베이지안 정리를 이용한 무선 ad hoc 네트워크에서의 침입탐지 기법 연구

An Intrusion Detection Method based on Bayesian Theory in Wireless Ad Hoc Networks

김현우[1], 한영구[2], 김기성[2], 김세헌[2]
[1]LG데이콤 종합연구소 NMS연구팀 (hwkim@lgdacom.net)
[2]한국과학기술원 산업공학과 ({yghan, kskim, shkim}@tmlab.kaist.ac.kr)

초록

무선 ad hoc 네트워크는 기존의 네트워크와 달리 인프라가 존재하지 않은 상태에서 각 노드들 간에 직접적인 통신을 수행하는 네트워크이다. 무선 ad hoc 네트워크는 불안정한 링크 특성과 동적인 네트워크 구조로 인해 데이터 위변조 문제와 노드 정보 도청 등의 공격에 매우 취약하다. 싱크홀 공격은 무선 ad hoc 네트워크에서 발생하는 대표적인 공격 유형으로서 라우팅 정보를 변경하여 공격자의 노드로 모든 데이터들이 지나가도록 한다. 본 연구는 무선 ad hoc 네트워크에서 노드들 간의 안전한 통신을 위협하는 싱크홀 공격을 효과적으로 탐지하기 위하여 베이지안 정리를 이용한 침입탐지 기법을 제안한다.

1. 서론

무선 ad hoc 네트워크는 기존의 네트워크와 다르게 고정된 중재자의 도움 없이 자율적으로 망의 구성이 가능한 네트워크이다. Ad hoc 네트워크에서는 고정된 라우터가 존재하지 않아 이동 노드간의 협력에 의한 라우팅 기능이 제공되며, 특정 서비스 제공자 없이 단말에서 서비스가 해결되어야 한다는 특징을 가진다. 또한, ad hoc 네트워크에서는 노드들의 네트워크 참여와 이탈이 자유로워서 네트워크 토폴로지가 수시로 변하는데, 이러한 동적인 네트워크 구조로 인해 불법 노드에 의한 네트워크 자원소비 및 경로방해, 노드 정보 도청 등이 손쉽게 일어날 수 있다 [1].

Ad hoc 네트워크에 불법 노드가 참여할 경우 불법노드는 경로요구 및 경로응답 패킷을 변경하여 데이터의 오전송을 유발할 수 있으며, 라우팅 트래픽을 범람시켜 통신을 거절할 수도 있다. 이러한 불법 노드의 고의적 행동들은 네트워크 동작을 불가능하게 할 뿐 아니라 노드사이의 경로를 탐색하는데 있어 지연을 발생시킨다. 그러므로 ad hoc 네트워크에서 불법 노드를 탐지하고, 라우팅 경로 상에서 배재하는 일은 매우 중요하다 [2].

불법 노드에 의해 발생하는 대표적인 공격으로는 정당한 데이터 패킷을 잘못된 경로로 가도록 하는 경로방해 공격과 네트워크 자원인 전력, 메모리 그리고, 대역폭을 소비하도록 네트워크에 패킷을 주입하는 자원소비 공격이 있다. 특히, 싱크홀 공격은 ad hoc 네트워크에서 발생하는 대표적인 공격 유형 중 하나로서, 공격노드가 주변의 노드에 잘못된 라우팅 정보를 제공하여 공격반경 안에 있는 노드들의 데이터를 모두 전송받은 후 잘못된 경로로 재전송하거나 중간에 데이터를 소실시키는 공격이다 [3].

본 연구에서는 ad hoc 네트워크의 취약점을 이용하여 노드들 간의 안전한 통신을 위협하는 싱크홀 공격을 효과적으로 탐지하기 위하여 베이지안 정리를 이용한 침입탐지 기법을 제안한다. 2장에서는 본 연구의 이론적 배

본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음

경인 Dynamic Source Routing과 싱크홀 공격에 대해서 살펴보고, 3장에서는 싱크홀 공격 탐지를 위한 특성 파라미터에 대하여 설명한다. 4장에서는 베이지안 정리를 이용한 싱크홀 공격 탐지 알고리즘을 제안하고, 5장에서 결론 및 향후 과제에 대하여 언급한다.

