킷의 존재를 탐지하는 방법과 디코딩 결과 중에 유효한 메시지를 식별하는 방법을 제안한다. 이 방법은 목적지가 위장패킷의 존재에도 불구하고 높은 확률로 재전송없이 수신된 패킷만으로 유효한 메시지를 구할 수 있으므로 네트워크 코딩의 성능이 향상됨을 보여준다.

ABSTRACT

Work to improve network throughput has been focused on network coding as the utilization of IoT-based application services increases and network usage increases rapidly. In network coding, nodes transform packets received from neighboring nodes into a combination of encoded packets for transmission and decoding at the destination. This scheme is based on trust among nodes, but in the IoT environment where nodes are free to join, a malicious node can fabricate the packet if it legally participates in the configuration. It is difficult to identify the authenticity of the encoded packet since the packet received at destination is not a single source but a combination of packets generated by several nodes. In this paper, we propose a method to detect "look-like-valid" packets that have been attacked and disguised in packets received at destination, and to identify valid messages in the reconstructions. This method shows that network coding performance is significantly improved because the destination can reconstruct a valid message with only received packets without retransmission with a high probability, despite the presence of disguised packets.

Keywords: Network Coding, IoT, Disguised Packet, Valid Reconstruction

Received(11. 29. 2019), Modified(02. 06. 2020), Accepted(02. 06. 2020)

† 주저자, yleehyun@gmail.com
‡ 교신저자, yleehyun@gmail.com(Corresponding author)

I. 서 론

최근 사물 인터넷(IoT, Internet of Things) 기반의 응용 서비스의 활용증가로 인한 센싱정보가 늘어나면서 네트워크의 데이터 처리량이 급격히 증가하고 있다. 따라서 네트워크 효율성을 향상시키는 연구에 대한 관심이 집중되고 있고, 이에 네트워크 코딩를 적용하는 연구도 증가하고 있다[1,2,3]. 네트워크 코딩은 중간 노드가 이웃 노드로부터 수신한 패킷을 결합하고, 목적지에서 디코딩될 수 있는 인코딩 패킷으로 변환하여 전송한다[4,5]. 이 방식은 네트워크에 대한 견고성과 오류허용 기능을 제공하기 위해 소스에서 싱크까지 멀티홉으로 여러 경로가 있는 네트워크 구조에서 주로 사용되므로 사물인터넷 환경에 특히 적합하다. 사물인터넷에서는 웨어러블 기기나 센서노드들이 데이터를 수집하고 전송, 릴레이하는 역할을 하게 된다. 인코딩 패킷은 여러 소스들로부터 수신한 패킷들을 조합한 것이기 때문에 정량적으로 합성된 정보가 전송되는 것이다[1,3-5]. 이렇게 정보가 혼합되고 정보량이 증가하는 방법은 실질적으로 전송 효율을 높이고, 네트워크가 통신, 하드웨어 또는 릴레이 에러에 대해서도 유연성을 가질 수 있다는 장점을 갖는다[1,3]. 이런 구조는 신뢰할 수 있는 노드들로 네트워크 토폴로지를 구성할 수 있는 경우에만 올바르게 작동할 수 있다. 악의적인 노드가 합법적으로 네트워크 구성에 참여하는 경우 중간 노드로서 위조된 인코딩 패킷을 삽입할 수 있게 된다. 목적지에서 수신한 패킷은 한 소스로부터 오는 것이 아니라 여러 소스의 패킷들이 결합된 것이므로 수신된 패킷이 정당한지를 쉽게 인식할 수 없다. 따라서 네트워크 코딩을 사용하는 네트워크는 구조적으로 악의적인 노드에 의한 정보 위변조의 위험이 높다.

합법적인 내부 노드에 의한 위협의 경우 공격이 신뢰할 수 있는 노드로부터 발생할 것이기 때문에 진위를 파악하거나 공격자를 식별하기가 더 어렵다. 서명이나 암호화와 같은 암호 알고리즘이 네트워크 코딩에 적용될 수 있지만, 노드들은 합법적으로 네트워크 구성에 합류하는 악의적인 중간 노드가 패킷을 의도적으로 조작, 위조하고 합법적인 서명이나 암호 등을 적용하여 내부 공격을 수행한다는 것을 쉽게 알아차릴 수 없다.

목적지는 전자서명 검증의 실패와 같이 암호화의 오류를 파악했을 때 공격받은 패킷을 구분하여 폐기하지만, 검증에 성공하면 수신한 패킷을 신뢰하고 이 패킷들에 대한 메시지 복구 메커니즘을 수행하게 된다. 올바른 패킷들로부터 디코딩된 결과와 내부 공격에 의해 유효한 패킷으로 위장된 패킷을 포함한 경우에 생성된 디코딩 결과는 서로 다른 복구 메시지를 보여줄 수 있다. 예를 들어 한 노드가 메시지를 b개의 패킷으로 분해하고 리던던시(redundancy)를 추가하여 n개의 인코딩 패킷으로 변환하여 전송한다고 가정하자. 모든 n개의 패킷이 목적지에 도착하지만 중간 노드가 내부 공격을 수행할 경우 이 패킷들중에는 *"유효하게 보이는(look-like-valid)"* 그러나 *"공격받은"* 패킷이 포함될 수 있다. 모든 패킷들은 유효한 서명이나 암호 형태를 가질 수 있으므로 목적지에서 수신한 패킷의 진위를 식별하는 방법은 중요한 이슈가 된다. 목적지는 패킷이 유효한지 여부를 전자서명 검증이나 디코딩만으로 확인할 수 없게 된다.

