

비지도 학습 기법을 사용한 RF 위협의 분포 분석

김철표¹⁾ · 노상욱^{*1)} · 박소령²⁾

¹⁾ 가톨릭대학교 컴퓨터정보공학부

²⁾ 가톨릭대학교 정보통신전자공학부

Analysis on the Distribution of RF Threats Using Unsupervised Learning Techniques

Chulpyo Kim¹⁾ · Sanguk Noh^{*1)} · So Ryoung Park²⁾

¹⁾ School of Computer Science and Information Engineering, The Catholic University of Korea

²⁾ School of Information Communication and Electronic Engineering, The Catholic University of Korea

(Received 16 September 2015 / Revised 12 February 2016 / Accepted 20 May 2016)

ABSTRACT

In this paper, we propose a method to analyze the clusters of RF threats emitting electrical signals based on collected signal variables in integrated electronic warfare environments. We first analyze the signal variables collected by an electronic warfare receiver, and construct a model based on variables showing the properties of threats. To visualize the distribution of RF threats and reversely identify them, we use k-means clustering algorithm and self-organizing map (SOM) algorithm, which are belonging to unsupervised learning techniques. Through the resulting model compiled by k-means clustering and SOM algorithms, the RF threats can be classified into one of the distribution of RF threats. In an experiment, we measure the accuracy of classification results using the algorithms, and verify the resulting model that could be used to visually recognize the distribution of RF threats.

Key Words : RF Threats(레이더 위협), Unsupervised Learning(비지도 학습), Self-Organizing Map(자기 조직화 지도), K-Means Clustering Algorithm(K-평균 군집화 알고리즘), Integrated Electronic Warfare(통합 전자전)

1. 서론

다양하며 복잡한 위협이 존재하는 미래의 통합

전자전 상황에서 아군이 다중/복합 스펙트럼 위협을 식별하기 위하여 지휘통제 모듈의 여러 장비가 운용되고, 이들 장비는 각각의 특성에 따른 위협을 분석하는 임무를 수행한다. 아군의 생존 확률을 높이고 주어진 임무를 성공적으로 수행하기 위해서는 이러한 전자전 장비에서 수신한 데이터를 종합하여 상호 연관

* Corresponding author, E-mail: sunoh@catholic.ac.kr
Copyright © The Korea Institute of Military Science and Technology

성을 분석하고, 분석한 데이터를 기반으로 위협에 대한 정보를 추출하여 임무 수행에 필수적인 정보를 생산할 수 있는 역추정 모델링 기법에 대한 연구가 필요하다.

본 논문에서는 수신된 데이터가 나타내는 위협의 종류가 주어지지 않은 상황에서 수집된 신호변수의 범위 값에 대한 패턴을 바탕으로 위협의 종류를 역으로 추정할 수 있는 비지도 학습 방식(unsupervised learning)에 대하여 알아본다. 사용된 비지도 학습 방식은 K-평균 군집화(K-means clustering) 알고리즘과 자기 조직화 지도(Self-Organizing Map, SOM) 알고리즘이며, 두 가지 알고리즘을 활용하여 전자전 상황에서 RF(radio frequency) 위협의 속성에 대한 분포를 분석하고자 한다.

논문의 구성은 다음과 같다. 2장에서는 본 논문과 관련한 비지도 학습 기법의 응용분야 및 관련연구에 대하여 알아본다. 3장에서는 위협체의 속성을 분석하고, 위협체를 역으로 추정할 수 있는 역추정 모델을 비지도 학습 기법으로 생성할 수 있는 방안을 제시한다. 4장에서는 비지도 학습 기법으로써 K-평균 군집화 알고리즘과 자기 조직화 지도 알고리즘을 구현하며, 통합 전자전 상황에서 두 가지 알고리즘의 성능을 비교 및 분석한다. 5장에서는 구현한 두 가지 알고리즘을 통합 전자전 시뮬레이터에 통합시키기 위하여 모듈 별로 결과를 보여주는 시각화 테스트 프로그램에 대하여 설명한다. 결론에서는 수행한 연구를 요약 및 정리하며, 앞으로의 연구방향에 대하여 서술한다.

2. 관련 연구

전자전(electronic warfare)은 전자기 스펙트럼을 제어하거나 적을 공격하기 위하여 전자기 에너지나 지향성 에너지를 사용하는 군사적 행동으로 정의되어 있다^[1]. 전자전과 관련된 선행연구로는 수신 측정된 변수들을 활용하여 관련없는 신호를 분리하여 오분석의 확률을 줄이는 연구가 진행되었다^[2]. 또한 항공기가 자신을 위협하는 위협에 대하여 확인하고 최적의 대응방안을 선정하기 위한 방법론을 제시한 연구^[3]가 수행되었다. 본 연구에서는 RF 위협의 특징을 나타내는 속성을 분석하여 수집신호변수로 정리하였으며, 이를 기반으로 위협체에 대한 분포를 시각적으로 보여줄 수 있는 기법에 중점을 둔다.

