Journal of The Korea Institute of Information Security & Cryptology ISSN 1598-3986(Print)
VOL.27, NO.3, Jun. 2017 ISSN 2288-2715(Online)
https://doi.org/10.13089/JKIISC.2017.27.3.563

동적 악성코드 분석 시스템 효율성 향상을 위한 사전 필터링 요소 연구

윤 광 택,† 이 경 호‡
고려대학교

Study of Pre-Filtering Factor for Effectively Improving Dynamic Malware Analysis System

Kwang-Taek Youn,† Kyung-Ho Lee‡
Korea University

요 약

인터넷과 컴퓨터의 발달로 인해 신종 변종 악성코드가 하루에 약 1백만 개씩 출현하고 있다. 더욱이 기업을 대상으로 하는 표적공격의 경우 알려지지 않은 악성코드를 통해 공격이 진행되므로 전통적인 시그니처에 의한 탐지 방법은 대응에 대한 효율성이 낮게 되어 많은 기업들은 새로운 샌드박스와 같은 동적 분석 시스템을 도입하였다. 그러나 실행 파일뿐만 아니라 워드문서 또는 PDF 형태의 악성코드도 지속적으로 증가하고 있으며 새로운 악성코드 또한 동적 분석 시스템을 우회하는 기술을 포함하고 있어 효율적인 운영을 위해 많은 자원이 필요하고 새로운 기술이 필요하게 되었다. 본 연구에서는 효율적인 동적 분석 시스템을 위해 사전 필터링 기술을 사용하여 효율성을 향상시키기 위한 사전 필터링 기술 선정 요소를 도출하고 기술 도입 시 합리적인 선택을 할 수 있도록 AHP(Analytics Hierarchy Process)를 사용하여 의사 결정 모델을 제시하고, 도입 시 활용할 수 있도록 공식을 제시하고 검증하였다.

ABSTRACT

Due to the Internet and computing capability, new and variant malware are discovered around 1 Million per day. Companies use dynamic analysis such as behavior analysis on virtual machines for unknown malware detection because attackers use unknown malware which is not detected by signature based AV effectively. But growing number of malware types are not only PE(Portable Executable) but also non-PE such as MS word or PDF therefore dynamic analysis must need more resources and computing powers to improve detection effectiveness. This study elicits the pre-filtering system evaluation factor to improve effective dynamic malware analysis system and presents and verifies the decision making model and the formula for solution selection using AHP(Analytics Hierarchy Process)

Keywords: Dynamic Analysis, Malware, Effectiveness, AHP, APT

I. 서 론

1.1 연구 배경 및 목적

인터넷과 컴퓨팅 파워의 발달로 인해 악성코드의 증가는 폭발적으로 증가하고 있는 상태이다. 시만텍사의 2017년 인터넷 보안 위협 보고서에서 의하면 2016년 한해 발견된 악성코드는 약 3억5천7백만 개이다. 2015년 대비 약 0.5% 증가한 것으로 보고되었다[1]. 악성코드가 매해 마다 증가하는 가장 큰 이유 는 실행 압축 기술이나 여러 가지 난독화 기술 등을 이용한 자동화 툴의 보급으로 악성코드 변종 생성을 기하급수적으로 증가시키기 때문이며, 결과적으

Received(05. 22. 2017), Modified(06. 07. 2017), Accepted(06. 09. 2017)
† 주저자, patrick_youn@naver.com
‡ 교신저자, kevenlee@korea.ac.kr(Corresponding author)

로 안티바이러스 제조사 또는 기업에서는 기존의 전통적인 기술로는 안티바이러스 대응 효율성이 낮아지게 되어 기존의 탐지 기술의 효율성을 향상 시키거나 또는 새로운 탐지 기술이 필요하게 되었다.

2014년에 보고된 미국 기업 타겟의 보안 사고의 경우 300명 이상의 보안 인력과 수조원대의 보안관제 시스템을 운영하고 있었으며, 신종 변종 악성코드 탐지를 위한 동적 분석 기술로 파일을 행동기반으로 분석하는 시스템까지 운영하였으나 너무나 많은 탐지로 인해 침해사고 대응 팀에 경고가 과다하게 통지되었고 그 결과 침해 사고 대응 팀이 해당 경고를 무시하여, 그 결과 약 4천개의 신용카드 정보가 유출되는 사고가 발생하였다[2].

기업 보안 담당자에게 있어서 악성코드를 통한 공격은 얼마나 빠르고 효율적으로 대응할 수 있는가가 침해로 인한 피해를 얼마나 줄일 수 있는 가로 귀결되게 되었다. 특히 최근과 같이 기하급수적으로 발생하는 악성코드 대응을 위해서, 전통적인 시그니처에 의한 탐지 기술을 사용하는 경우, 신종 악성코드가 안티바이러스 회사에 접수되어 악성코드 분석가에 의해 분석 후 탐지 시그니처를 생성한 후 다시 해당 제품을 사용하는 고객에게 전달할 때까지는 많은 시간이 소요되므로, 빠른 탐지 효율성을 위한 동적 분석 기술이 각광을 받게 되었고 그 의존도가 점점 높아지고 있다.

그러나 동적 분석 악성코드 분석 시스템 도입 시, 철저한 사전준비 및 운영계획 없이는 성공적인 도입 및 활용이 불가능하다. 예를 들어 최근 공격은 파일 형식이 실행파일(Portable Executable) 형식뿐만 아니라 비 실행형(Non-PE) 파일 형식 즉 오피스 형태의 파일[1]로 제작되고 유포되고 있어 해당 애플리케이션의 취약점이 존재한 상태이어야 악성코드 탐지가 가능하며, 최근의 출현되는 악성코드의 경우는 동적 분석 기술에 의한 탐지를 회피하는 기술을 포함하는 악성코드는, 전체 알려진 악성코드의 약 20%를 차지하고 있고 지속적인 증가세에 있어 단지 동적 분석에 의한 탐지 방법에 의존하는 것도 탐지 효율성이 낮아지고 있다.[3].

Fig 1은 이제까지 연구되어온 동적 분석 시스템의 효율화를 위해 머신러닝을 접목하여 탐지 효율을 높이거나[15] 모든 행위를 다 분석 후 악성 유무를 판단하지 않고 악성으로 판정할 만한 일정 임계치에 도달하면 악성으로 판정하는[13] 등의 여러 가지 기술을 투입하여 효율화를 증대시키는 방안들이 연구

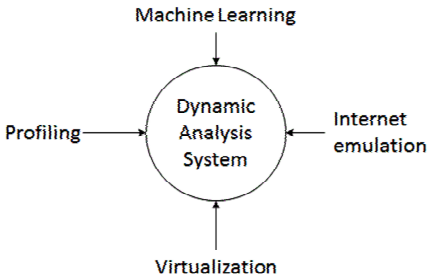

Fig. 1. Dynamic Analysis

중이다. 동적 분석 시스템의 효율성을 향상 시키는 방안이 동적 분석 시스템 자체에 새로운 기술을 접목하여 효율성을 향상 시키는 방안에만 연구되었으나, 사전 필터링 기술을 활용하여 동적 분석 시스템에 유입되는 파일의 수를 감소 시켜 동적 분석 시스템의 효율성을 향상 시키는 방안은 연구되지 않았다.

Fig 2는 동적 분석 시스템 효율성 향상을 위한 사전 필터링 예이다. 사전 필터링을 사용하는 경우, 이미 탐지되는 알려진 악성코드 및 정상 파일은 더 이상 동적 분석시스템에 분석되지 않고, 신종 악성파일만 분석되므로 동적 분석 시스템의 부하를 감소시킬 수 있다.

