Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.24, NO.1, Feb. 2014 ISSN 2288-2715(Online)
http://dx.doi.org/10.13089/JKIISC.2014.24.1.87

k-SVM을 이용한 Rogue AP 탐지 기법 연구

이 재 욱,[1†] 이 시 영,[1] 문 종 섭[2‡]
고려대학교 정보보호대학원

Detecting Rogue AP using k-SVM method

Jae-wook Lee,[1†] Si-young Lee,[1] Jong-sub Moon[2‡]
Graduate School of Information Security, Korea University

요 약

인가된 AP(Access Point)에 대해서만 사용을 허용하는 무선 통신 환경에서, 스마트폰의 테더링(tethering) 기능에 의한 로그 AP(rogue AP) 사용은 자료 유출과 같은 심각한 보안 문제를 발생시킨다. 이에 본 논문은 홉(hop) 간 RTT 값을 특징점(Feature)으로 하고 분류기로 k-SVM(Kernel Support Vector Machine)를 사용하여 로그 AP를 탐지하는 방법을 제안한다. 실험을 통해 유선 네트워크를 이용하여 설치한 일반 AP와 LTE 망을 활용하여 설치된 로그 AP 간의 구분이 가능함을 보였다.

ABSTRACT

Under only authorized AP is allowable environment, rogue AP which is generated by a smartphone tethering can be a serious security breach. To solve rogue AP problem, this paper proposes classifying algorithm of Kernel Support Vector Machine using features of RTT data. Through our experiment, we can detect rogue AP from LTE mobile network.

Keywords: Rogue AP, Network, SVM, LTE

I. 서 론

업무 효율성과 편의성을 극대화시키기 위해 많은 기업들이 AP(Access Point)를 이용해 무선 인터넷을 통한 네트워크 접근환경을 제공한다. 무선 인터넷 환경에서는 네트워크상에서 전달되는 패킷을 외부에서 모을 수 있기 때문에 이를 대비해 IEEE 802.11i[1] 또는 WPA와 같이 보안을 고려한 통신 환경을 사용한다.

IEEE 802.11i 프로토콜은 인가되지 않은 사용자들이 무선 네트워크를 사용하는 것으로부터 막기 위해 통신 내용의 암호화 및 접근을 위한 사전 인증 과정을 제공한다[2]. 하지만, 이런 보안 메커니즘을 사용하는 환경에서도 로그 AP와 같은 우회방법을 통하면 사용자의 AP 접근을 제어할 수 없다. 이 논문에서는 로그 AP를 모바일 기기에 AP 기능을 제공하는 소프트웨어를 사용하여 AP의 기능을 수행하는 단말기를 의미한다.

로그 AP는 무선 네트워크에서 중대한 보안 취약점을 제공한다. 로그 AP를 설치한 공격자는 로그 AP에 접근한 사용자의 기기들을 대상으로 패킷을 수집하여 개인 정보를 모으고 중간자 공격(man-in-the-middle attack)을 수행하여 웹 페이지 변조, 세션 하이재킹, 인증서 도난 같은 피해를 발생 시킬 수 있다[3].

로그 AP를 설치하기 위해서는 과거에는 별도의 무선 네트워크를 구축해야 하므로 공격이 어려웠지만, 현재 스마트폰과 태블릿 PC 의 확산으로 3G 또는 LTE 네트워크를 사용해서 손쉽게 모바일 AP를 설치

접수일(2013년 9월 27일), 수정일(2013년 12월 26일), 게재확정일(2013년 10월 21일)
† 주저자, cakel@korea.ac.kr
‡ 교신저자, jsmoon@korea.ac.kr (Corresponding author)

할 수 있게 되었으며 이렇게 만들어진 모바일 AP는 기존의 AP들과 달리 내부 인증 체제를 우회하여 외부 네트워크와 쉽게 연결 할 수 있으므로 외부로의 자료 유출을 차단하는 보안 정책 속에서도 모바일 AP를 사용하여 제한 사항들을 무력화 시킬 수 있다. 또한 사용자가 악의적으로 모바일 AP를 생성하지 않았더라도 잠재적인 보안 취약성들로 인해 공격 대상이 될 수 있다[4].

현재 상용화 된 로그 AP 탐지 방법들 대부분은, 서버에서 로그 AP를 탐지하는 방법이여서 AP 와 클라이언트를 통제하기 위해 네트워크 전반에 걸쳐 적용하기 힘든 요구사항들이 발생되었다. 그리고 로그 AP는 서버의 통제를 받지 않고 사용자를 공격할 수 있기 때문에 클라이언트 측면에서의 로그 AP를 탐지할 수 있는 해결책이 필요하다.

