Journal of Korea Multimedia Society Vol. 13, No. 5, May 2010(pp. 754-762)

악성코드 은닉 문서파일 탐지를 위한 이메일 백신 클라우드 시스템

박 춘 식†

요 약

최근 악성 문서파일이 첨부된 이메일을 특정인에게 발송하여 중요자료를 절취하는 형태의 해킹사고가 지속적으로 발생하고 있다. 이러한 공격에는 공격 성공률 향상과 바이러스 백신의 탐지회피를 위해 주로 제로데이 취약점이 이용되고 있으며, 적절한 사회공학적 기법이 병행되는 것이 일반적이다. 본 논문에서는 조직으로 유입되는 이메일 첨부 문서파일에 대한 행위기반 악성문서 탐지기술이 적용된 이메일 백신 클라우드 시스템을 제안한다. 이메일에 포함된 문서파일을 추출하여 이메일 백신 클라우드 시스템에 전달하면, 백신 클라우드에서 시그니쳐 기반 분석 및 행위기반 분석을 통해 악성코드 포함 여부를 판단 후 악성코드를 제거한다. 행위분석 과정에서 의도하지 않은 실행파일 생성, 프로세스 실행, 레지스트리 엔트리 접근, 인터넷 접속시도 등이 발견되면 악성문서로 판단하게 된다. 본 논문에서 제시된 이메일 백신 클라우드 시스템은 악성문서 첨부 이메일의 유입을 효과적으로 차단함으로써 중요자료 유출 등의 각종 사이버테러 예방에 도움이 될 것으로 기대된다.

An Email Vaccine Cloud System for Detecting Malcode-Bearing Documents

Choon Sik Park†

ABSTRACT

Nowadays, email-based targeted attacks using malcode-bearing documents have been steadily increased. To improve the success rate of the attack and avoid anti-viruses, attackers mainly employ zero-day exploits and relevant social engineering techniques. In this paper, we propose an architecture of the email vaccine cloud system to prevent targeted attacks using malcode-bearing documents. The system extracts attached document files from email messages, performs behavior analysis as well as signature-based detection in the virtual machine environment, and completely removes malicious documents from the messages. In the process of behavior analysis, the documents are regarded as malicious ones in cases of creating executable files, launching new processes, accessing critical registry entries, connecting to the Internet. The email vaccine cloud system will help prevent various cyber terrors such as information leakages by preventing email based targeted attacks.

Key words: VM-Based Behavior Analysis(가상머신 기반 행위 분석), Malicious Document Detection (악성문서 탐지), Email Vaccine Cloud(이메일 백신 클라우드)

※ 교신저자(Corresponding Author) : 박춘식, 주소 : 서울시 노원구 화랑로 623(139-774), 전화 : 02)970-5752, FAX: 02)970-5981, E-mail : csp@swu.ac.kr
접수일 : 2009년 12월 22일, 수정일 : 2010년 2월 9일
완료일 : 2010년 3월 3일
† 정회원, 서울여자대학교 정보보호학과
※ 본 연구는 2010학년도 서울여자대학교 교내학술연구비의 지원으로 수행되었음

1. 서 론

최근 증가하고 있는 특정인을 대상으로 하는 사이버공격은 이메일의 송신자를 이메일 수신자가 이미 알고 있는 지인으로 가장하여 악성 메일을 보냄으로써, 메일 수신자가 아무런 의심 없이 해당 메일을 보도록 유도하여, 특정인의 컴퓨터를 감염시켜 정보를 절취하는 형태의 공격이다[1]. 절취되는 정보에는 특정인의 개인정보를 포함하여 개인용 컴퓨터에 저장되어 관리되고 있는 중요정보까지 다양하다. 이러한 특정인 대상의 공격 위협은 한국인터넷진흥원(구 한국정보보호진흥원)의 2009년 전망에서도 예측되었다[2]. 또한 메시지랩의 2008년 연차보고서에서도 특정인 대상의 공격이 그림 1과 같이 2008년에도 꾸준히 발생하고 있음을 지적하고 있다[3]. 특정인을 대상으로 하는 공격은 주로 알려지지 않은 취약점을 이용하며, 특히 윈도우 등의 운영체제나 서비스보다는 마이크로소프트의 오피스, PDF 문서, 한글 등의 문서 관련 응용프로그램의 취약점을 이용하여 악성코드를 전파한다[2]. 일반적으로 알려지지 않은 취약점이란 소프트웨어 개발사나 보안 업계에서는 인지하지 못하고 있지만, 해커와 같은 공격자가 발견하여 이미 활용하고 있는 취약점을 의미한다. 특히 조직 내의 방화벽 등에 의해 직접적인 공격이 대부분 차단됨에 따라 조직 내부의 사용자 컴퓨터를 공격하기 위한 방안으로써, 문서 내에 악성코드를 은닉시켜 이메일에 첨부하여 전달하는 방법이 많이 활용되고 있다.