이에 본 연구에서는 동적 악성코드 분석 시스템의 효율성 증대를 위해 선택될 수 있는 사전 필터링 기술을 조사하고 이 기술 선정 시 고려되어야 하는 중요 요소를 도출하여 대안 선정 시 체계적인 접근법을 제시한다.

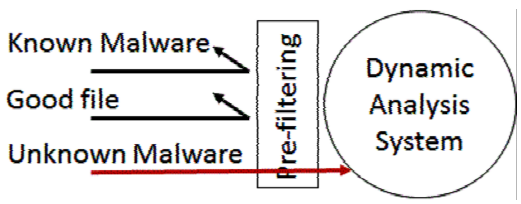

Fig. 2. Pre-filtering System

1.2 연구 방법과 구성

이에 본 연구에서는 효율적인 동적 악성코드 분석 시스템의 사전 필터링 기술 선정 시 다양한 고려사항에 직면하게 되는데, 사전 연구를 통해 고려되어야 하는 중요 요소를 도출하고, 이 도출된 각각의 요소의 중요도 선정을 다기준의사결정(Multiple

criteria decision making: MCDM) 문제로 간주하고, 대표적인 MCDM문제를 해결 위해 계층분석과정(Analytic Hierarchy Process: AHP)을 적용하여 도출된 요소의 중요도를 산출하며 향후 사전 필터링 선정 시 활용되도록 공식을 도출한다.

이를 위해 악성코드 동적 분석 시스템 효율성 향상을 위한 평가 기준을 바탕으로 수립된 AHP 모델을 이용하여 국내 동적 악성코드 분석 시스템 도입 및 그와 관련된 보안 업무 관련자 30명을 대상으로 설문조사를 수행하였다. 그중 유효성 평가를 통해 유효한 설문 응답자 21명의 설문 조사 결과를 바탕으로 관리자 그룹, 운영자 그룹 그리고 전문가 그룹으로 분류하여 사용자 그룹별 사전 필터링 요소 선정의 특성을 분석하였으며, 대안 기술 선정 공식을 도출하기 위해 3개의 그룹을 하나로 통합하여 종합적으로 특성을 분석하여 공식을 도출하였다.

다음 제2장에서는 선행 연구를 통해 대표적인 악성코드 분석 기술은 정적 분석과 동적 분석에 대한 개념과, 악성코드 분석의 효율성을 위한 사전 연구와 해외 선진 악성코드 탐지를 위한 모델을 알아보고 AHP 모델을 연구한다. 제3장에서는 악성코드 동적 분석 시스템 운영 현황 및 제안 모델을 제안한다. 제4장에서는 AHP 조사기법을 통해 동적 악성코드 분석 시스템의 효율성 향상을 위한 사전 필터링 기술 선정 요인과 요인별 대안의 중요도를 평가하여 분석 결과를 검토하며 향후 제안 모델을 적용하여 활용하기 위한 공식을 도출하고 실제 사례를 통해 공식을 증명한다. 마지막 제5장에서는 본 연구의 결론과 연구를 수행함에 있어서 도출된 한계점을 제시한다.

II. 선행 연구 검토

2.1 악성코드 분석의 효율성을 위한 연구

Christian Gorrecki [7]등은 최근의 악성코드는 동적 분석 시스템에서 분석을 회피하기 위해서 또는 인터넷이 연결되어 특정 도메인에 연결되는 경우에만 악성 행위가 실행되도록 설계되었기 때문에 인터넷이 연결된 것처럼 에뮬레이션 하여 악성행위를 하도록 유인하여 탐지 효율을 높이도록 제시하였다. Tamas K. Lengyel [19]등은 동적 분석 시스템에 의해 분석되어야 하는 악성코드의 수가 지속적으로 증가하고 악성코드가 진화하여 동적 분석 기술을 인식하고 동적 분석 시스템의 모니터링 영역 이외에 숨김 기능을 가지고 있기 때문에 최신의 가상화 기술과 Xen 하이퍼바이저의 기술을 활용하여 동적 분석 시스템의 모니터링 기능의 흔적을 남기지 않고 분석 할 수 있는 기술을 제시하였다. Ulrich Bayer 등[13]는 동적 분석 시스템은 악성코드가 실행 된 후 파일의 생성, 레지스트리의 생성 네트워크 사용, 서비스의 활동 등을 모니터링 하도록 설계되어 있는데, 최근 악성코드는 실행되고 수 초 동안에는 악성행위를 하지 않도록 설계되어 있어 동적 분석 시스템은 분석을 위한 많은 자원이 필요하게 되었다. 악성코드 제작자는 다형성 기술을 사용하거나 또는 실행파일 압축 기술을 사용하여 우회기술을 사용하므로, 악성코드 행동에 대한 데이터베스를 구축한 후 악성코드 실행시 식별되는 악성행위를 데이터베이스와 비교한 후 악성 여부를 판단하는 기술을 제시하였다. Konrad Rieck 외[15]는 동적 분석 시스템의 효율성 향상을 위해 머신러닝 알고리듬을 사용한 프레임워크를 제시하였다. 이는 유사한 행위들의 집합(Clustering)과 알려지지 않은 악성코드를 사전에 분류된 악성코드 집합에 분류(Classification)하는 알고리즘을 활용하고 증분 분석(Incremental Analysis)를 통해 분석에 투입되는 자원 효율성을 극대화 하였다[15]. 권종훈 외[23] 변종 신종 악성코드를 빠르게 탐지하기 위해 행위 그래프 분석을 통한 악성코드 모듈별 유사도 분석 기법을 제시하였는데, 악성코드들이 사용하는 2,400개 이상의 API등을 분석하여 128개의 추상화 하였고 그 결과 모두 탐지가 가능했다.

2.2 해외 선진 악성코드 탐지를 위한 모델 분석

악성코드 분석 기술은 정적 분석 또는 동적 분석에 의해 악성 유무를 판단하는 기술이 있으나, 이 부분은 악성을 판단하기 위한 기술이다. 많은 보안 회사는 이러한 분석 기술을 사용하여 여러 가지 제품 및 서비스로 악성코드 탐지 기술을 제공하고 있다.

2.2.1 인텔리전스 기반 솔루션

인텔리전스 기반의 악성코드 탐지 기술은 일반적으로 악성코드 유포 도메인, 공격자 IP, 명령 제어 시스템, 봇 등과 같은 위협요소의 IP나 도메인을 주기적으로 정보를 서비스 형태로 제공해주는 형태를 가지고 있다.

S사의 경우 이러한 인텔리전스 정보를 특정 주기

로 제공하며 기존의 보안 솔루션이 방화벽, 차세대 방화벽, 보안 관제 분석 시스템에 연동하거나 별도의 개방형 API를 통해 정보를 제공하고 있다[26].

G사의 경우 다수의 보안 회사들이 제공하는 안티바이러스 엔진의 연동을 통해 악성코드를 분석의뢰 시 다수의 안티바이러스 엔진의 분석 결과를 활용하거나 또는 해시값을 조회하면 탐지 이력에 대한 값을 제공해 주는 서비스를 제공하고 있다[27].

2.2.2 안티바이러스 솔루션

S사의 안티바이러스 탐지 기술은 기존의 시그니처 기반의 악성코드 탐지뿐만 아니라, 조금 더 진보한 휴리스틱 탐지 기술 등을 포함하고 있으며, 그 이외의 여러 가지 기존의 탐지 기술을 보완하기 위한 파일 평판, 행동 기반 탐지 기술 등을 통해 악성코드 탐지의 효율성을 높이고 있다[26].

