SVM을 이용한 Rogue AP 탐지 기법

강성배* 최진춘** 변정훈*** 양대헌****

*인하대학교 정보보호연구실

Rogue AP detection using SVM

Sung-Bae Kang* Jin-Chun Choi** Jung-Hun Byun*** Dae-Hun Nyang****

*Information Security Research Laboratory, Inha University.

요 약

스마트기기 들이 널리 보급되고 무선랜의 사용이 보편화 되면서 Rogue AP의 위험성도 더욱 높아지고 있다. 본 논문에서는 이러한 Rogue AP 들을 탐지해 내는 기존의 기법들을 파악해 보고 그중에 H. Han등이 제시한 Rogue AP의 추가적인 무선 구간에 의한 전송 지연을 이용한 탐지 기법을 분석 한 후 제시한 기법의 취약점을 보완하고 Rogue AP그룹 과 Legitimate AP그룹을 SVM을 이용하여 구분 함으로써 Rogue AP 탐지율을 향상 시킨다.

I. 서론

H. Han등은 Rogue AP 와 Legitimate AP 사이의 추가적인 구간의 지연 시간을 이용하여 Rogue AP를 식별하는 방법을 제안하였다. 이 기법은 채널 사용량(Channel Utilization)이 증가할 경우 Rogue AP를 식별해낼 확률이 급격히 떨어지는 단점이 있으며, 고정된 상수를 이용하여 Rogue AP 집단 과 Legitimate AP 집단을 구분해 내는 것이 현실에서 적용되지 않을 수 있다는 단점이 있다.

이 연구에서는 채널 사용량에 관계없이 Rogue AP를 높은 확률로 찾아낼 수 있는 방법을 고안하고, 기존의 무선 네트워크(IEEE 802.11)를 이용하는 Rogue AP뿐만 아니라 이동통신망(HSDPA, Wibro, LTE)을 이용하는 Rogue AP까지 탐지할 수 있도록 설계하였다. 2장 에서는 H. Han등 의 기법[5]을 간단하게 소개하고 해당 기법에 대해 고찰한다. 3장 에서는 SVM을 이용하여 Rogue AP를 탐지하는 기법을 소개하고, 우리의 연구가 기존 기법에 비해 우수함을 실험 결과를 통해 보인다. 마지막으로 4장에서는 이 연구의 결론과 앞으로의 연구 계획을 담는다.

II. 연구 동향

2.1 기존 기법들

2006년 P. Bahl 등은 AP의 SSID와 MAC 주소를 이용하여 정상적인 AP와 로그 AP를 분류하는 기법을 제안 하였다[1]. 이 기법은 정상적인 AP의 SSID와 MAC 주소를 데이터베이스에 저장해 두고, 새로운 AP의 SSID와 MAC 주소가 발견될 경우, 이를 로그 AP로 분류하는 방식이다. 이 기법은 짧은 시간에 로그 AP를 탐지할 수 있는 장점이 있지만, SSID와 MAC 스푸핑(MAC spoofing) 공격에 취약하며, 대규모 네트워크에는 적합하지 않으며, 장비의 이동이 잦은 조직에서는 적용하기 힘든 단점이 존재한다.

2007년 D. Schweitzer 등은 AP의 위치정보를 이용하여 로그 AP를 탐지해내는 기법을 제안하였다[2]. 이 기법은 정상적인 AP의 위치정보를 사전에 기억하고 있으며, 사전에 지정되지 않은 위치에서 AP가 발견될 경우, 이를 로그 AP로 분류하는 방식이다. 하지만 이 기법도 P. Bahl[1] 등이 제안한 기법과 마찬가지로 대규모 네트워크나 장비의 이동이 잦은 경우에 적용하기 힘들다.

2007년 L. Watkins 등은 유선망과 무선망의 응답 시간 차이를 이용하여, 로그 AP를 찾아내는 기법을 제안하였다[3]. 이 기법은 유무선망을 통해 전송되고 있는 패킷들을 수집하고, 수집된 패킷들의 시간 간격을 비교하여 로그 AP를 판별 한다.

2007년 W. Wei등은 TCP의 syn-ack 쌍을 이용하여, 로그 AP를 탐지하는 기법을 제안하였다[4]. 이 기법은 TCP의 발신자가 syn 패킷을 발송하면 수신자가 이에 대한 응답으로 ack을 보내온다는 점을 이용하였다. 발신한 syn 패킷과 응답으로 받은 ack 패킷간의 시간차를 측정하는 방법으로 로그 AP를 판별한다.

