JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 25 No. 11, pp. 123-129, November 2020
https://doi.org/10.9708/jksci.2020.25.11.123

The Next Generation Malware Information Collection Architecture for Cybercrime Investigation

Ho-Mook Cho*, Chang-Su Bae**, Jaehoon Jang**, Sang-Yong Choi***

*Principal Researcher, Cyber Security Research Center, KAIST, Daejeon, Korea
**Researcher, APEX ESC, Inchon, Korea
**CEO, APEX ESC, Inchon, Korea
***Assistant Professor, Dept. of Cyber Security, Yeungnam University College, Daegu, Korea

[Abstract]

Recently, cybercrime has become increasingly difficult to track by applying new technologies such as virtualization technology and distribution tracking avoidance. etc. Therefore, there is a limit to the technology of tracking distributors based on malicious code information through static and dynamic analysis methods. In addition, in the field of cyber investigation, it is more important to track down malicious code distributors than to analyze malicious codes themselves. Accordingly, in this paper, we propose a next-generation malicious code information collection architecture to efficiently track down malicious code distributors by converging traditional analysis methods and recent information collection methods such as OSINT and Intelligence. The architecture we propose in this paper is based on the differences between the existing malicious code analysis system and the investigation point's analysis system, which relates the necessary elemental technologies from the perspective of cybercrime. Thus, the proposed architecture could be a key approach to tracking distributors in cyber criminal investigations.

▸ **Key words**: Malware, Cyber criminal, Intelligence, Cyber investigation, Trace

[요 약]

최근 사이버범죄는 가상화 기술, 유포지 추적 회피 등 다양한 기술 등의 새로운 기술을 적용하여 추적이 점점 어려워지고 있다. 따라서 전통적인 악성코드 분석방법인 정적분석, 동적 분석 등 방법은 악성코드 유포자를 추적하는 데 한계가 있다. 또한, 사이버 수사 분야에서는 악성코드 자체에 대한 분석보다 악성코드 유포자를 추적하는 것이 더욱 중요하다. 이에 따라, 본 논문에서는 악성코드 유포자를 효율적으로 추적하기 위해 전통적인 분석방법과 OSINT, Intelligence 등 최근의 정보수집 방법을 융합한 차세대 악성코드 정보수집 아키텍처를 제안한다. 본 논문에서 제안하는 아키텍처는 기존의 악성코드 분석체계와 수사관점의 분석체계의 차이점을 기반으로 사이버범죄의 관점에서 필요한 요소기술을 연관시킴으로 인해 사이버 범죄 수사에서 유포자 추적을 위한 핵심적인 접근 방법이 될 수 있다.

▸ **주제어**: 악성코드, 사이버범죄, 지능, 사이버 수사, 추적

• First Author: Ho-Mook Cho, Corresponding Author: Sang-Yong Choi
*Ho-Mook Cho (chmook79@kaist.ac.k), Cyber Security Research Center, KAIST
**Chang-Su Bae (changsu.bae@apexesc.com) APEX ESC
**Jaehoon Jang (jh.jack@apexesc.com) APEX ESC
***Sang-Yong Choi (csyong95@gmail.com), Dept. of Cyber Security, Yeungnam University College
• Received: 2020. 10. 05, Revised: 2020. 10. 30, Accepted: 2020. 11. 02.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

악성코드는 최근 몇 년 동안 인터넷 환경에서 가장 심각한 위협으로 대두되고 있다. 공격자들은 악성코드를 유포하기 위해 사용자들의 접속이 많은 웹 애플리케이션을 공격하여 악성코드 유포지로 연결되는 링크를 삽입한다. 이러한 공격을 Drive-by Attack이라 부르며, 이 공격을 이용하여 불특정 다수의 사용자는 악성코드에 감염된다[1]. 이렇게 유포되는 악성코드를 효과적으로 분석하고, 대응하기 위해 다양한 분석 방법론이 연구되고 있다. 대표적으로 정적분석을 포함한 보안인텔리전스를 기반 악성코드의 유형을 프로파일링하는 방법[2-6], 악성코드를 분석함에 있어 머신러닝 기술을 적용하는 분석 방법[7-11], 에뮬레이터와 같은 독립적인 실행환경에서 실제 실행을 통해 악성코드의 행위를 분석하는 방법[12-14] 등 악성코드 자체에 관한 많은 연구가 이루어지고 있으며, 특히 분석환경을 우회하는 악성코드를 효과적으로 분석하기 위한 방법 또한 연구되고 있다[15].