K사의 경우도 기존의 전통적인 시그니처 기술과 휴리스틱 탐지 기술을 포함하고 있으며, 이를 보완하기 위한 여러 가지를 포함하여 탐지 효율을 보완하고 있다[28].

2.2.3 화이트리스팅

화이트리스팅이란 일반적으로 애플리케이션 화이트리스팅을 의미한다. 이것은 여러 가지 정보를 나열하여 해당 애플리케이션만 호스트에서 실행하고 그 이외의 파일 등은 실행하지 못하도록 하여 악성코드나 비인가된 애플리케이션을 차단하는 목적을 가지고 있다. 이러한 애플리케이션의 형식이란 파일 경로, 파일 이름, 파일 크기, 디지털 서명, 해시값 등을 제공하여 구현한다[24].

미국의 C사의 또는 M사의 경우 호스트 시스템에서 상기의 제공되는 애플리케이션 속성을 이용하여 호스트 상에서 프로세스의 실행을 통제하는 기능을 제공한다[29][30].

2.2.4 머신러닝에 의한 탐지 솔루션

기존의 안티바이러스 엔진이 악성코드 분석에 의한 시그니처 탐지로 악성유무를 판단하는 것은 신종 변종에 대한 대응이 효율적이지 않은 단점을 보완하기 위해 인공지능인 머신러닝 알고리즘을 악성코드 탐지에 적용한 솔루션이다. C사, S사 및 I사는 대표적인 머신러닝 기술을 활용하여 악성코드 탐지에 적용한 솔루션을 제공한다[31][26][32].

악성코드의 집합과 악성코드가 아닌 안전한 파일의 집합 모두를 학습하여 악성유무를 판단하게 된다. 악성의 유무를 머신러닝 알고리즘에 의한 신뢰도로 악성 유무를 판단하게 된다. 유사한 형태의 악성코드의 신종 변종 탐지에 좋은 효과가 있다.

2.2.5 각 솔루션별 장단점 비교

선진 악성코드 탐지 솔루션 및 서비스의 대한 장단점은 Table 1에 나와 있는 것과 같이 장단점이 존재한다.

인텔리전스 기반의 솔루션은 업체의 빅데이터 분석과 신뢰성 검증을 통해 제공되는 데이터로 빠른 탐지가 가능하며, 기존에 도입된 방화벽과 같은 네트워크 시스템 또는 보안 관제 분석 시스템에 연동하여 차단 및 분석에 활용이 가능하다. 단점으로는 비용이 비싸며, 기존에 도입된 솔루션에 연동하여 적용할 경우 시스템 성능 저하의 문제가 존재하며 제공하는 업체도 소수에 불과하다.

안티바이러스 솔루션은 오랫동안 안티바이러스 탐지에 기여해온 솔루션으로 많은 업체가 존재하고, 낮은 오탐지율을 제공하며, 운영비용이 적게 소요 되는 장점이 있다. 반면 신규 또는 변종 악성코드에 대한 탐지율이 낮은 단점이 존재한다.

화이트리스팅 솔루션은 제공하는 업체가 작고, 낮은 오탐지율 및 신종 변종 탐지율이 좋은 장점이 있으나, 도입 비용이 높고, 운영시 비용도 많이 소요되는 편이다.

머신 러닝 기반의 솔루션은 신종 변종 탐지에 대한 탐지 효율이 좋으나 도입비용이 높고 오탐지에 의한 사후 처리 비용이 높게 소요되며 제공하는 업체가 많지 않다.

Table 1. Anti-malware Solution's Pros and Cons

Solution	Pros	Cons
Intelligence-based	• Good integration with existing solution or API • Fast detection	• High cost • Scalability issue on top of existing solution • Small number of vendors

Anti-Virus (Blacklisting)	• Large number of vendors • Low false positive • Low CAPEX	• Weak for unknown malware
Whitelisting	• Small number of vendors • Low false positive • Good for unknown malware at host level	• High CAPEX • High OPEX
Machine Learning-based	• Good for unknown malware	• High Cost • High false positives • High OPEX for managing false positives • Small number of vendors

2.3 선행연구와 차이점

앞에서 살펴본 바와 같이, 기존의 연구는 악성코드 탐지를 동적 분석 시스템에서 빠른 탐지를 위한 여러 가지 기술적인 보완을 통해 효율성 증대 방안을 제시하고 있다. 그러나 새로 출현되는 악성코드 중 약 20%는 동적 분석 시스템이 주로 이용하는 가상 시스템을 인식하여 악성 행위를 하지 않거나 실행을 지연하여 동적 분석 시스템의 많은 자원을 차지하고 소모하고 있다. 동적 분석 시스템 자체의 탐지 빠른 탐지와 자원의 효율적인 분배를 위해 머신러닝, 프로파일링 등의 여러 가지 보완 기술을 통해 부족한 부분을 개선하려는 노력을 제안하였다.

본 연구에서는 동적 분석 시스템 자체의 성능 개선이나 기술 개선을 통한 효율성 증대의 측면이 아니라, 동적 분석 시스템에 분석을 위해 유입되는 파일이 기존에 알려져 있는 정상파일이나 또는 이미 탐지되고 있는 악성코드의 경우 사전 필터링을 통해 동적 분석 시스템에 의존도를 낮추어 어려워지고 있는 동적 분석 시스템의 효율성도 증대하고 더 나아가 탐지 결과가 알려지지 않은 악성파일에 대한 분석으로 침해사고 대응 측면에 있어서도 운영상의 효율성을 증대 할 수 있을 것이라 판단한다.

그러나 사전 필터링 기술이 하나만 존재하지 않고 여러 가지 기술이 존재하게 되는데, 이러한 다양한 기술의 선택 시 어떠한 부분이 중요한 선택 요소이고 고려사항을 체계적으로 제시하려고 한다.

III. 악성코드 분석 시스템 운영 현황 및 제안 모델

3.1 동적 분석 시스템 운영 현황

동적 분석 시스템은 일반적으로 악성코드가 유입되는 위치에서 파일의 유입시 악성 유무를 판단하는 역할을 하게 된다. Table 2는 동적 분석 시스템의 위치 및 분석 대상이다. 네트워크의 경우 사용자가 인터넷 사용시 유입되는 파일을 분석하기 위해 인터넷 관문에 위치한다. 이메일의 경우, 외부로부터 메일 수신시 첨부 파일에 대한 동적 분석 또는 이메일에 포함에 되어 있는 URL 방문시 유입되는 파일에 대한 동적 분석을 수행한다. 엔드포인트에서는 엔드포인트 저장되어 있는 파일 중 의심스러운 파일에 대해 동적 분석을 수행하게 된다.

Table 2. Dynamic Analysis location and objectives

Location	Objectives
Network	Malware analysis for inbound files via WEB
Email	Malware Analysis for email attachment or URL
Endpoint	Malware Analysis for file on the host

3.2 동적 분석 시스템 문제점 및 개선방안

동적 분석 시스템은 기본적으로 통제된 상태의 가상 시스템에서 파일을 실행하고 시스템의 변경 부분을 모니터링하며 악성행위를 탐지하는 시스템이다. 악성코드의 형태가 기존의 실행파일에서 오피스 문서와 같은 비실행 파일의 형태도 존재하고 있는데, 이러한 비실행형 문서 파일의 경우 해당 애플리케이션을 구동하여야 하며, 파일 열기시 해당 애플리케이션의 취약점이 악용되어 악성행위로 이루어져야 하는 환경을 갖추어야 한다. 결과적으로 취약한 애플리케이션을 설치 운영해야 하는 어려움이 존재한다.

