확장된 증거수집 및 사건연관분석을 기반으로 한 컴퓨터 포렌식

정 일 옥○
고려대학교 컴퓨터정보통신대학원
okkida@korea.ac.kr

Comprehensive Computer Forensics based on Event Correlation with Extended Evidence Scope

Jung il ok○
Department of Computer Science and Engineering, Korea University

요 약

진화되고 위협적인 사이버공격 및 피해가 증가함에 따라 기업이나 기관의 정보보호에 대한 책임도 증가하게 되었다. 이에 종합적인 컴퓨터 범죄 재현과 정확한 침입경로 및 피해규모, 정보의 신뢰성을 파악하기 위한 컴퓨터 포렌식에 대한 연구가 활발해 지고 있다. 이에 대부분의 기업이나 조직에서 이기종의 보안장비에서 발생하는 다량의 경보와 이벤트를 효과적으로 수집통합하고 상호연관분석 할 수 있는 통합보안관리시스템(ESM)을 도입하여 운영하고 있으나 많은 경보발생으로 인해 적절한 판단이나 분석 및 효율적인 대응이 이루어지고 있지 않다. 이에 본 논문에서는 수집되는 증거의 범위를 재 정의하고, 이벤트 상관분석을 통해 발생된 침해경보에 대해 경보검증을 적용하여 경보의 오탐율을 감소시켰으며, 검증된 경보에 대해서 신속히 분석 및 대응이 이루어지는 포렌식 모델을 제안한다. 이를 통해 오탐율 감소는 물론 신속하고 신뢰성 있는 탐지 및 침해 분석이 가능하다

1. 서 론

최근 몇 년 동안 지속적으로 발생하고 증가하는 보안사고는 이제 기업이나 기관의 정보보호에 대한 책임도 요구하고 있다. 금전적인 피해가 있는 사이버 범죄와 같은 형사소송뿐만 아니라 개인정보유출, DOS 공격으로 인한 고객서비스 중단 등의 업무상 과실을 묻기 위한 민사상의 소송영역까지 영향을 미칠 수 있으므로 침해 대응뿐만 아니라 증거자료 수집이나 추적 등의 컴퓨터 포렌식 기술의 중요성이 증가하고 있다.

이에 대부분의 기업이나 조직에서는 기 도입된 서버네트워크장비, 보안장비(침입차단시스템, 침입탐지시스템, 웹방화벽 등)에서 발생하는 다량의 경보와 이벤트를 수집 통합하고 상호연관분석 할 수 있는 통합보안관리시스템(ESM)을 도입하여 운영하고 있으나 이 또한 한정된 이벤트를 근간으로 하기 때문에 경보에 대한 검증없이 보안경보를 발생시켜 여전히 관리자의 부담이 되고 있어, 수집된 이벤트를 통한 포렌식에도 적절하게 이용되지 못하고 있다.

이러한 이유들로 인해 수집되는 증거(이벤트)의 범위를 단순한 이벤트뿐만 아니라 서버 취약점 정보, 패치정보, 휘발성 있는 정보 등으로 확장하여 수집하고, 보안경보와 확장된 이벤트를 연관분석하여 침입을 확인하는 보안검증단계를 추가하여 적절한 판단과 분석, 효율적인 대응이 가능하도록 해야 한다.

본 논문에서는 2장에서 컴퓨터 포렌식에 대해서 정의한 후 침입과 포렌식의 관계, 기존에 정의된 ECF(Event Correlation for Forensic), 상관분석과 경보검증(Alert Verified)에 대해서 살펴본다. 3장에서는 범위가 확장된 이벤트를 기반으로 보안검증단계가 추가된 포렌식 모델을 제안한다. 마지막으로 4장에서는 시나리오 기반으로 한 공격을 통해 경보검증을 통한 이벤트 감소와 침해분석을 살펴본다.

2. 관련연구

2.1. 컴퓨터 포렌식의 정의와 중요성

컴퓨터 포렌직은 1991년 미국 Oregon Portland에서 개최된 IACS에서 처음 사용되었으며, "컴퓨터를 매개로 이루어지는 범죄행위에 대한 법적 증거자료 확보를 위해 컴퓨터 저장매체 등의 컴퓨터 시스템과 네트워크로부터 자료(정보)를 수집, 분석 및

보존절차를 통해 법정 증거물로써 제출할 수 있도록 하는 일련의 행위'라고 정의되었다.