2. 연구 배경

2.1 Dynamic Source Routing

ad hoc 네트워크에서 라우팅 프로토콜은 라우팅 정보의 유지 여부에 따라 table-driven 방식과 on-demand 방식으로 분류된다 [4]. Table-driven 방식은 각 이동 노드들이 전체 경로에 대한 라우팅 정보를 table에 저장하는 방식으로 destination sequence distance vector (DSDV) 라우팅이 대표적인 예이다. Table-driven 방식은 라우팅이 신속하게 이루어진다는 장점이 있지만 모든 이동 노드들이 라우팅 정보를 유지해야 하는 문제가 발생하게 된다. 반면 on-demand 방식은 노드가 통신을 원할 때 경로 탐색 프로세스를 통해 라우팅을 설정하는 방식으로 각 노드가 전체 라우팅 정보를 유지할 필요가 없기 때문에 상대적으로 부하가 적다는 장점이 있다.

Dynamic Source Routing (DSR)은 ad hoc on-demand distance vector routing (AODV)와 더불어 대표적인 on-demand 방식의 라우팅 프로토콜이다 [5]. DSR에서 이동 노드들은 source route를 가진 route cache를 유지하고 있다가 새 경로가 입력될 때마다 지속적으로 cache를 갱신한다. 이동 노드가 전송해야 할 패킷이 있으면, route cache를 확인 후 유효기간이 지나지 않았다면 이 경로를 이용한다. 이 때 cache에 경로가 존재하지 않는다면 경로를 탐색하는 과정을 거친다.

DSR의 경로 탐색 과정은 route request (RREQ) 패킷과 route reply (RREP) 패킷을 통해 수행된다. RREQ 패킷은 경로를 설정하기 위해 source 노드에 의해 broadcast되는 패킷으로서 <소스 노드 IP, 목적지 IP, sequence number>로 이루어진 필드를 가지고 있다. 소스 노드로부터 RREQ를 받은 노드들은 자신의 cache를 확인하고 경로가 존재하지 않는다면 RREQ를 계속 전파한다. RREQ가 목적지에 도달하거나 그에 대한 cache를 가진 중간 노드에 도달하면 RREP가 생성된다. 이 때, RREP를 만든 곳이 목적지라면 RREQ에 포함된 route record를 RREP에 넣어서 source로 보내게 되며, 중간 노드의 경우에는 cache된 경로를 route record에 덧붙여 RREP를 반송하게 된다. 그림 1은 RREQ가 네트워크를 통하여 전파되는 과정에서의 route record의 형식을 보여준다.

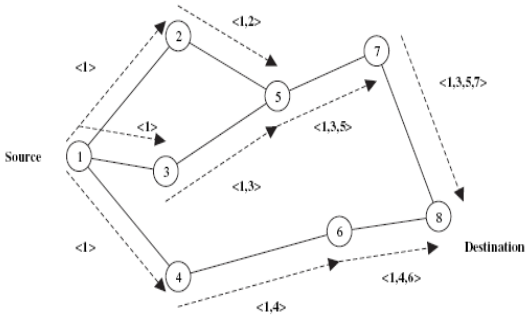

그림1. RREQ 전파 과정 [5]

그림 1에서 목적지(노드 8)은 자신에게 도착한 RREQ 중에서 하나를 선택하여 RREP를 보냄으로써 경로를 설정한다. 그림 2는 그림 1에서 <1,4,6> 경로가 선택되었을 때 RREP가 반송되는 과정을 보여준다.

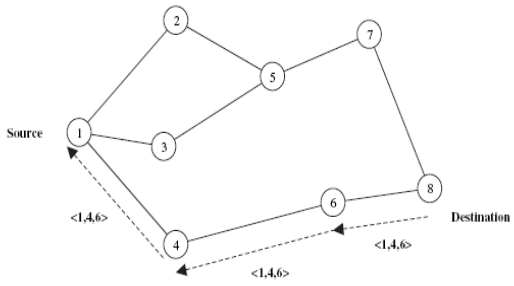

그림 2. RREP 전파 과정 [5]

2.2 싱크홀 공격

싱크홀 공격은 ad hoc 네트워크에서 발생하는 대표적인 공격 유형 중 하나로서, 공격노드가 주변의 노드에 잘못된 라우팅 정보를 제공하여 공격반경 안에 있는 노드들의 데이터를 모두 전송받은 후 잘못된 경로로 재전송하거나 중간에 데이터를 소실시키는 공격이다. 이러한 공격은 네트워크 데이터의 엿듣기나

서비스 거부를 유발할 뿐만 아니라 전체 네트워크의 오버헤드를 가중시켜 네트워크 수명을 줄이게 한다 [6].

