https://doi.org/10.5392/JKCA.2017.17.03.231

보안 인텔리전트 유형 분류를 위한 다중 프로파일링 앙상블 모델

Ensemble Model using Multiple Profiles for Analytical Classification of Threat Intelligence

김영수
배재대학교 사이버보안학과

Young Soo Kim(experkim@gmail.com)

요약

최근 기업의 보안 시스템으로부터 수집되는 보안 인텔리전스 수는 악성코드의 확산으로 인해 기하급수적으로 증가하고 있다. 빅 데이터 환경이 도래하면서 기업들은 침해사고에 대한 다양한 정보를 이용할 수 있게 되면서 기업이 수집할 수 있는 침해사고 정보가 다양해지고 있다. 이에 따라 보안 인텔리전스를 구성하고 있는 침해사고의 다양한 속성을 사용하여 보다 정확하게 유사침해사고를 그룹별로 분류할 필요성이 요구되고 있다. 본 연구에서는 유사도 비교 분석 이론에 근거하여 침해사고를 공격유형과 침해자원을 고려한 다중 프로파일을 개발하고, 이를 활용하여 보안 인텔리전스를 구성하고 있는 침해사고 유형 분류의 정확성을 개선하는 다중 프로파일 기반 앙상블 모델을 제안한다. 제안 모델은 침입탐지시스템에서 수집된 계층적 침해자원에 대한 유사도 분석을 통해 새로운 침해사고를 효과적으로 분석할 수 있다. 사실적이고 의미 있는 침해사고의 구성을 통한 유형 분류는 새로운 침해사고에 대한 유사 침해사고를 정확하게 분류 제공함으로써 분석의 실용성을 향상시킨다.

■ 중심어 : | 빅데이터 | 보안 | 인텔리전스 | 침해사고 | 프로파일 | 앙상블 모델 | 기계학습 |

Abstract

Threat intelligences collected from cyber incident sharing system and security events collected from Security Information & Event Management system are analyzed and coped with expanding malicious code rapidly with the advent of big data. Analytical classification of the threat intelligence in cyber incidents requires various features of cyber observable. Therefore it is necessary to improve classification accuracy of the similarity by using multi-profile which is classified as the same features of cyber observables. We propose a multi-profile ensemble model performed similarity analysis on cyber incident of threat intelligence based on both attack types and cyber observables that can enhance the accuracy of the classification. We see a potential improvement of the cyber incident analysis system, which enhance the accuracy of the classification. Implementation of our suggested technique in a computer network offers the ability to classify and detect similar cyber incident of those not detected by other mechanisms.

■ keyword : | Big Data | Threat Intelligence | Cyber Incident | Profile | Ensemble Model | Machine Learning |

접수일자 : 2016년 11월 25일
수정일자 : 2016년 12월 26일
심사완료일 : 2016년 12월 26일
교신저자 : 김영수, e-mail : experkim@gmail.com

I.서 론

최근 사이버 침해사고의 대부분은 일련의 공격단계를 통해서 공격자는 다양한 방법으로 시스템에 공격을 시도한다. 인터넷 사용자가 웹 서핑하는 과정에서도 공격이 시도되고 있다. 웜이나 바이러스과 같은 악성코드를 통해 다양한 침해사고가 발생하고 침해사고 정보를 담고 있는 보안 인텔리전의 데이터가 기하급수적으로 증가하고 있고 예측할 수 없는 공격자들의 다양한 변종 패턴과 넘쳐 나는 데이터의 양으로 인해 빅데이터와 머신 러닝 기술을 사이버 보안에 접목하려는 연구가 확대되고 있다. 따라서 침해사고를 일으킨 사이버 공격 기법 분석과 연구 등을 통해 피해를 줄이거나 이를 원천 봉쇄할 수 있는 지능형 보안기술로써 공격로그 이벤트 및 네트워크 패킷을 수집하여 유사 침해사고를 분석함으로써 공격자의 의도를 신속하게 인지하고 차단할 수 있는 침해사고 분석 기술이 요구되고 있다. 과거 침해사고에 대한 보안 데이터의 외부 노출과 공개가 제한적이었던 환경과는 달리 글로벌한 침해사고공유센터로부터 보안 인텔리전스를 수집할 수 있는 빅 데이터 환경이 도래하면서 잘 고안한 침해사고 빅데이터 분석 알고리즘을 통해 침해사고에 대해서 신속하고 정확하게 대응할 필요가 있다. 보안 관련 빅 데이터를 분석하고 침해사고를 예측할 수 있는 다양한 기술이 개발되고 있다[1-5]. [그림 1]과 같은 침해사고 식별자와 침해자원으로 구성된 보안 인텔리전트를 수집하고 분석에 적합한 형태로 침해사고를 구성하고 정보를 유지한다.