이와 같은 다양한 연구방법은 악성코드 자체에 대한 효과적인 분석은 가능하지만, 사이버범죄 수사의 측면에서는 한계가 있다. 사이버범죄 수사의 경우 단순히 악성코드의 행위를 분석하는 데서 그치지 않고 악성코드를 유포한 유포자를 식별하고, 유포자에 대한 특징 분석을 통해 유사한 악성코드를 분류하고 사이버 수사에 필요한 정보를 프로파일링하여야 하지만, 알려진 악성코드 분석방법은 기술적으로 악성코드 자체 또는 악성코드의 유포 방법을 효과적으로 식별하고 탐지하는데 관점이 있으므로 유포자에 대한 정보를 수집하여 수사에 활용하는 데 한계가 있다.

본 논문에서는 이와 같은 한계점을 해결하고 보다 효과적으로 수사에 적용하여 악성코드로 인한 피해를 최소화하고 유포자를 추적하기 위해 기술적인 악성코드 분석방법과 OSINT(Open-source intelligence)기법 등을 적용하여 악성코드에 대한 정보뿐만 아니라 악성코드를 통해 악성코드를 유포한 유포자에 대한 정보를 연관 분석할 수 있는 차세대 악성코드 정보수집 시스템을 제안한다. 제안하는 시스템의 아키텍처는 사이버 수사 업무프로세스에 기반을 두어 사이버 수사 각 단계에 따라 어떤 기술과 알고리즘을 활용하여 악성 정보를 수집할 것인지를 명확하게 함으로 차세대 사이버 수사의 방향성을 제안하고자 한다.

이를 위해 본 논문은 2장에서 악성코드 분석방법을 조사하고, 사이버 수사 업무에 악성코드 분석방법 및 정보분석 방법 등을 효과적으로 적용하기 위한 방안을 도출한다. 그리고 3장에서 제안하는 차세대 사이버 수사 업무프로세스 기반 악성코드 분석 아키텍처를 설계하고, 아키텍처의 각 단계를 설명한다. 그리고 4장에서는 아키텍처의 각 단계에 대한 효과를 기술하고, 결론을 맺는다.

II. Preliminaries

1. Related works

1.1 Malware distribution methods

악성코드를 유포하는 방법 중 대표적인 방법은 웹브라우저나 브라우저 확장프로그램, 웹사이트, CMS(Content Management System), 웹서비스 등을 통해 악성코드를 유포하는 방법으로, 웹브라우저나 확장프로그램의 취약점을 이용하여 악성코드 유포사이트로 연결되는 링크를 삽입하여 피해자를 공격자가 만들어놓은 악성코드 유포사이트로 유도한다[1]. 만약 피해자의 컴퓨터에 설치된 프로그램 또는 운영체제의 버전이 취약하다면 공격자는 피해자 컴퓨터의 취약한 소프트웨어를 악용하여 악성코드를 피해자 컴퓨터에 설치하게 된다. 세부적으로 알려진 유형은 drive-by, watering-hole, redirection, man-in-the-browser 등이 있으며, 이러한 유형을 공통으로 Drive-by download 공격이라 한다. 최근 수년간 다양한 위협이 되는 APT(Advanced Persistent Threat) 공격과 같은 공격이 이와 같은 방법으로 유포되는 악성코드로부터 시작한다.

Drive-by download 공격을 위해 공격자는 악성코드 유포사이트로 연결되는 Bad Web Site를 다수 만들게 되며, 이와 같은 악성 사이트는 서로 다른 악성코드에 대해 공격자를 식별하는데 필요한 정보일 수 있다.