DSR 방식의 ad hoc 네트워크에서 싱크홀 공격은 RREQ에 첨부되는 sequence number를 조작함으로써 이루어진다. Sequence number는 RREQ의 broadcasting 과정에서 서로 다른 경로의 비교 및 갱신을 위해 사용되는 값으로 sequence number의 값이 높은 RREQ의 경로를 더 새로운 경로로 판단한다. 이를 이용하여 네트워크 상의 불법 노드는 매우 높은 sequence number를 가진 거짓 RREQ를 발생시켜 다른 노드들에게 소스 노드로부터 목적지 사이에 불법 노드를 포함시키는 경로가 기존의 경로보다 나은 것처럼 잘못된 판단을 하도록 만든다. 그림 3은 싱크홀 공격에서 발생하는 거짓 RREQ의 전송과정을 보여준다. 그림 3에서 불법 노드 1은 노드 6에서 노드 8로 가는 경로에 sequence number를 999로 높게 설정하여 거짓 RREQ를 발생시킨다. 이러한 거짓 RREQ를 받은 다른 노드들은 sequence number를 보고 기존의 경로를 노드 1을 포함한 새로운 경로로 갱신한다. 이러한 과정을 반복하면 전체 노드들 간의 모든 통신에 노드 1이 포함되도록 라우팅을 구성할 수 있게 된다.

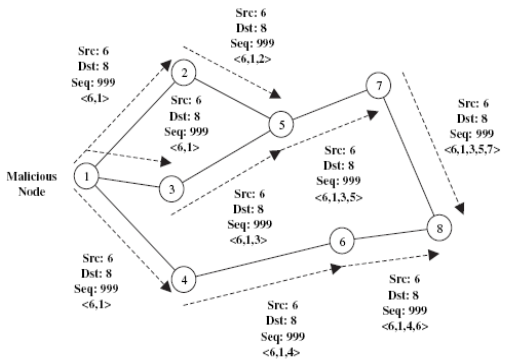

그림 3. 거짓 RREQ 전파과정 [5]

3. 싱크홀 공격 탐지를 위한 특성 파라미터

싱크홀 공격을 효율적으로 탐지하기 위해서는 싱크홀 공격의 특성을 반영하고 탐지할 수 있는 파라미터를 선정해야 한다. Tseng et al.은 싱크홀 공격을 탐지하기 위한 특성 파라미터로서 sequence number discontinuity, route add ratio를 사용하였다 [5]. 본 연구에서는 이에 number of hops을 추가하여 세 가지 파라미터를 싱크홀 공격 탐지에 사용한다.

3.1 Sequence number discontinuity

Sequence number discontinuity (SeqN_D)는 각 노드에서 최근에 갱신된 sequence number와 이전의 sequence number와의 차이를 의미한다. 이에 추가로 중복된 sequence number의 수가 페널티로서 합산된다. 일반적으로 DSR에서 정상적인 노드 사용자들만 있을 경우에 RREQ의 sequence number는 1씩 차례로 증가하게 된다. 하지만 싱크홀 공격을 발생시키려는 공격자가 있을 경우, 공격자는 그림 3과 같이 sequence number를 전체 경로 상의 기존 sequence number보다 월등히 높게 설정을 하게 되므로 sequence number 상의 차이가 커지게 된다. 따라서 sequence number 간의 차이를 관찰하는 것은 싱크홀 공격의 탐지를 위한 중요한 단서를 제공한다.

이를 피해가기 위해 공격자는 sequence number를 싱크홀 공격이 가능할 정도로만 약간 높여서 발생시키는 방식을 사용할 수 있다. 이러한 행위는 sequence number 간의 차이를 계산하는 것만으로 확인이 어렵다. 하지만 이러한 경우 싱크홀 공격으로 인한 sequence number와 기존 sequence number 간의 차이가 크지 않기 때문에 어느 정도 시간이 흐르면 sequence number가 동일해지게 된다. 따라서 중복된 sequence number의 수를 추가적으로 관찰함으로써 탐지가 가능하다.

3.2 Route add ratio

Route add ratio는 전체 노드 경로 상에 특정 노드가 포함되어 있는 비율을 의미한다. 싱크홀 공격은 노드들 간의 모든 통신 경로에 공격자 자신을 포함시키기 때문에 공격자가 네트워크 경로에 포함되는 비율이 다른 노드에 비해서 크게 높아지게 된다. 따라서 각 노드들의 라우팅 테이블을 관찰하고 노드 별로 라우팅 테이블에 포함된 횟수를 계산하면 싱크홀 공격 발생의 유무를 확인할 수 있을 뿐 아니라 공격자의 위치도 판단 가능하다. 본 연구에서는 싱크홀 공격의 유무를 파악하는 것을 목적으로 하고 있으므로 전체 노드의 평균

route add ratio를 계산하여 공격 탐지를 위한 파라미터로 사용한다.