일반적으로 비지도 학습 방식은 입력된 데이터의 클래스 정보가 주어지지 않은 상황에서 유용한 패턴을 찾기 위하여 사용된다. [4]에서는 비지도 학습 방식을 사용하여 14개의 지형을 분류하기 위한 영상 분류 모델을 제안하였으며, 문서분류^[5-7] 분야 및 영상처리/분류^[8-10] 분야 등에 비지도 학습 방식을 사용하였다. 이러한 연구에서 사용한 비지도 학습방식을 본 논문에서는 RF 위협을 역으로 확인하고 분류할 수 있는 기법으로 사용한 점에서 유사하며, 각각의 역추정 알고리즘 결과인 RF 위협의 분포를 시각적으로 보이려고 한다는 점에서 차이점을 갖는다.

군집화 알고리즘은 여러 분야의 연구에서 많이 사용되는 대표적인 비지도 학습 방법이다. 군집화 알고리즘은 분할 군집화 알고리즘(partitional clustering), 계층 군집화 알고리즘(hierarchical clustering), 밀도 기반 군집화 알고리즘(density-based clustering) 및 격자 기반 군집화 알고리즘(grid-based clustering) 등 군집화 방식에 따라 다양한 종류의 군집화 알고리즘이 존재한다^[15]. 계층 군집화 알고리즘은 데이터가 증가할수록 학습시간이 오래 걸리기 때문에 새로운 모델에 적용하는데 분할 군집화 알고리즘보다 많은 시간이 필요로 한다. 따라서, 본 연구에서는 위협 모델이 추가될 경우에 이를 빠르게 반영할 수 있는 분할 군집화 알고리즘을 사용하기로 결정하였다. 가장 일반적으로 사용되는 군집화 알고리즘은 분할 군집화 알고리즘의 일종인 K-평균 군집화 알고리즘으로 정확도는 다른 알고리즘에 비해 낮지만 상대적으로 구현이 쉽고 효율적으로 동작한다. 본 연구에서는 위협체에 대한 역추정 모델을 군집으로 형상화할 수 있는 기법으로 K-평균 군집화 알고리즘을 사용한다.

자기 조직화 지도 알고리즘은 대표적인 비지도 학습 신경망 알고리즘이다. 인공 신경망 알고리즘은 학습 방식에 따라 전방향(feedforward) 신경망, 되먹임(feedback) 신경망 및 경쟁(competitive) 신경망으로 나뉜다. 자기 조직화 지도는 세 방식 중 경쟁 신경망 방식에 속한다^[11]. 자기 조직화 지도는 입력층(input layer)과 출력층(output layer)의 두 개의 층으로 구성된다. 또한 학습 결과 2차원 및 3차원으로 구성된 형상지도(feature map)로 표시되어, 고차원의 벡터 데이터를 2차원 및 3차원으로 사상하여 효율적인 시각화가 가능하다. 본 연구에서는 위협 속성을 입력층의 입력으로 하여, 학습결과인 위협체의 유형을 2차원 형상지도로 표시한다.

3. 위협체 모델링 및 역추정 기법

적 위협에 대한 역추정 및 분류를 위하여 레이더 위협체의 속성 및 클래스를 모델링하였다. 본 논문에서는 모델링한 속성과 클래스를 바탕으로 알려지지 않은 위협체를 역으로 추정하는 역추정 기법으로 K-평균 군집화 알고리즘과 자기 조직화 지도 알고리즘을 선택하였다.

3.1 위협체 속성 분석

문헌 조사를 통하여 현재 사용되는 RF 위협체를 목록화 하였으며, RF 위협체를 분류하기 위한 주요 속성들을 구체화하였다^[12]. 구체화 한 속성들은 주파수 대역, 펄스 폭, 펄스 반복 주기 및 펄스 전력이다. Table 1은 주요 속성의 유형 및 단위를 나타낸다.

Table 1. Attributes of threat

속성	유형	단위
주파수 대역 (RF)	numeric	Hz
펄스폭 (PW)	numeric	μ s
펄스 반복 주기 (PRI)	numeric	μ s
펄스 전력 (PP)	numeric	kW

구체화 한 속성을 기준으로 RF 위협체의 클래스를 탐지 레이더, 추적 레이더 및 레이더 유도 미사일 탐색기의 세 가지 클래스로 정의하였다. Fig. 1부터 Fig.

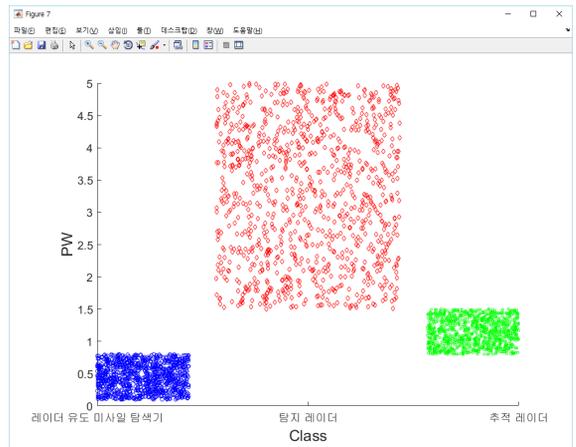


Fig. 2. The distribution of pulse width(PW)