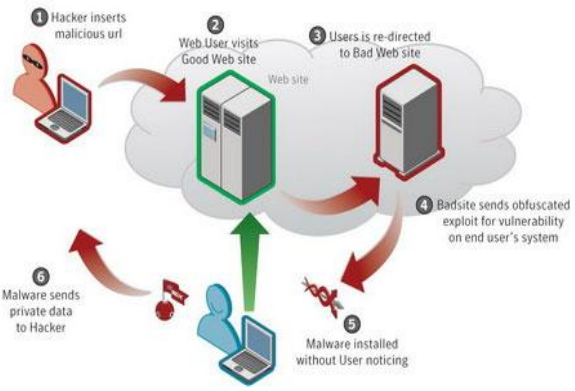


Fig. 1. Drive-by download attack

1.2 Malware analysis - intelligence

보안 인텔리전스를 악성코드 분석에 활용하기 위한 연구의 배경은 인터넷 사용자의 경우 웹서핑을 하는 과정에

서도 악성코드를 통해 다양한 침해사고가 발생하고 있으며, 침해사고 정보를 담고 있는 보안 인텔리전스의 데이터가 기하급수적으로 증가하고 있고, 예측하기 힘든 공격자의 다양한 변종 패턴과 넘치는 데이터의 양으로 인해 기존의 정적 및 동적 분석방법만을 가지고는 분석의 효과성이 떨어진다는 가정하에 침해사고의 다양한 속성을 사용하고 더욱 정확하게 유사 침해사고를 그룹별로 분류하기 위해 침해사고 프로파일의 IP주소, Domain, HASH의 각각에 대한 거리계산 알고리즘을 통해 다중 프로파일 기반 앙상블 알고리즘의 의사결정을 통한 유사도 분석 알고리즘을 적용하는 시도가 있었다[2]. 또 다른 접근 방법으로는 사이버보안 지표, 악성코드 분석정보 등을 조직 간에 공유하기 위한 공유 플랫폼을 활용하여 악성코드를 분석하는 방법이다[3-4, 6]. 이 외에도 운용 중인 각종 정보보호체계 및 통합보안관제시스템에서 발생한 많은 양의 위협 이벤트를 수집, 융합하여 전처리하며, 위협경보에 대한 통계 기반의 상관관계 분석과 위협경보 간 상호 인과관계를 학습하는 모델을 구축하여 지능형 분석 모델을 구축, 머신러닝 기반의 위협경보 상관분석과 OSINT를 통해 인텔레전스와의 자동화된 연관 분석을 통해 정보를 수집하고 예상되는 위험을 예측하는 기술이 연구되었다[5]

이러한 접근 방법은 기존의 정적분석과 동적 분석에서 요구되는 악성코드 자체에 대한 특성뿐만 아니라 이러한 악성코드로 인한 침해사고의 정보를 분류의 기준에 적용하였다는 점에서 의미가 있다. 이러한 정보를 활용함으로 악성코드 유포자와 악성코드 간의 연관 관계를 통해 악성코드 유포자를 추적하는 데 도움을 줄 수 있다.

1.3 Malware analysis - machine learning

최근 머신러닝을 악성코드 분석에 적용하기 위한 다양한 연구가 이루어지고 있다. 대표적인 이유는 전통적인 정적분석 방법의 경우 난독화되있는 코드를 정확하게 분석하기 힘들며, 동적 분석의 경우 악성코드가 잘 동작할 수 있도록 악성코드가 실행되는 환경을 구성해야 하는 어려움이 있다. 머신러닝을 이용한 방법의 하나로 악성코드를 실행시키지 않고 패밀리를 분류하는 방법으로 악성코드 파일을 8-bit gray-scale 이미지로 시각화하고 이미지 인식 분야에서 널리 쓰이고 있는 convolutional neural network를 통해 악성코드의 악성 특징을 기준으로 분류해내는 기법을 적용한 결과 9개의 악성코드 패밀리로 분류해 내는 실험을 통해 예측 정확도는 각각 96.2%, 98.7%를 기록하였고 27개의 패밀리를 분류하는 실험에서는 82.9%, 89% 악성코드 패밀리를 분류는 성과를 보였다[7] 또한, 악성코드 분류를 위해 CNN 알고리즘을 이용하여 바이너리 데이터를 이미지화하여 악성코드를 분석하는 연구가 진행되었다[8, 10].

이와 같은 연구는 머신러닝을 사이버보안 분야에 적용하기 위한 시도로써 의미가 있으며, 대량으로 수집되는 바이너리 파일을 효율적으로 분류할 수 있는 방법으로 의미가 있다. 하지만 사이버 수사의 측면에서는 대량의 바이너리에서 악성코드를 분류하는 것보다 하나의 악성코드와 연관된 데이터를 기반으로 프로파일링하여 유포자를 추적하는데 관점이 있으므로 머신러닝 기술은 직접 적용하기에 한계가 있다.

1.4 Malware analysis - dynamic method

동적 분석방법은 악성코드 정적분석의 한계점을 해결하기 위해 연구되고 있는 분야이며, 특히 탐지 회피기술이 적용된 악성코드를 분석함에 있어 정적분석의 난독화 기술에 대응할 수 있으며, 실제 악성코드가 실행될 때 시스템에 어떤 영향을 주는지 등을 세부적으로 분석할 수 있으므로 의미가 있다. 동적 분석방법의 경우 에뮬레이터를 이용하거나[12], 가상머신(Virtual Machine)을 이용하는 방법[13] 등이 주로 사용되고 있다. 동적 분석에서는 분석 회피기술에 대응하기 위해 분석환경을 실제 환경으로 하거나, 마우스, 키보드 등 사용자의 행위를 자동화하여 입력하는 방법으로 악성코드의 분석 회피기술에 대응한다.