따라서 본 논문에서는 센서노드의 추가가 자유로운 사물인터넷 환경에서 센싱정보를 교환할 때, 목적지에서 수신한 패킷들 중에서 공격받은 패킷의 존재를 탐지하고 이를 재전송하는 것이 아니라 위장된 "유효하게 보이는" 패킷을 포함하는 수신 패킷들로부터 유효한 메시지를 복구하는 방법을 제안한다.

II. 관련 연구

임의의 디바이스의 참여가 가능한 사물인터넷 환경에 네트워크 효율성을 높이기 위해 네트워크 코딩을 적용할 경우 오염(pollution) 공격을 방지하거나 탐지하기 위한 네트워크 코딩 보안을 적용하는 것은 중요한 이슈이다.

네트워크 코딩에서 오염된 패킷을 식별하기 위하여 암호 알고리즘이나 전자서명 방식을 적용하는 연구가 활발히 진행되고 있다. Peralta 등[1]은 사물인터넷에서 데이터의 기밀성 보장 등 단대단 보안기능을 높이기 위해 동형(Homomorphic) 암호기술을 네트워크 코딩에 적용하는 모델을 제시하였다. Boneh 등[4]은 네트워크에서 임의의 노드에 의한 공격을 방지할 수 있는 동형 서명방식을 제안했으며 목적지는 이 서명방식을 사용하여 손상된 패킷을 필터링하고 중간 노드도 손상된 패킷을 폐기할 수 있다고 주장했다. Yu 등[6]은 키사전분배 및 메시지 인증코드를 사용하여 오염공격을 필터링할 수 있는 XOR 네트워크 코딩 보안 체계를 제안했다.

Li 등[7]은 노드의 추가가 용이한 사물인터넷에서는 키를 공유하기가 어려우므로 데이터의 진위를

보장하기 위해 네트워크 코딩 서명 기법을 적용하고 여러 디바이스들이 수집한 데이터를 자신의 인증키를 사용하여 서명할 수 있는 방안을 제안하였다. Wu 등[8]은 VANET에서 네트워크 코딩을 사용하는 경우 오염공격으로 인해 차량들이 메시지를 복구하지 못하는 경우에 대비하여 서명 알고리즘을 적용하고 유효한 패킷인지를 검증하는 방식을 제안하였다. Cheng 등[9]은 네트워크 코딩에서 오염공격에 대응하기 위하여 동형의 부공간 서명 방식을 사용하는 경우에 다중-생성 오염공격이 있음을 보이고, 키분배 방식을 개선하는 알고리즘을 설명하였다.

Li 등[10]은 인코딩 패킷에 오류 탐지와 오류 수정 기술을 적용하여 전송하는 방안을 제안하였다. 이 논문에서는 중간노드들이 전송하는 패킷에 대하여 오류 수정을 할 수 있도록 하여 처리율이 높아짐을 보여준다. Mamidwar 등[11]과 Wang 등[12]은 오염공격이 네트워크로 빠르게 전파되는 것을 막기 위하여 악의적인 노드를 지역화하거나 신뢰성있는 노드에 대해 가중치를 주는 방안을 제안하였다.

사물 인터넷에서는 데이터수집센터에 많은 데이터를 저장할 때 정보의 검색이나 안전성 보장들을 위해 데이터에 네트워크 코딩을 적용한다. Oliveira 등은 데이터의 저장에 네트워크 코딩을 적용하고 추가되는 리던던시를 최적화함에 따라 신뢰성 있는 데이터 저장 기능을 가지고 되고 최소의 비용으로 더 많은 패킷을 검색할 수 있게 하였다[13]. Lei 등[14]은 사물인터넷 응용을 제공하기 위해 NDN(Named Data Networking) 모델에서 데이터 생산자와 소비자 간에 네트워크 코딩을 적용하여 많은 양의 데이터 전송을 효율적으로 처리할 수 있음을 보여주었다.

III. IoT 네트워크 코딩 모델

이 논문에서는 전송오류는 고려하지 않고 악의적인 공격자가 생성하는 오염을 오류로 의미한다. 노드는 랜덤 네트워크 코딩(random network coding)을 위해 수신된 패킷에 랜덤 선형 방정식(random linear combination)을 적용하여 변환된 패킷을 송신한다. 여기서 메시지는 소스가 생성한 임의의 크기의 데이터를 의미하고, 패킷은 전송을 위해 고정된 크기의 길이로 나뉘어진 메시지의 조각을 의미한다.