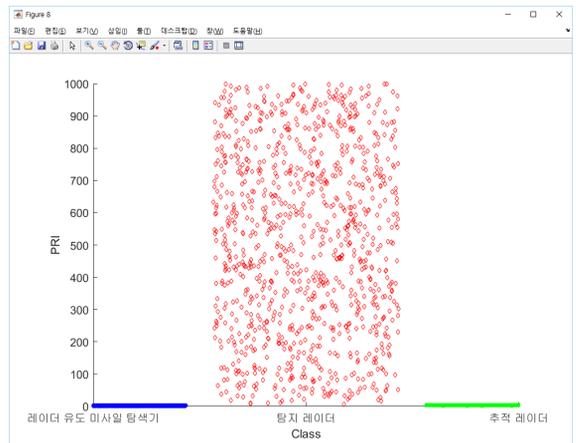


Fig. 3. The distribution of pulse repetition interval(PRI)

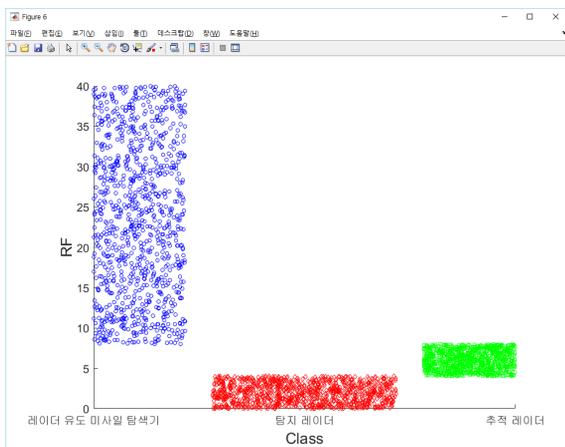


Fig. 1. The distribution of radio frequency(RF)

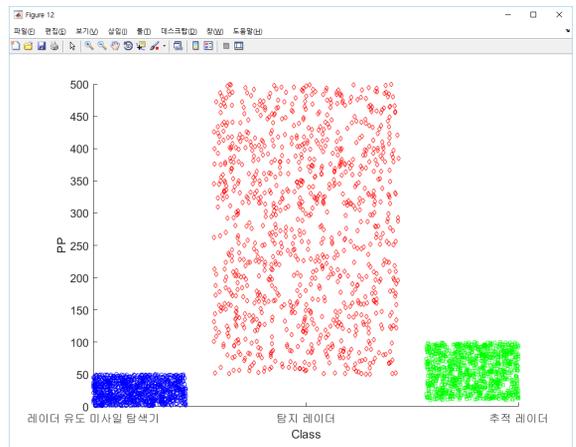


Fig. 4. The distribution of pulse power(PP)

4는 클래스별 속성의 분포를 나타낸다. 각 그래프에서 세로 축은 각각 RF, PW, PRI, PP를 나타내며, 가로 축은 RF 위협체의 유형을 나타낸다. 가로축이 나타내는 위협은 순서대로 ‘레이더 유도 미사일 탐색기’, ‘탐지 레이더’ 및 ‘추적 레이더’ 이다.

3.2 수집신호를 이용한 역추정 기법

아군의 전자전 수신기를 통하여 수집된 신호를 바탕으로 적의 RF 위협체를 역으로 추정하는 기법으로 비지도 학습인 K-평균 군집화 알고리즘과 자기 조직화 지도 알고리즘을 사용하였다. K-평균 군집화 알고리즘은 주어진 데이터를 k 개의 군집으로 분류하는 것을 목표로 하며, 자기 조직화 지도 알고리즘은 고차원의 데이터를 2차원 및 3차원으로 사상하여 시각화가 가능하다.

3.2.1 K-평균 군집화 알고리즘

K-평균 군집화 알고리즘은 주어진 데이터를 k 개의 군집으로 묶는 알고리즘이다. 주어진 데이터를 k 개의 군집으로 분류하는 알고리즘은 다음과 같다.

Algorithm 1. K-Means algorithm

```

Function K-Means(T, k)
Input: 데이터의 수가 m인 집합 T, 군집의 수 k
Output:  $c_j \in R^n, (0 \leq j \leq k-1)$ 를 원소로 하는 군집의 중심 집합 C

Random Initialize(C)
Initialize with Zero(A)

while isChanged(C) and L times do
  for i = 0 to m-1 do
    j ← argminj d(ti, cj), (0 ≤ j ≤ k-1)
    aij ← 1 (aij ∈ A)
  end_for
  for j = 0 to k-1 do
    cj ←  $\frac{\sum_{i=0}^{m-1} a_{ij} t_i}{\sum_{i=0}^{m-1} a_{ij}}$ 
  end_for
end_while

return C
end_function K-Means
    
```

K-평균 군집화 알고리즘은 n 차원의 데이터 집합 T 와 군집의 수 k 를 매개변수로 갖는다. 알고리즘의 시

작 부분에서 군집의 중심을 나타내는 배열 C 를 임의의 n 차원 데이터로 초기화 하고, 입력된 데이터가 어떤 군집의 집합에 할당되었는지를 나타내는 배열 A 를 0으로 초기화 한다. 군집의 중심 집합을 계산하기 위하여 반복을 시작한다. 군집의 중심을 계산하기 위하여 배열 T 에 포함된 모든 입력 데이터 t_i 에 대하여 배열 C 에 포함된 가장 가까운 군집의 중심 c_j 에 할당한다. 모든 데이터가 각각 하나의 군집에 할당되면 할당된 입력 데이터의 평균을 계산하여 군집의 중심 c_j 로 설정한다. 이렇게 설정된 군집의 중심이 이전과 동일하거나 최대 반복횟수 L 만큼 반복했다면 종료조건을 만족하여 반복을 종료하고 군집의 집합을 반환한다.