동적 분석방법은 악성코드가 실제 동작할 때 접속하는 외부 IP(통상적으로 C&C 서버), 도메인 정보, 레지스트리 정보, 파일시스템, API 정보 등을 다양하게 수집할 수 있어, 악성코드를 개발한 개발자가 선호하는 특성을 분류하는 데 도움을 줄 수 있다.

2. Information Required for Cyber Investigation

악성코드 분석 측면에서 사이버 수사를 위해서는 악성코드 자체보다는 사이버 수사를 위한 종합적인 정보가 필요하다. 사이버 수사의 목적이 악성코드를 유포한 유포자를 식별하고 추적하는데 그 의미가 있기 때문이다. 하지만 유포자를 특정하기 위해서는 유포자가 만든 악성코드를 분석하여 악성코드에 포함된 정보를 활용하여야 한다. 이러한 측면에서 사이버 수사에는 악성코드에 대한 정보 및 악성코드와 관련된 인텔리전스 정보, 유포자의 특성에 관한 정보가 필요하며, 이러한 정보는 악성코드 자체에서 얻을 수도 있고, 포털 검색엔진 등 인터넷에 공개된 정보와 인텔리전스 정보를 모두 활용하여야 한다. 이번 절에서는 사이버 수사 측면에서 필요한 정보에 대해 살펴본다.

2.1 Malware Information

사이버 수사에서 필요한 악성코드 정보는 악성코드를 작성할 때 제작자의 코딩 특성, 제작자가 선호하는 행위적 특성 등이 필요하다. 즉, 통상적으로 프로그램을 제작할 때에는 동일한 기능을 사용하더라도 함수의 종류, 변수에 이름을 부여하는 방법, 개발자가 즐겨 쓰는 특정 단어, 함수의 순서 등이 개발자에 따라 다른 특성을 나타낼 수 있다. 이러한 정보를 수집하기 위해서는 악성코드의 코드를 살펴보는 정적분석의 방법을 활용하거나 악성코드를 실행하여 악성코드가 실행될 때 호출되는 API, 등 행위정보를 분석하여야 한다. 그리고, 전통적인 악성코드 분석방법에서 이와 같은 정보는 Ollydbg[16], IDA pro[17]등과 같은 분석 툴과 Cuckoo Sandbox[18]등과 같은 동적분석 툴에서 획득할 수 있다.

Ollydbg, IDA pro와 같은 분석 도구는 악성코드의 코드의 특성인 바이너리 자체에 대한 분석과 바이너리를 라인 단위로 실행과 인터럽트를 반복하면서 진행함으로 분석가가 각 코드가 어떻게 동작하고 전체적인 프로그램이 어떻게 실행되는지의 구조를 파악할 수 있도록 해준다. 특히, 이러한 도구는 악성코드가 실행될 때 레지스트리의 변화, 메모리의 변화 등을 세부적으로 살펴볼 수 있도록 해주어 다양한 정보를 획득할 수 있다.

Cuckoo Sandbox는 실제 윈도우 환경에서 악성코드를 실행하면서 악성코드가 동작할 때 시스템의 변화와 호출하는 API의 내용, 네트워크 접근정보, 패킷 정보, 파일시스템의 변화 등을 살펴볼 수 있도록 해주는 격리된 환경으로 가상환경 내에 Windows 머신을 설치하고 Cuckoo Core가 머신을 통제하면서 분석정보를 습득할 수 있다.

Cuckoo Sandbox의 경우 내부에 Snort, Suricata와 같은 탐지 엔진을 탑재할 수 있고, tcpdump 등과 같은 패킷 덤프 툴을 포함하고 있어 악성코드에 대한 종합적인 분석이 가능하고, 외부의 인텔리전스인 Virus-total과 연동을 할 수 있어 악성 행위를 더욱 정확히 분석할 수 있다.