더욱이 네트워크에서 분석한 파일이 이미 알려진 악성코드의 경우 네트워크 상에 있는 동적 분석 시스템도 악성으로 분류하나 사용자의 엔드포인트에 전달된 악성코드는 이미 알려진 악성파일이므로 차단되어 더 이상의 침해사고 대응이 필요하지 않으므로, 보안 관제에 있어서는 상관관계 분석을 통해 침해사고 대응 결정이 필요하다.

그러나 Fig 3과 같이 유입되는 모든 파일을 동적 분석 시스템에 의존하는 경우 분석되는 파일과 하나의 파일 분석 시 소요되는 시간을 고려하여 동적 분석 시스템을 준비하여야 한다.

이러한 문제를 해결하기 위해 Fig4와 같이 동적 분석 시스템에 분석하기 전에 사전에 필터링 기술로 알려진 파일에 대해 필터링한 후, 알려지지 않은 파일에 대해서만 동적 분석을 하는 경우 과탐에 의한 문제도 해소되며 고가의 동적 분석 시스템을 효율적으로 운영이 가능하다.

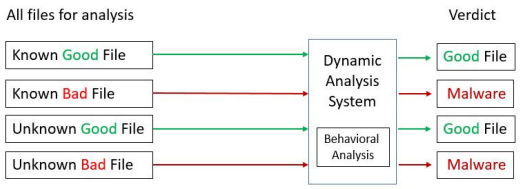

Fig. 3. As-is model for Dynamic Analysis

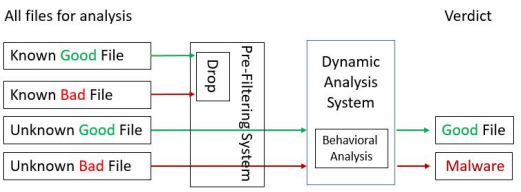

Fig. 4. To-be model for Dynamic Analysis

IV. 효율적인 동적 악성코드 분석 시스템을 위한 기술 선정 모델

본 장에서는 선행 연구를 바탕으로 효율적인 동적 분석 시스템 효율성 향상을 위한 기술 평가요인을 선정 및 정의하고 AHP(Analytic Hierarchy Process)를 사용하여 중요도 평가를 수행하고, 효율적인 동적 분석 시스템을 위한 기술 운영방식에 대한 대안을 선정 모델을 제시한다.

4.1 효율적인 악성코드 분석 결정을 위한 방법론

AHP는 Saaty[4]에 의해 개발된 의사결정 방법기법으로 복수의 대안에 대한 복수의 평가기준이 존재하는 다기준 의사결정(MCDM: Multiple Criteria Decision Making) 문제를 해결하기 위한 대표적인 의사결정 방법으로 다양한 분야의 의사결정 문제에 사용되어 왔다. 본 연구에서는 효율적인 악성코드 동적 분석 시스템 운영을 위한 사전 필터링 기술의 주요 요소 선정에 AHP를 이용하여 목표, 평가요인, 대안으로 이루어지는 계층 모델을 정의하고 쌍대비교(pairwise comparison)방식을 통해 각 계층별 의사결정 요인들 사이의 중요도를 평가하고 최종 우선 선위 도출에 사용한다.

4.2 효율적인 동적 분석 시스템의 사전 필터링 기술 요소 도출 절차

AHP는 계층 모델 설계, 쌍대비교, 부분 우선 순위 도출, 일관성 평가, 최종 우선순위 도출 및 대안의 선택의 총 5단계로 이루어진다[6].

각 단계별 수행 내용은 다음과 같다.

(1) 사전 분석: 사전 필터링 기술 선정을 위한 평가 기준 수립을 위한 선행연구
(2) 평가 요인 선정 및 계층 모델 수립: 사전 분석 결과를 바탕으로 최종 평가 기준 선정 및 AHP를 이용한 계층 모델 수립
(3) 설문 조사 및 유효성 평가: 앞 단계에서 수립된 AHP 계층 모델을 바탕으로 설문조사를 수행하고 그 결과 데이터에 대한 유효성을 평가
(4) 평가 요인별, 대안별 가중치 도출: 사용자 유형별로 그룹별로 구분한 후 평가 요인별, 대안별 중요도를 평가하여 최종 가중치를 도출
(5) 사전 필터링 기술 선정: 최종 도출된 결과값을 이용하여 중요평가 요인 및 대안을 바탕으로 사전 필터링 기술 선정

4.3 1단계: 사전 필터링 평가 요인의 선정

AHP를 사용하기 위해서는 우선 사전 필터링을 결정하는 평가 요인을 선정해야 한다.

평가 요인에 대한 연구를 살펴보면, Manish Dodse[20]는 서비스로서의 소프트웨어 선정시 8가

지의 평가 원칙으로 라이선싱, 유지보수, 연동, 확장성, 신뢰성, 보안성, 제조사의 고객 수, 비젼 등을 제시하였다.

Chun-Chi Wei[5]는 ERP시스템 선정시 6가지의 평가원칙으로 라이선싱, 유지보수, 운영비용, 보안, 제조사의 고객 수, 비전 등을 제시하였다.

신상필 외[6]는 모바일 오피스 구현방식 선정에 있어서 3가지의 중요 평가 원칙으로 라이선싱, 유지보수, 운영비용, 보안 등을 제시하였다.

Table 3은 상기의 선행연구를 비교 검토하고 효율적인 동적 악성코드 분석 시스템의 사전 필터링 기술 요소 선정을 위한 총 8가지 요소를 최종 선정하였다. 요소 선택시 제조사에 대한 고객 수 및 비젼 등은 동일한 제품 선정시 차별화 요인으로 선정하는 요인으로 이번 평가요인에서는 제외를 하였으며 대안의 기술 및 동적 악성코드 분석 시스템의 상호 운영을 위한 선정이므로 아키텍처 상의 연동성, 확장성, 신뢰성 등을 선정하였고, 오탐지 또는 과탐지에 의한 문제도 본 효율적인 운영에 있어 중요한 요소이므로 선정하였다.

Table 3. Pre-filtering System Evaluation Criteria

Principles	Manish Godse	Chun-ChiWei	Sangphil Shin	Fre-quency	Sele-ction
Licensing	V	V	V	3	O
Maintenance	V	V	V	3	O
Operation cost		V		1	O
Integration	V			1	O
Scalability	V			1	O
Reliability	V			1	O
Security	V	V	V	3	O
False positive				0	O
Number of client	V	V		2	X
vision	V	V		2	X

4.4 2단계: AHP 계층 모델 구현

위에서 선정된 총 8가지 평가 요소는 유사한 항목별로 분류하여 상위 평가 기준과 하위 평가기준으로 세분화하여 각 상위 평가 요소별 상대비교 할 수 있도록 Table 3과 같이 분류하여 각 항목별로 Table 4와 같이 상위평가에는 비용, 아키텍처, 위험의 3가지로 분류하고, 비용 상위평가에는 라이선스 비용, 유지보수 비용, 운영비용의 3가지 하위평가로 아키텍처 상위평가에는 연동성, 확장성, 신뢰성 3가지의 하위평가로, 위험 상위평가에는 보안성과 오탐율을 하위 평가요소로 분류하였다.

Table 4. Pre-filtering System Evaluation Criteria & Definition

Top criteria	Sub-criteria	Definition
Cost	Licensing	Purchasing Cost for official permission to use or own product/service. Product include Software/Hardware
	Maintenance	Maintenance Cost for licensed product/service such as upgrade, patch or technical support
	Operation cost	Operation cost for administrator training, shift, or education
Architecture	Integration	Function/feature to cooperate/communicate with 3rd party solution
	Scalability	Ability/attribute to grow or manage increased demand
	Reliability	Consistency and validity of test results
Risk	Security	Computer system is protected from data corruption, destruction, interception or unauthorized access
	False positive	A result that indicate a false result as a malware but which is not malware

도출된 주요 요인은 상호 운영방식에 따라 효율성이 달라지므로, 인텔리전스 기반 솔루션, 안티바이러스, 화이트 리스팅, 머신러닝 안티바이러스의 총 4가지 대안을 선정하는 것을 목적으로 한다.

