안전한 국방 빅데이터 프레임워크를 위한 Learned MAPE-K 기반 자료교환 시스템

조준하[1], 유진용[1,2], 김영갑[1,2†]
[1]세종대학교 정보보호학과
[2]세종대학교 지능형드론융합전공
wearegoodman@naver.com, instrol30@gmail.com, alwaysgabi@sejong.ac.kr

Data Exchange System Based on Learned MAPE-K for a Secure Defense Big Data Framework

Jun-Ha Cho[1], Jin-Yong Yu[2], Young-Gab Kim[2,†]
[1]Department of Computer and Information Security, Sejong University
[2]Dept. of Computer and Information Security, and Convergence Engineering for Intelligent Drone, Sejong University

요 약

국방 각급 부대는 망연계 자료교환 시스템에 의해 인터넷과 국방망을 연계하여 데이터를 수집하고 있다. 또한, 안전한 국방 데이터수집과 빅데이터 환경조성을 위해 악성코드를 내재한 데이터들을 차단 및 분류하는 데이터 검열을 수행한다. 그러나 수집되는 데이터들이 새로운 악성코드를 내재할 경우, 현재 운용되고 있는 국방 시스템으로 식별하는 것이 불가능하여 외부로부터의 보안위협이 존재한다. 따라서 본 논문에서는 새로운 악성코드 위협에도 대응할 수 있는 Learned MAPE-K 기반 자료교환 시스템을 제안한다.

1. 서론

최근 4차 산업혁명이 가속화됨에 따라 국방 분야에서도 민간영역의 클라우드 컴퓨팅(Cloud Computing), 빅데이터(Big Data) 시스템, 사물인터넷(Internet of Things; IoT), 메타버스(Metaverse) 등의 신기술 도입 및 발전을 위한 계획을 수립 중이다. 특히, 육군은 비밀데이터 유통과 빅데이터 분석 기능을 제공하는 전장지능화센터를 추진하고 있다.

군은 인터넷과 폐쇄망으로 구성된 국방망에 대해 국방 망연동 보안 가이드라인을 준수한 자료교환 시스템을 통해 데이터를 수집하고 있다. 그러나 수집되는 데이터가 시그니처 기반의 알려진 악성코드에 대해서만 식별하고, 최신 악성코드에 대해서는 식별이 제한되는 단점이 있어 이에 대한 기술적인 보안대책이 요구되는 상황이다. 특히 앞으로 추진될 육군의 전장지능화센터에서는 공개 데이터와 비밀 데이터가 융합되어 활용범위를 확장할 것이다. 이에 따라, 안전한 국방 빅데이터 플랫폼을 구성하기 위해 데이터수집 시, 기존의 식별하지 못한 최신의 악성코드를 탐지 및 검열할 수 있어야 한다.

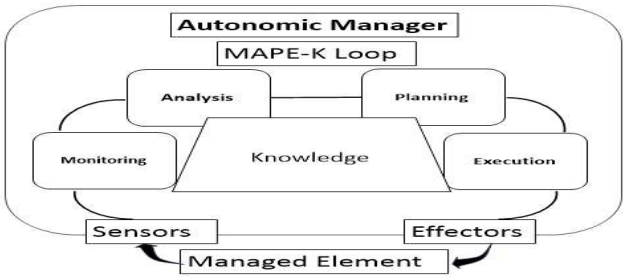
따라서 본 연구에서는 안전한 국방 빅데이터 환경을 조성하기 위해 Learned MAPE-K 기반 자료교환 시스템을 제안한다. 제안된 자료교환 시스템은 식별되지 않은 악성코드를 내재하고 있는 데이터를 분류 및 검열함으로써 안전한 국방 빅데이터 환경을 조성할 수 있다. 또한, 국방 망연동 보안 가이드라인을 준수하며 Learned MAPE-K를 적용하여 데이터 분류를 시도한 첫 사례이다.

본 논문의 구성은 다음과 같이 구성된다. 2장에서는 기존 MAPE-K 시스템과 관련된 연구들을 분석한다. 3장에서는 국방 망연계 자료교환 시스템에 Learned MAPE-K를 적용해 안전한 데이터수집을 할 수 있음을 서술하고, 활용방안으로 분석된 데이터를 국방망 및 전장망 사용자에게 제공해 주기 위한 국방 빅데이터 프레임워크를 제시한다. 마지막으로 4장에서는 결론과 향후 연구에 대해 서술한다.