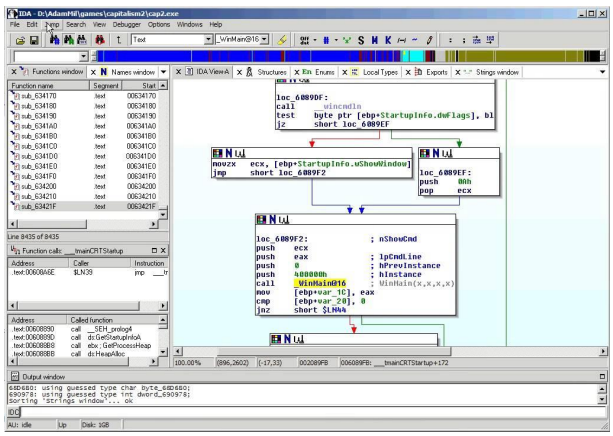


Fig. 2. IDA pro

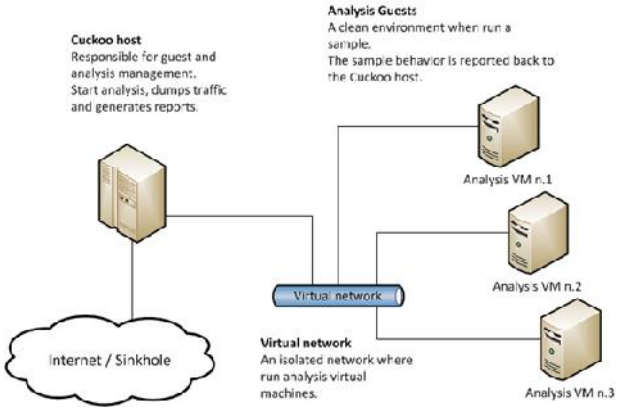


Fig. 3. Cuckoo Sandbox

2.2 OSINT and Intelligence

OSINT는 Open Source Intelligence의 약자로 온라인 소스로부터 공개 가능한 정보를 수집하고 분석하는 것을 말한다. 따라서, 많은 양의 디지털 데이터는 OSINT 분석을 어렵게 하는 문제 중의 하나이다. 하지만 OSINT를 활용하게 되면 예를 들어 Domain Name으로부터 subdomain 을 추출하고, 이를 분석함으로 도메인 간의 연결구조라든지, 악성코드에 관련된 정보로부터 제작자를 유추할 수 있는 연결고리를 찾는 것과 같은 지능적인 분석기 가능하다.

사이버 수사의 측면에서 OSINT를 통해 정보를 얻기 위한 툴과 방법으로는 IP를 기반으로 사용자의 지리적 위치, 데이터 및 연결속도, ISP 및 도메인 이름을 추출할 수 있는 IP2location[19], MaxMind[20], GeoByte[21], NetAcuity[22], 도메인에 대한 각종 정보를 획득할 수 있는 Whois, Domaintools[23], Bing, Google 등과 같은 웹사이트 등이 있으며, 이러한 OSINT 툴에서 수집할 수 있는 정보를 분류하고 정리하여 악성코드의 특성을 분류하는데 활용할 수 있다.

사이버 수사를 위한 인텔리전스로는 전 세계의 대부분의 백신을 통합하여 악성 바이너리 또는 악성 URL에 대한 정보를 분석해 주는 Virustotal[24]과 같은 인텔리전스 정보를 활용할 수 있으며, 국내 대표적인 사이버위협 인텔리전스로 C-TAS[25]와 같은 인텔리전스를 활용하여 연관분석에 사용할 수 있다.

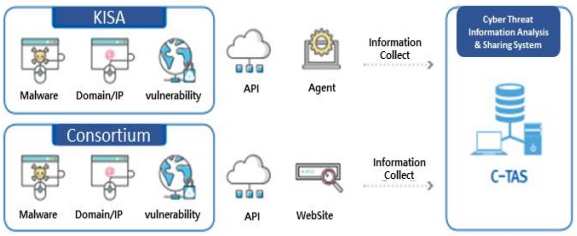


Fig. 4. C-TAS

III. The Proposed Scheme

사이버수사를 위해서는 앞서 언급한 바와 같이 단순히 악성코드에 대한 정보만을 분석하는 것 보다는 악성코드가 나타내는 다양한 자체의 특성과 악성코드의 행위적인 특성을 연관분석하여 유포자를 특정지을 수 있도록 하여야 한다. 이를 위해서는 먼저 1단계에서 악성코드 자체에 대한 분석정보를 수집한 후 2단계 스마트 필터링 작업을 거쳐 불필요한 정보를 걸러내고, 악성코드 자체의 특성 간에 연관분석을 통해 데이터를 정규화하여야 한다. 이후 3단계에서 OSINT와 Intelligence를 통해 공개 출처정보와의 연계 정보 및 빅데이터 연계 정보를 수집한 다음, 악성코드 자체의 분석 결과와 OSINT&Intelligence 정보를 연관분석하여 사이버 수사에 필요한 필수 의미정보를 수집하고 도출하여야 한다. 이 단계를 살펴보면 Table. 1과 같이 표시할 수 있다.