소스노드는 Fig. 1과 같이 전송될 메시지를 b개의 패킷들로 분해하며, 이 패킷들을 $P_i, i=1,2,\cdots,b$로 정의한다. 소스노드는 분해된 패킷들에 선형방정식을

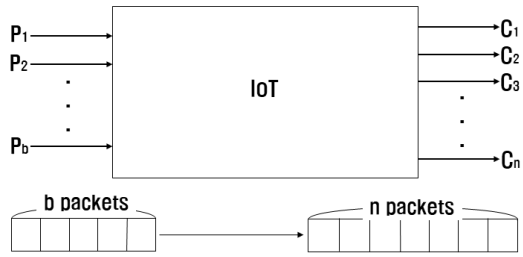

Fig. 1. IoT Network Coding Model.

적용 후 식(1)과 같은 인코딩된 패킷 C_j, $j=1,2,\cdots,n$으로 변환하여 전송한다[15,16].

$$C_j = \sum_{i=1}^{b} r_{ji} P_i \quad (1)$$

식(1)에서 r_{ji}는 갈로아 필드(Galois Field), $GF(2^q)$에 대하여 임의의 계수를 가지는 덧셈과 곱셈으로 구성되고 패킷 C_j의 헤더에 내장되어 목적지에서 패킷을 복구하기 위해 디코드할 때 사용된다[3,15.16].

목적지노드에서는 인코딩 패킷을 수신하면 인코딩 계수를 사용하여 메시지를 복구한다. 인코딩 패킷은 원본 패킷, b에 대한 선형방정식으로 표현되므로 목적지노드는 선형방정식을 사용하여 수신된 패킷에 디코딩 메커니즘을 수행하고 메시지를 복구할 수 있다. 하나의 복구과정에 사용되는 b개의 인코딩된 패킷은 수신한 n개 패킷의 부분집합이며 선형적(lineraly)으로 독립이다[15,16].

IV. IoT에서 공격 탐지 및 공격 수정 알고리즘을 적용한 네트워크 코딩

네트워크 코딩을 사용하는 네트워크 구조에서는 어떤 공격으로 인해 목적지노드가 수신한 인코딩 패킷에 오류가 있는 경우 원본 메시지를 올바르게 복구할 수 없다. 목적지노드는 오류의 존재를 인식하지 못하고 잘못된 메시지를 복구한 후에 정상적인 디코딩을 수행한 것으로 착각할 수 있다. 목적지노드가 수신한 패킷 중 공격으로 위장된 패킷이 있음을 감지하면 메시지를 복구하는데 도움이 될 것이다. 이 장에서는 네트워크 코딩에서 수신한 패킷 중에 공격으로 위장된 패킷이 존재하더라도 이를 탐지하고 목적지노드가 정상적으로 디코딩된 유효한 메시지를 식별

할 수 있도록 한다.

4.1 가정과 기호

제안하는 알고리즘은 다음과 같은 가정을 적용한다.

- 소스노드에서는 메시지를 b개의 패킷으로 분해하고 네트워크 코딩을 적용하여 리던던시 m을 가진 n개의 인코딩된 패킷으로 변환하여 전송한다
- 인코딩된 패킷 n개 중 b개의 패킷이 목적지에서 유효한 메시지를 복구하는 데 필요하다.
- 각 패킷은 다른 패킷과 독립적으로 전송된다고 가정한다.
- 악의적인 노드는 *"유효하게 보이는"* 조합으로 변조된 패킷을 생성할 수 있고 이 *"유효하게 보이는"* 패킷을 정상적인 패킷 대신 전송할 수 있다. 이것을 위장패킷이라고 한다.
- 네트워크에 전송오류는 없고 소스노드가 전송하는 패킷은 목적지노드가 모두 수신한다고 가정한다.

이 논문에서 아래와 같은 기호를 사용한다.

- b: 데이터 메시지가 분해된 평문 패킷의 수
- n: 평문 패킷에 리던던시를 적용하여 인코딩된 패킷 조합의 수, $n=b+m$
- m: 인코딩된 패킷을 생성하기 위해 평문 패킷에 추가된 리던던시의 수
- e: n개의 인코딩 패킷 중에서 위장패킷의 수
- r: n개의 인코딩 패킷 중에서 공격받지 않은 유효패킷의 수, $r=n-e$.
- $(P_1,P_2,...P_b)$: 원본 메시지로부터 분해된 b개의 유효한 평문 패킷들
- $(P_1^e,P_2^e,...P_b^e)$: 위장패킷을 포함하는 인코딩 패킷들로부터 복구된 유효하지 않은 패킷들
- $(C_i,\ i=1,2,\cdots n)$: 유효한 평문 패킷과 리던던시를 적용하여 인코딩된 패킷 조합
- $C_i^e,\ i=1,2,\cdots n$: 공격에 의해 변조된 위장패킷
- $\|\{C_1,C_2,\cdots,C_i,i>b\}\|$: 그룹 $\{C_1,C_2,\cdots,C_i,i>b\}$에 포함된 패킷의 수
- $coeff$: 인코딩 계수

4.2 그룹과 일치성

목적지노드가 위장패킷을 포함하여 수신한 패킷들에 대해 디코딩을 수행할 경우 하나 이상의 복구결과를 얻을 수 있고(수신된 모든 패킷이 유효하다면 유효한 복구결과는 하나만 존재하게 될 것이다.) 복구결과에 따라 수신한 패킷을 분류할 수 있게 된다. 두 개 이상의 복구결과가 일치하는 값을 가지면 이 복구과정에 포함된 패킷들을 하나의 그룹으로 분류한다.