3.3 Number of hops

Number of hops은 소스 노드로부터 목적지까지 데이터를 전송할 때 거치는 중간 노드의 수를 의미한다. 싱크홀 공격이 발생하면 네트워크의 라우팅이 공격자를 포함하도록 설정된다. 따라서 공격자와 멀리 떨어져 있는 두 노드 간의 통신 시에도 공격자를 포함하도록 하므로 number of hops가 상승하는 현상이 발생한다. 그림 3의 예를 보면 노드 6과 노드 8 간의 통신이 이루어질 때도 노드 1을 포함하는 경로가 사용되기 때문에 number of hops가 높아지게 되는 것을 관찰할 수 있다. 따라서 전체 네트워크 경로들의 평균 number of hops를 계산함으로써 싱크홀 공격을 탐지할 수 있다.

4 싱크홀 공격 탐지 알고리즘

본 장에서는 베이지안 정리를 이용하여 싱크홀 공격을 효과적으로 탐지하기 위한 알고리즘을 제안한다. 베이지안 정리는 잘 알려진 통계적 이론 중 하나로서 조건부 확률과 사전확률을 이용해 사후확률을 계산할 수 있다 [7].

베이지안 정리에 의해 싱크홀 공격을 탐지하기 위해서는 3장에서 거론한 sequence number discontinuity, route add ratio, number of hops의 세 가지 파라미터 값이 계산되어야 한다. i 시점에서 각 파라미터들을 계산하였을 때 그 결과를 SeqN_D = α_i, route add ratio = β_i, number of hops = γ_i라고 하자. 이 때 싱크홀 공격 발생 확률 $P(S|\alpha_i,\beta_i,\gamma_i)$는 베이즈 룰(bayes rule)에 의해 다음과 같이 계산할 수 있다.

$$P(S|\alpha_i,\beta_i,\gamma_i)=\frac{P(\alpha_i,\beta_i,\gamma_i|S)P(S)}{\sum_{t=1}^{N}P(\alpha_t,\beta_t,\gamma_t|S)P(S)} \quad (1)$$

식 (1)에서 $P(S)$는 사전 확률로서 일반적인 네트워크 상태에서 싱크홀 공격이 발생할 확률을 의미한다. $P(\alpha_i,\beta_i,\gamma_i|S)$는 싱크홀 공격이 발생했을 때 세 가지 파라미터의 값에 따른 조건부 확률을 의미한다. 이러한 사전 확률들은 쉽게 알기 어려우며 훈련 데이터를 통한 샘플 지식을 통해서 유추해야만 한다. 이 문제에 대한 가장 간단한 접근 방식은 샘플들을 이용해서 미지의 확률들과 확률 밀도들을 추정하고, 결과로 얻은 추정들을 참 값으로 사용하는 것이다. 이러한 방식은 복잡한 계산이 필요하지 않기 때문에 실시간으로 확률을 계산하고 침입을 탐지해야 하는 상황에서 유용하게 사용할 수 있다. 따라서 본 연구에서는 사전 확률을 구하기 위해 훈련 데이터로부터 얻은 확률을 그대로 이용한다.

식 (1)의 $P(\alpha_i,\beta_i,\gamma_i|S)$을 계산하기 위해서는 α,β,γ 세 가지 파라미터의 모든 가능한 값에 대한 확률이 계산되어야 한다. 하지만 α,β,γ의 조합은 경우의 수가 너무 많기 때문에 이를 모두 계산하는 것은 불가능하다. 따라서 α,β,γ 값에 따라 구간을 나누어 각 구간에 따른 조건부 확률을 계산하는 것이 효율적이다.

본 연구에서는 α,β,γ의 구간을 나누기 위해 K-평균 군집분석을 이용한다 [8]. 군집분석은 개체들 간의 유사성 또는 비유사성의 정도를 측정하여 개체들을 동질적인 몇 개의 군집으로 분류하는 기법이다. 이 중에서도 K-평균 군집분석은 군집의 수를 알고 있을 때 유용하게 사용할 수 있다. K-평균 군집분석을 이용하면 α,β,γ값을 미리 정의한 K개의 구간으로 분류할 수 있다. K-평균 군집분석에서 유사성의 측정은 다음과 같은 유클리드 거리가 주로 사용된다.

$$Distance(x,y)=\sqrt{\sum_{i=1}^{n}(x_i-y_i)^2} \quad (2)$$

식 (2)에서 x와 y는 유사성을 비교하고자 하는 두 개체이며 n은 각 개체가 가지고 있는 파라미터의 수를 나타낸다.