각 단계는 유기적으로 연결되어 이루어지며, ① 1단계에서는 Cuckoo Sandbox와 같은 악성코드 분석 엔진을 커스터마이징 하여 악성코드 행위정보 및 아티팩트 추출하고, ② 2단계에서 사이버 수사 업무기반 전문 필터를 통해 수집 악

Table 1. Composition of a step-by-step malicious code analysis architecture based on cyber investigation

	Step #1	Step #2	Step #3	Step #4
	Analysis Engine	Smart Filtering	OSINT, Intelligence	Correlation analysis
Concept	· Collection of malicious activity logs using malicious code analysis platform · Advanced Automatic Malicious Code Execution Environment	· Professional Filters based on cyber investigation · Screening and filtering semantic data among collected malicious behavior information	· Collection of public source information	· Phase 1 to Phase 3 Correlation Analysis · Extract essential information for cyber investigation
Content	· Static analysis · Dynamic analysis · Artifact analysis · Virtual C&C Interworking and Analysis · Automatic injection of user interaction	· Define the information required for cyber investigation operations. · Regularization and lightening of information · Extracting investigation-based information	· Location, Analyze Service Provider Link Information · Analysis of information linking domain service information and registration information · Malicious code reputation analysis · Portal sites such as domain, IP, etc.	· Analysis Result - IP - Domain - Country. - Secondary Domains. - IP Class - Domain owner - Malicious code hash - Malicious code diagnosis name - Malicious code string - Open Port - Received - Mail From - Mail To - URLs in Body - About C&C - Signature
SW·Algorithm	· Analytic platform: (Customized)Cuckoo, Norman, Joe Sandbox, etc · Tools - String - Yara - Static Analyzer - API Analyzer - Process Analyzer - tcpdump - volatility - m2crypto - suricata - snort - hosts - Pre-filter - FakeNet-NG - iNetSim - HIEH	· Filter - Host - IP - DNS address - HTTP URL - mail address - copyright - author - account - MAC - Time stamp - signature - GeoIP	· OSINT - IP2Location - MaxMind - GeoBytes - NetAcuity - Akamai - Quova - Whois - C-Class - ISP - Virustotal - DomainTools - Bing - google · Intelligence - Virustotal - C-TAS - CTI - Mandiant Threat Intelligence	

성 행위정보 중 사이버수사에서 필요한 Host, IPAddress, URL, Email, Author, GeoIP와 같은 의미있는 데이터를 선별하고 필터링한다. ③ 이후 3단계에서 OSINT와 Intelligence를 활용하여 공개된 정보 및 위협 빅데이터 연계 정보를 수집하여 1단계에서 수집된 정보와 2단계에서 필터링 된 정보의 연관관계를 추출한다. ④ 마지막 4단계에서 1~3단계 정보 연관분석을 통해 사이버수사에 필요한 도메인, 국가, 도메인 소유자, 악성코드 정보, 악성코드에 감염된 컴퓨터의 활성정보, C&C정보 등 사이버 수사에 필요한 의미 추출한다. 이와 같은 전체 과정에 포함되는 도구 및 추출 정보의 관계는 Fig. 5 와 같이 설명할 수 있다.

이러한 과정을 통해 사이버 수사관은 단순한 악성코드 자체에 대한 정보를 활용하여 유포자를 추적할 수 있는 기반을 마련할 수 있다.