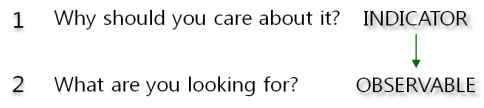


그림 1. 보안 인텔리전스의 구성요소

축적된 보안 인텔리전스를 분석함으로써 공격자의 공격 패턴이나 공격을 예방할 수 있는 기술을 습득할 수 있다. 보안 인텔리전스를 통한 침해사고 유형 분류는 서로 다른 특성으로 침해사고를 나눈다는 개념이며, 이를 위해서는 침해사고의 구별을 위한 기준이 되는 유사한 특성을 도출하고 침해사고를 유형 분류하는 다양한 방법과 연구들이 진행되고 있다[6-12].

침해사고를 구성하는 실제 관측치를 가지고 계산적 실험적 모델을 교차 검증하고 다양한 데이터 소스를 함께 가져와서 분석하는 것이 단일 데이터 소스들을 이용하는 것보다 더욱 효과적이다[13].

빅데이터는 많은 데이터 속에서 유효한 정보를 찾고 데이터 속에 숨어 있는 의미를 발견해 내는 과정인 반면 머신러닝은 여기에 데이터를 이용하여 검증과 학습의 과정을 통해 특정 조건에서 예측 값을 얻는 과정이다. 이와 같이 빅데이터와 머신 러닝은 현재의 데이터를 가지고 미래를 예측하는 방안을 가지고 있다. 이는 데이터와 인접한 현재 상황에 대한 예측은 서로 잘 융합되나, 만일, 현재 데이터를 가지고 미래를 예측한다면 정확성은 낮은 결과를 가져올 수 있다. 따라서 시간의 경과에 따른 현재의 data를 계속적으로 학습시키는 증분 학습(incremental learning)이 필요하다. 또는 표 과 같이 시간 흐름에 따라 데이터를 분리하여, 각각 학습하고 따로 여러개의 모델을 만들고. 이런 여러 개의 모델을 가지고, 앙상블 투표를 하여, 가장 높은 항목으로 결정하는 방식이 응용될 필요가 있다.

[표 1]와 같이 임의의 단일 데이터를 모델 A~D까지 적용하여 평가된 각각의 예측에 대해서 앙상블 다수투표의 논리라면 최종 예측은 No가 된다.

표 1. 앙상블 알고리즘의 미래예측 방식

연도	데이터	모델	예측	투표	최종예측
2000	Data 1	A	No	다수 투표	No
2005	Data 2	B	No		
2010	Data 3	C	Yes		
2015	Data 4	D	No		

본 논문의 구성은 다음과 같다. 2장에서 침해사고의 프로파일 구성 모델에 대하여 기술하고, 3장은 침해사고의 유사도 평가 모델에 대하여 설명한다. 4장에서는 다중 프로파일 기반 앙상블 모델을 제안하고 검증을 위해 침해사고 사례를 적용해 결과를 분석하고, 5장에서는 결론 및 향후 연구 방향에 대해 살펴본다.

II. 침해사고 프로파일 구성

분석하고자 하는 보안 인텔리전스에 포함되어 있는 침해사고 정보에 대한 데이터 셋은 이종(homogeneous) 열이 수직으로 파티션 되어(vertically partitioned) 구성한다. 서로 다른 특성을 가진 두 개의 데이터 셋은 침해사고라는 공통된 유일 식별자에 의해서 연결되어 있다. 침해사고라는 같은 식별자에 의해서 연결되어 있기 때문에 다른 데이터 셋을 연결하여 분석가능하다. [표 2]와 [표 3]은 침해사고를 구성하는 특성이 공격유형과 침해자원유형에 따라 계층적으로 구성되는 두 개의 데이터 셋을 보여주고 있다.