이에 본 논문은 클라이언트에서 개선된 RTT(Round Trip Time)을 원 데이터로 사용하여, 특징점들을 도출하여 수학적으로 가장 뛰어난, 기계학습의 일종인 k-SVM을 이용해 로그 AP를 탐지하고자 한다.

본 논문은 다음의 내용으로 구성되어 있다. 2장에서 관련연구를 통해 현재까지 진행되어 오고 있는 로그 AP 의 탐지 방법과 3장에서 제안 방법으로 사용되는 기계학습 알고리즘인 k-SVM에 대해 알아보고 4장에서 제안하는 알고리즘을 통해 효과적 로그 AP를 탐지하는 방법과 5장에서 로그 AP를 탐지한 결과를 확인하고, 마지막 6장에서 향후 과제를 정리한다.

II. 기존 연구 - 로그 AP를 탐지하는 방법

비 인가된 로그 AP를 찾는 이전 연구들의 대부분은 MAC 주소와 같은 AP 자체의 정보를 이용하는 방법[5-13]과 유무선 네트워크 간의 특성을 이용한 방법[14-17], 두 가지 방법으로 나누어진다.

2.1 AP 의 정보를 이용하는 방법

로그 AP를 찾는 방법으로 MAC 주소, SSID 나 제조업체와 같은 AP에 기재된 하드웨어 정보를 이용한 방법이 있다. 이 방법은 인가된 AP들에 대해 그들의 하드웨어 정보들을 저장하고 이 정보들을 바탕으로 확인되고 있는 시스템에 연결된 AP들을 주기적으로 검사하여 로그 AP를 탐지한다. [5]에 의하면 무선 네트워크를 위한 프레임워크인 DAIR(Dense Array of Inexpensive Radios) 시스템에서 에어 모니터(Air Monitor)라고 불리는 USB 감지기를 사용해 AP 정보를 수집하고 이를 데이터베이스(DB)에 저장하여 관리하는 방법을 제안하였다. [6]에서는 MAC 주소, SSID, 채널, 그리고 신호 세기와 같은 정보로 비 인가된 로그 AP를 탐지하는 방법을 제시하였다.

현재 에어디펜스(Airdefence)[7], 넷스텀블러(Netstumber)[8], 로그스캐너(Roguescanner)[9]과 같은 하드웨어 정보 기반 스캐너들이 존재하며 이들은 사용자에게 위협이 될 수 있는 AP들의 목록을 생성하여 사용자에게 제공한다. 이들의 장점은 짧은 시간에 정보를 제공할 수 있고 낮은 오탐율을 가지고 있다는 것이다. 하지만, 이런 네트워크 스캐너들이 가지는 단점으로 AP들을 주기적으로 검사를 해야 하며 AP에 저장된 하드웨어 정보는 악의적인 공격자에 의해 변형되어 검사 시 이를 우회할 수 있다.

다른 방법으로 [10-12]에서와 같이 감지기를 통한 위치 기반 정보로 로그 AP를 탐지하는 방식이다. AP의 존재가 예상하지 못한 곳에서 확인 되었을 때, 이를 로그 AP로 탐지하는 것이다. 현재의 위치 판단 기술을 사용하면 3-5m 의 정확도를 가지고 무선 AP 위치를 판단 할 수 있다[13]. 문제는 위치 측정을 위해 추가적인 탐지 장비가 필요하며 AP의 위치 이동이 활발한 대규모 네트워크나 조직에서는 이를 적용하기가 힘들다. 그리고 단순히 한 위치에서만 국한되어서 동작하지 않은 모바일 AP의 특성상 위치 기반 정보를 가지고 로그 AP를 탐지한다는 것은 탐지의 의미가 없다.

2.2 유선과 무선네트워크의 특징을 이용하는 방법

네트워크의 특정 정보를 사용해 로그 AP를 탐지하는 다른 방법으로 대개의 로그 AP들은 이미 존재하는 일반 AP나 라우터에 연결되어 사용하므로, 추가적인 홉(hop)이 존재하는지의 여부를 검사하는 방식이다. 추가 홉이 있는지 여부를 판단하기 위해 [14]에서는 실제 네트워크상에서 추출한 2개의 패킷 간의 시간차를 이용했다. 이와 유사하게 [15,16] 에서는 TCP-ACK 패킷 쌍의 시간차를 사용하였다. 좀 더 탐지율을 높이기 위해 [17]에서는 DNS (Domain Name System) 서버로 전달하는 임의의 패킷들로 인해 발생되는 RTT값을 계산하여 특정 기준을 넘었을 때 이를 로그 AP로 탐지하는 방법을 사용했다. 이 방법을 통하면 AP의 정보를 기록하는 데이터베이스를 유지할 필요가 없지만, 지정된 DNS 서버와 같은

특정 측정 대상 서버를 지정해야 하는 부담이 생긴다. 그리고 RTT값을 조작해 탐지를 우회하는 것을 어렵게 할 수 있으나, LTE 망과 802.11ac 과 같이 고속 무선 네트워크를 사용하는 모바일 AP 사용 환경에서 RTT값을 효과적으로 선정, 계산하지 않으면 로그 AP를 탐지할 수 없는 문제점이 존재한다.