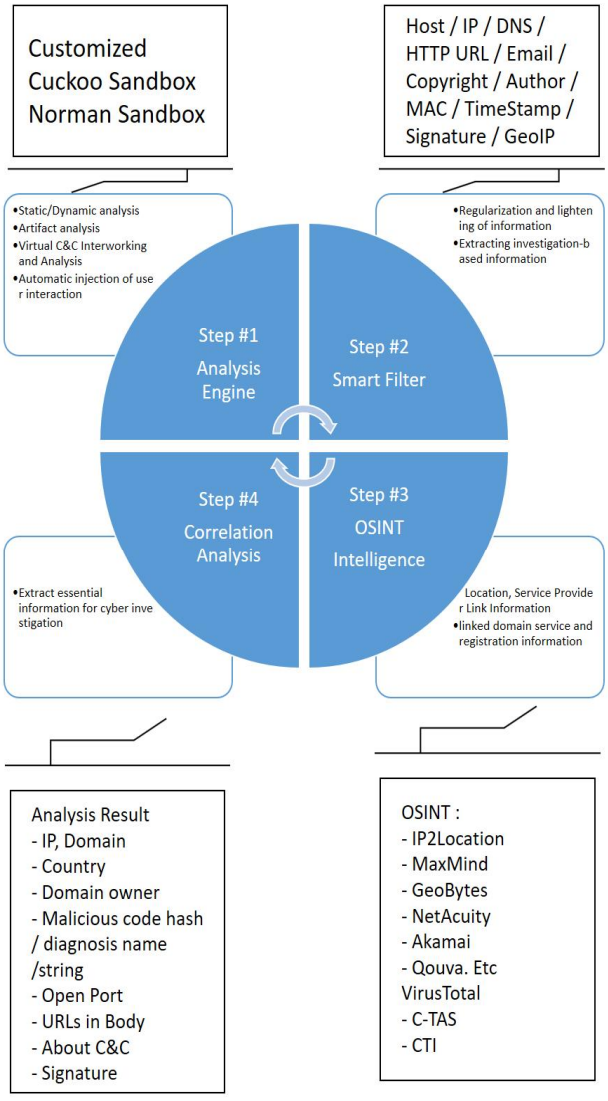


Fig. 5. Analyzer Structure

IV. Conclusions

본 논문에서는 악성코드 분석을 사이버 수사에 활용하기 위한 접근 방법을 살펴보았다. 최근 사이버범죄의 대부분은 악성코드로 이루어지는데, 이러한 이유로 악성코드를 분석하는 것은 사이버범죄 수사의 대부분을 차지한다. 하지만 단순히 악성코드의 행위를 분석하는 것은 악성코드로 인한 피해를 예방하는 데는 도움을 주지만, 사이버 수사의 관점에서 유포자를 찾아내고 추적하는 데는 한계가 있다. 이러한 한계를 극복하기 위해 본 논문에서는 악성코드 자체의 정보와 OSINT 및 Intelligence를 연계 분석하여 사이버 수사관이 악성코드를 유포하는 범죄자를 찾아내기 위한 차세대 분석 플랫폼 아키텍처를 제안한다. 본 논문에서 제안한 아키텍처를 활용하면 단순한 악성코드 행위정보뿐만 아니라 이를 통해 악성코드 유포자를 특징지을 수 있는 핵심적인 정보를 찾는 데 도움을 줄 수 있을 것으로 기대한다. 다만, 본 논문에서 제안한 아키텍처는 그 범위가 방대하여 구현을 통한 실험에는 한계가 있어 본 논문에서는 이론적인 부분에서 제안한 아키텍처의 효율성을 설명하였다. 향후 제안한 시스템을 실제 구축하는 과정을 통해 미비점을 보완하고 제안하는 아키텍처가 차세대 사이버 수사 플랫폼으로 가치가 있음을 검증할 예정이다.

ACKNOWLEDGEMENT

This work was supported by the Supreme Prosecutors' Office ("Architectural design research for the establishment of an intelligent malicious code analysis system").

REFERENCES

[1] ENISA, “ENISA Thread Landscape Report 2018”, Jun, 2019

[2] Y.S.Kim, "Ensemble Model using Multiple Profiles for Analytical Classification of Threat Intelligence", JOURNAL OF THE KOREA CONTENTS ASSOCIATION, Vol.17, No.3, pp.231-237, 2017.03, 10.5392/JKCA.2017.17.03.231

[3] Open Threat eXchange(OTX), *https://otx.alienvault.com/*

[4] Malware Information Sharing Platform(MISP), *https://www.misp-project.org/*

[5] Changwan Lim, Youngsup Shin, Dongjae Lee, Sungyoung Cho, Insung Han, Haengrok Oh "Real-time Cyber Threat Intelligent

Analysis and Prediction Technique, KIISE Transactions on Computing Practices, Vol.25, No.11, pp.565-570, 2019.11, 10.5626/KTCP.2019.25.11.565

[6] Choi Wonseok, Kim Jinsoo, "A System for Generating and Sharing Cyber Threat Intelligence on malicious code", Korea Software Congress 2018, pp.1035-1036, PeungChang, korea, Dec, 2018,

[7] Seonhee Seok, Howon Kim, "Visualized Malware Classification Based-on Convolutional Neural Network", Journal of the Korea Institute of Information Security & Cryptology, Vol.26, No.1, pp. 197-208, Feb. 2016, 10.13089/JKIISC.2016.26.1.197