SANS에서 발표한 2007년 보안위협 Top 20에 따르면, 엑셀, 워드, 비지오 등의 마이크로소프트 오피스 취약점 중 악성코드 설치와 같은 심각한 위협을 가할 수 있는 취약점의 발견건수가 증가하고 있으며, 이러한 현상은 그림 2에 잘 나타나 있다[4]. 이러한

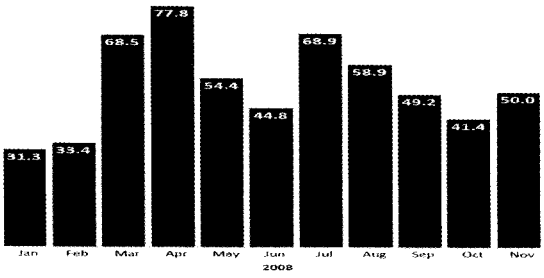

그림 1. 특정인 대상 공격의 일일 평균
(출처: MessageLabs Intelligence Report)

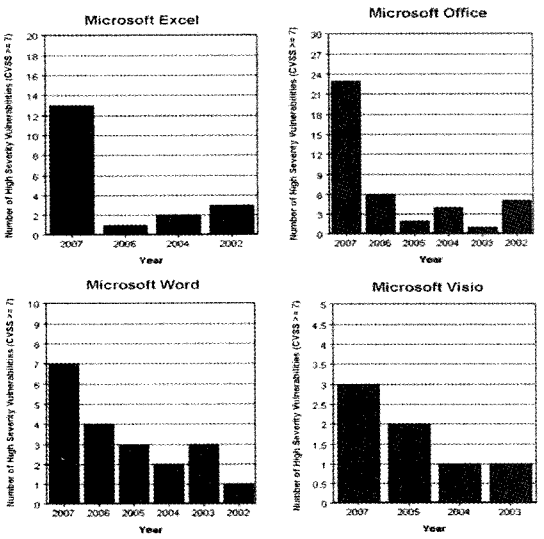

그림 2. 문서 응용프로그램 취약점 발견수
(출처: SANS 보안위협 Top20 2007)

응용프로그램 취약점의 증가로 인해 특정인 대상 공격 사례[5,6]가 언론에도 보도되고 있다.

본 논문에서는 응용프로그램의 알려지지 않은 취약점을 이용하여 악성코드를 은닉한 문서를 첨부한 이메일 공격을 예방하기 위한 이메일 백신 시스템을 제시한다. 알려지지 않은 취약점을 이용하여 악성코드를 문서에 은닉한 악성문서는 기존의 시그니쳐 방식으로는 탐지가 어렵기 때문에 본 논문에서는 행위기반의 탐지 기법을 도입했으며, 탐지의 효율성을 위해서 시그니쳐 기반의 백신도 활용하였다. 특히 수신자가 조직으로 유입되는 이메일을 열람하기 전에 첨부파일의 악성유무를 분석할 수 있도록 함으로써, 사용자는 기존의 메일 사용과 동일하여 시스템 도입에 따른 불편함이 없도록 하였다.

본 논문에서 제시하는 악성문서의 행위기반 탐지 기법은 응용프로그램 버전별 가상머신 환경을 구축하여 문서를 실행하고 모니터링함으로써, 악성문서가 특정 버전의 응용프로그램에서만 실행되어 탐지 회피하는 경우를 예방하도록 하였다. 악성문서 파일은 문서 편집 응용프로그램의 문서 파싱 취약점이나 스크립트 처리의 취약점을 주로 이용하며, 이러한 취약점은 응용프로그램의 버전 및 패치 상태에 따라 다르기 때문에 동일한 악성문서라 하더라도 모든 버전의 응용프로그램에서 악성행위가 수행 되는 것은 아니다. 따라서 문서 편집 응용프로그램 버전별로 문

서파일의 행위를 모니터링하지 않는다면 사용자 컴퓨터에 악성문서 파일이 전달될 수 있다. 기존 대부분의 연구는 실행파일 위주의 모니터링이었으며, 특정한 하나의 응용프로그램 버전 중심으로 문서파일을 모니터링으로써, 기관의 메일 서버 앞에 설치되어 악성코드를 사전에 예방하기 위한 목적에는 적합하지 않다.

본 논문은 2장에서 문서에 은닉된 악성코드의 탐지와 관련된 연구 동향을 살펴보고, 3장에서 이메일 백신 클라우드 시스템의 요구사항에 대해서 분석하며, 4장에서는 이메일 백신 클라우드 시스템의 구조에 대해서 설명한다. 5장에서는 이메일 백신 클라우드 시스템에서 문서에 은닉된 악성코드를 탐지하는 알고리즘에 대해서 설명하고 6장에 결론이 있다.