K-평균 군집화 알고리즘에서 입력된 데이터를 분류하기 위하여 입력된 데이터와 군집의 중심 사이의 거리를 비교하여야 한다. 이때 입력된 데이터가 n 개의 속성으로 이루어진다면 데이터는 n 차원 벡터로 나타낼 수 있으며, n 차원 벡터 상의 거리를 나타내기 위하여 유클리디안 거리 측정 방식과 코사인 유사도 측정 방식을 사용하였다. 즉, 알고리즘 1의 K-평균 군집화 알고리즘에서의 계산을 위하여 두 가지 측정 방식을 활용한다.

위협체를 나타내는 인스턴스 a, b 를 입력 받았을 때, 두 벡터 사이의 거리를 유클리디안 거리 측정 방식으로 계산하는 수식은 (1)과 같다.

$$d(\vec{a}, \vec{b}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (1)$$

코사인 유사도 측정 방식은 내적공간의 두 벡터 사이의 각의 코사인 값을 이용하여 입력된 벡터간의 유사도를 계산한다. 이렇게 계산된 코사인 유사도 측정 방식은 0 부터 1 까지의 값을 가진다. 두 벡터 a, b 를 입력받았을 때 코사인 유사도 측정 방식을 계산하는 수식은 (2)와 같다.

$$d(\vec{a}, \vec{b}) = \frac{\sum_{i=1}^n a_i \times b_i}{\sqrt{\sum_{i=1}^n a_i^2} \times \sqrt{\sum_{i=1}^n b_i^2}} \quad (2)$$

3.2.2 자기 조직화 지도 알고리즘

자기 조직화 지도 알고리즘은 주어진 데이터를 통해 임의의 크기를 가지는 2차원 및 3차원의 지도를 생성한다. 지도를 생성하기 위한 알고리즘은 다음과 같다.

Algorithm 2. Self-Organizing Map

```

Function SOM(T, M)
input: 데이터의 수가 n인 집합 T, l×m 크기의 지도 M
Output: l×m 크기의 지도 M

Random Initialize(M)
range ← r
size ← l×m
itr ← 0

while range > 0 do
  for i = 0 to size-1 do
    M = Update_Neurons(T, M, itr, range)
  end
  itr ← itr + 1
  range ← range - 1
end_while

return M
end_Function SOM
    
```

자기 조직화 지도 알고리즘은 n차원의 데이터 집합 T와 결과로 출력 할 지도의 크기를 나타내는 l과 m을 입력받아 수행된다. 우선 지도에 포함된 뉴런의 수를 나타내는 size변수를 l×m으로 초기화 하며, 이웃의 범위를 나타내는 range변수를 r로 초기화 한다. 또한 크기를 갖는 지도 M을 임의의 벡터로 초기화 한다. 그 후 range변수가 0이 될 때까지 뉴런을 갱신하기 위하여 Update_Neurons을 수행한다. Update_Neurons의 알고리즘은 다음과 같다.

Algorithm 3. Update_Neurons algorithm

```

Function Update_Neurons(T, M, itr, range)
Input: 데이터의 수가 n인 집합 T, l×m 크기의 지도 M, 현재 반복을 나타내는 itr, 현재 반복에서의 이웃의 범위 range
Output: l×m 크기의 지도 M

for i = 0 to n-1 do // 모든 입력 데이터에 대하여 반복 수행
  j ← arg minj d(ti, pj), (0 ≤ j ≤ l×m, ti ∈ T, pj ∈ M)

  for each ||pj - pneighbor|| < range do
    pneighbor ← pneighbor + h(r) • d(ti, pneighbor)
  end_for
end_for

return M
end_Function Update_Neurons
    
```

함수 Update_Neurons는 데이터의 집합, 지도, 현재 반복 횟수 및 현재 반복에서의 이웃의 범위를 입력받는다. 입력받은 데이터의 집합과 지도를 바탕으로 선택된 데이터 t_i와 가장 가까운 지도위의 뉴런 p_j를 찾아 승자로 설정하고, p_j와 지도상의 거리가 range이내의 뉴런 p_{neighbor}에 대해 값을 갱신한다. 값 갱신을 위해 사용되는 식 h(r)은 다음과 같다.

$$\alpha(itr) \cdot \exp\left(-\frac{\|r_j - r_{neighbor}\|^2}{2\sigma^2(r)}\right) \tag{3}$$

h(r)은 가우시안 함수로 표현된다. α(r)은 학습 비율로 0 < α(r) < 1의 값이 할당되며 반복될수록 값이 감소하며, 초기값은 0.5이다. r_j, r_{neighbor}는 지도상의 뉴런 좌표를 나타낸다.

자기 조직화 지도 알고리즘에서 좌표에 할당된 값과 입력된 RF 위협체 데이터 사이의 유사도는 3.2.1에서 설명한 유클리디안 거리 측정 방식(수식 (1) 참조)을 사용한다.