군집분석 방식은 네트워크 데이터 상에서 α,β,γ 값이 가지는 분포에 기반하여 구간을 정할 수 있는 장점을 가진다. 따라서 임의대로 구간을 나눌 경우 공격 특성을 반영하는 파라

미터 값과 정상 상태의 파라미터 값이 같은 구간 안에 섞일 수 있는 문제점을 해결할 수 있다.

훈련 데이터로부터 K-평균 군집분석을 이용해 나눠진 구간들을 $A_1,...,A_K$라 정의하면 새로운 i 시점에 계산된 파라미터 α,β,γ는 유클리드 거리를 이용해 가장 가까운 구간 A_i에 포함된다. 이에 따라 식 (1)은 다음과 같이 표현할 수 있다.

$$P(S|A_i) = \frac{P(A_i|S)P(S)}{\sum_{k=1}^{K} P(A_k|S)P(S)} \quad (3)$$

식 (3)에서 $P(A_i|S)$는 훈련 데이터 중 싱크홀 공격 데이터가 구간 A_i에 포함되어 있는 비율을 구함으로써 얻을 수 있다. 식 (3)을 통해 현재 네트워크 상에 싱크홀 공격이 일어나고 있을 확률을 계산할 수 있으며, 이렇게 얻어진 싱크홀 공격 발생 확률은 기존의 침입탐지가 공격과 정상으로만 판정을 하는 것에 비해 다양한 값을 가지기 때문에 더욱 세밀한 대처가 가능하다. 또한 공격 발생 확률 정보를 라우팅에 이용함으로써 라우팅의 보안성을 높이는 등 응용 가능성이 높다는 장점을 가진다.

5. 결론

본 연구에서는 베이지안 정리를 이용하여 싱크홀 공격을 탐지하기 위한 기법을 제안하였다. 싱크홀 공격은 공격자가 sequence number를 조작한 거짓 RREQ를 보냄으로써 전체 네트워크의 라우팅이 공격자를 거치도록 하는 방식을 사용한다. 본 연구에서는 싱크홀 공격의 특성을 반영하기 위해 sequence number discontinuity, route add ratio, number of hops의 세 가지 파라미터를 사용하였다. 이들 파라미터로부터 베이지안 정리를 이용해 싱크홀 공격의 발생 확률을 계산하는 알고리즘을 제안하였다. 싱크홀 공격 발생 확률은 싱크홀 공격의 탐지를 위해 사용될 뿐 아니라 ad hoc 네트워크의 안전한 라우팅 설계에 활용될 수 있다.

이후에는 네트워크 시뮬레이션을 통해 본 논문에서 제안한 알고리즘의 성능을 평가하고 보완할 것이다. 또한 사전확률 값 추정의 정확도를 높이기 위해 최대 우도 추정, 베이지안 추정 등의 다양한 방식을 활용할 것이다.

참고문헌

[1] 권혜연, 신재욱, 이병복, 최지혁, 남상우, “이동 Ad-Hoc 네트워크 기술동향”, 전자통신동향분석, 제18권, 제2호, pp.11-24, 2003.

[2] 박영호, 이경근, 이상곤, 문상재, “무선 Ad Hoc 네트워크에서의 안전한 라우팅 프로토콜에 관한 연구”, 한국정보보호학회지, 제15권, 제3호, pp.76-81, 2005.

[3] 김신효, 강유성, 정병호, 정교일, “U-센서 네트워크 보안 기술 동향”, 전자통신동향분석, 제 20권, 제1호, pp. 93-99, 2005.

[4] E. M. Royer, C. K. Toh, "A review of current routing protocols for ad hoc mobile wireless networks", IEEE Personal Communications, pp.46-55, 1999.

[5] H. C. Tseng, B. J. Culpepper, “Sinkhole intrusion in mobile ad hoc networks: The problem and some detection indicators”, Computers & Security, 24, pp.561-570, 2005.

[6] 이강현, “Hop-depth 알고리즘을 이용한 무선 센서 네트워크 상에서의 내부공격자 및 공모노드 검출”, 전자공학회 논문지, 제44권 제1호, pp.113-121, 2007.

[7] R. O. Duda, P. E. Hart, D. G. Stork, Patern Classification, John Wiley & Sons, Incs., 2000.

[8] L. Kaufman, P. J. Rousseeuw, "Finding groups in data: An introduction to cluster analysis", Wiley Series in probability and mathematical statistics, John Wiley & Sons, Incs., 1990.