III. SVM(Support Vector Machine)

3.1 기계학습

기계학습(machine learning)은 분류할 데이터를 사용하여, 파라미터를 설정하는 방법을 의미한다. 즉, 이들 데이터를 사용하여 학습하고 학습에 사용되지 않은 데이터를 탐지/분류하는 방법이다[18]. 기계학습은 데이터 학습 과정을 통해 시스템 모델을 구성하는 표상화(representation) 과정과 구성된 시스템이 확인되지 않은 데이터를 분류하는데 적절한지의 여부를 관여하는 일반화(generalization) 과정이 포함된다[19].

3.2 SVM 개요

전통적인 학습 접근 방법은 학습 데이터-셋(data-set)에 대한 오차(error)를 줄이는 것에 기반하는 경험적 위험 최소화(ERM)를 목표로 하고 있다. SVM은 이와 반대로 구조적 위험 최소화(SRM)를 기반으로 하는 통계적 학습 이론을 기반으로 하고 있으며, 또한 신경망 네트워크(neural network) 알고리즘에 비해 확인되지 않은 시험 데이터-셋에 대해 더 좋은 일반화 능력을 가지고 있다[20].

SVM의 기본 원리는 SVM 입력 패턴들을 교사학습방법(supervised learning)을 통해 두 개의 클래스인 {+1, -1} 로 분류하는 것으로 시작된다. 훈련 집단이 두 클래스로 분류가 되면, 이를 포함하는 훈련 패턴들을 분리하는 초월면(hyperplane)을 결정할 수 있다. 이때의 초월면이란 각 집단을 분리 할 수 있는 절단 평면을 말한다. 이 평면을 결정하는 경계 패턴들을 서포트 백터(support vector) 라고 한다. 모든 서포트 백터는 초월면으로부터 같은 최소 거리에 위치해야 하지만, 선형으로 분리되는 경우가 거의 드물기 때문에 이때의 초월면과 서포트 백터는 제약 식(constraint equation)을 갖는 최적 문제의 해로부터 구해낼 수 있다. 최적 해는 각 클래스의 서포트 백터 사이의 거리인 마진(margin)을 최대화 하는 것과 오차를 최소화 하는 것인데 둘 사이에는 트레이드-오프(trade-off) 관계를 가지고 있으며, 이는 정규화된 파라미터에 의해 조정된다. 각 훈련 과정은 제약식을 갖는 이차 최적 문제 (quadratic programming)를 해결하는 방법과 기본적으로 같다[21-23].

3.3 분류를 위한 SVM

훈련 데이터 $\{(x_i, d_i), i = 1, 2, \ldots, N\}$ 으로 주어졌다면, x_i 는 두 클래스 {+1 , -1} 중 하나에 속하며, $d_i \in \{-1, 1\}$ 은 해당 클래스를 표시하는 라벨 역할을 수행한다. SVM은 각 클래스들을 구분하는 최적의 분리 경계면에 인접한 점들과의 거리(margin)을 최대화 한다. 따라서 최적의 분리 경계면을 f(x) = $w^t x + b$ 로 지정하면, f(x) 와 서포트 백터 간의 거리를 $1/\|w\|$ 로 표시할 수 있다. SVM은 $\|w\|^2$ 값을 최소화 하여, 분리 간격을 최대화하는 최적 분리면을 찾는다. 이 문제는 다음과 같은 식으로 표현 할 수 있다.

$$\begin{cases} \text{minimize} & \frac{1}{2}\|w\|^2 \\ \text{subject to } d_i(w^T x_i + b) \geq 1 & \text{for } i = 1, \ldots, N \end{cases} \quad (1)$$

이 문제를 라그랑주 배수 (lagrange multipliers) 로서 쌍대화(dual problem) 시키면 아래의 이차 문제가 된다.