4. 알고리즘 구현 및 실험

실험을 위하여 K-평균 군집화 알고리즘과 자기 조직화 지도 알고리즘을 사용한 역추정 모델을 각각 구현하였다. 실험에서 총 3,000개의 RF 위협체를 생성하였으며, 각 클래스는 1,000개씩의 RF 위협체를 포함한다. 세 가지 알고리즘의 성능을 평가하기 위하여 두 알고리즘의 위협체에 대한 분류 정확도(accuracy)를 측정하였다. 실험 환경에 대한 구체적인 수치는 표 2로 정리하였다.

Table 2. Experimental environment

총 데이터 수	3,000개
클래스별 데이터 수	1,000개
실험에 사용한 데이터 집합	5개
학습에 사용한 데이터 비율	90%
정확도 측정에 사용한 데이터 비율	10%

가장 좋은 분류 성능을 보이는 데이터 집합의 개수를 정하기 위하여 각각 210, 420, 630, 840, 990개의

RF 위협체 집합을 생성하였다. 각 위협체 집합별 정확도는 Fig. 5와 같다. Fig. 5의 세로축은 위협체 클래스의 역추정 정확도를 나타내며, 가로축은 생성한 RF 위협체 클래스의 갯수를 나타낸다.

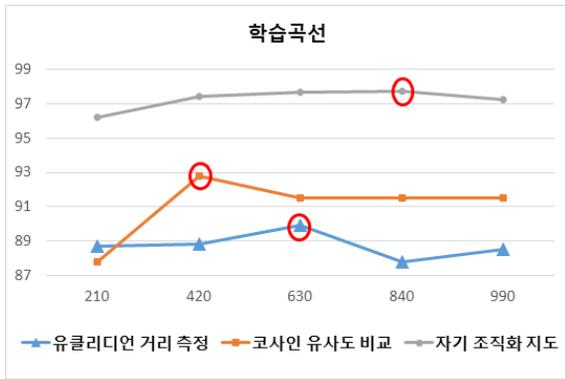


Fig. 5. Learning curve representing reverse extrapolation accuracy

역추정 결과를 바탕으로 각 알고리즘의 재현율과 정밀도를 측정하였다^[13]. 정밀도(precision)와 재현율(recall)에 대한 식은 다음과 같다.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (4)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (5)$$

식 (4)와 (5)에서 True Positive는 역추정한 위협체의 클래스가 학습에 사용된 클래스와 같은 경우를 의미한다. False Positive와 False Negative는 역추정한 위협체의 클래스가 학습에 사용된 클래스와 다른 경우다. False Positive의 경우 역추정에 사용된 위협체의 클래스가 ‘레이더 유도 미사일 탐색기’ 일 때, 역추정 모델을 통한 역추정 결과가 ‘탐지 레이더’ 혹은 ‘추적 레이더’인 경우를 의미한다. False Negative의 경우 역추정에 사용된 위협체의 클래스가 ‘탐지 레이더’ 혹은 ‘추적 레이더’ 일 때, 역추정 모델을 통한 역추정 결과가 ‘레이더 유도 미사일 탐색기’ 인 경우를 의미한다.

4.1 K-평균 군집화 알고리즘을 이용한 역추정 모델
K-평균 군집화 알고리즘은 Python 프로그래밍 언어 (Version 3.4)를 사용하여 MS Windows 환경에서 구현

하였다. K-평균 군집화 알고리즘을 이용한 역추정 모델은 거리 측정 방식에 따라 두 가지로 나뉜다.

4.1.1 유클리디언 거리 측정방식을 사용한 역추정

유클리디언 거리 측정방식을 사용한 역추정 모델에서 가장 좋은 성능을 나타낸 집합은 Fig. 5에 나타낸 바와 같이 각각의 클래스가 630개인 집합(총 데이터 개수: $630 \times 3 = 1,890$)이었다. 각각의 클래스 개수인 630개 RF 위협체 중에서 576개는 훈련에 사용하였으며, 64개는 역추정 정확도 측정에 사용하였다. 이러한 역추정 결과를 바탕으로 각 클래스 별 정확도, 정밀도 및 재현율을 정리한 결과는 Table 3과 같다.

Table 3. The resulting reverse extrapolation using euclidean distance method

	탐지 레이더	추적 레이더	레이더 유도 미사일 탐색기
정밀도	100 %	76.8 %	100 %
재현율	95.2 %	100 %	74.6 %
정확도	88.9 %		

정밀도는 역추정 모델을 통한 역추정 결과에서 실제 클래스와 역추정된 클래스가 동일한 비율을 나타낸다. 재현율은 하나의 클래스에 속한 위협체들 중 정확히 분류된 위협체의 비율을 나타낸다. Table 3의 추적 레이더의 경우 정밀도가 76.8 %로 추적 레이더로 클래스가 분류된 위협체중 23.2 %는 추적레이더가 아닌 탐지 레이더나 레이더 유도 미사일 탐색기에 속함을 알 수 있다. 또한 재현율이 100 %이므로 추적 레이더 클래스에 속한 모든 위협체는 역추정 모델을 통한 역추정 결과가 추적 레이더로 나타난다. 정확도는 모든 클래스에 대하여 역추정을 수행한 결과가 정확히 분류된 위협체의 비율을 나타낸다.