2. 관련연구

피드백 루프가 반복됨으로써 시스템 스스로가 변화하는 환경에 적응하는 것을 목표로 하는 시스템을 자가-적응 시스템이라고 한다.[1] 이러한 자가-적응

시스템 중에서 가장 많이 사용되고 있는 MAPE-K 루프는 그림 1과 같이 총 4단계로 구성된다. 모니터링은 센서 인터페이스를 사용하여 ME(Management Element)에서 수집된 세부 정보를 수집, 집계, 필터링 및 보고하는 메커니즘을 제공한다. 분석은 복잡한 상황을 상호 연관시키고 모델링하는 메커니즘을 제공한다. 이러한 메커니즘을 통해 AM(Autonomic Manager)은 IT 환경을 이해하고 미래의 상황을 예측할 수 있다. 계획은 목표와 목표를 달성하는 데 필요한 작업을 구성하는 메커니즘을 제공한다. 실행은 이펙터를 사용하여 동적 업데이트를 고려하면서 계획 실행을 제어하는 메커니즘을 제공한다. 마지막으로 Knowledge는 루프의 동작을 지원하기 위해 다양한 유형의 지식을 포함할 수 있는 지식 소스를 관리한다.[2]



(그림 1) MAPE-K loop 개요[2]

Kleber Vieira 외[3]는 대규모 분산시스템에서 사이버 보안 프로세스를 최적화하기 위해 MAPE-K를 사용해 자율 침입 감지 및 대응방법을 제시하였다. 이는 하이재킹, 중간자 및 DDoS 공격에 취약한 분산시스템에 대한 사이버 보안을 제공할 뿐만 아니라 복잡한 환경에서 효과성과 확장성이 크게 향상되었음을 증명하였다. Dimitrios Papamartzivanos 외[4]는 Misuse IDS(Intrusion Detection System)에 self-taught learning과 MAPE-K를 결합한 프레임워크를 제안하여 지속되고 급격한 사이버 공격에 대해 73.37%라는 높은 탐지 비율을 보여주는 동시에 IDS를 수동으로 업데이트하지 않아도 됨을 서술하였다. 그러나 실험에 사용한 데이터 셋이 적었고, 새로운 종류의 악성코드 공격탐지 효율성이 73.37%로 아주 높지는 않았다.

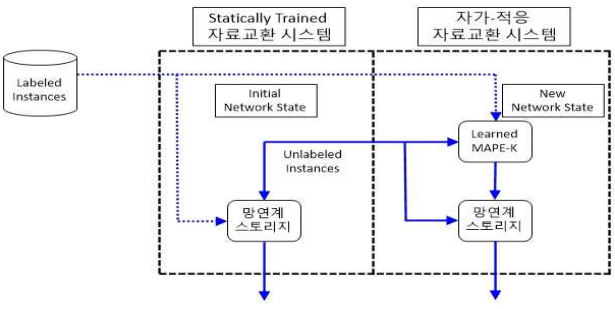
3. 시스템 제안

제안한 시스템은 국방 자료교환 시스템에 데이터 수집 시 Learned MAPE-K를 적용해 미리 알려지지 않은 악성코드에 대해서도 안전성을 제공하기 위한 시스템이다. 본 장의 3.1절에서는 Learned MAPE-K 기반 자료교환 시스템에 대한 동작 방식을 설명하고, 3.2절에는 국방 빅데이터 프레임워크 구성 시 Learned MAPE-K 자료교환 시스템 활용방안에 대해서 설명한다.

3.1. Learned MAPE-K 기반 자료교환 시스템 동작방식

국방 망연계 자료교환 시스템은 인터넷 공개자료 수집을 통해 빅데이터 환경을 조성한다. 이때 인터넷망 방향에서 데이터수집 시 TCP/IP 프로토콜이 아닌 전용 프로토콜로 변경되는 지점이 있다. 이 지점에 그림 2와 같이 Learned MAPE-K를 적용하여 새로운 악성코드에 대해서도 안전한 데이터수집을 한다. MAPE-K에 대한 동작 방식은 다음과 같다.

1) **모니터링(Monitoring)**: 인터넷에서 국방 자료교환 시스템을 경유하여 수집되는 데이터를 감시하고 정보를 수집한다. Knowledge에서는 네트워크 트래픽이 저장된다.
2) **분석(Analysis)**: 탐지된 모니터링 데이터를 Audit Tool을 이용하여 분석한다. 정적인 시그니처 기반의 자료교환 시스템과 달리 계획 단계에서 수립된 Deep Learning 정책을 통한 악성코드 분석을 실시한다.
3) **계획(Planning)**: Deep Learning 인공신경망(Artificial Neural Network)을 적용하여 새로운 트레이닝 데이터 셋을 구하여 Knowledge를 업데이트하고 자료교환 시스템에 대한 자가-적응 정책을 수립한다.
4) **실행(Execution)**: 학습하여 결정된 정책을 토대로 새로운 악의적인 패턴의 악성코드가 있다면 이것을 식별하여 안전한 데이터만 폐쇄망으로 구성된 빅데이터 플랫폼 영역 내부로 전송한다.