2. 관련연구

이메일 백신 클라우드 시스템 관련연구는 이메일 대상의 악성코드를 탐지하는 연구, 행위기반 악성코드 탐지 연구, 악성 이메일 첨부파일 탐지 연구 분야로 나누어 살펴본다. 먼저 이메일 대상의 악성코드를 탐지하는 연구에는 시그니쳐 기반의 탐지 기술이 있다. 시그니쳐 기반의 탐지 기술은 국내외의 많은 보안업체에서 연구되고 상품화되어 활용되고 있는 기술이다. 시그니쳐 기반의 이메일 악성코드 탐지는 탐지 속도가 빠르고, 오탐율이 낮은 장점이 있지만, 시그니쳐가 아직 없는 제로데이 취약점이나 알려지지 않은 취약점을 이용하는 악성코드 은닉 문서파일을 탐지할 수 없는 단점이 있다.

행위기반 악성코드 탐지 연구는 주로 호스트 기반의 침입탐지시스템이다. 공개소스 호스트 침입탐지시스템(Open Source Host-based Intrusion Detection System)은 각종 운영체제에서 동작하는 공개용 호스트기반 침입탐지시스템을 제공하고 있으며, 파일 및 레지스트리의 무결성 검사, 웹 서버, 데이터베이스 서버 등의 서버 로그 분석, 루트킷 탐지 방안에 대한 연구를 수행했다[7]. 애리조나 주립대에서 수행한 행위분석 기반 호스트 침입탐지시스템은 기계 학습을 통해 호스트 침입탐지시스템의 한계인 오탐을 최소화하기 위한 연구를 수행하였다[8]. 호스트 기반 침입탐지 시스템은 시스템에서 행해지는 모든 행위에 대한 모니터링을 통해 행위를 분석하여 악성행위를 판단하므로, 사용자 컴퓨터의 모든 행위 모니터링을 위해 많은 리소스가 요구되며, 사용자 컴퓨터에 설치할 경우 오탐이 증가하는 문제가 있다.

악성 이메일 첨부파일 탐지 연구는 컬럼비아 대학의 이메일 웜 백신 구조 연구[9]와 악성문서 하이브리드 탐지 모델 연구[10], 텍사스 대학의 악성코드의 하이브리드 탐지 모델 연구[11]가 있다. 이메일 웜 백신 구조 연구는 이메일의 실행 첨부파일을 추출하여 가상머신으로 전달 후 실행하고, 실행파일이 레지스트리에 대해 수행하는 정보를 이용하여 악성행위를 판단하는 연구를 수행하였다. 해당 연구는 실행파일 모니터링 위주로 수행되었으며, 악성 문서파일은 레지스트리 관련 행위를 하지 않을 수도 있으므로 악성 문서파일에는 적용이 어렵다[9]. 또 다른 연구로는 악성문서에 대한 하이브리드 탐지 모델 연구로서, 악성문서 파일에 대한 정적 분석과 동적 분석을 통해 악성여부를 판단한다. 정적 분석은 문서의 구조를 파싱하고 바이트 발생 횟수 및 엔트로피를 분석하여 악성코드 삽입 여부 및 악성코드가 삽입된 위치를 추정하는 연구로서 악성문서 분석 용도로는 적합하지만 이메일 서버 앞단에 적용하여 실시간으로 악성문서를 탐지하는 용도에는 적합하지 않다. 또한 동적 분석은 가상머신에서 특정 버전의 워드 응용프로그램만 설치하여 시험하였으므로, 설치된 버전 이외의 워드 응용프로그램에서 동작하는 악성코드는 탐지할 수 없다[10]. 텍사스 대학의 악성코드 하이브리드 탐지 모델 연구는 악성 실행파일을 탐지하기 위해서 바이너리 파일로부터 바이너리 n-gram 특징, 어셈블리 n-gram 특징, DLL 라이브러리 함수 호출관계 특징을 추출하여 SVM(Support Vector Machine)을 적용하여 악성여부를 탐지하는 연구이다[11]. 바이너리 n-gram 특징은 실행파일을 hexdump 도구를 이용하여 각 바이트를 16진수 값으로 변환하여 n 크기의 윈도우를 이용하여 인접하는 바이트들을 묶어서 특징으로 추출한다. 예를 들어 1F, 2E, 3C, 4B, 5A라는 16진수 값이 있을 때 n=2인 경우는 2-gram={1F2E, 2E3C, 3C4B, 4B5A}이고, n=3인 경우는 3-gram={1F2E3C, 2E3C4B, 3C4B5A}로 표현할 수 있다. 실행파일로부터 추출되는 n-gram 특징은 매우 많아서 해당 논문에서는 디스크 I/O를 통해 메인 메모리 부족 문제를 해결하고, 검색 지연 시간 감소를 위해서 AVL(Adelson Velsky Landis) 트리를 활용