특히, 컴퓨터 포렌식에 사용되는 디지털 자료는 복사나 조작이 쉽기 때문에 사건이 발생하였을 때에 법적증거자료로 사용가능하도록 무결성, 객관성, 정규화된 포렌식기법을 사용하여 분석하여야 한다.

2.2. 침입(Intrusion)과 포렌식(Forensic)의 정의 및 관계

침입(Intrusion)이란 비인가된 사용자가 자원의 무결성(Integrity), 기밀성(Confidentiality), 가용성(Availability)을 저해하는 일련의 행동들과 보안 정책을 위반하는 행위를 말한다.

이러한 침입을 탐지하고 대응하기 위해서 이기종의 보안장비와 분석방법이 필요하며, 현재 많은 연구가 진행되고 있다[7][8] 이에 대해서 아래와 같이 분류하여 보았다.

O 침입탐지 : 침입에 대한 탐지
O 경보확인 : 침입에 대해 탐지하고 서버정보(취약점, 자산정보)를 통해 침해여부를 판단
O 포렌식 : 침입에 대해 판단하고, 서버정보(취약점, 자산정보)를 통해 침해여부를 판단하며, 서버분석을 통해 침해 사황을 분석

침입탐지는 진행 중인 이벤트에 대해서 수행하며, 경보확인은 저장된 DB 와 진행 중인 이벤트를 분석하여 점검한다. 또한 포렌식은 발생한 이벤트에 대해서 분속하는 것이라 하겠다. 본 논문에서는 경보확인을 통해 진행 중인 이벤트와 발생한 이벤트사이의 간격을 최소화하는 것이 하나의 목적이라 할 수 있다.

2.3. ECF (Event Correlation Forensics)

Kevin.Chen 과 Andrew.Clack 은 Network & Information Forensics Conference 2003에서 ECF (Event Correlation for Forensics) 라는 포렌식을 목적으로 한 이벤트 상관분석에 대해서 제안하였다. [2]

ECF (Event Correlation for Forensics)란 시스템 및 기타 로그에서 제공되는 이벤트(시스템 이벤트, 감사로그, 접근로그 등)에 대해서 상호 연관분석을 통해 침입 등을 파악하는 기술이다.

처음에 수집되는 Event 로는 Apache Server Log, Windows 2000 Security Log, Door Log, Browser Cache Logs, UNIX, Linux SYSLOG(POP3, SSH, sendmail 등) 등을 사용하였다. 이후 J Abbott, J Bell, A Clark, O De Vel, G Mohay 은 2006년 ACM symposium에서 Auto-ECF 를 제안하면서 수집되는 이벤트를 아래와 같이 정의하였다.[3]

O Raw Event - 가공되지 않은 데이터, 해당 이벤트는 표준화가 되어 있지 않기 때문에 변조 및 수정이 어렵다. 또한, 해당 데이터는 특정 어플리케이션에서만 읽을 수 있다.
O Logical Event (or Event) - 데이터베이스안에 표준화된 형태로 저장된 이벤트
O Simple Logical Event (or Simple Event) - Log parser 에 의해 데이터베이스에 표준화된 형태로 저장된 이벤트이며, 해당 이벤트는 Raw Event 와 직접적이며, 1:1 관계를 맺고 있다.
O Constituent Events - Simple Logical Events가 아닌 Local Event 로 구성된 이벤트

하지만, Kevin.Chen과 Andrew.Clack이 제안한 ECF은 이벤트를 이용하여 시스템의 이상현상을 파악하거나 분석하는 한계가 있다

2.4 Event Correlation (상호연관분석) 및 Alert Verified (보안검증)

가) Event Correlation (상호연관분석)[5][6]

Correlation 이란 사전적인 의미로"사건과 사건, 현상과 현상 사이에 존재하는 어떤 종류의 관계"를 의미한다.