$$\Theta(a) = \sum_{i=1}^{N} a_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j d_i d_j \langle x_i, x_j \rangle$$
$$\text{subject to } a_i \geq 0 \;\; i = 1, \ldots, N \text{ and } \sum_{i=1}^{N} a_i d_i = 0 \quad (2)$$

선형 분리경계면으로 완전히 분리할 수 없는 서로 겹쳐져 있는 패턴의 경우에는 슬랙 변수(slack variable)(ξ)을 이용한다. 이를 통해 식(1)에서 아래의 모델과 같이 표현 될 수 있다.

$$\text{minimize} \tau(w, \xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \quad (3)$$
$$\text{subject to } d_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i, \xi_i \geq 0 \;\; , i=1, \ldots, N$$

위 식(3) 의 $d_i(\langle w, x_i \rangle + b) \geq 1 - \xi_i$ 에서 $\xi_i = 0 (\forall_i)$ 이면 모든 패턴을 완전하게 분리 할 수 있다는 것을 의미한다. 그러나 대부분의 패턴은 선형적(linear)으로 분리가 되지 않게 되어 있다. 따라서 비선형

(nonlinear) 패턴을 분리하기 위해 비선형 패턴의 입력 공간을 선형 패턴의 특징 공간(feature space)으로 전환한다.

즉, $x_i \Rightarrow \Phi(x_i)$ 에서 커널 함수(kernel function) $K(x_i, x_j) = \Phi(x_i) \bullet \Phi(x_j)$ 를 정의하면 비선형 패턴을 분리하기 위한 모델은 식(1),(2),(3)으로부터 아래와 같이 표현된다.

$$\Theta(a) = \sum_{i=1}^{N} a_i^N - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_i a_j d_i d_j K(x_i, x_j)$$

$$\text{subject to} \quad \sum_{i=1}^{N} a_i d_i = 0, \quad 0 \le a_i \le C, \forall i \qquad (4)$$

여기서 C 는 식 (3) 에서의 패널티 파라미터(penalty parameter) 이다. 위의 모델에서 라그랑주 배수 i 를 구하면 특징 공간에서 가장 평평한 함수인 아래 식 (5)를 구할 수 있다.

$$f(x) = sgn(\langle w, \Phi(x) \rangle + b) = sgn\left(\sum_{i=1}^{N} a_i d_i K(x_i, x_j) + b\right) \qquad (5)$$

k-SVM에서 사용하는 커널 함수로는 여러 가지가 있지만 일반적으로 다음과 같은 함수들 중 선택 될 수 있다[24].

이렇게 최적화된 시스템에서 w를 구하면 식 (5)에 의해, 임의 x는 {-1, 1} 중 하나로 분류 된다.

Table 1. Kernel function

커널 함수 (Kernel Function)	내적(Inner Kernel Product)
폴리노미널 커널 (Polynomial Kernel)	$K(x_i, x_j) = (x_i^T x_j + 1)^d$
가우시안 커널 (Gaussian Kernel)	$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{2\sigma^2})$
라플라시안 커널 (Laplacian Kernel)	$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|}{\sigma})$
다층 퍼셉트론 (Multi-layer perceptron)	$K(x_i, x_j) = \tanh(\beta_o x_i^T x_j + \beta_1)$ β_0와 β_1 값은 사용자가 결정

IV. 제안하는 로그 AP 탐지 방안

4.1 탐지 방법 개요

본 장에서는 일반 AP와 로그 AP 간에 존재하는 RTT값을 사용하여 추출한 특징값을 사용하고, 분류 방법으로는 k-SVM 학습 방법을 제안한다. 정상 AP의 RTT값과 로그 AP의 RTT값을 구분하기 위해 여러 SVM 커널 함수들을 적용하여 최적의 분리 경계면을 찾는다. 이 과정에서 로그 AP와 일반 AP를 구분하는 것과 같은 결과를 얻게 된다. 이후에 설명할 트레이스라우트(traceroute) 수행 정보에서 얻게 될 RTT값을 통해 정상 AP 의 RTT값과 로그 AP 의 RTT값을 구분할 수 있는 방법을 제시한다.