예 : 소스노드는 메시지를 패킷 P_1,P_2로 나누고, 리던던시 $m=2$를 추가하여 C_1,C_2,C_3,C_4의 인코딩된 패킷을 생성하고 전송한다. 전송 중에 공격을 받아 목적지노드는 위장패킷 C_4^e를 포함한 C_1,C_2,C_3,C_4^e를 수신하고 다음의 복구결과를 얻는다고 가정한다.

i) $\begin{pmatrix}C_1\\C_2\end{pmatrix}coeff^{-1}=\begin{pmatrix}C_2\\C_3\end{pmatrix}coeff^{-1}=\begin{pmatrix}C_3\\C_1\end{pmatrix}coeff^{-1}=(P_1\ P_2)$

ii) $\begin{pmatrix}C_1\\C_4^e\end{pmatrix}coeff^{-1}=\begin{pmatrix}C_2\\C_4^e\end{pmatrix}coeff^{-1}=(P_1^e\ P_2^e)$

iii) $\begin{pmatrix}C_3\\C_4^e\end{pmatrix}coeff^{-1}=(P_1^{e\prime}\ P_2^{e\prime})$

위의 결과에서 (C_1,C_2), (C_1,C_3), (C_2,C_3)이 일치하는 복구결과값 (P_1,P_2)을 생성하므로 (C_1,C_2,C_3)이 하나의 그룹이 된다. (C_1,C_4^e), (C_2,C_4^e)도 동일한 복구결과값 (P_1^e,P_2^e)를 보여주므로 (C_1,C_2,C_4^e)도 역시 하나의 그룹이 된다. 나머지 (C_3,C_4^e)는 단일 결과값이 되므로 그룹을 구성하지 못한다.

n개의 인코딩된 패킷이 목적지노드에서 수신될 때, r개의 유효패킷과 e개의 위장패킷으로 구성되어 있다고 가정하면 $n=b+m=r+e$가 된다. 따라서 이 패킷들은 $\{C_1,C_2,\cdots,C_r\}\{C_{r+1}^e,\cdots,C_n^e\}$로 표현할 수 있다. n개의 모든 패킷을 수신한 경우 이들에 의해 얻을 수 있는 모든 가능한 복구결과의 총 수는 $\left\|\binom{n}{b}\right\|$ 이다. 이중에서 유효패킷으로만 이루어진 복구결과의 수는 $\left\|\binom{r}{b}\right\|$ 이고 적어도 하나의 위장패킷 C_i^e를 포함하는 복구결과의 수는 $\left\|\binom{n}{b}-\binom{r}{b}\right\|$ 이 된다.

하나 이상의 위장패킷을 포함하는 복구결과가 모두 다른 결과값을 생성하고, $\{C_1,C_2,\cdots,C_r\}$이 유일한 그룹이면 유효패킷들로만 만들어진 이 결과값이 유일한 유효한 복구결과가 된다. 두 개 이상의 복구결과가 하나 이상의 위장패킷을 포함하면서 일치하는 값을 생성하면 이 복구과정에 포함된 패킷들도 하나의 그룹이 된다(위의 예에서 (P_1^e,P_2^e)). 그룹이 두 개

이상일 때, 그 중에서 유효한 결과를 구별해야 한다. 위의 예에서는 목적지노드가 C_4^e가 위장패킷임을 감지하지 못하므로 (C_1, C_2, C_3)과 (C_1, C_2, C_4^e)이 생성한 복구결과중에 유효한 결과값을 구별해야 한다.

다음은 일치성 및 앞에서 정의한 그룹을 이용하여 위장패킷의 탐지와 유효한 메시지 식별 알고리즘의 동작에 대하여 살펴볼 것이다. x를 한 그룹내 포함된 인코딩 패킷의 총 수라고 하고 y를 유효한 메시지를 복구하는 데 필요한 패킷의 수라고 하자. $\left\|\binom{x}{y}\right\|$는 x개의 패킷이 생성할 수 있는 총 복구결과의 수가 된다.

그룹을 구성하는 데 참여한 복구과정의 수가 $\left\|\binom{x}{y}\right\|$와 같을 때, 이 그룹은 *일치성*을 가진다고 한다. 앞의 예에서 그룹 i)는 3개의 복구과정이 참여하고 포함된 패킷의 수가 $\|\{C_1, C_2, C_3\}\|=3$으로, $x=3$, $y=2$이 되서 $\left\|\binom{x}{y}\right\|=\left\|\binom{3}{2}\right\|=3$ 이므로 일치성을 가진다. 그룹 ii)는 2개의 복구과정만 참여하므로 $\left\|\binom{x}{y}\right\|=\left\|\binom{3}{2}\right\|\neq 2$ 이 돼서 일치성을 갖지 못한다.