표 2. 공격 유형 기반의 침해사고

침해사고	스캔공격	DoS 공격	액세스 공격
#no_01	Satan	Land	Rootkit
#no_02	Saint	Smurf	Worm

표 3. 침해자원 기반의 침해사고

침해사고	IP	Domain	Hash
#no_01	1.1.1.1	www.abc.com	ab2c7defg
#no_02	2.2.2.2	www.zyx.com	z1yxw9vut

유사 침해사고의 속성을 규명하기 위해 공격유형 기준과 침해자원 기준의 두 개의 프로파일을 생성한다. 프로파일링 기반 침해사고분석시스템 연구에서 침해사고의 유사성을 계산하기 위한 유사도 측정 방법은 다양하며, 변수 및 매트릭스의 유형(type)별로 적합한 측정방법을 사용해야 한다. 새로운 침해사고와 유사한 침해사고를 분류해 내기 위해서는 유사도 산정의 기준이 되는 침해자원 프로파일을 생성하기 위한 알고리즘이 필요하다. 프로파일을 생성하기 위한 알고리즘은 [그림 2]와 같고 침해사고를 구성하고 있는 계층적 침해자원을 클러스터링하여 이를 대표할 수 있는 침해자원 프로파일을 도출한다. 프로파일 알고리즘의 구조에 대한 설명으로 침해사고를 대표할 수 있는 다중 침해자원에 대한 프로파일의 구성 기준은 침해사고와 연관된 침해자원이 최초인 또는 최근의 침해자원으로 구성하거나 또는 구성된 침해자원의 빈도수가 큰 순으로 침해사고를 구성하는 방식이 있을 수 있다. 프로파일 구성 알고리즘은 침해자원의 빈도수를 평가해서 결과값이 큰 순으로 정렬하여 설정된 프로파일의 개수만큼 구성하는 방식을 보여주고 있다.

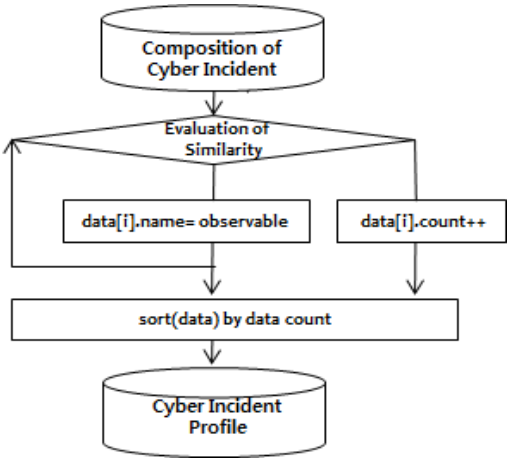


그림 2. 프로파일 알고리즘의 구조

침해자원이 수치로 표현이 불가능하기 때문에 침해자원의 중앙값을 프로파일로 지정하기 어렵다. 침해자원의 특성을 고려하여 침해자원 간의 비교를 통해서 가장 중복이 많이 된 침해자원을 프로파일로 지정한다. 침해사고의 대표 값인 프로파일을 1개로 지정하지만 분류의 정확률을 고려하여 침해사고 별로 다수의 프로파일을 생성할 수 있다.

III. 침해사고의 유사도 평가 모델

침해사고 분석시스템은 새로운 침해사고가 발행하였을 때 새로운 침해사고를 구성하는 침해자원과 침해자원 프로파일과의 유사도를 계산하는 데이터마이닝 기반의 알고리즘이 요구된다. 이의 알고리즘은 [그림 3]과 같다. 신규침해사고에 대해서 유사 침해사고를 분류 평가하는 방식으로 신규침해사고와 침해사고 프로파일의 IP, Domain, Hash의 각각에 대한 거리계산 알고리즘을

사용하여 유사도를 평가하고 각각의 가중치를 곱하여 평균한 값을 유사도 값으로 저장한다.