4.1.2 코사인 유사도 측정방식을 사용한 역추정

코사인 유사도 측정 방식을 사용한 역추정 모델에서 가장 좋은 성능을 나타낸 집합은 Fig. 5에 나타낸 바와 같이 각 클래스가 420개인 집합(총 데이터 개수: 1,260) 이다. 특정한 클래스의 개수가 420개인 RF 위협체 중에서 378개는 훈련에 사용하였으며, 42개는 역추정 정확도 측정에 사용하였다. 이러한 역추정 결과

를 바탕으로 각 클래스 별 정확도, 정밀도 및 재현율을 정리한 결과는 Table 4와 같다.

Table 4. The resulting reverse extrapolation using cosine similarity method

	탐지 레이더	추적 레이더	레이더 유도 미사일 탐색기
정밀도	100 %	84 %	100 %
재현율	80.9 %	100 %	100 %
정확도	93.6 %		

4.2 자기 조직화 지도 알고리즘을 이용한 역추정 모델

자기 조직화 지도 알고리즘을 이용한 역추정 모델은 Matlab®의 SOM tool box를 활용하여 구현하였다^[14]. 자기 조직화 지도를 통한 역추정 모델은 두 가지 과정을 거친다. 첫 번째, SOM tool box를 이용하여 자기 조직화 지도를 작성 후 각 좌표에 할당된 RF 위협체 데이터들을 바탕으로 클래스를 할당한다. 두 번째, 지도에 할당된 클래스를 사용하여 새로 입력된 RF 위협체를 분류하고 정확도를 측정한다. 두 가지 과정을 거쳐 측정된 정확도를 바탕으로 4.1에서 설명한 K-평균 군집화 알고리즘과 같은 방식으로 정밀도와 재현율을 측정하였다.

자기 조직화 지도 알고리즘을 사용한 역추정 모델에서 가장 좋은 성능을 나타낸 집합은 Fig. 5에 나타난 바와 같이 각 클래스가 840개인 집합(총 데이터 개수: 2,520)이다. 특정한 클래스의 개수가 840개인 RF 위협체 중에서 756개는 훈련에 사용하였으며, 84개는 분류 정확도 측정에 사용하였다. 이러한 역추정 결과를 바탕으로 각 클래스 별 정밀도와 재현율을 정리한 결과는 Table 5와 같다.

Table 5. The resulting reverse extrapolation using SOM

	탐지 레이더	추적 레이더	레이더 유도 미사일 탐색기
정밀도	100 %	99 %	100 %
재현율	100 %	100 %	99 %
정확도	99.7 %		

4.3 ANOVA 테스트

ANOVA 테스트는 통계학에서 다수의 집단을 비교할 때 집단내의 분산, 총 평균과 각 집단의 평균 사이에 생긴 집단 간 분산의 비교를 통해 만들어진 F 분포를 이용하여 가설을 검정하는 방법이다. 본 논문에서는 유클리디안 거리 측정 방식을 사용한 K-평균 군집화 알고리즘, 코사인 유사도 측정 방식을 사용한 K-평균 군집화 알고리즘 및 자기 조직화 지도 알고리즘에 대한 분류 정확도의 통계적인 유의성을 검증하기 위하여 ANOVA 테스트를 수행하였다.

알고리즘에 대한 정확도는 각 알고리즘별로 10회 측정하였으며, 알고리즘 정확도의 평균은 Table 6과 같다.

Table 6. The average of reverse extrapolation accuracy for three algorithms

	평균
유클리디안 거리 측정	88.9 ± 1.88 %
코사인 유사도 측정	92.7 ± 2.06 %
자기 조직화 지도	98.55 ± 0.21 %

이때, ANOVA 테스트를 수행하기 위한 속성으로 $k = 3$, $n = 10$, $V_1 = k - 1 = 2$, $V_2 = k(n - 1) = 27$ 을 사용하였다. 결과적으로 계산된 ANOVA 테스트의 값은 $f = 232.42$ 이며, F 분포표를 참조하여 $f_{.05, 2, 27} = 3.35$ 이고 $f_{.01, 2, 27} = 5.45$ 임을 알 수 있으며, 최종적으로 $f > f_{.01, 2, 27} > f_{.05, 2, 27}$ 임을 보아 세 가지 측정 방식간의 통계적인 유의성이 존재함을 알 수 있다. 따라서 주어진 도메인에서 자기 조직화 지도의 성능이 다른 두 알고리즘보다 우월함을 결론지을 수 있다.

5. 알고리즘별 역추정 모델 시각화

구현된 알고리즘별 역추정 모델을 바탕으로 역추정 모델을 시각화 하고 새로 입력된 RF 위협체에 대하여 역추정을 수행하기 위하여 각 역추정 모델별 시각화 모듈을 구현하였다. K-평균 군집화 알고리즘은 Javascript를 사용한 웹 어플리케이션이며, 자기 조직화 지도 알고리즘은 Matlab®의 SOM tool box에 포함된 그래프 기능을 이용하여 구현하였다.