[8] Taejin Lee "Trend of intelligent malicious code analysis technology using machine learning", REVIEW OF KIISC, Vol.28, No.2, pp.12-19, Apr, 2018

[9] Jun-ho Hwang, Tae-jin Lee, "Study of Static Analysis and Ensemble-Based Linux Malware Classification", Journal of the Korea Institute of Information Security & Cryptology, Vol.29, No.6, pp.1327-1337, Dec. 2019,10.13089/JKIISC.2019.29.6.1327

[10] Jun-ho Hwang, Tae-jin Lee, "Malware Packing Analysis Based on Convolutional Neural Network with 2-Dimension Static Feature Set", The Journal of Korean Institute of Communications and Information Sciences, Vol.43, No.12, pp.2089-2099, Dec. 2018, 10.7840/kics.2018.43.12.2089

[11] Seongmin Jeong, Hyeonseok Kim, Youngjae Kim, Myungkeun Yoon, "V-gram: Malware Detection Using Opcode Basic Blocks and Deep Learning", Journal of KIISE, Vol.46, No.7, pp.599-605, Jul, 2018, 10.5626/JOK.2019.46.7.599

[12] M. Sharif, A. Lanzi, J. Giffin, W. Lee, “Automatic Reverse Engineering of Malware Emulators”. 2009 30th IEEE Symposium on Security and Privacy. pp. 94-109, May. 2009.

[13] Soon-Gohn Kim, "Code Automatic Analysis Technique for Virtualization-based Obfuscation and Deobfuscation", Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol.11, No.6, pp.724-731, Dec. 2018, 10.17661/JKIIECT.2018.11.6.724

[14] Ki-Hwan Kim, Woo-Jin Joe, Hyong-Shik Kim, "A Malware Variants Detection Method using Malicious Behavior Signature", Korea Software Congress 2019, pp. 1633-1635, Dec. 2019

[15] Jinung Ahn, Hongsun Yoon, Souhwan Jung, "An Enhancement Scheme of Dynamic Analysis for Evasive Android Malware", Journal of the Korea Institute of Information Security & Cryptology, Vol.29, No.3, pp.519-529, Jun, 2019, 10.13089/JKIISC.2019.29.3.519

[16] Ollydbg, *http://www.ollydbg.de/*

[17] IDA pro, *https://www.hex-rays.com/products/ida/*

[18] Cuckoo Sandbox, *https://cuckoosandbox.org/*

[19] IP2Location, *https://www.ip2location.com/*

[20] MaxMind, *https://www.ip2location.com/*

[21] GeoByte, *https://geobytes.com/iplocator/*

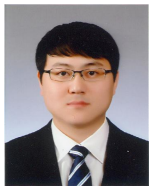
[22] NetAcuity, *https://www.digitalelement.com/solutions/*

[23] DomainTools, *https://www.domaintools.com/*

[24] Virustotal, *https://www.virustotal.com/gui/*

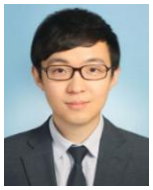
[25] C-TAS, *https://www.krcert.or.kr/data/noticeView.do?bulletin_writing_sequence=25824*

Authors



Ho-Mook Cho received his M.S. degree in Information Security from Ajou University in 2006, and Ph.d degree in Interdisciplinary of Information Security from Chonnam National University in 2018, Dr. Cho is a principal researcher in KAIST Cyber Security Research Center, KAIST, Daejeon, Korea. His research interests are in web security, malware analysis, XAI.

Chang-Su Bae received his B.S. degree in Cyber Security from Yeungnam University College in 2020. Bae is currently working as a researcher at APEX ESC Inc. He is interested in malware analysis, AI in cyber security, penetration testing, and DevOps.



Jaehoon Jang received his B.S. degree in Computer Science from Seokyeong University in 2013, and M.S. degree from Sungkyunkwan University EMBA in 2019. JaeHoon Jang established APEX ESC Inc. in 2016. He is the founder and current CEO of Zipida Inc. from 2019. He is interested in Business Development, AI, Big Data Visualization, and Information Security.

Sang-Yong Choi received his B.S. degree in Mathematics and M.S. degree in Computer Science, both from Hannam University in 2000 and 2003, and Ph.d degree in Interdisciplinary of Information Security from Chonnam National University in 2014, Dr. Choi is a assistant professor at the Dept. of Cyber Security in Yeungnam University College, Daegu, Korea. His research interests are in web security, network security and cloud computing security.