새로운 침해사고에 대해 이미 분류가 된 침해사고의 침해자원의 거리를 평가하여 유사도를 계산한다. 공격 타입이나 침해자원에 따라 가중치를 부여하여 유사도를 계산함으로써 공격 타입에 따른 위험도를 반영하고, 침해자원의 특성을 고려하여 유사도 평가를 할 수 있다. 우리가 제안한 알고리즘은 다수 프로파일에 대한 단일 거리평가 알고리즘의 결과 값을 평균하여 유사도를 계산하는 대신 다수의 거리평가 알고리즘을 적용하여 산출된 다수의 유사도에 대해서 다수투표 알고리즘으로 앙상블하여 유사도를 평가한다.

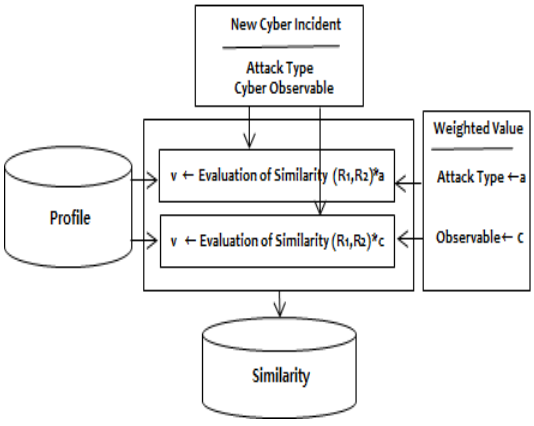


그림 3. 유사도 평가 알고리즘의 구조

Ⅳ. 다중 프로파일 기반 앙상블 모델

클러스터링 앙상블(clustering ensemble)방법은 클러스터링 결과들의 장점을 결합하여 최종 클러스터링 결과의 정확성과 신뢰성(reliability)을 향상시킨다[14-17].

또한 하나의 클러스터링 알고리즘을 데이터 셋에 적용한 클러스터링 결과를 최종 결과물로 제시하는 것보다는 다양한 특성을 가지는 클러스터링 알고리즘을 데이터 셋에 적용한 클러스터링 결과들의 장점을 결합하는 것이 더 나은 클러스터링 결과를 보이고자 하는 접근 방법이다. [그림 4]는 앙상블 알고리즘의 구조를 보여주고 있고 해를 결합하여 최종 해를 구하고 있다.

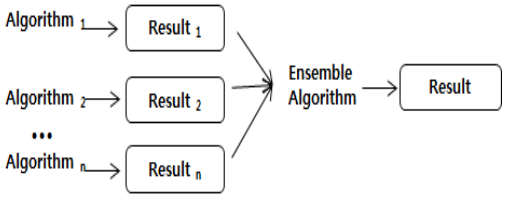


그림 4. 앙상블 알고리즘의 구조

결합 알고리즘의 의사결정 방식은 [그림 5]와 같고 다수 투표와 가중 다수 투표 방식을 보여주고 있다. 부류의 개수가 M=3이고 분류기의 개수가 T=5라 하면 미지의 패턴 x에 대해 다섯개의 분류기가 출력하는 표지 벡터를 보여주고 있다. 다수투표 결합 알고리즘에서는 두번째 부류가 최다 득표했으므로 w2로 분류한다. 하지만 가중치αt를 고려한 가중다수투표 결합 알고리즘의 경우에는 세번째 부류가 가장 높은 값을 획득하였으므로 w3으로 분류한다. 이 결과는 신뢰도가 높은 c3=0.4이 다른 것에 비해 높은 세번째 부류기 c3가 w3을 선택하였기 때문에 발생한다.