2011년 H. Han등은 Rogue AP 와 Legitimate AP 사이의 추가적인 구간의 지연 시간을 이용하여 Rogue AP를 찾아내는 기법을 제안하였다[5]. 그림 1은 H. Han등이 제안한 기법의 기본 아이디어를 보여주고 있다. Rogue AP는 Legitimate AP와 무선으로

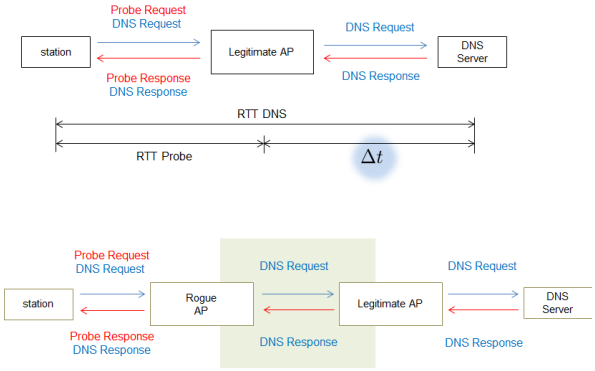

그림 1 H. Han등이 제안한 기법의 기본 아이디어

연결되어 있기 때문에 사용자가 정상적으로 Legitimate AP에 연결된 것과 비교하여 무선 구간히 하나 더 존재한다. 이는 사용자가 알고 있는 무선구간을 제외한 구간의 RTT값을 나타내는 Δt의 값이 Legitimate AP의 Δt값보다 더 높게 나올 수 있음을 의미한다. 이 논문에서는 사용자와 사용자가 연결되어 있는 AP까지의 구간의 RTT를 측정한 RTT_{probe}와 사용자와 Local DNS 까지의 RTT를 측정한 RTT_{dns}값을 이용하여 구한 Δt = (RTT_{dns} - RTT_{probe}) 값을 이용하여 Rogue AP를 식별 한다.

이 기법에서 제안한 Rogue AP 탐지 알고리즘은 그림 2와 같다. 사용자는 AP에 n개의 ICMP패킷 과 DNS 패킷을 전송하여 RTT_{probe}와 RTT_{dns}를 구한다. RTT_{probe}와 RTT_{dns}의 평균을 구하여 사용자가 연결된 AP와 DNS 서버 구간의 RTT 평균값인 Δt = (RTT_{dns} - RTT_{probe}) 값을 구하고 RTT_{probe}와 RTT_{dns}의 표준편차를 구하여 σ_{probe} (RTT_{probe}의 표준 편차)와 σ_{dns}(RTT_{dns}의 표준 편차)를 구한다음

Algorithm 1 Detecting Rogue AP (AP_x)

1: Connect and associate with AP_x
2: **for** $i = 1$ to n **do**
3: Send *unicast* probe request to AP_x, record round trip time $RTT_{probe} = RTT_{probe} - T_{data}(probe)$.
4: Send DNS lookup to local DNS server, record round trip time $RTT_{dns} = RTT_{dns} - T_{data}(dns)$.
5: **end for**
6: Filter out outliers
7: $\overline{RTT_{probe}}$ = Mean of remaining RTT_{probe}
8: $\overline{RTT_{dns}}$ = Mean of remaining RTT_{dns}
9: σ_{probe} = Standard deviation of remaining RTT_{probe}
10: σ_{dns} = Standard deviation of remaining RTT_{dns}
11: $\Delta t = \overline{RTT_{dns}} - \overline{RTT_{probe}}$
12: $\theta = f(\sigma_{probe}, \sigma_{dns})$
13: **if** $\Delta t > \theta$ **then**
14: AP_x is a rogue AP
15: **end if**

그림 2 Rogue AP 탐지 알고리즘

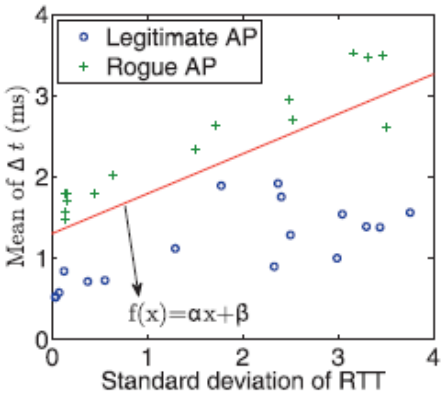

그림 3 구분 방법

$\frac{\sigma_{probe}+\sigma_{dns}}{2}$ 값을 x값으로, Δt 값을 y값으로 갖는 2차원 그래프를 그린다. 그리고 σ_{probe}와 σ_{dns}를 이용하여 직선의 방정식($\alpha \cdot \frac{\sigma_{probe}+\sigma_{dns}}{2}+\beta$)으로 직선을 그림으로써 Rogue AP와 Legitimate AP를 구분한다. 2차원 그래프 상에서 직선보다 위쪽에 있는 AP 의 집합을 A전부 Rogue AP로 식별한.