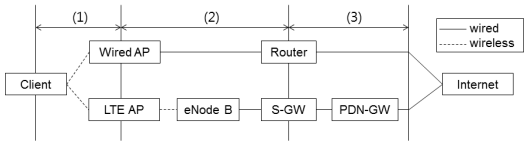

Fig.1. Diagram of difference between normal AP and rogue AP using LTE

Fig. 1.은 유선 네트워크로 연결된 일반 AP와 로그 AP로 정의한 LTE 망을 사용한 모바일 AP 간의 차이를 보여준다. 두 AP 모두 공통적으로 (1) 영역에서 와이파이(Wi-Fi)를 통해 클라이언트와 AP 간에 무선으로 연결 된다. (2) 영역에서 일반 AP 는 유선을 사용하는 네트워크(통상적으로 라우터)에 연결되는 반면, 모바일 AP는 기지국과 게이트웨이(gateway) 등 두 개의 노드를 걸쳐 일반 공중 네트워크에 바로 연결되어 있다[25]. 이로 인해 IEEE 802.11 무선 네트워크(Wi-Fi) 의 지연시간(latency) 은 약 1ms 이며 이어서 연결되는 유선 네트워크 또한 비슷한 지연 시간을 가진다[26]. 반면 LTE 망에서의 지연 시간은 10-20ms 정도[27]로 와이파이 망과 차이가 있다. 또한 Fig. 1.에서 LTE 망에서의 eNodeB 과 같은 무선 기지국이 존재함으로 인해 이런 차이는 더욱 분명해진다. 따라서 (2) 영역에서 유선망으로 연결된 AP에 비해 로그 AP는 인터넷 프로토콜(IP) 를 지원하는 일련의 노드들을 지나면서 지연시간이 증가하게 된다.

4.2 특징점 선정(Feature Selection)

4.1 에서 본 것과 같이, 로그 AP 와 일반 AP 간의 지연시간 차이는 평균값과 분산이 로그 AP 가 일반 AP 보다 크다는 것을 의미한다. 다시 말해, RTT값이 무선 기지국을 통해 증가하고, 신호 상태에 따라 일정치 않게 되어 로그 AP는 일반 AP와 구분 된다.

RTT값은 트레이스라우트의 수행결과[28]에서 얻을 수 있다. 이 과정에서 각 홉당 패킷 수신 시간을 n회 측정 하여 이들의 평균과 분산을 특징점으로 한다. 로그 AP가 일반 AP 와 달리 무선 기지국에 의해 영향 받기 시작하는 첫 번째 홉의 RTT값과 두 AP를 통해 공통으로 응답 받는 최소 홉 위치에서 측정되는 RTT값의 평균과 분산 각각의 차이를 특징점으로 지정한다.

4.3 RTT 측정 알고리즘

Table 2. RTT estimation and data-set generation algorithm

1. Execute **Traceroute** to **Public Domain** M times
2. Find **common least hop count** (N) except 1
3. for each i in *traceroute results* (M) do
4. for j ∈ n (n: try time)

Extract $\overline{RTT_{i\,hopN}} = \sum_{j=1}^{n} RTT_{(i,j)hopN} / n$

Extract $\overline{RTT_{i\,hop1}} = \sum_{j=1}^{n} RTT_{(i,j)hop1} / n$

end of for j

$\Delta Avg_i = \overline{RTT_{i\,hopN}} - \overline{RTT_{i\,hop1}}$,

$\Delta Var_i = \sum_{j=1}^{n} (RTT_{(i,j)hopN} - \overline{RTT_{i\,hopN}})^2/n - \sum_{j=1}^{n} (RTT_{(i,j)hop1} - \overline{RTT_{i\,hop1}})^2/n$

end of for i

5. Data-Set $\{(\Delta Avg_i, \Delta Var_i) \mid i \le M\}$

*$RTT_{(i,j)hopN}$= M 개의 트레이스라우트 결과들에서 i 번째 결과 중, 홉 N 에서 n회 시간측정 하여 얻은 RTT 값들에서의 j 번째 측정된 RTT 값

Table 3. Feature selection using RTT value

항목	특징점 개수	특징점 설명요약
RTT	2	$\Delta Avg = \sum(RTT_{hopN} - RTT_{hop1})/n$ $\Delta Var = \sum(RTT_{hopN} - \overline{RTT_{hopN}})^2/n - \sum(RTT_{hop1} - \overline{RTT_{hop1}})^2/n$

LTE 망을 사용하는 로그 AP를 탐지하기 위한 알고리즘은 Table 2.과 같다. 먼저 공인된 IP를 대상으로 트레이스라우트를 수행하여 해당 호스트로 전달되는 Packet의 경로와 응답 시간을 기록한다. 라우터에서 ICMP TTL expired packet을 전달하지 않는 홉이 있을 수 있으므로 홉 1 이후로 이 패킷을 받을 수 있는 최소 홉 N을 찾는다. 일반 AP 와 로그 AP 간 공통으로 확인되는 최소 홉 N 을 찾았으면, 홉 1과 홉 N간에 측정된 n개의 RTT값의 평균과 분산 차이를 각각 구한다. 이 두 홉 간의 평균과 분산의 차이를 학습하여 로그 AP를 탐지한다.