현재, 대부분의 상호연관분석은 오탐이 가장 많이 발생하는 침입탐지시스템의 경보에 대해서 연관분석을 통해 경보를 축약하고, 오탐을 제거하는 연구를 수행하고 있다.

아래는 기존에 연구된 경보분석 방법을 살펴본다.

O Similar Attack Attributes

SRI International 의 Valdes 는 경보간의 연계성(probabilism) 에 비중을 두고 유사한 경보들을 그룹화하는 방법을 사용하였다. 새로운 공격을 탐지했을 때 경보들을 비교하여 유사그룹이 있으면 해당그룹에 포함시키고 그렇지 않으면 새로운 경보 쓰레드를 생성한다. 이때 유사도(Similarity) 를 평가하는 방법은 두 경보간의 특징(feature)을 비교하는 것이다. 아래는 침입탐지 경보들 간의 종합적인 유사도를 계산하는 방법이다.

$$\sim(X, Y) = \frac{\sum_j E_j \sim (X_j, Y_j)}{\sum_j E_j}$$

X = 매칭을 위한 후보 메타경보
Y = 새로운 경보
j = 경보속성에 대한 인덱스
E_j = 속성 J에 대한 유사도 기대치
X_j, Y_j = 각 경보 X와 Y의 속성 J의 값

O Pre-definded Attack Scenarios

Debar 와 Wespi 가 IBM Tivoli Enterprise Console(TEC) 에 구현된 방법으로 ACC (Aggregation and Correlation Component) 의 개념을 가진다. 기본개념은 유사공격동향을 집합화하는 것이고 속성이 연관된 경보들을 묶어 Security Level 을 평가하고 추론한다. 경보를 집합화할 때 공격자 ip, 공격대상 ip, 공격클래스 3 가지 속성을 기준으로 하며 다음과 같은 시나리오를 이용할 수 있다.

(시나리오 1) 한 공격자가 한 공격대상을 공격할 때(같은 공격자 ip, 공격대상 ip, 공격클래스) 항목이 동일하다.
(시나리오 2) 한 공격자가 한 공격대상을 여러 방법으로 공격할 때 (같은 공격자 ip, 공격대상 ip) 항목이 동일하다.
(시나리오 3) 한 공격대상에 분산된 공격할 때(같은 공격대상 ip, 공격클래스) 항목이 동일하다.
(시나리오 4) 한 공격자가 여러공격대상에 같은 공격을 할 때 (같은 공격자 ip, 공격클래스) 항목이 동일하다.
일련의 경보들은 시간대별로 제약조건을 가지고 연관 되어진다. 통합보안관리시스템(ESM)에서 룰기반의 방법과 유사하다.

O Pre & Post Condition of Individual Attack
경보가 발생하기 위해 필요한 조건(Prerequisites)과 공격으로 인해 발생 가능한 결과(Consequences)를 선행관계(Prepare-For)로 연결하는 방법이다. 이 분류는 조건별로 제약을 주어 단순화 시켜야 한다. 그렇지 않으면 선행조건에 따라서 결과가 방대해 질 수 있고 이 방대한 결과들을 모두 표현하기 위해서는 현실적으로 구현이 어려워진다.

나) Alert Verified (보안검증)[5]
Alert Verified 는 2004년 Kuregel이 제안한 방법인데 상관분석을 통해 경보를 축약하였음에도 불구하고 오탐이 발생하여 이를 검증하는 단계를 두어 경보를 최소화 하는 방법이다. 실제 Kuregel은 표1에서와 같이 검증을 통해 오탐의 확률을 0%로 줄이기까지 하였다. 보안 검증을 위해서는 해당 네트웍에 존재하는 취약점 및 자산정보(OS, 어플리케이션등)에 대한 정보가 저장된 DB가 존재하여야 하며 발생된 Alert 에 대해서 경보검증을 통해 경보를 축소시킬 수 있다.