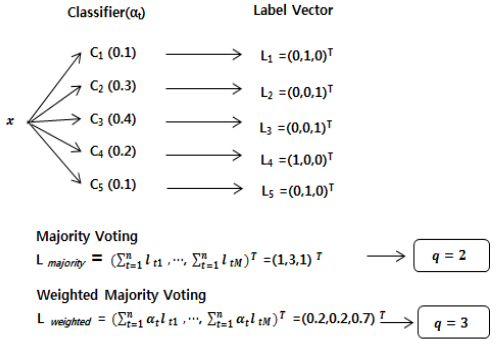


그림 5. 앙상블 알고리즘의 의사결정 방식

앙상블 결합 알고리즘은 다중 분류기의 출력을 결합하여 하나의 분류 결과를 만드는 과정으로 다수 투표 알고리즘을 적용한다. 본 연구에서 제안하는 앙상블 알고리즘의 적용 방법은 [그림 6]와 같이 프로파일 전체를 통합하여 하나의 유사도를 계산하는 방법과 [그림 6]과 같이 개별 프로파일별 유사도를 계산하여 결합하는 방법을 사용한다.

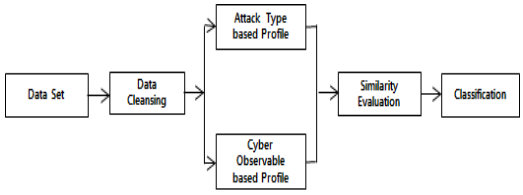


그림 6. 통합 프로파일기반 침해사고 분류모델

[그림 7]은 다양한 속성정보를 일률적으로 처리하는 대신 속성정보의 특성에 맞게 그룹핑하여 각 그룹의 유사도를 계산하여 결합하는 형태로 유사도를 계산한다.

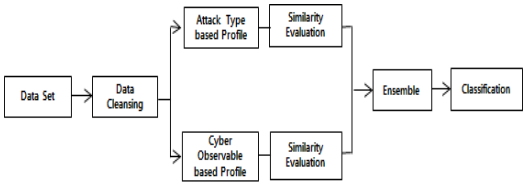


그림 7. 개별 프로파일기반 침해사고 분류모델

[그림 8]은 [그림 6]의 단일 프로파일을 사용하는 분류기 모델에서 사용하는 프로파일로 프로파일의 특성값을 하나로 통합하여 유사도를 계산하고 평균하여 하나의 유사도를 계산하는 방법이다.

Cyber Incident	Scan Attack	DoS Attack	Access Attack	IP	URL	Hash
#no_03	Satan	land	rootkit	1.1.1.2	test.abc.com	ab2c7defg

Profile	Scan Attack	DoS Attack	Access Attack	IP	URL	Hash
#1	Satan	land	rootkit	1.1.1.1	test.abc.com	ab2c7defg
#2	Saint	Smurf	Worm	2.2.2.2	www.zyx.com	z1yxw9vut

Similarity

그림 8. 통합 프로파일기반 유사도 평가방법

[그림 9]는 [그림 7]의 다중 프로파일을 사용하는 앙상블 모델에서 사용하는 프로파일로 프로파일의 특성값을 분류하여 각각의 프로파일에 대해서 유사도를 계산하는 방법이다.

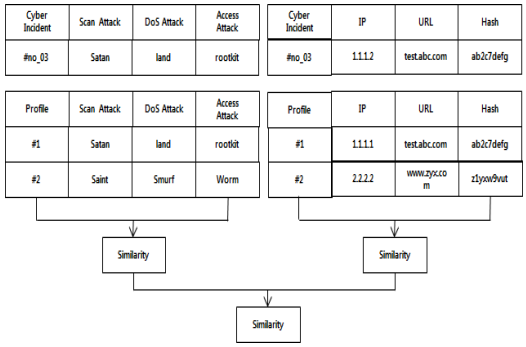


그림 9. 개별 프로파일기반 유사도 평가방법

검증을 위한 데이터 셋 구성은 [표 4]와 같고 학습과 테스트에 사용할 데이터는 653개의 침해사고 데이터를 사용하였다. 653개 의 데이터는 6개의 속성(Attribute)들로 이루져 있으며, 295개의 데이터는 양 (Positive)으로, 358개의 데이터는 음(Negative)으로 레이블 되어있다. 학습과 테스트의 비율은 7:3으로 하였다.