5.1 군집화 알고리즘을 이용한 역추정 모델 시각화
 K-평균 군집화 알고리즘을 이용하여 구현한 역추정 모델의 시각화 모듈은 Fig. 6 및 Fig. 7과 같다. 시각화 모듈은 시각화를 위한 그래프 부분과 새로운 RF 위협체 입력을 위한 입력 부분으로 구성된다. 그래프 부분에는 3.1에서 세분화한 총 4개의 속성이 축으로 존재하며, 클래스별로 총 3개의 가로방향 선이 기본으로 표시된다. 새로운 RF 위협체를 입력할 경우 해당 RF 위협체가 속하는 클래스와 동일한 색의 새로운 선이 표시되며 해당하는 두 선의 두께가 두껍게 표시된다. 입력 부분에는 총 5개의 입력 부분과 한 개의 출력력이 존재하며 각각 주파수 대역, 펄스 폭, 펄스 반복 주기, 펄스 전력 및 거리 측정 방식을 입력하는 양식과 역추정 결과를 표시하는 부분으로 구성되어 있다.

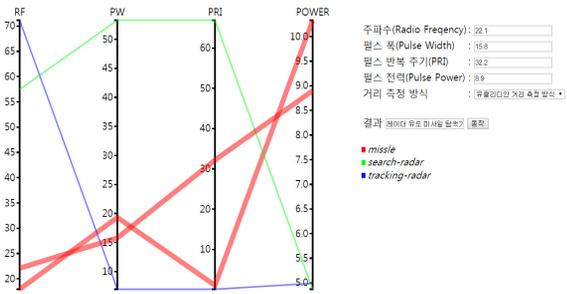


Fig. 6. The visualization of resulting reverse extrapolation using euclidean distance metric

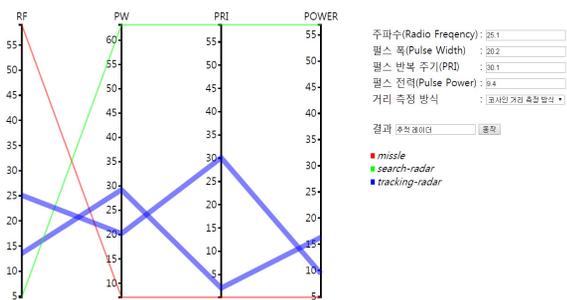


Fig. 7. The visualization of resulting reverse extrapolation using cosine similarity metric

5.2 자기 조직화 지도 알고리즘을 이용한 역추정 모델 시각화

자기 조직화 지도를 통한 역추정 모델의 시각화는 Fig. 8과 같다. 자기 조직화 지도에서 지도란 2차원 평

면에 가로 및 세로로 임의의 크기를 갖는 데이터 배열을 표시한 시각화 방법을 말한다. 지도의 가로 및 세로의 크기는 역추정 모델 학습시 학습한 데이터의 크기를 바탕으로 정해진다. 작성된 지도를 바탕으로 지도의 각 좌표 사이의 거리를 나타내는 U-Matrix, 각 속성들의 분포 지도 4개, 위협체의 클래스 분포 지도 등 총 6가지의 지도를 표시한다. 또한 지도 오른쪽에 표시된 d 의 경우 해당 지도에 표시된 속성의 값에 따른 색을 표시한다.

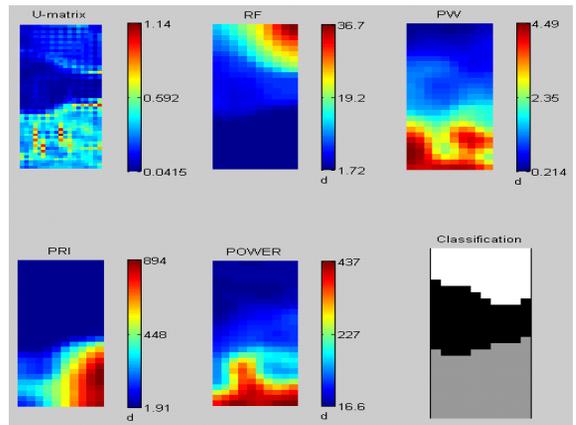


Fig. 8. The visualization of resulting reverse extrapolation using SOM

알고리즘의 결과로 작성된 지도 중 RF 위협체의 클래스 분포를 나타내는 분류지도(Classification map)를 RF 위협체의 역추정을 시각화하기 위한 지도로 사용한다. 해당 지도의 가로 및 세로 축은 분류지도의 위치 좌표를 나타낸다. 위치 좌표를 통해 표시되는 하나의 점은 하나의 RF 위협체를 나타내며 세 가지 색으로 위협체의 클래스를 구분한다. 지도에 표시된 데이터중 흰색은 레이더 유도 미사일 탐색기, 회색은 탐지 레이더, 검은색은 추적 레이더를 의미한다.

Fig. 9는 새로운 RF 위협체를 입력할 경우 시각화 결과를 나타낸다. 새로운 RF 위협체가 입력될 경우 지도에 표시된 모든 RF 위협체와 거리 비교를 수행한다. 결과적으로, 새로운 RF 위협체는 가장 가까운 거리에 위치한 한 점에 할당되며, 그 점에 해당하는 RF 위협체의 클래스로 분류된다. Fig. 9의 우측 분류지도와 같이 새로 입력된 위치를 나타내기 위하여 해당 좌표를 붉은색의 사각형으로 표시하며, 예제에서는 새로운 위협체가 탐지레이더로 역추정되는 것을 보여 준다.