표 1. Alert Verification - Evaluation Results

	Alerts	True Positives	False Positives
Stand-alone	6,659	24	99.64%
Verification enable	24	24	00.00%

Kuregel이 경보에 대해서 아래와 같이3가지로 정의하였다.
(타입 1) True Positive : 센서가 올바르게 성공한 공격으로 구분한 경우
(타입 2) Non-Relevant Positive : 센서가 올바르게 공격으로 구분했지만, 실패한 경우
(타입 3) False Positive : 센서가 정상적인 접속을 공격으로 구분한 경우 (오탐)

하지만, Kuregel이 제시한 방법은 침입탐지시스템에 한정되어 검증하였으며, 또한 sync flooding, icmp flooding 과 같은 네트워크와 관련된 공격에 대해서는 처리하기 어려운 단점이있다.

3. 제안된 ECF(Event Correlation for Forensic) 모델

기존에 Kevin.Chen 과 Andrew.Clack 이 제시한 ECF 모델은 시스템에 대해서 실시간으로 수집 분석하여 침입을 판단하는 한계점을 가지고 있었다면 본 논문에서 제안하는ECF 모델은 실시간으로 서버, 네트워크장비, 보안장비(침입차단시스템, 침입탐지시스템, 웹방화벽등)의 이벤트를 수집, 분석하여 침입을 파악하는 확장된 ECF 모델을 제시한다.

제안된 ECF 기법은 단순한 이벤트를 수집 후 이에 대한 분석이 아니라 그림 1과 같이 [로그수집]-[Alert 발생]-[Alert 축약]-[Alert 검증]-[침입판단] 으로 이루어진 포렌식 모델이다. 이는 실시간으로 이기종의 보안장비나 시스템을 통해서 확장된 이벤트를 수집하고, 수집된 이벤트를 연관분석을 통해 분석 후 보안검증을 통해 침입이라 판단된 공격에 대해서는 대상서버에 대해 포렌식을 실시한다.

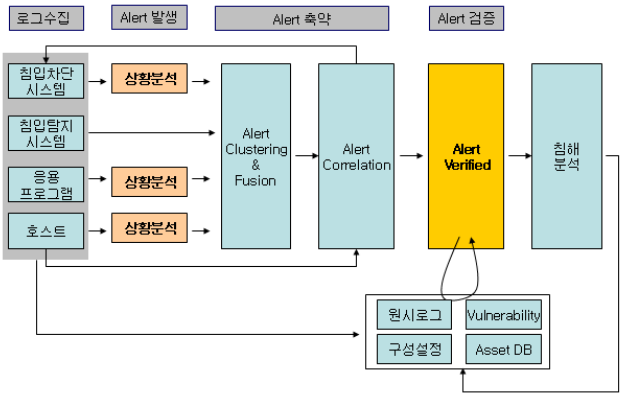

그림 1 제안된 ECF 모델

3.1 로그 수집기

제안 모델에서 로그 수집은 가장 기본이며 중요하다.이는 Agent 형태로 해당 시스템에서 작동하며 수집된 정보를 실시간으로 Manager에 전달한다. 로그 수집은 두 가지 방식으로 행해지는데 하나는Event 형태로 이미 정의한 정형화된 형태로 실시간으로 수집 전달된다. 다른 하나는 서버에 대한 정보와 원시형태의 바이너리파일 휘발성이 있는 정보에 대해서는 정기적 또는 요청 시 전달한다. (5분, 10분, 1시간 등) 이는 컴퓨터 포렌식의 증거수집 중 무결성과 가용성 신속성을 보장하기 위해서이다.
특히, 침입탐지시스템의 경우 발생된 이벤트는 이미 이벤트를 분석한 경보의 형태이므로 상세한 분석이 어렵기 때문에 판단의 근거가 되는 RAW DATA 의 수집이 필요하다.
아래 표2는 각 단위보안시스템의 정형화된 보안이벤트와 포렌식에 필요한 항목을 비교한 것이다.[1]