AHP의 계층 모델은 효율적인 동적 악성코드 분석 시스템의 사전 필터링을 통해 효율적인 운영방식을 선정하는 것이므로 상위 평가 기준간의 쌍대비교를 진행과 하위평가 기준간의 쌍대비교를 수행하게 되며, 하위 평가기준별 대안의 쌍대비교를 수행하여 대안의 우선순위 점수를 도출한다. Fig 5는 목표, 상위 평가요소, 하위 평가요소 및 대안을 AHP 모델로 구현한 것이다.

4.5 3단계: 설문 조사 및 유효성 평가

Fig 5에 제시된 AHP 계층 모델을 바탕으로 설문지를 작성하여 동적 분석 시스템을 도입하여 운영 중인 운영자 및 관리자 그리고 관련 업무 종사 전문가로 구성된 3개의 유형의 사용자 그룹을 대상으로 설문지를 구성하였으며, 9점 척도를 사용하여 쌍대비교를 수행하였다. 쌍대비교를 통해 도출된 가중치가 논리적으로 유효한지 확인하는 것이 중요하다. 이에 Saaty[6]는 일관성 비율(CR Consistency Ratio)을 사용하여 일관성 비율이 0.1 이하인 경우에 쌍대비교행렬이 일관성이 있다고 하였고 설문결과 일관성 비율이 0.1 미만인 경우에는 일관성이 있다고 판단하고 0.1 이상의 일관성 비율을 가지고 있는 설문의 경우는 합리적 판단이 불가능하다고 판단하여 이번 분석에는 포함하지 않았다. 설문지는 총 30부가 회수되었고 그중 9부는 일관성 비율이 0.1 이상으로 제외하여 총 21부를 분석하였다.

21개의 설문 중 전문가 그룹은 7부, 운영자 5부, 그리고 관리자 9개로 분류 되었다. 이 설문은 Expert Choice 2000 교육용 소프트웨어를 사용하여 분석하였고, Saaty[6]가 검증한 행렬의 역수성을 유지하는 기하평균(Geometric mean)을 이용하였다.

4.6 4단계: 평가 요인별 대안별 우선순위 도출

본 연구에서는 AHP모델을 사용하여 Saaty[6]의 일관성 비율(CR)이 0.1 이하인 설문지를 신뢰할 수 있는 설문지로 분류하였으며 Fig 5 과 같이 AHP 계층 모델을 통해 총 9개의 주요 평가요인들을 3개의 그룹으로 분류하여 쌍대비교를 수행하였다. 그룹1은 관리자 그룹, 그룹2는 운영자 그룹, 그룹3은 전문가(컨설턴트) 그룹으로 구분하였다. 그리고 본 모델의 검증 및 향후 대안 솔루션 도입을 위해 각 그룹을 통합하여 산술 평균으로 그룹4를 도출하여 가중치 도출하였다.

4.6.1 그룹1 - 관리자 그룹 분석 결과

Table 5는 관리자 그룹의 평가 요인별 가중치와 최종 가중치를 정리한 것이다. 관리자 그룹의 분석 결과를 보면 상위평가 기준의 중요도는 위험(0.525), 아키텍처(0.272), 비용(0.203)의 가중치로 분석되었다.

그 중 상위 평가 가중치가 높은 순서로 하위평가의 가중치를 분석해 보면, 첫 번째의 위험은 보안성(0.521), 오탐지(0.479)의 순서로 분석되었고, 두 번째의 아키텍처는 신뢰성(0.408), 확장성(0.346), 연동성(0.245)로 분석되었고, 세 번째의 비용은 유지보수비용(0.377), 운영비용(0.360), 라이선싱(0.254)로 분석되었다.

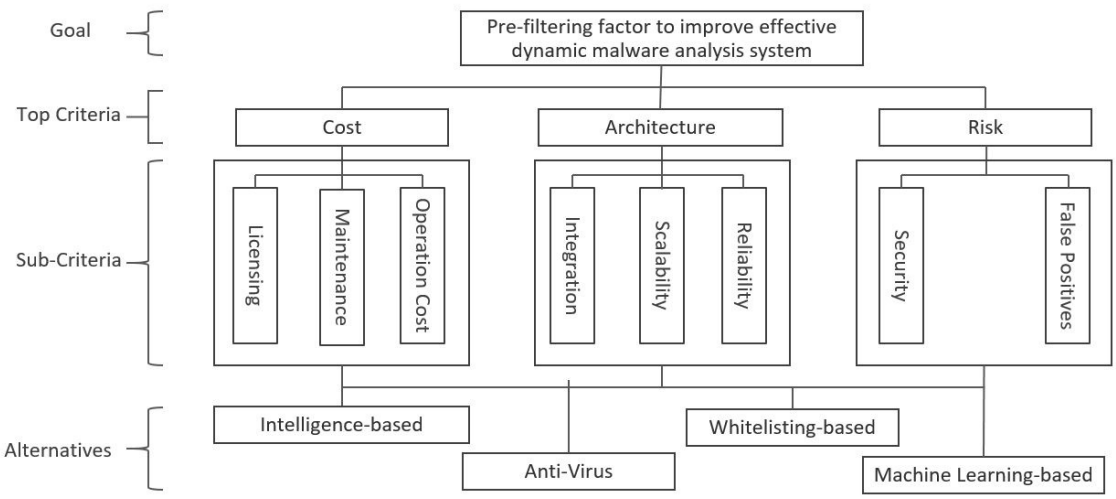

Fig. 5. AHP Model for pre-filtering system

각 평가 요인들의 최종 가중치는 상위 평가 가중치 값과 하위평가 가중치의 값을 곱하여 계산하는데, Table 5에 표시되어 있는 것과 같이 보안성, 오탐지 및 신뢰성이 가장 높은 평가 요인으로 분석되었다.

Table 5. Weight of Group 1's Evaluation Factors

Criteria	Weight	Sub-Criteria	Weight	Final Weight	Priority
Cost	0.203	License	0.254	0.052	8
		maintenance Cost	0.377	0.077	5
		Operation cost	0.360	0.073	6
Architecture	0.272	Integration	0.245	0.067	7
		Scalability	0.346	0.094	4
		Reliability	0.408	0.111	3
Risk	0.525	Security	0.521	0.274	1
		False Positive	0.479	0.251	2

Table 6. Weight of Group 1's Alternatives

Criteria	Sub-Criteria	Final Weight	Intelligence-based solution	Antivirus	Whitelisting	Machine Learning Antivirus
Cost	License	0.052	0.012	0.009	0.008	0.022
	Maintenance	0.077	0.022	0.016	0.008	0.031
	Operation	0.073	0.015	0.009	0.014	0.035
	Total	0.201	0.050	0.034	0.030	0.087
Architecture	Integration	0.067	0.018	0.010	0.008	0.030
	Scalability	0.094	0.024	0.016	0.008	0.030
	Reliability	0.111	0.032	0.020	0.026	0.033
	Total	0.272	0.074	0.045	0.042	0.110
Risk	Security	0.274	0.079	0.044	0.066	0.084
	False Positive	0.251	0.044	0.051	0.077	0.080
	Total	0.525	0.123	0.095	0.143	0.164
Final weight of alternatives			0.246	0.174	0.215	0.362
Final Priorities of alternatives			2	4	3	1

Table 6의 대안에 대한 최종 우선순위를 살펴보면, 머신러닝(0.362), 인텔리전스(0.246), 화이트리스팅(0.215), 안티바이러스(0.174)로 분석되어 머신러닝에 의한 상호 운영이 가장 적합한 대안으로 평가 되었다.