하였으며, 많은 특징 중 정보 이득(Information Gain) 이론을 적용하여 500개의 최상의 특징을 추출하였다. 어셈블리 n-gram 특징과 DLL 라이브러리 함수 호출관계 n-gram 특징은 바이너리 n-gram 특징과 동일한 방식으로 진행되며, PEDisassem 도구를 통해 역어셈블된 어셈블리 명령어 단위의 n-gram 특징 추출과 DLL 함수호출 관계의 n-gram 특징을 추출한다. 이렇게 추출된 바이너리 n-gram 특징, 어셈블리 n-gram 특징, DLL 함수호출 관계 n-gram 특징을 SVM의 입력으로 하여 대상 파일의 악성여부를 판정한다. 실험결과 데이터 셋의 특징에 따라 최적의 n의 값이 달라지며 96%~99%의 탐지 정확도를 나타내었다. 해당 연구는 실행파일에만 적용되어 문서 파일에 적용할 수 없으며, 학습 데이터에 포함되지 않은 신규 취약점 이용 공격은 탐지가 어렵다. 또한 해당 연구는 사후 분석 용도로는 활용이 가능하지만, 사전 탐지를 위한 용도로 적용하기에는 특징 추출에서 많은 시간이 소요되어 적합하지 않다.

3. 요구사항

이메일 백신 클라우드 시스템은 현대의 조직적 협업에 가장 기초가 되는 이메일 시스템에 적용되어야 한다. 따라서 알려지지 않은 취약점 공격에 대응하면서도 이메일 시스템 본연의 임무를 수행하기 위해서는 다음의 요구사항이 만족되어야 한다.

첫째, 제로데이 취약점 또는 알려지지 않은 취약점을 이용하는 악성문서 파일을 탐지 할 수 있어야 한다. 시그니쳐 기반의 백신은 제로데이 취약점에 대하여 대응할 수 없다. 따라서 시그니쳐 기반이 아닌 악성문서를 실행하여 동작 과정을 모니터링함으로써 악성여부를 판단하는 행위기반 모니터링이 필요하다.

둘째, 응용프로그램의 취약점을 공격하는 악성 문서 파일은 모든 버전의 문서 편집 응용프로그램에서 동작하도록 만드는 것이 어렵다. 따라서 이러한 악성문서 파일은 특정 버전의 문서 프로그램에서만 동작할 가능성이 높기 때문에 이에 대한 대비가 필요하다. 본 논문에서는 이러한 제약사항을 해결하기 위해서 여러 버전의 문서 편집 프로그램을 준비하여 입력으로 전달되는 문서 파일을 각 버전별 응용 프로그램에서 실행해보고 악성 여부를 판단하는 백신 클라우드 구조를 도입하였다.

셋째, 이메일 시스템은 현대 조직의 가장 기본적인 협업 시스템으로 사용자가 기대하는 시간 이내에는 송신한 메일을 수신자가 확인할 수 있어야 한다. 따라서 악성 여부 분석을 위해 장시간이 소요될 경우는 이메일의 악성코드를 분석하는데 적합하지 않다. 본 논문에서는 시그니쳐 기반 백신 시스템으로 사전 점검 후 악성으로 판정한 경우는 즉시 해당 파일을 격리조치하고, 정상인 파일은 여러 버전의 응용프로그램에 동시에 실행 후 정해진 시간 동안만 모니터링 후 결과를 판단하도록 하였다. 조직 내에서는 시그니쳐 기반 백신 시스템이 사용자에게 모두 설치되어 있지만 백신의 업데이트 문제가 있다. 또한 다양한 백신 회사의 시그니쳐 업데이트 특성이 다르므로 여러 회사의 제품을 활용하는 것이 용이하나 조직 내에서의 비용 문제와 유사 제품의 동시 설치로 인한 충돌문제가 있다. 여러 회사의 시그니쳐 기반 백신 시스템을 클라우드로 구축하여 운영함으로써 백신 시그니쳐 업데이트 문제와 충돌 문제를 해결하도록 하였다.

넷째, 악성코드로 탐지된 분석 시스템이 초기화를 일정시간 내에 완료하여 다음 분석이 가능하도록 해야 한다. 또한 복수 버전의 응용프로그램에서 동시 분석이 가능하도록 해야 분석 시간을 줄일 수 있다. 이를 위해서 본 논문에서는 가상화 기술을 이용하여 분석 시스템을 가상머신에 구축하여 악성으로 탐지 시 가상머신의 재시작으로 빠르게 다음 문서를 분석할 수 있으며, 여러 버전의 응용프로그램을 각각 가상머신으로 구성하여 동시에 분석이 가능하도록 하였다. 가상머신의 재시작은 실제 시스템의 재시작보다 시간이 매우 짧으며, 분석 시스템이 문서를 분석하기 이전이 상태로 초기화 시킬 수 있다.

4. 이메일 백신 클라우드 시스템

본 장에서는 악성코드가 은닉된 악성 문서파일을 이메일로 전달하여 특정 대상을 공격하는 것을 예방하기 위한 이메일 백신 클라우드 시스템의 구조에 대해서 설명한다.

4.1 운영환경

알려지지 않은 취약점을 이용하여 특정 대상을 공격하는 악성 이메일을 예방하기 위한 이메일 백신

클라우드 시스템의 운영 환경은 그림 3과 같다. 이메일 백신 클라우드 시스템은 조직 내의 DMZ 구간에 설치된 메일 서버와 연동되어 운영된다. 메일서버에서 메일을 수신하면, 수신된 신규 메일로부터 API를 이용하여 첨부파일을 추출 후 분석을 수행하며, 분석결과 악성으로 판단되면 해당 메일을 격리시키게 된다. 이메일 백신 클라우드 시스템의 연동은 첨부파일의 추출 방식에 따라 메일 서버 앞부분과 연동될 수도 있다.