한 그룹이 r개의 인코딩 패킷을 갖고 $r>b$라 가정하면, $\left\|\binom{r}{b}\right\|>1$ 이 된다. 이 그룹은 r개의 인코딩 패킷이 모두 유효패킷이고 이들이 $\left\|\binom{r}{b}\right\|$ 개의 복구과정을 통해 모두 하나의 일치하는 결과값을 만들기 때문에 일치성을 가지게 된다. 다른 예로 $e(e>b)$개의 위장패킷만을 가지는 그룹을 가정하면, $\left\|\binom{e}{b}\right\|>1$ 이 된다. 이 그룹의 모든 인코딩 패킷들이 $\left\|\binom{e}{b}\right\|$ 개의 복구과정에 참여하고 하나의 일치하는 결과값을 만든다면, 이 그룹도 일치성을 가지게 된다.

이번에는 r'개의 유효패킷과 e'개의 위장패킷을 포함하는 그룹을 가정하고 $r' \geq b$, $e' \geq 1$라고 하자. 이 그룹이 일치성을 갖기 위해서는 $r'+e'$개의 패킷이 $\left\|\binom{r'+e'}{b}\right\|$ 개의 복구과정에 참여하여야 한다. 그러나, r'개의 인코딩된 패킷은 이미 $\left\|\binom{r'}{b}\right\|>1$ 을 복구과정을 통해 공통의 일치하는 결과값을 생성하고 유효패킷들로만 구성된 그룹에 속해 있으므로 이 그룹에 참여하는 복구과정의 수는 $\left\|\binom{r'+e'}{b}\right\|$ 보다 적게 되어 이 그룹은 일치성을 갖지 못한다.

다음 b보다 작은 $r'(r' \leq b-1)$개의 유효패킷과 위장패킷 e개를 가지는 그룹에 대해 $r'+e>b$라고 하고 $r'+e$개의 인코딩된 패킷으로 수행할 수 있는 모든 복구과정이 공통의 결과값을 생성한다면 이 그룹은 일치성을 가지게 된다.

4.3 제안 알고리즘

앞에서 살펴본 그룹과 일치성의 특성을 이용하여 목적지노드에서 인코딩 패킷에 포함된 위장패킷을 탐지하고 유효한 복구결과를 식별하기 위해 다음과 같은 알고리즘을 얻을 수 있다.

```
for all received encoded packets
    Perform all recovery processes
for all matching results
    Set group
Select the consistent group
for all consistent group
    Find the largest group
```

먼저 유효한 복구결과가 존재하는 조건을 고려하면 복구결과가 일치성있는 그룹을 구성할 때, 이 그룹이 생성하는 메시지는 유효한 복구결과의 후보가 된다. $e<m$일 때, $r>b$이므로 $\left\|\binom{r}{b}\right\|>1$ 를 얻을 수 있다. r개의 인코딩 패킷은 동일한 결과를 생성하는 2개 이상의 복구과정에 참여하고 일치성있는 그룹이 된다. 그러므로 유효한 복구결과가 존재하기 위한 조건은 $e<m$이 된다.

$\left\|\binom{n}{b}-\binom{r}{b}\right\|$ 개의 다른 복구과정이 하나의 공통 결과를 가지고 그룹을 구성하더라도 이 그룹은 일치성을 갖지 못하게 된다. 그룹에 포함된 패킷의 수는 n이고 이 패킷들이 참여하는 이 그룹의 복구과정은 $\left\|\binom{n}{b}\right\|$ 보다 작은 $\left\|\binom{n}{b}-\binom{r}{b}\right\|$ 이 되기 때문이다.

예 : $b=2$, $m=2$, $n=2$인 인코딩패킷이 전송되고 목적지노드는 $e=1$로 위장패킷을 가진 상태로 패킷을 수신한다고 가정한다. $r=3$이므로 목적지노드는 이 그룹으로부터 $\left\|\binom{r}{b}\right\|=\left\|\binom{3}{2}\right\|=3$ 개의 공통 결과를 얻고 일치성을 확인할 수 있다. 다른 복구과정은 $\left\|\binom{4}{2}-\binom{3}{2}\right\|=3$ 개이고 공통 결과를 얻을지라도 $\|\{C_1, \cdots, C_4\}\|=4$가 되어 일치성을 갖지 못하게 된다. 따라서 유효한 결과를 가지는 그룹을 찾을 수 있게 된다. 만약 목적지노드가 $e=2$, $r=2$인 상태로 패킷을 수신한다면 $r=2$을 통해서 하나의 유효한 결과값만을 구할 수 있으므로 이것이 유효한 메시지인지를 확인할 수 없게 된다.