2.2 H. Han등이 제안한 기법에 대한 고찰

해당 기법에서 제안한 알고리즘에서 RTT_{probe}와 RTT_{dns}를 보내는 시간이 완전 동일하지 않다. 실제 DNS까지의 RTT에서 AP까지의 RTT를 빼야 하는데 두 개의 RTT를 재는 것이 동시에 이루어지지 않는 것처럼 보인다. 또 Rogue AP 그룹과 Legitimate AP 그룹을 구분 하는 직선의 방정식을 구하는데 있어서 사용되는 α=0.49 , β=1.3 이라는 상수 값이 실제로 두 그룹을 정확하게 구분 할 수 있는 상수 값인지 알 수 없다. 만약 해당 상수 값이 두 그룹을 구분 할 수 있는 기준 값이라 하더라도 두 그룹을 구분 하는 기준이 단순히 직선의 방정식으로 가능 할 것인가 하는 의문이 남는다. 따라서 우리는 이러한 RTT_{probe}와 RTT_{dns}를 보내는 시간을 최대한 동기화 하여 같은 네트워크 상황에서의 시간을 측정하며, Δt와 $\frac{\sigma_{probe}+\sigma_{dns}}{2}$ 값에 대한 상관관계에 대하여 SVM을 이용하여 훈련을 시키고 두 그룹을 구분하게 한다.

III. SVM을 이용한 Rogue AP 탐지 기법

3.1 실험환경

그림 4에서 보듯이 Legitimate AP 는 항상 최적의 채널을 검색 하여 준다고 가정 한다. Rogue AP 도

그림 4

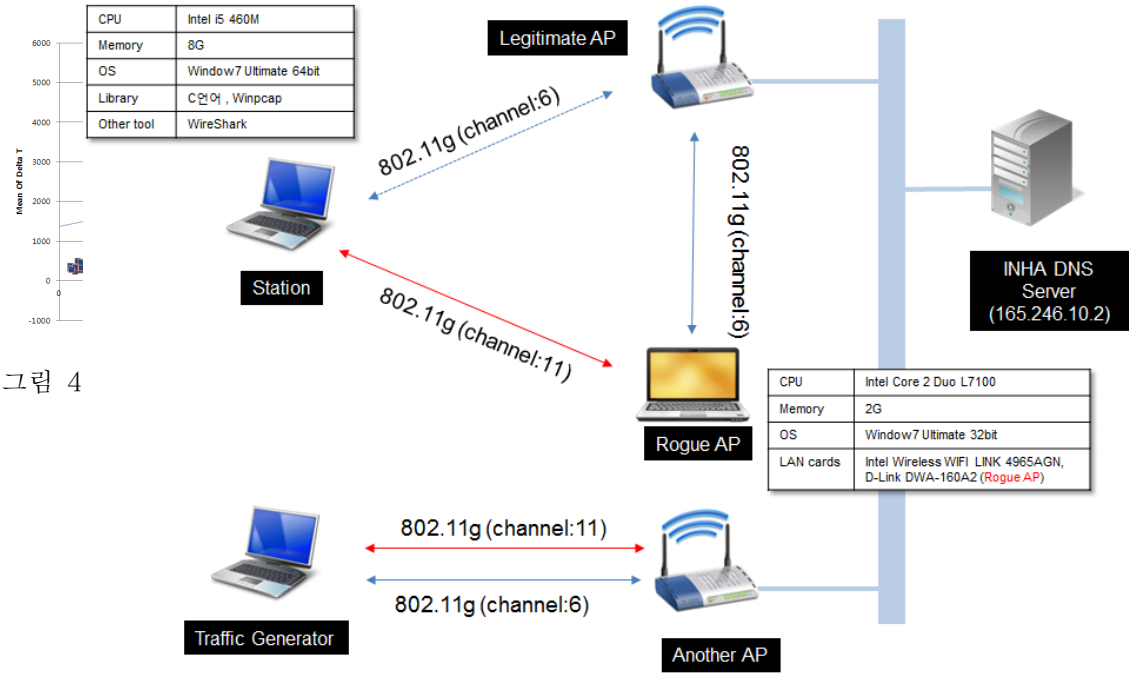

그림 5 실험 환경

사용자의 연결을 유도하기 위하여 최적의 채널을 검색하여 인터넷 서비스를 제공하려고 할 것이다. 그러나 사용자가 연결되어 있는 AP 구간의 Channel 이 연결 후에 혼잡하게 되면 Legitimate AP를 Rogue AP 로, Rogue AP를 Legitimate AP로 잘못 탐지 하게 되는 긍정오류와 부정오류 가 발생 할 수 있다. 따라서 이러한 Channel 이 혼잡하게 되는 상황까지 가정 하며 실험하며, Traffic을 혼잡하게 만드는 상황은 같은 채널을 사용하는 또 다른 AP를 사용함으로써 사용자가 접속한 AP 에 처리량이 증가함으로써 생기는 전송 지연을 피하게 한다. DNS쿼리를 처리하는 DNS서버로는 Local DNS서버를 사용하도록 한다.