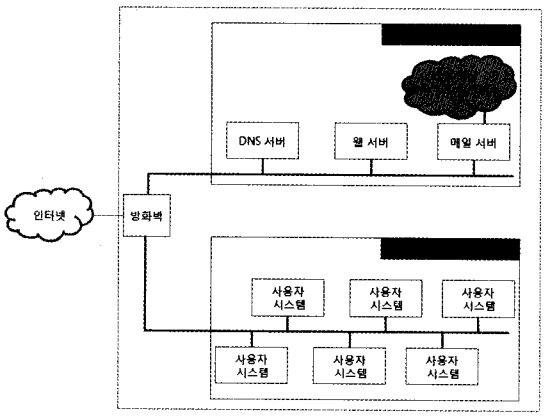

그림 3. 운영환경

4.2 시스템 구조

이메일 백신 클라우드 시스템의 세부 구조는 그림 4와 같다. 이메일 백신 클라우드 시스템은 클라우드 관리 서브시스템, 시그니쳐 기반 탐지 클라우드, 행위기반 탐지 클라우드로 구성된다. 클라우드 관리 서브시스템은 수신된 메일로부터 첨부파일을 추출하여 시그니쳐 기반 및 행위기반 탐지 클라우드로 전달하여 악성여부를 판정하고 그 결과를 원본 메일에 반영하는 역할을 수행한다. 시그니쳐 기반 탐지 클라우드는 탐지 성능을 향상시키기 위해서 이미 시그니쳐가 생성되어 있는 첨부파일에 대한 탐지를 수행하며, 행위기반 탐지 클라우드는 첨부된 문서 파일을 응용프로그램 버전별로 나뉘어진 각각의 가상머신에서 실행하여 악성여부를 판정하는 시스템이다.

클라우드 관리 서브시스템의 주요 역할은 다음과 같다.

- 탐지 처리 모듈 : 이메일 백신 클라우드 시스템 전체의 운영을 관리하며, 첨부파일의 추출 요청, 추출된 첨부파일의 시그니쳐 및 행위 기반의 탐지 클라우드에서의 분석을 담당한다.
- 전처리 모듈 : 수신된 메일로부터 첨부파일을 추출하여 확장자를 이용해 분석대상 문서가 포함되어 있는지를 확인한다.
- 후처리 모듈 : 분석결과 악성으로 판정된 첨부파일이 포함된 메일을 격리하고 해당 메일 수신자에게 악성 첨부파일이 포함된 메일이 수신되었음을 통보한다.

클라우드 관리 서브시스템의 동작모델은 그림 5와 같다.

시그니쳐 기반 탐지 클라우드는 여러 백신 회사의 제품을 각기 다른 가상머신에 설치하여, 분석대상 첨부파일을 각각의 가상머신에 전달하여 백신에서 제공하는 API를 이용하여 분석을 수행한다.

행위기반 탐지 클라우드는 여러 버전의 문서 응용프로그램을 각기 다른 가상머신에 설치하여, 분석대상 첨부파일을 각각의 서로 다른 버전의 응용프로그램에서 동시에 분석하도록 함으로써 분석시간을 절약하고 특정 버전의 응용프로그램에서만 실행되는

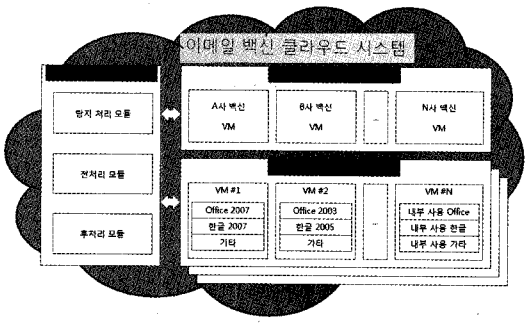

그림 4. 이메일 백신 클라우드 시스템 구조

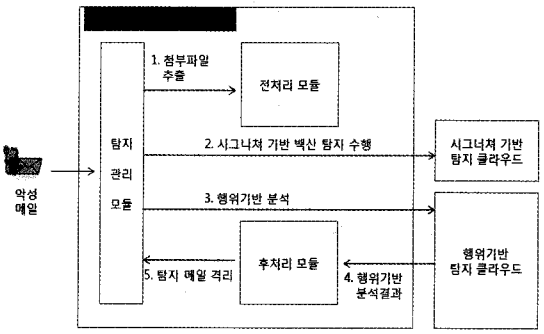

그림 5. 클라우드 관리 서브시스템 동작 모델

악성코드를 탐지할 수 있다. 특히 행위기반 분석을 위해서는 문서를 실행 후 일정시간 대기하면서 시스템의 행위를 모니터링 해야 하므로 시그니쳐 기반 시스템에 비해서 많은 시간이 소요된다. 이러한 문제점을 해결하기 위해서 행위기반 분석 시스템은 물리적으로 여러대의 시스템에 동일한 분석 환경을 구축하여 운영함으로써 분석에 소요되는 시간으로 인한 비효율성 문제를 극복할 수 있다.