V. 로그 AP 및 일반 AP 탐지 실험 및 결과

5.1 특징점 추출(Feature Extraction)

학습 및 테스트에 사용되는 RTT값은 네트워크를 사용하는 동안 지속적으로 습득할 수 있다. 데이터 습득이 많아지면 실험의 정확성이 높아지지만 공간 효율이 떨어지며 SVM의 특징상 작은 데이터에도 효과적으로 동작하므로[29] Table. 4.와 같이 학습 및 테스트를 위해 각 AP 당 지정된 데이터들을 수집하여 이를 데이터-셋으로 구성하고 실험하였다. 평균 및 분산을 계산하기 위해 각 데이터당 추출 회수 n은 3을 사용하였다.

Table 4. SVM test data-set using RTT

데이터-셋 / AP 종류	학습 데이터-셋	테스트 데이터-셋
정상 AP	U+Zone 유선 인터넷을 연결한 와이파이에서의 트레이스라우트 정보 (300개)	지역방송 인터넷망을 연결한 와이파이에서의 트레이스라우트 정보 (250개)
로그 AP	SKT LTE 모바일 AP를 사용한 트레이스라우트 정보 (300개)	각 통신사별 LTE 모바일 AP를 사용한 트레이스라우트 정보 (250개)

5.2 로그 AP 탐지 실험 환경

로그 AP 탐지 실험에 앞서 k-SVM을 실행하기 위한 데이터-셋을 구성하기 위해 Table 4.의 데이터 수집방안과 Fig. 2.의 실험 방안을 정했으며, 데이터 수집 과정에 로그 AP와 일반 AP 모두 MS 윈도우(MS Windows) 에서 제공하는 트레이스라우트 도구인 tracert.exe 를 사용하였다.

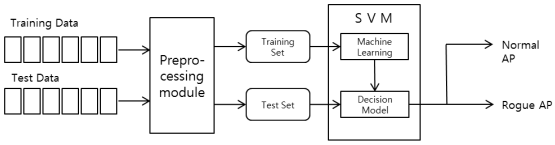

Fig.2. Rogue AP detection using SVM

본 실험에서는 구성된 데이터-셋을 대상으로 k-SVM을 수행하기 위해 Accord.Net[30] 프레임워크(framework)에서 제공하는 SVM 라이브러리(library)를 사용하였으며, 커널 함수에서 사용하는 파라미터는 폴리노미널(polynomial)을 제외하고[31]를 이용하여 계산하였다.

실험 데이터는 아래 Table 5.와 같은 장비를 사용했으며 로그 AP는 LTE 스마트폰의 테더링 기능을 이용해 만들어진 모바일 AP를 통해 수집된 데이터를 사용하였다.

Table 5. Testing device

AP 분류	장비	세부 정보
일반 AP	Goldstar (U+Zone AP)	802.11n
	IPTime N104M	802.11g
LTE 망을 사용하는 AP	삼성 갤럭시 노트2 (SKT)	Android 4.1.2 802.11n
	삼성 갤럭시 S3 (KT)	Android 4.1.2 802.11n
	LG Optimus LTE2 (U+)	Android 4.1.2 802.11n

5.3 실험결과

홉 간 응답시간의 연관성을 고려하여 일반 AP와 로그 AP 간의 특성을 학습한 뒤, 이후 각 시험 데이터를 대상으로 탐지 과정을 수행 하였다. 실험 결과 라플라시안 커널(Laplacian kernel)을 사용하여 형성된 k-SVM의 Fig. 3.과 같이 오류 없이 100% 두 그룹을 분리할 수 있는 경계면이 생성되었으며, Fig. 4.과 같은 초월면을 가질 수 있었다.

또한, Table 6.은 국내 대표 통신사들의 데이터를 사용하여 k-SVM으로 탐지 시험한 결과를 보여 준다.이 때 사용한 커널은 학습 데이터 생성 시 사용했던 파라미터를 사용하였으며 이 결과에서도 라플라시안 커널이 비교적 잘 맞는 것으로 확인되었다.