3.2 실험결과

3.2.1 기존 논문의 알고리즘 적용 결과실험 결과

값$\triangle t$(사용자가 연결된 AP 뒤쪽의 유선랜 구간의 RTT) 와 $\frac{\sigma_{probe}+\sigma_{dns}}{2}$ 에 대한 값으로 기존 논문의 알고리즘을 그대로 적용한 결과 그림 5 와 같이 단순한 직선으로 Rogue AP 그룹과 Legitimate AP 그룹을 구분 하는 방법은 83% 의 확률로 두 그룹을 구분 하였다.

3.2.2 SVM을 이용한 Rogue AP 탐지 결과

기존 논문의 알고리즘에 적용한 데이터와 동일한 데이터를 사용한다. Legitimate AP로부터 측정한 결과 값 74개 중 64개의 데이터는 SVM 트레이닝용으로 사용 되었고 10개의 데이터는 구분용으로 사용 하였다. Rogue AP로부터 측정한 결과 값 79개 중 69개의 데이터는 SVM 트레이닝용으로 사용되었고 10개의 데이터는 구분용으로 사용 되었다. 실험결과 SVM 은 두 그룹을 구분하는데 90% 확률로 구분 하는 것을 볼 수 있다.

```
jin@ubuntu:~/libsvm-3.12/test/2$ svm-train training_data_scale
*
optimization finished, #iter = 66
nu = 0.970149
obj = -110.925393, rho = 0.123143
nSV = 131, nBSV = 129
Total nSV = 131
jin@ubuntu:~/libsvm-3.12/test/2$ svm-predict predict_scale training_data_scale.model output
Accuracy = 90% (18/20) (classification)
jin@ubuntu:~/libsvm-3.12/test/2$ cat output
1
1
1
1
1
-1
1
1
-1
1
-1
-1
-1
-1
-1
-1
-1
-1
-1
jin@ubuntu:~/libsvm-3.12/test/2$
```

그림 6 SVM을 이용한 Rogue AP 탐지 결과

IV. 결론

지금까지 Rogue AP 탐지 기술과 기존 기법과 H. Han등이 제안한 기법의 구현과 분석을 통해 해당 기법에서 Rogue AP와 Legitimate AP를 구분하는데 사용되는 알고리즘 이 실용적이지 못함을 발견하였다. 이러한 문제를 해결하기 위해 이 연구에서는 SVM을 이용한 Rogue AP 탐지 기법을 개발하였다. 기존의 기법의 경우 무선 채널 사용량이 증가할 경우 샘플링n 의 개수를 늘림으로써 탐지율을 80%까지 높일 수 있음을 보이고 있지만, 이 연구에서 제안한 바와 같이 SVM의 커널 함수를 활용할 경우에는 채널 사용량과 관계없이 약 90%의 탐지율을 보여주었다.

이 연구의 목적은 기존의 기법의 문제점을 발견하는 것이 아닌, Rogue AP를 효율적으로 탐지할 수 있는 방법론은 개발하는 것이다. 앞으로 심층적인 분석을 통해 Rogue AP를 식별하는데 유용하게 사용될 수 있는 정보들을 도출하고 이를 SVM의 입력 Factor로 적용하여, 무선 채널의 사용량이 높은 환경에서도 90%이상의 탐지율을 보이는 기법으로 발전시킬 것이다.

[참고문헌]

[1] P. Bahl, R. Chandra, J. Padhye, L.Ravindranath, M. Singh, A. Wolman and B. Zill, “Enhancing the security of corporate Wi-Fi networks using DAIR”, MobiSys, pp. 1-14, Jun. 2006

[2] D. Schweitzer, W. Brown and J. Boleng, “Using visualization to locate rogue access points”, Journal of Computing Sciences in Colleges, pp. 134-140, Oct. 2007

[3] L. Watkins, R. Beyah, and C. Corbeet, “a passive approach to rogue access point detection”, IEEE Global Telecommunications Conference, pp. 355-360, Nov. 2007

[4] W. Wei, K. Suh, B. Wang, Y. Gu and J. Kurose,“Passive Online Rogue Access Point Detection Using Sequential Hypothesis Testing with TCP ACK-Pairs”, The 7thACMinternetmeasurementconference,pp.365-378,Oct.2007

[5] H. Han, B. Sheng, C. C. Tan, Q. Li and S. Lu, “A Timing-Based Scheme for Rogue AP Detection”, IEEE Transactions on Parallel and Distributed Systems, pp. 1912-1925, Nov. 2011