표 4. 침해자원 기반의 침해사고

	학습 데이터	테스트 데이터
Positive	244	102
Negative	213	94
Total	457	196

제안한 다중프로파일 기반 앙상블 모델은 [표 5]와 같이 단일 프로파일 기반의 통합 프로파일 방식은 침해사고 #1를 구성하는 침해자원 IP에 대해서 유사 침해사고를 A침해사고로 평가하고 다중프로파일 기반의 앙상블 평가 방식은 침해사고 #1를 구성하는 침해자원 IP에 대해서 평가한 유사침해사고의 빈도가 큰 결과값인 B 침해사고로 평가한다.

표 5. 다중 프로파일 기반의 앙상블 모델의 사례

모델	침해사고	침해자원	분류 평가	최종 평가
통합프로파일	001	10.10.10.10	A	A
분리프로파일	001	10.10.10.10	A	B
	001	100.20.10.10	B	
	001	100.10.10.10	B	

본 논문이 제시한 침해사고 유형 분류를 위한 유사도 평가 기반의 앙상블 모델은 다중 프로파일의 개수가 증가할수록 분류 정확률이 높아지고 있음을 [그림 10]과 같은 실험 결과로 확인 할 수 있다. 전반적으로 separated profile 이 combined profile 보다 분류 정확률이 높음을 알 수 있다.

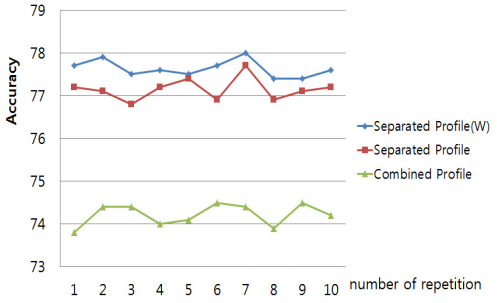


그림 10. 분류기의 학습 반복에 따른 분류의 정확도

[그림 10]에서 통합된 프로파일(Combined Profile)는 [그림 6]의 individual classification model 의 측정값이고, 분리된 프로파일(Seperated Profile)은 [그림 7]의 Ensemble classification model을 기반으로 측정한 결과이다. Separated profile은 [그림 3]의 결합 알고리즘을 적용한 것으로 가중 다수 투표(Weighted Majority Voting)가 다수 투표(Majority Voting)보다 분류 정확률이 높음을 알 수 있다.

V. 결 론

본 논문은 빅데이터 환경에서 수집된 보안 인텔리전스의 침해사고 정보를 분석하는 다중 프로파일 기반 앙상블 모델을 사용하여 침해사고 프로파일을 생성하기 위해서 어떤 정보를 사용해야 하고 그것들을 효과적으로 결합하고 활용하는 방법에 중점을 두고 있다. 제안 앙상블 모델은 보안 인텔리전스로 부터 침해사고를 구성하고, 침해사고 구성으로 부터 침해사고 프로파일을 도출하여 유사한 침해사고에 대한 분류의 정확도와 신뢰성을 향상시킨다. 향후 침해사고 유형 분류 앙상블 알고리즘의 개선을 위해서 다른 침해사고 분석 알고리즘과의 비교를 통해서 분류의 정확도를 비교하는 추가적인 연구가 필요하다. 또한 제안 알고리즘은 침해사고 간의 유사성을 찾는 데는 좋은 결과를 보이지만 침해사고의 변칙성을 찾는 데는 그다지 효과적이지 못하다. 따라서 전 세계적으로 빅데이터 활용도를 극대화하는 연구 경향에 맞추어, 현재 정교해지는 사이버 위협에 대응 하는 전통적인 대응 방법들의 한계를 벗어나 기계학습 알고리즘의 접목을 통한 침해사고 간 유사성과 함께 침해사고의 변칙성을 찾는 알고리즘의 연구와 트레이닝 데이터와 테스트 데이터를 사용한 일반화된 알고리즘의 연구가 필요하다.

참 고 문 헌

[1] 김영수, 문형진, 조혜선, 김병익, 이진혜, 이진우, 이병엽, “계층적침해자원기반의 침해사고 구성 및 유형 분석,” 한국콘텐츠학회논문지, 제16권, 제11호, pp.139-153, 2016.