Table 6. Rogue AP detection result

커널	테스트 데이터-셋 1 (SKT)			테스트 데이터-셋 2 (KT)			테스트 데이터-셋 3 (U+)		
	FP	FN	TC	FP	FN	TC	FP	FN	TC
가우시안	0.0	8.0	92.0	0.0	8.0	92.0	0.0	8.0	92.0
폴리노미널	7.2	7.0	85.8	0.8	7.0	92.2	1.0	7.0	92.0
라플라시안	0.0	6.6	93.4	0.0	6.6	93.4	0.0	6.6	93.4

* Negative: Rogue AP, Positive: Normal AP, Polynomial Kernel의 Degree = 3, FP: False Positive(%), FN: False Negative(%), TC: Total Correctness(%)

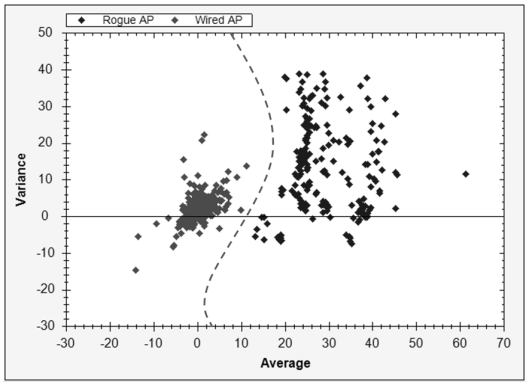

Fig.3. Result of learning data

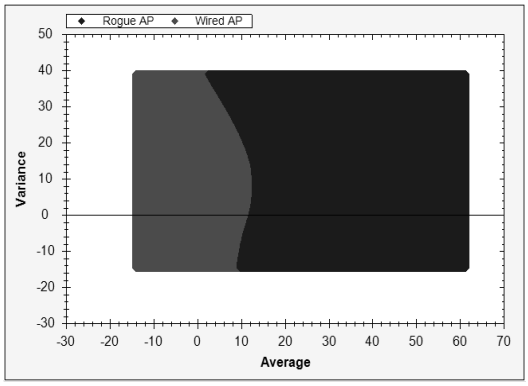

Fig.4. Generated hyperplane by SVM using Laplacian kernel

VI. 결 론

본 논문에서는 LTE망을 사용하는 로그 AP를 탐지하기 위해 RTT값을 특징점으로 하는 k-SVM 알고리즘 제안하였다. 제안한 알고리즘을 통해 수집한 데이터를 정해진 분류기를 이용하여 분류 및 학습을

수행하였고 학습 과정을 수행한 뒤 실험 데이터를 통해 로그 AP 탐지 할 수 있음을 확인 하였다. 향후 보다 다양한 통신 환경에서 생성되는 로그 AP에 대응하기 위해 라우팅 알고리즘과 SVM이외에 신경망과 같은 기법에 대한 비교 검증 실험과 학습 및 테스트 셋을 확장하고 제약 조건 인자에 대해 더 정확하고 유연한 결과 도출을 위한 연구가 필요하다.

References

〔1〕 IEEE P802.11i/D10.0, "Medium Access Control(MAC) security enhancements, amendment 6 to IEEE Standard for local and metropolitan area networks part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications," Apr. 2004.

〔2〕 Changhua He and John C. Mitchell, "Security Analysis and Improvements for IEEE 802.11i," The 12th Annual Network and Distributed System Security Symposium (NDSS'05) Stanford University, pp. 90-110, Feb. 2005.

〔3〕 Henry Paul S and Hui Luo, "WiFi: what's next?," Communications Magazine IEEE, Vol. 40, Issue 12, pp. 66-72, Dec. 2002.

〔4〕 Tarek S. Sobh, "Wi-Fi Networks Security and Accessing Control," IJCNIS, Vol. 5, No. 7, pp.9-20, Jun. 2013.

〔5〕 Paramvir Bahl et al., "Enhancing the security of corporate Wi-Fi networks using DAIR," Proceedings of the 4th international conference on Mobile systems, applications and services, pp. 1-14, Jun. 2006.

〔6〕 Daisuke Takahashi et al., "IEEE 802.11 user fingerprinting and its applications for intrusion detection," Computers & Mathematics with Applications, Vol. 60, Issue 2, pp. 307-318, Jul. 2010.

〔7〕 AirDefense Services Platform - Motorola Solutions USA, http://www.airdefense.net

〔8〕 The award-winning wireless networking tool and the best source for your daily Wi-Fi, WiMAX, 3G and VoIP news. NetStumbler, http://www.netstumbler.com/

〔9〕 RogueScanner Free System Administration software downloads at SourceForge.net, http://sourceforge.net/projects/roguescanner/

〔10〕 Dino Schweitzer, Wayne Brown and Jeff Boleng, "Using visualization to locate rogue access points," Journal of Computing Sciences in Colleges, Vol. 23, Issue 1, pp. 134-140, Oct. 2007.

〔11〕 Payal Bhatia, Christine Laurendeau, and Michel Barbeau, "Solution to the wireless evil-twin transmitter attack," Risks and Security of Internet and Systems (CRiSIS2010), pp. 1-7, Oct. 2010.