다음은 여러 개의 일치성있는 그룹이 존재할 때 유효한 결과를 생성하는 찾기 위해 다수의 법칙을 적용할 수 있는 조건을 알아보기로 하자. 먼저 $r>e$ $(r\geq b,\ e\geq b)$인 경우를 생각해보자. $\left\|\binom{r}{b}\right\|$ 개의 복구과정은 $\left\|\binom{e}{b}\right\|$ 개의 복구과정보다 더 많으므로 다수의 법칙을 적용하여 유효한 결과를 구별해 낼 수 있게 된다.

다음은 $e\leq \lfloor \frac{m}{2} \rfloor$인 경우를 생각해보자. 한 그룹이 $\{C_1,\cdots,C_{r'}\}\{C_1^e,\cdots,C_e^e\}$인 패킷으로 구성되고 $r'\leq b-1$, $r'+e>b$라고 가정한다. $r'+e$개의 인코딩 패킷이 하나의 공통 결과를 생성하면 이 그룹은 일치성을 갖는다. $r'\leq b-1$이므로 r'개의 인코딩 패킷은 결과를 생성하지 못한다. $r'+e<r$이라면 $\left\|\binom{r'+e}{b}\right\|<\left\|\binom{r}{b}\right\|$ 가 되고 다수의 법칙을 적용하여 유효한 결과를 찾을 수 있게 된다. 이 그룹이 일치성을 갖기 위한 r'의 최대값은 $b-1$이 된다. $r'=b-1$이라고 가정하면 $r'+e<r$로부터 $b=n-m$, $r=n-e$을 대입하여 $e\leq \lfloor \frac{m}{2} \rfloor$이 된다. $e\leq \lfloor \frac{m}{2} \rfloor$일 때, 위장패킷을 가지는 그룹이 하나의 공통 결과를 생성하고 일치성을 갖더라도 이 그룹은 다수의 법칙에서 탈락하게 된다.

앞에서 $e<m$, $r>e$, $e\leq \lfloor \frac{m}{2} \rfloor$의 3가지 조건을 얻게 되고, 유효한 복구결과를 찾기 위한 조건은 $r>e$와 $e\leq \lfloor \frac{m}{2} \rfloor$이 된다.

V. 성능분석 및 결과

5.1 성능분석

인코딩 패킷이 전송 중에 공격받아 위장패킷이 될 확률이 p라고 가정하자. 위장패킷의 수 e가 i와 같을 확률은 다음과 같다.

$$P(e=i)=\binom{n}{i}p^i(1-p)^{n-i}. \tag{2}$$

유효한 복구결과를 찾을 수 있는 조건들을 적용하면 $e\leq \lfloor \frac{m}{2} \rfloor$일 확률은

$$p(e\leq \lfloor \frac{m}{2} \rfloor)=\sum_{i=0}^{\lfloor \frac{m}{2} \rfloor}\binom{n}{i}p^i(1-p)^{n-i}, \tag{3}$$

이 되고, $e<r$일 확률은

$$p(e<r)=\sum_{j=0}^{r-1}\binom{n}{j}p^j(1-p)^{n-j}, \tag{4}$$

이 된다.

따라서 유효한 복구결과를 식별할 수 있는 확률은 다음과 같이 표현될 수 있다.

$$P_s=\sum_{k=0}^{\min(\lfloor \frac{m}{2} \rfloor,r-1)}\binom{n}{k}p^k(1-p)^{n-k}. \tag{5}$$

이를 이용하여 단일 메시지의 성공적인 전송 비용 T_s를 다음과 같이 구할 수 있다.

$$T_s=\frac{1}{P_s}\cdot n\cdot Co_{tx} \tag{6}$$

여기서 Co_{tx}는 한 패킷에 대한 전송비용이고, n은 소스노드가 전송한 총 인코딩 패킷의 수이다.

5.2 결과

이 절에서는 평문 패킷의 수 b와 리던던시의 수 m의 변화와 위장 패킷이 발생할 확률 p의 변화에 따라 식(5)의 유효한 복구결과를 식별할 수 있는 확률이 어떻게 변화하는지를 그래프로 나타내었다.

Fig. 2는 패킷이 공격받을 확률이 높아짐에 따라 유효한 복구결과를 찾을 수 있는 확률이 감소하는 것을 보여준다. 또한 리던던시가 증가함에 따라 확률 P_s