5. 악성문서 탐지 메커니즘

5.1 시그니쳐 기반 탐지 클라우드

시그니쳐 기반 탐지 클라우드에서의 악성문서 탐지 과정은 그림 6과 같다. 클라우드 관리 모듈은 분석요청을 수신하여 첨부파일별로 백신이 설치된 가상머신으로 차례대로 전달한다. 가상머신 내에서는 분석관리 모듈이 백신과 연동되어 분석대상 파일을 스캔하여 악성여부를 판정 후 결과를 전달한다. 여러 백신 가상머신 중에서 하나의 가상머신이라도 해당 파일을 악성으로 판정하는 경우, 해당 첨부파일은 악성으로 판정한다. 왜냐하면 시그니쳐 기반의 경우 특정 회사의 시그니쳐가 먼저 발표되어 업데이트 될 수 있기 때문이다.

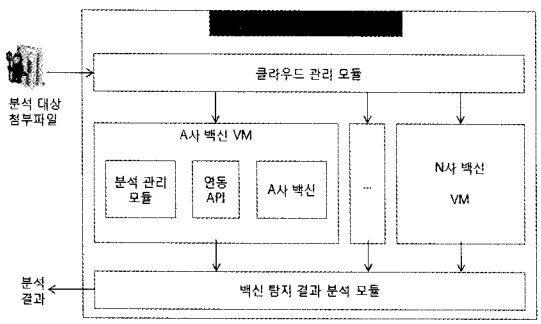

그림 6. 시그니쳐 기반 탐지 클라우드 동작 모델

5.2 행위 기반 탐지 클라우드

행위기반 탐지 클라우드는 시그니쳐 기반 탐지 클라우드에 비해 악성여부를 판정하는 것이 복잡하다. 클라우드 관리 모듈은 분석요청 및 대상 파일을 관리하며, 요청된 파일을 여러 버전의 응용프로그램이 설치된 가상머신으로 전달하여 분석을 수행한다.

가상머신 내의 문서 실행 관리 모듈은 분석 요청된 문서파일을 응용프로그램을 실행하여 오픈하며, 가상머신 내의 파일, 프로세스, 레지스트리, 네트워크, 커널, 추가 모니터링 모듈 등을 통해 시스템의 행위를 모니터링 한다. 가상머신은 불필요한 서비스 및 응용프로그램을 제거했기 때문에 일반 사용자의 컴퓨터에 비해서 모니터링을 통해 악성여부를 판정하는 것의 정확성이 높다. 문서 실행을 종료 후 각각의 모니터링 모듈에 의해서 로깅된 정보를 분석해서 악성여부를 판정한다. 이와 같은 과정은 그림 7의 행위기반 탐지 클라우드의 동작 모델과 같다.

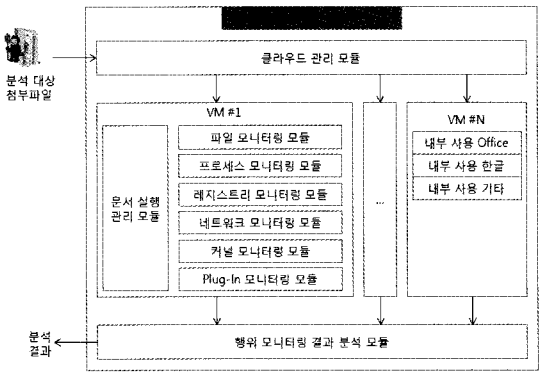

그림 7. 행위기반 탐지 클라우드 동작 모델

5.3 악성문서 탐지 메커니즘

본 절에서는 가상머신 내에서 악성문서를 탐지하는 방안을 설명한다. 악성문서 파일의 분석과정은 그림 8과 같다. 먼저 분석대상 문서 파일을 각각의 분석 가상머신으로 전달하면 문서 실행 관리 모듈이 응용프로그램을 실행 후 해당 문서를 오픈하게 된다. 오픈된 문서는 일정시간 대기한 후 응용프로그램을 종료시키게 된다. 일정시간 대기하는 이유는 문서가 열린 후 악성행위를 하는지 모니터링을 하기 위함이다. 대부분의 문서 파일은 실행이 되면서 해당 문서의 파싱과정의 취약점으로 악성행위를 수행하므로 일정시간 동안 시스템의 행위를 모니터링한다. 가상머신 내의 악성행위 모니터링 모듈은 각각의 모니터링 결과를 생성하게 되고, 각각의 가상머신에서 생성된 모니터링 결과 로그를 수집하여 분석한다. 악성 문서파일은 많은 경우에 악성코드를 하드디스크에 복사하고 실행하므로 파일시스템에 실행파일의 생성 여부 및 신규 프로세스의 실행, 레지스트리 접근 여부 등을 통해 문서의 악성여부를 판정한다.