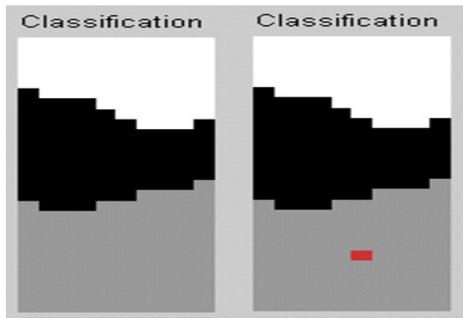


Fig. 9. The visualization of resulting reverse extrapolation for new RF threat

6. 결론

RF 위협체의 특징을 나타내는 속성을 분석하여 신호변수로 정의하였으며, 아군의 전자정보수집시스템에서 수집된 신호변수를 기반으로 레이더 위협을 역추정하였다. 분석한 속성과 클래스를 바탕으로 K-평균 군집화 알고리즘과 자기 조직화 지도 알고리즘을 사용한 역추정 모델을 제시하였다. 각 역추정 모델의 성능 측정을 위해 정확도, 정밀도 및 재현율을 측정하였으며, ANOVA 테스트를 통해 각 역추정 모델간 통계적 유의도를 측정하였다. 성능 평가 결과 자기 조직화 지도 알고리즘이 평균 정확도 98.55%로 가장 좋은 성능을 보였다. 또한 각 역추정 모델별 시각화 모듈을 구현하여 RF 위협체의 결과 분포를 시각적으로 확인할 수 있었다.

앞으로 지속적인 연구를 통하여 구현한 비지도학습 모델을 위협체 역추정 통합 시뮬레이터에 이식할 예정이다. 또한, 비지도학습 알고리즘이 제시한 각각의 결과를 융합하여 하나의 위협체로 제시할 수 있는 방안을 제시할 계획이다. 이러한 연구는 미래의 통합 전자전 상황에서 운용자의 개입없이 자율적으로 진장상황을 인식할 수 있도록 하며, 아군의 생존율을 향상시키는 궁극적인 방안이 될 수 있을 것이다.

후 기

본 연구는 '13년 기초연구 국방특화연구센터의 “전자전 수집신호변수를 사용한 역변환모델링 연구(EW-41)” 과제의 지원으로 수행하였음.

References

- [1] US DOD, “Electronic Warfare”, Joint Publication 3-13.1, 2012.
- [2] D. Lee, J. Han, W. Lee, “A Kernel Density Signal Grouping Based on Radar Frequency Distribution,” The Institute of Electronics Engineers of Korea - Signal Processing, Vol. 48, No. 6, pp. 124-132, November, 2011.
- [3] S. R. Park, H. Park, J. Ha, C. Choi, U. Jeong, S. Noh, “Designing Operational Effectiveness of Autonomously Decided Countermeasures,” Journal of Korean Society for Internet Information, Vol. 13, No. 4, pp. 11-21, August, 2012.
- [4] J. W. Choi, S. T. Noh, S. K. Choi, “Unsupervised Classification of Landsat-8 OLI Satellite Imagery Based on Iterative Spectral Mixture Model,” Journal of the Korean Society for Geospatial Information System, Vol. 22, No. 4, pp. 53-61, December, 2014.
- [5] S. Jun, “Spares Document Data Clustering Using Factor Score and Self Organizing Maps,” Journal of The Korean Institute of Intelligent Systems, Vol. 22, No. 2, pp. 205-211, April, 2012.
- [6] S. Jung, S. Lim, J. Jeon, B. Kim and H. Lee, “Web Search Result Clustering using Snippets,” Journal of KIISE : Databases, Vol. 39, No. 5, pp. 321-331, October, 2012.
- [7] K. Lee, K. Kim, M. Lee, W. Kim and J. Hong, “Post Clustering Method using Tag Hierarchy for Blog Search,” The Journal of Society for e-Business Studies Vol. 16, No. 4, pp. 301-319, November, 2011.
- [8] J. O. Kim, “Bio-mimetic Recognition of Action Sequence using Unsupervised Learning,” Journal of Korean Society for Internet Information, Vol. 15, No. 4, pp. 9-20, August, 2014.
- [9] K. Kim, K. Lee, “Hand Shape Detection and Recognition usin Self Organized Feature Map (SOMF) and Principal Component Analysis,” The Journal of the Korea Contents Association, Vol. 13, No. 11, pp. 28-36, November, 2013.
- [10] H. Noh, J. Min, “A Study of Hybrid Neural Network to Improve Performance of Face

- Recognition,” Journal of the Korea Institute of Information and Communication Enginee, Vol. 14, No. 12, pp. 2622-2627, November, 2010.
- [11] T. Kohonen, “The Self-Organizing Map,” Neurocomputing Vol. 21, Issues 1-3, 6, 1998. 11.
- [12] S. Noh and U. Jeong, “Intelligent Command and Control Agent in Electronic Warfare Setting,” International Journal of Intelligent Systems, Vol. 25, Issue 6, pp. 514-528, June, 2010.
- [13] I. H. Witten, E. Frank and M. A. Hall, “Data Mining 3/e,” Morgan Kaufmann, pp. 174-175, 2010.
- [14] J. Vesanto, J. Himberg, E. Alhoniemi and J. Parhankangas, “Self-organizing Map in Matlab : the SOM Toolbox,” Proceedings of the Matlab DSP Conference, pp. 35-40, 1999. 11.
- [15] J. Han, M. Kamber, J. Pei, “Data Mining : Concepts and Techniques 3/e,” Morgan Kaufmann, pp. 444-454, 2011.