4.6.2 그룹2 - 운영자 그룹

Table 7은 운영자 그룹의 평가 요인별 가중치와 최종 가중치를 정리한 것이다. 운영자 그룹의 분석 결과를 보면 상위평가 기준의 중요도는 위험(0.457), 아키텍처(0.321), 비용(0.222)의 가중치로 분석되었다.

그 중 상위 평가 가중치가 높은 순서로 하위평가의 가중치를 분석해 보면, 첫 번째의 위험은 오탐지(0.557), 보안성(0.443)의 순서로 분석되었고, 두 번째의 아키텍처는 신뢰성(0.599), 확장성(0.221), 연동성(0.180)로 분석되었고, 세 번째의 비용은 운영비용(0.439), 라이선싱(0.318), 유지보수비용(0.243)로 분석되었다.

각 평가 요인들의 최종 가중치는 Table 7에 표시되어 있는 것과 같이 오탐지, 보안성 및 신뢰성이 가장 높은 평가 요인으로 분석되었다.

Table 8의 대안에 대한 최종 우선순위를 살펴보면, 머신러닝(0.412), 안티바이러스(0.231),인텔리전스(0.186), 화이트리스팅(0.171)로 분석되었다.

Table 7. Weight of Group 2 Evaluation Factors

Criteria	Weight	Sub-Criteria	Weight	Final Weight	Priority
Cost	0.222	License	0.318	0.070	6
		maintenance Cost	0.243	0.054	8
		Operation cost	0.439	0.097	4
Architecture	0.321	Integration	0.180	0.058	7
		Scalability	0.221	0.071	5
		Reliability	0.599	0.192	3
Risk	0.457	Security	0.443	0.202	2
		False Positive	0.557	0.254	1

Table 8. Weight of Group 2's Alternatives

Criteria	Sub-Criteria	Final Weight	Intelligence-based solution	Antivirus	Whitelisting	Machine Learning Antivirus
Cost	License	0.007	0.011	0.014	0.013	0.033
	Maintenance	0.054	0.011	0.015	0.013	0.015
	Operation	0.097	0.014	0.034	0.010	0.039
	Total	0.222	0.036	0.063	0.036	0.087
Architecture	Integration	0.058	0.012	0.013	0.009	0.025
	Scalability	0.071	0.015	0.015	0.013	0.028
	Reliability	0.192	0.057	0.043	0.025	0.068
	Total	0.321	0.084	0.071	0.046	0.120
Risk	Security	0.202	0.042	0.050	0.038	0.073
	False Positive	0.254	0.025	0.047	0.050	0.131
	Total	0.457	0.067	0.097	0.088	0.204
Final weight of alternatives			0.186	0.231	0.171	0.412
Final Priorities of alternatives			3	2	4	1

4.6.3 그룹3 - 전문가 그룹

Table 9는 전문가 그룹의 평가 요인별 가중치와 최종 가중치를 정리한 것이다. 운영자 그룹의 분석 결과를 보면 상위평가 기준의 중요도는 위험(0.592), 아키텍처(0.216), 비용(0.192)의 가중치로 분석되었다.

그 중 상위 평가 가중치가 높은 순서로 하위평가의 가중치를 분석해 보면, 첫 번째의 위험은 보안성(0.502), 오탐지(0.498)의 순서로 분석되었고, 두 번째의 아키텍처는 신뢰성(0.570), 연동성(0.229), 확장성(0.201)로 분석되었고, 세 번째의 비용은 라이선싱(0.387),운영비용(0.349), 유지보수비용(0.264)로 분석되었다.

Table 9. Weight of Group 3 Evaluation Factors

Criteria	Weight	Sub-Criteria	Weight	Final Weight	Priority
Cost	0.192	License	0.387	0.074	4
		maintenance Cost	0.264	0.051	6
		Operation cost	0.349	0.067	5
Architecture	0.216	Integration	0.229	0.050	7
		Scalability	0.201	0.043	8
		Reliability	0.570	0.123	3
Risk	0.592	Security	0.502	0.297	1
		False Positive	0.498	0.295	2

각 평가 요인들의 최종 가중치는 Table 9에 표시되어 있는 것과 같이 보안성, 오탐지, 및 신뢰성이 가장 높은 평가 요인으로 분석되었다.

Table 10의 대안에 대한 최종 우선순위를 살펴보면, 인텔리전스(0.298),머신러닝(0.275), 안티바이러스(0.219), 화이트리스팅(0.208)로 분석되었다.

Table 10. Weight of Group 3's Alternatives

Criteria	Sub-Criteria	Final Weight	Intelligence-based solution	Antivirus	Whitelisting	Machine Learning Antivirus
Cost	License	0.074	0.019	0.011	0.013	0.031
	Maintenance	0.051	0.017	0.010	0.009	0.015
	Operation	0.067	0.021	0.012	0.015	0.019
	Total	0.192	0.057	0.034	0.036	0.065
Architecutture	Integration	0.050	0.018	0.009	0.007	0.016
	Scalability	0.043	0.015	0.008	0.009	0.012
	Reliability	0.123	0.037	0.031	0.031	0.024
	Total	0.216	0.069	0.049	0.047	0.051
Risk	Security	0.297	0.084	0.071	0.052	0.091

	False Positive	0.295	0.088	0.066	0.073	0.068
	Total	0.592	0.172	0.137	0.125	0.159
Final weight of alternatives			0.298	0.219	0.208	0.275
Final Priorities of alternatives			1	3	4	2

4.6.4 그룹4 - (그룹1+그룹2+그룹3)

관리자 그룹인 그룹1은 '보안성'과 '오탐지'의 우선순위가 높게 평가되었고 운영자 그룹인 그룹2는 '오탐지'와 '보안성'을 높게 평가되었으며, 전문가 그룹인 그룹3의 경우는 '보안성'과 '오탐지'를 가장 높이 평가하는 것으로 분석이 되었다. 악성코드 동적 분석 시스템의 효율성을 위한 사전 필터링 기술 대안 선정 시 비용, 아키텍처, 위험 측면이 고려된 모델을 수립하기 위해 3개의 그룹을 기하평균 하였다. Table 11은 그룹1, 그룹2, 그룹3의 기하평균을 산술평균하여 그룹4의 평가 요인별 가중치를 정리한 것이다. 상위평가에는 위험(0.544), 아키텍처(0.255), 비용(0.201)의 우선순위로 가중치가 분석되었으며 최종 가중치를 보면 오탐지가 가장 높은 1순위이며, 보안성이 2순위 그 다음이 신뢰성이 3위의 순서로 분석되었다.