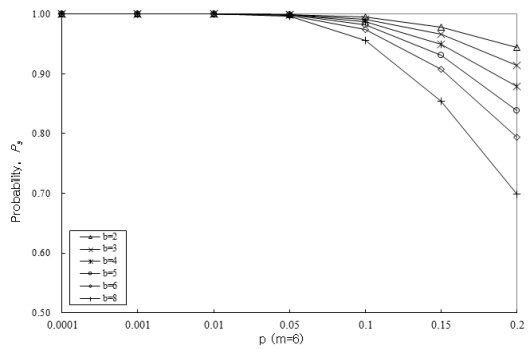

Fig. 2. The probability P_s as the number of packets increases when $m=6$

도 증가함을 알 수 있다. $m=4$인 예를 살펴보면 식 (5)에 따라 e의 값이 ($\lfloor \frac{m}{2} \rfloor$, $e<r$를 만족하는 $r-1$) 중에서 최소값이 되므로 위장패킷이 2개 발생할 때까지는 높은 확률로 복구결과를 찾을 수 있음을 알 수 있다. 또한 리던던시 m이 짝수일 때 유효한 복구결과를 구할 확률이 홀수일 때 보다 높음을 알 수 있다. 이는 $\lfloor \frac{m}{2} \rfloor = \lfloor \frac{2q}{2} \rfloor = \lfloor \frac{2q+1}{2} \rfloor (q=1,2,\cdots)$이기 때문이다. 그러나 위장패킷이 발생할 확률 p가 충분히 작을 때는 $(1-p)^{n-k} \approx 1$이 되므로 확률 P_s는 리던던시의 이런 특성에 의해 영향을 받지 않는다.

Fig. 3에서는 $m=6$일 때 평문패킷의 수에 대한 확률 P_s의 변화를 보여준다. b의 값에 상관없이 $p=0.05$가 될 때까지 확률 P_s가 거의 1을 유지함을 알 수 있다. 평문패킷의 수 b가 클수록 p의 증가에 따라 P_s가 급격히 감소함을 알 수 있다.

Fig. 4는 $b=4$일 때 위장패킷이 발생할 확률 p가 증가함에 따라 전송비용 T_s의 변화를 보여준다. $p=0.05$가 될 때까지는 전송 비용 T_s는 m에 상관없이 안정적으로 유지되는 것을 알 수 있다. m이 홀수일 때 전송비용의 증가율은 m이 커짐에 따라 감소하고 p가 증가함에 따라 급격하게 증가함을 보여준다.

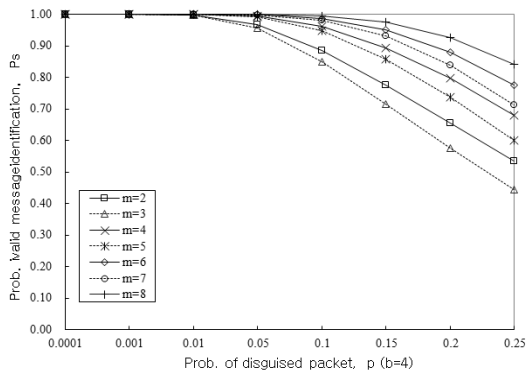

Fig. 3. Probability of valid message identification P_s as p increases when $b=4$

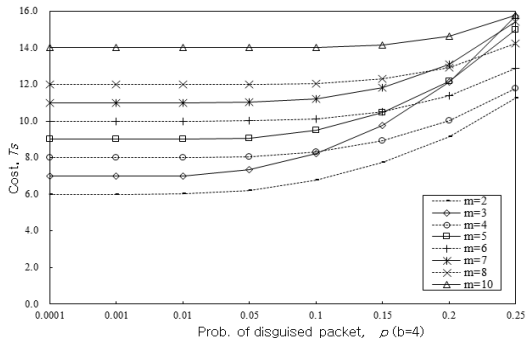

Fig. 4. Cost, T_s as p increases when $b=4$

2장에서 살펴본 네트워크 코딩의 안전성을 보장하기 위한 연구는 [1,4,8,9]의 암호화나 전자서명과 같은 암호기술을 적용하는 방법과 [6,10,11]에서 제안한 안전한 네트워크 코딩을 이용하는 방법, [12.13]의 신뢰성 있는 노드를 구분하여 공격에 대비하는 방법 등이다. 암호기술 방식은 키교환 등의 오버헤드가 발생하고 내부자가 패킷을 변조한 후 정상적인 키로 암호화하여 보내면 목적지에서는 이를 복호화하여 패킷을 신뢰하고 네트워크 코딩에 대한 복구를 실행함으로써 패킷의 위변조를 알아차릴 수 없게 된다. 안전한 네트워크 코딩의 적용도 오염패킷의 발생을 막는 것을 목적으로 한다. 신뢰 방식의 경우 신뢰성을 쌓은 노드가 변심하는 내부적 공격은 대응할 수가 없게 된다. 기존의 논문들이 오염패킷의 발생을 막는데 중점을 두는 반면에 본 논문에서 제안하는 방식은 내부적 공격에 의해 위장된 패킷이 발생하더라도 이의 존재를 탐지하고 메시지를 복구할 수 있는 해결방안이 될 수 있다.

VI. 결 론

네트워크 코딩은 사물인터넷 환경에서 네트워크의 처리량을 개선하는데 기여할 수 있다. 라우팅의 중간 노드가 이웃 노드로부터 수신한 패킷을 조합하고 인코딩하여 전송하는 네트워크 코딩에서는 악의적 노드에 의해 위변조되거나 손상된 패킷의 조합이 발생하는 것은 피할 수는 없는 문제이다. 또한 정체를 숨긴 악의적인 노드에 의하여 유효한 서명과 유효한 암호화를 가진 것처럼 보이는 *"유효하게 보이는"* 위장패킷의 존재가 가능하다.