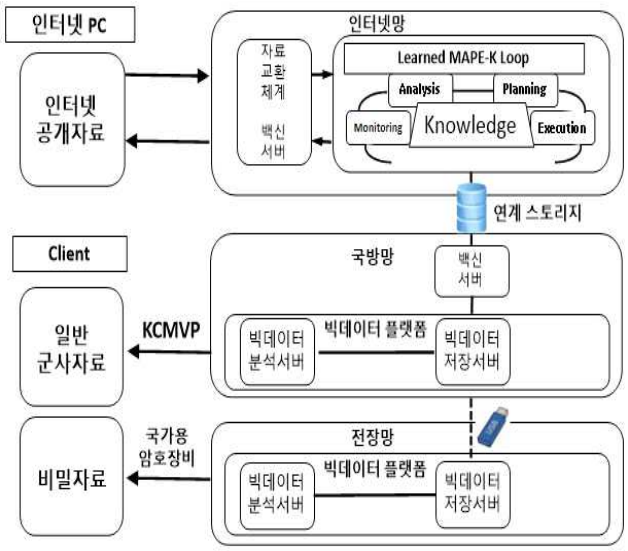


(그림 2) Learned MAPE-K 자료교환 시스템

3.2. Learned MAPE-K 기반 자료교환 시스템을 활용한 국방 빅데이터 프레임워크

안전한 국방 빅데이터 환경조성을 위해 Learned MAPE-K 기반 자료교환 시스템 적용 시 빅데이터 프레임워크 활용방안은 다음과 같다.

1) 그림3과 같이 국방 망연계 스토리지를 경유하여 수집된 데이터인 인터넷 공개 데이터와 국방망 내부에서 수집된 일반군사자료는 국방망에서, 비밀 데이터는 전장망에서 각각 빅데이터 플랫폼을 구성한다. 이때 국방망에 있는 공개 데이터 및 일반군사자료는 전장망으로 데이터를 전송할 수 있지만, 전장망에 있는 데이터는 비밀 데이터이므로 국방망으로 전송할 수 없다.
2) 국방망과 전장망 각각의 빅데이터 플랫폼 내부에서는 저장된 데이터가 복호화된 상태에서 빅데이터 분석을 한다.
3) 권한이 있는 사용자에게 분석자료가 제공될 때에는 일반군사자료는 KCMVP로 암호화하고, 비밀 데이터는 국가용 암호 장비로 암호화하여 제공한다.



(그림 3) Learned MAPE-K 자료교환 시스템 활용방안

4. 결론 및 향후 연구

국방 망연계 자료교환 시스템을 통해 데이터수집을 할 때, 국방 분야 최초로 Learned MAPE-K를 적용해 미리 알려지지 않은 악성코드에 대해서도 안전한 자료교환 시스템을 제안하였다. 제안한 시스템을 활용하면 안전한 국방 데이터수집과 빅데이터 환경조성을 할 수 있다.

향후 연구는 망연계 자료교환 시스템과 유사한 실험환경을 구성하고 다양한 데이터 셑을 준비한 후, 어떤 Learned MAPE-K 기반 알고리즘이 새로운 악성코드에 대해 안전하면서도 가장 효율성이 높은지 실험 및 비교하는 연구를 할 계획이다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1A2C2012635)

참고문헌

[1] 서영덕, 김영갑, 이의종, 설광수, 백두권, “스마트 온실을 위한 온톨로지 지식저장소 설계 및 시뮬레이션”, 한국경영과학회 2017년 춘계공동학술대회 논문집, 4753-4757p, 2017.

[2] IBM, “An architectural blueprint for autonomic computing”, IBM White Paper, 2005.

[3] Kleber Vieira, Fernando L. Koch , João Bosco M. Sobral, Carlos Becker Westphall, and Jorge Lopes de Souza Leão, “Autonomic Intrusion Detection and Response Using Big Data.”, IEEE Systems Journal, 14(2), 8p, 2019.

[4] Papamartzivanos, Dimitrios, Félix Gómez Mármol, and Georgios Kambourakis, “Introducing Deep Learning Self-Adaptive Misuse Network Intrusion Detection Systems.”, IEEE Access, 7, 13546-13560p, 2019.

[5] Mehdi Mohammadi, Ala Al-Fuqaha, Sameh Sorour, Mohsen Guizani, “Deep Learning for IoT Big Data and Streaming Analytics: A Survey.”, IEEE Communications Surveys & Tutorials, 20(4), 2923-2960p, 2018.