표 12는 Table 11 그룹4의 분석된 평가 기준별 가중치를 사용하여 4가지 대안에 대한 우순선위를 산출하였다. 대안의 최종 가중치는 머신러닝

Table 11. Weight of Group 4 Evaluation Factors

Criteria	Weight	Sub-Criteria	Weight	Final Weight	Priority
Cost	0.201	License	0.329	0.066	5
		maintenance Cost	0.298	0.060	7
		Operation cost	0.373	0.075	4
Architecture	0.255	Integration	0.225	0.057	8
		Scalability	0.253	0.065	6
		Reliability	0.522	0.133	3
Risk	0.544	Security	0.497	0.270	2
		False Positive	0.503	0.273	1

Table 12. Weight of Group 4's Alternatives

Criteria	Sub-Criteria	Final Weight	Intelligence-based solution	Antivirus	Whitelisting	Machine Learning Antivirus
Cost	License	0.066	0.015	0.011	0.011	0.028
	Maintenance	0.060	0.018	0.013	0.010	0.020
	Operation	0.075	0.018	0.015	0.014	0.027
	Total	0.201	0.051	0.039	0.035	0.075
Architecture	Integration	0.057	0.018	0.010	0.008	0.022
	Scalability	0.065	0.018	0.012	0.010	0.024
	Reliability	0.133	0.039	0.030	0.030	0.034
	Total	0.255	0.075	0.052	0.048	0.080
Risk	Security	0.270	0.073	0.058	0.054	0.086
	False Positive	0.273	0.060	0.057	0.070	0.086
	Total	0.544	0.133	0.115	0.124	0.172
Final weight of alternatives			0.259	0.206	0.207	0.328
Final Priorities of alternatives			2	4	3	1

(0.328), 인텔리전스(0.259), 화이트리스팅(0.207), 안티바이러스(0.206)으로 분석되었다. 분석결과 비용, 아키텍처, 위험 모두 머신러닝이 가장 높은 가중치로 분석되었다.

4.7 악성코드 동적 분석 시스템 효율성 향상을 위한 사전 필터링 대안 선정 공식

여러 가지 대안 기술 선정시, 본 연구에서 도출된 중요 요소와 도입의 중요도에 따라 합리적인 선택을 할 수 있도록 공식을 도출한다. Table 11은 그룹4에서 도출된 상위 평가 요인의 가중치와 Table 12에 있는 것과 같이 대안별 가중치를 바탕으로 Expert Choice 2000에서 제공되는 민감도 분석 기능 및 삼차연립방정식을 이용하여, 상위 평가요인인 비용, 아키텍처 및 위험의 각각의 상수 값을 도출하였다. 이 상수 값을 사용하여 Table 13과 같이 대안을 선정할 수 있는 공식을 도출하였다.

Table 13. Alternative Selection Criteria Formula

Alternatives	Formula
Intelligence	*I = (26xC) + (29xA) + (24xR)*
Antivirus	*V = (20xC) + (18xA) + (21xR)*
Whitelisting	*W = (17xC) + (19xA) + (23xR)*
Machine Learning	*M = (36xC) + (30xA) + (31xR)*
I : Weight of Intelligence V : Weight of Antivirus W : Weight of Whitelisting M : Weight of Machine Learning C : Weight of Cost A : Weight of Architecture R : Weight of Risk	

이 공식의 검증을 위해 그룹4에 도출된 상위 평가 요인의 가중치인 위험(0.201), 아키텍처(0.255), 위험(0.544)을 Table 13에서 제시하는 공식에 대입하면 Table 14와 같이 공식에 의해 계산 값이 Table 12의 대안의 최종 가중치 값과 일치하는 것을 검증하였다.

Table 14. Verification Formula using Result Value

Alternatives	Cost		Architecture		Risk		Total	Priorities
	Constant	Weight	Constant	Weight	Constant	Weight		
Intelligence	26	0.201	29	0.255	24	0.544	26%	2
Antivirus	20	0.201	18	0.255	21	0.544	21%	4
Whitelisting	17	0.201	19	0.255	23	0.544	21%	3
Machine Learning	36	0.201	30	0.255	31	0.544	32%	1

4.8 사례를 통한 의사 결정 모델 활용

본 연구에서 AHP 의사결정 모델과 대안 선정 공식을 적용하여 검증을 실시하였다. 사전 필터링 시스템의 선정 의사결정은 Table 15의 3단계와 같이 검증을 실시하였다.

검증을 위한 단계는 다음과 같다.

1 단계: Table 16과 같이 총 8가지의 하위 요인에 대하여 Table 17과 같이 9점 척도로 중요도를 계산한다.

Table 15. Alternative Selection Steps

Step	Description
Step 1	Calculate importance factor for each sub-criteria from 1 to 9 scale
Step 2	Calculate average and weight of top criteria for Cost, Architecture and Risk
Step 3	Calculate the alternative weight using Table 13 Alternative Selection Criteria Formula

2 단계: Table 16의 중요도 평가 데이터를 기반으로 상위 요인에 대한 가중치 값을 하위 요인에 대안 중요도를 사용하여 Table 18과 같이 계산한다.

3 단계: Table 13과 같이 대안 선정 공식에 1단계와 2단계에서 도출된 대안의 가중치 값을 이용하여 계산한다.

사례검증에는 국내 금융A를 사용하였다. A사는 약 15,000명의 규모의 금융회사이고 동적 악성코드 분석에 5명 정도가 운영 중이며 인터넷 사용의 증가로 추가로 동적 분석 시스템의 증설이 요구되고 있으나, 동적 분석 시스템 자체의 오탐 및 과탐으로 많은 인력이 투입되고 있어 사전 필터링 기술을 활용하여 효율성 향상을 고려하고 있다.

Table 16. Importance Rating for Sub-Criteria from Company A

Criteria	Sub-criteria	Importance rating(1-9)								
		1	2	3	4	5	6	7	8	9
Cost	License					V				
	Maintenance					V				
	Operation					V				
Architecture	Integration							V		
	Scalability							V		
	Reliability							V		
Risk	Security									V
	False Positive									V

Table 17. Rating for Relative Importance

Rating	Definition
1	Not important
3	Some Important
5	Important
7	Very Important
9	The Most Important
2,4,6,8	Median of determine the 2 adjacent

Table 18. Top Criteria Weight Calculation based on A Company

Cost		Architecture		Risk	
Average	Weight	Average	Weight	Average	Weight
5	0.238	7	0.333	9	0.429

검증에 사용된 회사는 위험이 가장 높은 가중치가 높았으며, 이 가중치를 도출된 공식에 의해 계산하여 Table 19와 같이 머신러닝의 의한 사전 필터링 기술을 선정이 가장 효과적이라는 결과를 얻었다.

Table 19. Result of Alternative Solution Selection using Formula for Company A

Alternatives	Cost		Architecture		Risk		Total	Priorities
	Constant	Weight	Constant	Weight	Constant	Weight		
Intelligence	26	0.238	29	0.333	24	0.429	27%	2
Antivirus	20	0.238	18	0.333	21	0.429	20%	4
Whitelisting	17	0.238	19	0.333	23	0.429	20%	3
Machine Learning	36	0.238	30	0.333	31	0.429	33%	1

4.9 분석 결과 요약

본 연구에서는 동적 악성코드 분석 시스템의 효율적인 운영을 위해 사전 필터링 기술 선정에 있어 가장 중요한 요소를 도출하고 이에 따른 최적의 대안 기술을 선정하는 요인은 분석하였다. CR값이 0.1이인 값을 가진 21개의 설문지만 유효성이 있는 것으로 판정하고 직무별로 그룹으로 나눠 직무별로 분석하였다.

우선 상위 평가요소의 경우 3개의 그룹 모두 위험을 가장 가중치가 높은 것으로 선정하고 그 다음으로 아키텍처와 비용의 순서였다.

하위의 평가 요소는 각 그룹별로 조금 상이한 것으로 분석이 되었다. 첫 번째 관리자 그룹의 경우 보안성을 가장 높은 우선순위를 두고 그 다음으로 오탐지 신뢰성 순이었다. 두 번째 그룹인 운영자 그룹의 경우는 관리자와 달리, 오탐지를 가장 주요한 요소로 선정하고 있었으며, 세 번째 그룹인 컨설턴트의 경우 관리자 그룹과 같이 보안성을 가장 높은 순위로 선정하였다.