〔12〕 Kuo-Fong Kao et al., "A location-aware rogue AP detection system based on wireless packet sniffing of sensor APs," Proceedings of the 2011 ACM Symposium on Applied Computing, pp. 32-36, Mar. 2011.

〔13〕 Paramvir Bahl, and Venkata N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, Vol. 2, pp. 775-784, Mar. 2000.

〔14〕 Beyah Raheem et al., "Rogue access point detection using temporal traffic characteristics," Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE, Vol. 7, pp. 2271-2275, Dec. 2004.

〔15〕 Wei Wei et al., "Passive online rogue access point detection using sequential hypothesis testing with TCP ACK-pairs," Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, pp. 365-378, Oct. 2007.

〔16〕 Lanier Watkins, Raheem Beyah, and

Cherita Corbett, "A passive approach to rogue access point detection," Global Telecommunications Conference 2007 (GLOBECOM'07), pp. 355-360, Nov. 2007.
〔17〕 Hao Han et al., "A measurement based rogue AP detection scheme," INFOCOM 2009, pp. 1593-1601, Apr. 2009.
〔18〕 Phil Simon, "Too Big to Ignore: The Business Case for Big Data," Wiley and SAS Business Series, pp 89-90, 2013.
〔19〕 John Robert Anderson et al,. "Machine Learning: An artificial intelligence approach," Morgan Kaufmann, pp. 19, 1986.
〔20〕 Bernhard E. Boser, Isabelle M. Guyon, and Vladimir N. Vapnik, "A training algorithm for optimal margin classifiers," Proceedings of the fifth annual workshop on Computational learning theory, pp. 144-152, Jul. 1992.
〔21〕 Vladimir N. Vapnik, "The nature of statistical learning theory," Springer, pp. 133-136, 2000.
〔22〕 Colin Campbell and Nello Cristianini, "Simple learning algorithms for training support vector machines," Technical report, University of Bristol, 1998.
〔23〕 Massimiliano Pontil and Alessandro Verri, "Properties of support vector machines," Neural Computation, Vol. 10, No. 4, pp. 955-974, Mar. 1998.
〔24〕 Simon Haykin, "Neural networks: a comprehensive foundation," Prentice Hall PTR, 1994.
〔25〕 Pyeong-jung Song and Yeon-seung Shin, "LTE Mobility Management Technology for Network Convergence," Electronics and Telecommunications Trends, 25(6), Dec. 2010.
〔26〕 Mathieu Lacage, Mohammad Hossein Manshaei and Thierry Turletti, "IEEE 802.11 rate adaptation: a practical approach," Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems, pp. 126-134, Oct. 2004.
〔27〕 Holma H. and Toskala A., "LTE for UMTS : OFDMA and SCFDMA Based Radio Access," Wiley, pp. 244-245, 2009.
〔28〕 Ram Periakaruppan and Evi Nemeth, "GTrace - A Graphical Traceroute Tool," LISA '99 Proceedings of the 13th USENIX conference on System administration, pp 69-78, 1999.
〔29〕 Kenneth Jonsson et al., "Support vector machines for face authentication," Image and Vision Computing, Vol. 20, Issue 5-6, pp. 369-375, Apr. 2002.
〔30〕 accord - Accord.NET Framework, http://code.google.com/p/accord/
〔31〕 B. Caputo et al., "Appearance-based Object Recognition using SVMs: Which Kernel Should I Use?," Proc of NIPS workshop on Statistical methods for computational experiments in visual processing and computer vision, Whistler, Dec. 2002.

〈저자소개〉

이 재 욱 (Jae-wook Lee) 정회원
2008년 2월: 고려대학교 전자 및 정보공학부 졸업
2011년 3월: 고려대학교 정보보호대학원 정보보호학과 석사과정 수료
2011년 6월~현재: LG전자 TV Commercial 연구소 근무
〈관심분야〉 정보보호, 소프트웨어 공학

이 시 영 (Si-young Lee) 학생회원
2013년 2월: 수원대학교 산업정보공학과 졸업
2013년 3월~현재: 고려대학교 정보보호대학원 금융보안학과 석사과정
〈관심분야〉 정보보호, 무선/모바일 네트워크 보안

문 종 섭 (Jong-sub Moon) 종신회원
1981년 2월~1985년: 금성 통신 연구소 연구원
1991년: Illinois Institute of technology 졸업(전산학 박사)
1993년~현재: 고려대학교 전자 및 정보공학부 교수
〈관심분야〉 생체인식, 침입탐지, 운영체제