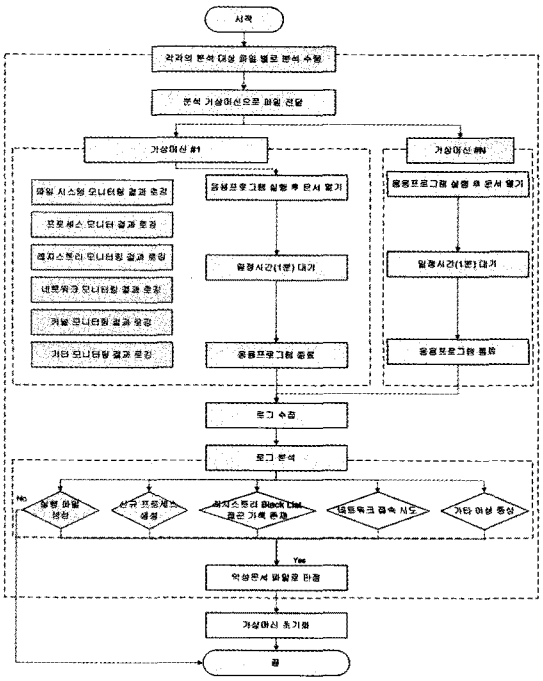

그림 8. 행위기반 악성문서 파일 탐지 순서도

5.4 비교 분석

본 절에서는 이메일 백신 클라우드 시스템과 관련 연구들의 특징을 관련연구와 비교 분석하고, 3장에서 제시하였던 요구사항 만족 여부를 분석한다.

관련연구에서 제시했던 유사 시스템과의 특징 비교는 표 1과 같다. 이메일 백신 클라우드 시스템은 문서에 은닉된 악성코드 탐지가 가능하며, 특정 문서 응용프로그램 버전에서 동작하는 악성코드 탐지가 가능하다. 기존 악성문서의 행위기반 탐지 연구에서 사용한 탐지 기법은 탐지시스템에 설치된 문서의 응용프로그램에서 실행되는 악성코드를 포함한 악성문서만 탐지 가능하다. 일례로서 워드 2003에서 여는 경우 악성행위를 수행하는 문서가 워드 2007에서 여는 경우는 악성해위를 수행하지 않거나, 워드 2007에서는 여는 경우 악성행위를 수행하는 문서가 워드 2003에서 여는 경우는 악성행위를 수행하지 않을 수 있다. 이는 문서 편집 응용프로그램의 버전에 따라서 해당 프로세스의 스택 포인트(SP)를 포함한 레지스트리의 값들이 달라져서, 문서 내에 은닉된 악성코드의 위치 계산 결과가 달라져서 실행이 되지 않기 때문이다. 본 논문에서 제시한 시스템은 여러 버전의 문서 편집 응용프로그램에서 문서를 모니터링하기 때문에, 특정 버전의 문서 응용프로그램에서 실행되는 악성코드를 포함하는 악성문서도 탐지가 가능한 장점이 있다. 또한 수분 이내에 첨부파일의 악성여부를 분석할 수 있으며, 시그니쳐 기반 백신 클라우드와 연계하여 분석 시간을 최소화할 수 있도록 하였다.

행위기반 분석 시스템에서 분석결과 악성으로 판정될 경우 분석 시스템은 다음 파일을 분석하기 전에 초기화가 필요하다. 본 논문에서 제안한 시스템에서는 가상화 기술을 이용하여 분석 시스템을 가상머신에 구축하고, 악성으로 탐지 시 가상머신의 재시작으

표 1. 관련 연구와 제안 시스템의 비교

	실행파일 탐지	악성문서 탐지	제로데이 취약점 또는 알려지지 않은 취약점 공격 탐지	특정 버전에서 동작하는 악성문서 탐지	탐지 시간
Email Vaccine Cloud System (본 연구에서 제안한 시스템)	○	○	○	○	수분 이내 첨부파일 분석 가능
악성문서 탐지 모델 연구[10]	×	○	○	△	문서의 구조를 파싱하고 바이트 발생 횟수 및 엔트로피를 분석해야 하므로 수분내 탐지 불가능
시그니쳐 기반 탐지 기술	○	○	×	○	수초내 시그니쳐 매칭 가능
Email Worm Vaccine Architecture [9]	○	×	○	-	수분내 분석 가능
택사스 대학 연구 [11]	○	×	△	-	특징 추출에 소요되는 시간으로 인해 수분내 탐지 불가능
호스트기반 침입탐지시스템[7]	○	×	△	-	수분내 분석 가능

로 분석시스템을 초기화하여 다음 문서를 분석할 수 있도록 하였다. 가상머신의 재시작은 실제 시스템의 재시작보다 매우 짧은 시간 내에 분석환경의 초기화가 가능하며, 자동화도 가능하다.