[2] Y. S. Kim, H. J, Mun, H. S. Cho, B. I. Kim, J. H. Lee, J. W. Lee, and B. Y. Lee, "Analysis Model of Cyber Incident based Threat Intelligence," International Conference on Convergence Content 2016, pp.351-352, Dec. 10, 2016

[3] C. Ten, G. Manimaran, and C. Liu, Cybersecurity for Critical Infrastructures : Attack and Defense Modeling, IEEE TRANSACTIONS ON SYSTEMS, Vol.40, No.4, pp.853-865, 2000.

[4] M. A. Faysel and S. S. Haque, "Towards Cyber Defense: Research in Intrusion Detection and Intrusion Prevention Systems," IJCSNS, Vol.10, No.7, pp.316-325, 2010.

[5] H. D. Nguyen and Q. Cheng, An Efficient Feature Selection Method For Distributed Cyber Attack Detection and Classification, 2011 45th Annual Conference on Information Sciences and Systems (CISS), pp.1-6, 2011.

[6] B. K. Mishra and H. Saini, Cyber Attack

Classification using Game Theoretic Weighted Metrics Approach, World Applied Sciences Journal 7(Special Issue of Computer & IT), pp.206-215, 2009.

[7] H. Du, C. Murphy, J. Bean, and S. J. Yang, “Toward Unsupervised Classification of Non-uniform Cyber Attack Tracks,” International Conference on Information Fusion, pp.1919-1925, 2009.

[8] A. Jain and A. K. Singh, “Distributed Denial Of Service (Ddos) Attacks - Classification And Implications," Journal of Information and Operations Management, Vol.3, No.1, pp.136-140, 2012.

[9] B. Dharamkar and R. R. Singh, "Cyber-Attack Classification Using Improved Ensemble Technique Based On Support Vector Machine and Neural Network," International Journal of Computer Application, Vol.103, No.11, pp.1-7, 2014.

[10] P. Amudha, S. Karthik, and S. Sivakumari, "An Experimental Analysis of Hybrid Classification Approach for Intrusion Detection," Indian Journal of Science and Technology, Vol.9, No.13, April, 2016.

[11] M. Sharma, S. K. Singh, P. Agrawal, and V. Madaan, "Classification of Clinical Dataset of Cervical Cancer using KNN," Indian Journal of Science and Technology, Vol.9, No.28, July, 2016.

[12] S. R. Suganthi and M. Hanumanthappa, "Classification of Event Image Set Using Mining Techniques," Indian Journal of Science and Technology, Vol.9, No.22, June, 2016.

[13] P. E. Jouve and N. Nicoloyannis, A New Method for Combining Partitions, Applications for Distributed Clustering. Proc. of the International Workshop on Parallel and Distributed Machine Learning and Data Mining, pp.69-76, 2003.

[14] A. Verma, I. Kaur, and A. Kaur, "Algorithmic Approach to Data Mining and Classification Techniques," Indian Journal of Science and Technology, Vol.9, No.28, July, 2016.

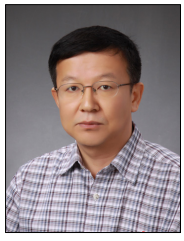
[15] S. Vega-Pons and J. Ruiz-Shulcloper, "A survey of clustering ensemble algorithms," International Journal of Pattern Recognition and Artificial Intelligence, Vol.25, No.3, pp.337-372, 2011.

[16] S. Singh and S. Silakari, An Ensemble Approach for Cyber Attack Detection System: A Generic Framework, Proc. 14th ACIS Int. Conf. Softw. Eng. Artif. Intell. Netw. Parallel/Distrib. Comput., pp.79-84, 2013.

[17] D. Rathore and A. Jain, "Design Hybrid method for intrusion detection using Ensemble cluster classification and SOM network," International Journal of Advanced Computer Research, Vol.2, No.5, pp.181-186, 2012.

저 자 소 개

김 영 수(Young Soo Kim) 정회원



- 2003년 8월 : 국민대학교정보관리학(정보관리학박사)
- 현재 : 충남 재활IT 융합 기술원 대표 컨설턴트

<관심분야> : 빅데이터보안, 정보 보안, 기계학습