이 논문에서는 목적지노드가 이런 종류의 위장패킷의 존재를 감지하지 못하고 메시지를 복구하는 경우에도 유효한 복구결과를 찾고자 하였다. 목적지노드가 위장패킷을 포함하는 패킷들로 메시지를 복구한다면, 모든 복구결과가 일치하지 않고 목적지노드는 잘못된 결과를 얻을 수 있다. 본 논문은 수신 패킷 중에서 위장패킷의 존재를 탐지하고 위장패킷이 존재하더라도 유효한 메시지를 복구할 수 있는 알고리즘을 제안하였다. 이 알고리즘은 위장패킷이 발생할 확률이 0.1이 될 때까지는 비용의 증가없이 높은 확률로 유효한 메시지를 복구할 수 있음을 보여주었다.

References

[1] G. Peralta, R. Fuentes, J. Bilbao and P. Crespo, "Homomorphic Encryption and Network Coding in IoT Architectures: Advantages and Future Challenges," Electronics, vol. 8, no. 827, July 2019

[2] A. Fuqaha, M. Guizani, M.Mohammadi, M. Aledhari and M. Ayyash, "Internet of Things:A Survey on Enabling Technologies, Protocols and Applications," IEEE Communication Surveys and Tutorials, vol. 17, no. 4, pp.2347-2376, June 2015

[3] J. Li, Y. Liu, Z. Zhang, J. Ren and N. Zhao, "Towards Green IoT Networking: Performance Optimization of Network Coding Based Communication and Reliable Storage," IEEE Access, vol. 5, pp. 8780-8791, May 2017

[4] D. Boneh, D. Freeman and J. Katz, "Signing a Linear Subspace: Signatures for Network Coding," PKC 2009, Mar. 2009

[5] R. Ahlswede, N. Cai, S. Li, W. Yeung, "Network Information Flow," IEEE Trans. Information Theory, vol. 46, no. 4. pp. 1204-1216, July 2000

[6] Z. Yu, Y. Wei, B. Ramkumar and Y. Guan, "An Efficient Scheme for Securing XOR Network Coding against Pollution Attacks", IEEE INFOCOM 2009, pp. 406-414, 2009

[7] T. Li, W. Chen, Y. Tang and H. Yan, "A Homomorphic Network Coding Signature Scheme for Multiple Sources and its Application in IoT," Security and Communication Networks, vol. 2018, June 2018

[8] G. Wu, J. Wang, Y. Wang and L. Yao, "Pollution Attack Resistance Dissemination in VANETs Based on Network Coding," Procedia Computer Science, vol. 83, pp. 131-138, 2016

[9] C. Cheng, J. Lee, T. Jiang and T. Takagi, "Security Analysis and Improvements on Two Homomorphic Authentication Schemes for Network Coding", IEEE Trans. on Information Forensics and Security, vol. 11, no. 5, pp. 993-1002, May 2016

[10] J. Li, T. Li, J. Ren and H. Chao, "Enjoy the Benefit of Network Coding: Combat Pollution Attacks in 5G Multihop Networks," Wireless Communications and Mobile Computing, vol. 2018, Dec. 2018

[11] N. Mamidwar and D. Gothawal, "Schemes against Pollution Attack in Network Coding: A Survey," Intl. Journal of Computer Science and Information Technologies, vol. 6, no. 6, pp. 5085-5089, 2015

[12] W. Cheng, L. Yu, F. Xiong and W. Wang, "Trusted Network Coding in Wireless Ad Hoc Networks," IEEE Globecom 2010, Oct. 2010

[13] C. Oliveira, Y. Doudane, C. Brito and S. Lohier, "Optimal Network Coding -Based In-Network Data Storage and Data Retrieval for IoT/WSNs," IEEE 14th NCA, Sep. 2015

[14] K. Lei, S. Zhong, K. Xu and H. Zhang, "An NDN IoT Content Distribution Model With Network Coding Enhanced Forwarding Strategy for 5G," IEEE Trans. on Industrial Informatics, vol. 14, no. 6, pp. 2725-2735, Jun. 2018

[15] M. Rabin, "Efficient Dispersal of Information for Security, Load Balancing, and Fault Tolerance," Journal of the ACM, vol. 36, no. 2, Apr. 1989

[16] J. Barros, "Mixing Packets: Pros and Cons of Network Coding," The 11th WPMC, Sep. 2008

〈저자소개〉

이 용 (Yong Lee) 종신회원
1997년 8월: 연세대학교 컴퓨터과학과 석사
2001년 2월: 연세대학교 컴퓨터과학과 박사
2001년~2003년: 한국정보보호진흥원 선임연구원
2005년~2007년: 삼성전자 통신연구소 책임연구원
2007년~2011년: 충주대학교 전자통신공학전공 조교수
〈관심분야〉 네트워크 보안, 차세대 인터넷, IoT보안, 이동통신망 보안, 정보보호