6. 결 론

최근 응용프로그램 취약점이 증가하고 알려지지 않은 응용프로그램 취약점을 이용한 특정 대상 공격으로 인해 중요 자료가 절취되는 사이버 침해 사고가 발생하고 있다. 본 논문에서는 이러한 응용프로그램의 취약점을 악용하여 문서 내에 악성코드를 은닉한 첨부파일을 이메일로 전달하여 시스템을 장악하는 공격을 예방하기 위한 이메일 백신 시스템을 제안하였다. 이메일 백신 시스템은 시그니쳐가 아직 준비되지 않은 악성코드를 탐지하기 위하여 행위기반 모니터링 기법을 사용하였으며, 특히 특정 응용프로그램에서만 동작하는 악성코드를 탐지하기 위하여 여러 버전의 응용프로그램을 실행할 수 있도록 가상머신을 이용하여 행위기반 모니터링 클라우드를 구성하였다. 또한 기존 시그니쳐 기반 시스템의 장점인 빠른 탐지 성능을 활용하고 패턴 업데이트 미비로 감염되는 사례를 예방할 수 있도록 하기 위하여 시그니쳐 기반 클라우드를 함께 구성하였다. 분석대상을 시그니쳐 기반 클라우드에서 점검 후 행위기반 모니터링을 수행함으로써 탐지의 효율성을 높였다.

본 논문에서 제시한 이메일 백신 클라우드 시스템은 수신자가 조직으로 유입되는 이메일을 열람하기 전에 첨부파일의 악성유무를 탐지함으로써 첨부파일 열람 과정에서 악성코드에 감염되는 것을 예방하고, 이를 통해 중요 자료의 유출을 예방할 수 있다. 이를 통해 해당 조직의 사이버보안 수준을 향상시킬 수 있을 것으로 기대된다.

본 논문에서 제시한 이메일 백신 클라우드 시스템은 응용프로그램의 버전별로 분석 가상머신이 요구되며, 여러 가상머신에서 동시에 분석을 수행하기 위해서는 높은 성능의 시스템이 요구되는 단점도 있다. 따라서 향후 연구로는 하나의 분석 시스템에서 악성문서를 버전과 관계없이 탐지할 수 있는 방안에 대한 연구가 필요하다. 또한 탐지된 악성문서의 분석을 자동화하여 악성코드의 은닉위치 등을 판단할 수 있는 자동화된 악성문서 분석 연구도 필요하다.

참 고 문 헌

[1] Targeted Trojan Email Attacks, http://www.us-cert.gov/cas/techalerts/TA05-189A.html

[2] 한국정보보호진흥원, "인터넷 침해사고 동향 및 분석월보-2008년 침해사고 동향 및 2009년 전망 특별 보고서," 한국정보보호진흥원 인터넷침해사고대응지원센터, 2008, 12.

[3] MessageLabs Intelligence: 2008 Annual Security Report, 2008, http://www.messagelabs.com/mlireport/MLIReport_Annual_2008_FINAL.pdf

[4] SANS Top-20 Security Risks, Nov., 2007, http://www.sans.org/top20/2007/top20.pdf

[5] 군·주한미군 장성 겨냥 北 해커 해킹메일 살포, 서울신문, 2009, 6, 17, http://www.seoul.co.kr/news/newsView.php?id=20090617004007

[6] 이명박 대통령 사칭 해킹메일 조심!, 동아일보, 2008, 3, 14, http://news.donga.com/fbin/output?n=200803140410

[7] OSSEC Homepage, http://www.ossec.net/main/

[8] Haiyan Qiao, Jianfeng Peng, Chuan Feng, and Jerzy W. Rozenblit, "Behavior Analysis-Based Learning Framework for Host Level Intrusion Detection," Proc. of the 14 th IEEE Intl. Conference and Workshops on the Engineering of Computer Based Systems (ECBS '07), pp. 441-447, Tucson, Arizona, March 2007.

[9] S. Sidiroglou, J. Ioannidis, A. D. Keromytis, and S. J. Stolfo, "An Email Worm Vaccine Architecture," In Proceedings of the 1st Information Security Practice and Experience Conference (ISPEC), pp. 82-101, April 2005.

[10] M. Masud, L. Khan, and B. Thuraisingham, "A hybrid model to detect malicious executables," IEEE International Conference on Communications 2007(ICC '07), pp. 1443-1448, 2007.

[11] W. Li, S. Stolfo, A. Stavrou, E. Androulaki, and A. Keromytis, "A Study of Malcode-Bearing Documents," In Conference on

Detection of Intrusions and Malware & Vulnerability Assessment(DIMVA), pp. 231-250, 2007.

박 춘 식

1995년 일본동경공업대 공학박사
1982년~1999년 한국전자통신연구원 책임연구원
2000년~2008년 국가보안기술연구소 책임연구원,소장
2009년 3월~현재 서울여자대학교 클라우드컴퓨팅연구센터 정보보호학과 교수,
관심분야 : 개인정보보호기술, 클라우드컴퓨팅보안