대안 기술에 대한 우선순위는 관리자 그룹과 운영자 그룹의 경우 머신러닝을 활용한 악성코드 탐지에 가능 높은 우선순위를 두고 있었고, 전문가 그룹의 경우는 인텔리전스에 의한 악성코드 탐지를 가장 높은 순위로 분석이 되었다.

V. 결 론

본 연구에서는 동적 분석 시스템 효율성 향상을 위한 사전 필터링 기술 선정을 위한 요인들을 8개의 주요 평가 요소로 도출하였다. 또한 도출된 요인들은 비용(Cost), 아키텍처(Architecture), 위험(Risk)의 세 가지 상위 평가 기준으로 분류하여 AHP 계층 모델을 구축 및 제시하였으며, 사전 필터링 기술로는 인텔리전스 기반 솔루션, 안티바이러스, 화이트리스팅 및 머신러닝에 의한 솔루션의 대안 기술을 제시하였다.

사전 필터링 기술을 활용하는 대안 솔루션에 경우에는 머신러닝에 의한 솔루션이 가장 높은 가중치를 보였다. 머신러닝에 의한 탐지 기술은 탐지율과 오탐지의 적절한 균형이 필요하므로 향후 머신러닝에 의한 탐지 기술의 시장 형성과 확대 가능성을 제시하였다.

향후 연구에서는 여러 가지 사전 필터링 기술 선정 시 본 연구와 같이 라이선스 비용, 유지보수비용, 운영비용, 연동성, 확장성, 신뢰성, 보안성, 오탐지의 총 8가지 요소를 도출하여 합리적 선택을 할 수 있도록 하였으나 이 평가 요소 이외에도 다양한 연구를 통해 다양한 고려 요소를 선정하는 것이 필요하며, 향후 실무에서 얻은 실측 데이터를 기반으로 AHP 이외의 분석 기법을 사용하여 사전 필터링 기술의 효율성을 향상 측면을 연구할 것이다.

References

[1] Symantec, "Internet Security Threat Report." https://www.symantec.com/content/dam/symantec/docs/reports/istr-22-2017-en.pdf" Vol 22, Apr. 2017

[2] Michael Riley, Benjman Elgin, Dune lawrence and Carol Matlack, "Missed Alarms and 40 Million Stolen Credit Card Numbers: How Target Blew It." Bloomberg, "https://www.bloomberg.com/news/articles/2014-03-13/target-missed-warnings-in-epic-hack-of-credit-card-data," Mar 2014.

[3] Candid West, "Threats to Virutal Environments," Symantec, http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/threats_to_virtual_environments.pdf" 2015

[4] Saaty T. L., "The Analytic Hierchy Process," McGraw-Hill, New York, 1980

[5] Chun-Chn Wei, chen-Fu Chien, Mao-Jiun. Wang, "An AHP-Based approach to ERP System Selection," Elsevier, pp.47-62, 2004

[6] Shin-Pil Shin, "An analytics hierarchy process(AHP) approach to selection of implementation mode of mobile office system," Seoul National University of Science and Technology, July. 2013

[7] Gorecki, Christian, et al. "Trumanbox: Improving dynamic malware analysis by emulating the internet." Symposium on Self-Stabilizing Systems. Springer Berlin Heidelberg, 2011.

[8] Eric M. Hutchins, MichaelJ. Cloppert, Rohan M Amin, Ph.D. "Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains," Leading Issues in information Warfare & Security Research, 2011

[9] Moshchuk, Alexander, et al. "A Crawler-based Study of Spyware in the Web," NDSS. Vol. 1. 2006.

[10] Pareek, Himanshu, Sandeep Romana, and P. R. L. Eswari. "Application whitelisting: approaches and challenges." International Journal of Computer Science, Engineering and Information Technology (IJCSEIT) 2.5 (2012).

[11] Chen, Xu, et al. "Towards an understanding of anti-virtualization and anti-debugging behavior in modern malware." Dependable Systems and Networks With FTCS and DCC, 2008. DSN 2008. IEEE International Conference on. IEEE, 2008.

[12] Bayer, Ulrich, et al. "Dynamic analysis of malicious code." Journal in Computer Virology 2.1 (2006): 67-77.

[13] Bayer, Ulrich, Engin Kirda, and Christopher Kruegel. "Improving the efficiency of dynamic malware analysis." Proceedings of the 2010 ACM Symposium on Applied Computing. ACM, 2010.

[14] Egele, Manuel, et al. "A survey on automated dynamic malware-analysis techniques and tools." ACM Computing Surveys (CSUR) 44.2 (2012): 6.

[15] Rieck, Konrad, et al. "Automatic analysis of malware behavior using machine learning." Journal of Computer Security 19.4 (2011): 639-668.

[16] Dinaburg, Artem, et al. "Ether: malware analysis via hardware visualization extensions." Proceedings of the 15th ACM conference on Computer and communications security. ACM, 2008.

[17] Grance, Timothy, Marc Stevens, and Marissa Myers. "Guide to selecting information technology security products." Network Security (2003).

[18] Mamaghani, Farrokh. "Evaluation and selection of an antivirus and content filtering software." Information management & computer security 10.1 (2002): 28-32.

[19] Lengyel, Tamas K., et al. "Scalability, fidelity and stealth in the DRAKVUF dynamic malware analysis system." Proceedings of the 30th Annual Computer Security Applications Conference. ACM, 2014.
[20] Godse, Manish, and Shrikant Mulik. "An approach for selecting software-as-a-service (SaaS) product." Cloud Computing, 2009. CLOUD'09. IEEE International Conference on. IEEE, 2009.
[21] Bayer, Ulrich, Christopher Kruegel, and Engin Kirda. "TTAnalyze: A tool for analyzing malware." na, 2006.
[22] Gorecki, Christian, et al. "Trumanbox: Improving dynamic malware analysis by emulating the internet." Symposium on Self-Stabilizing Systems. Springer Berlin Heidelberg, 2011.
[23] Kwon Jonghoon, et al. "Metamorphic Malware Detection using Subgraph Matching." Korea Institute of Information Security & Cryptology, 2011, 21.2: 37-47.
[24] http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-167.pdf
[25] Suk-Won Lee, "Decision Making Model for Selecting Financial Company Server Privilege Account Operations," Korea Institute of Information Security & Cryptology, 25(6), p1607-1620, Dec. 2014
[26] https://www.symantec.com/products/endpoint-hybrid-cloud-security/endpoint/endpoint-protection
[27] https://www.virustotal.com/
[28] https://usa.kaspersky.com/enterprise-security/endpoint
[29] https://www.carbonblack.com/products/cb-protection/
[30] https://www.mcafee.com/us/solutions/dynamic-endpoint-threat-defense.aspx
[31] https://www.cylance.com/en_us/products/our-products/protect.html
[32] https://www.invincea.com/solution-overview/

〈저자소개〉

윤 광 택 (Kwang-Taek Youn) 정회원
1997년 2월: 단국대학교 무역학과 졸업
2014년 3월 ~ 현재: 고려대학교 정보보호대학원 석사과정
〈관심분야〉 정보보호 정책, 위험관리, 네트워크 보안, 엔드포인트 보안

이 경 호 (Kyung-Ho Lee) 종신회원
1989년 8월: 서강대학교 수학과 학사
1997년 8월: 서강대학교 정보통신대학원 석사
2009년 8월: 고려대학교 정보보호대학원 박사
2011년 9월 ~ 현재 : 고려대학교 정보보호대학원 부교수
〈관심분야〉 위험관리, 정보보호 컨설팅, 정보보호 및 개인정보보호정책