표 2. 장비별 정형화 & 포렌식에 필요한 항목 비교

유형	장비명	정형화	포렌식
네트워크유형	침입차단시스템	발생시간, 발생장비, 근원지 IP, Port 목적지 IP, Port, Protocol, 이벤트종류 중요도, 제품명, 발생횟수	시간, Version 통제정책(Rule) 원시로그
	침입탐지시스템	발생시간, 발생장비, 근원지 IP, Port, 목적지 IP, Port, Protocol, 위험도, 공격종류, 공격명(Signature), 제품명, 조치내역, 발생횟수	시간, Version 탐지정책 침입탐지시의 RAW DATA
호스트유형	서버	발생시간, 발생장비, 발생이벤트 처리내용 근원지 IP, Port, 목적지 IP, Port, 프로토콜, 패킷크기, 제품명, 이벤트발생규칙, 발생횟수	시간, Version Connection Log 사용자 Process 중요 Process
웹유형	웹로그	발생시간, 발생장비, 근원지 IP, 결과코드, Method, 쿼리값, 쿠기값, 발생횟수	시간, Version Connection Log 사용자 Process

3.2 Alert 발생

Alert 는 각각의 보안장비별로 수집된 이벤트에 대해서 상황분석을 적용하여 발생시킨다. 하지만, 침입탐지시스템의 이벤트에 대해서는 오탐의 여지가 너무 많고, 이미 침입탐지 시스템에서 이벤트에 대한 점검을 통해 발생된 경보이기 때문에 상황분석을 적용하지 않는다. 이때 사용되는 연관분석기법은 전자통신연구원에서 개발한 NASA (Network Attacks Situation Analysis) "사전에 정의한 공격시나리오에 의한 분류 방법을 준용하기로 한다.

3.3 Alert 축약

발생된 Alert 에 대해서는 Clustering 이나 Correlation 과정을 거친 후 검증을 수행한다. 이때의 Correlation 은 Alert 발생에서 발생된 경보와 수집된 이벤트와의 상관분석을 통해 이루어 진다. 이는 다양한 보안이벤트간의 상호연관요소를 기반으로 침입을 추론하는 방법이다. 아래는 시나리오를 통한 판단 분석방법이다[5].

O 두 가지 디바이스간의 관계적인 요소를 기반으로 분석
- 네트워 트래픽이 임계치를 초과하고 그 시간대에 트래픽을 유발하는 바이러스가 발견된 경우

O 두 가지 디바이스간의 인과적 요소를 기반으로 분석
- 침입탐지시스템에서 불법접근시도에 관련된 이벤트가 탐지되고 그때 동일한 근원지IP, 목적지 IP와 관련된 방화벽의 이벤트가 거부 또는 허용된 경우

O 세 가지 디바이스간의 인과적 요소 분석
- 침입탐지시스템에서 불법접근시도에 관련된 이벤트가 탐지되고 그때 동일한 근원지IP, 목적지 IP와 관련된 방화벽의 이벤트가 허용되고 있으며 해당 목적지 시스템의 주요 파일에 변경이 발생한 경우

3.4 보안 검증

보안검증단계에서 Alert 축약과정을 거친 경보에 대해서 출발지 IP에 대한 예외처리를 먼저 확인한 다음 목적지IP에 대한 취약점 및 자산정보 (OS, 어플리케이션 정보, 취약점 패치여부)를 연관분석을 통해 대상 서버가 취약점에 영향이 있는지 없는지를 판단한다. 또한, 취약점에 대한 영향이 존재하면 서버의 원시로그와 Server 이벤트설정 정보, 스크립트를 통해 수집된 정보를 통해 침해를 확인한다. 보안검증을 통해서 침해의 위험성이 존재하더라도 서버의 상태나 설정에 따라 침해가 일어날 가능성이 달라진다.

4. 실험 및 결과

본 실험은 공격시나리오를 통해 침해가 발생하였을 때 각 장비별, 제시된 모델단계별로 발생하는 이벤트의 양을 조사하였으며, 특히, 보안검증을 통해 오탐을 최소화하고, 성공한 공격을 판단할 수 있었으며, 침입탐지시스템의 RAW DATA, 서버정보, 시스템 접속로그 등을 통해 공격방법 등에 대해 정확한 분석이 가능하였다.

4.1 실험 구성 및 시나리오

실험구성은 아래의 그림2와 같이 일반적인 네트웍을 구성하였으며, 공격자는 SQL Injection 공격을 통해 서버에 침입한 후 서버에 사용자를 추가하고 데이터를 삭제하는 시나리오를 작성하였다. 해당 이벤트 수집은 이글루시큐리티의 SPiDER-TM을 이용하여 수집하였다.

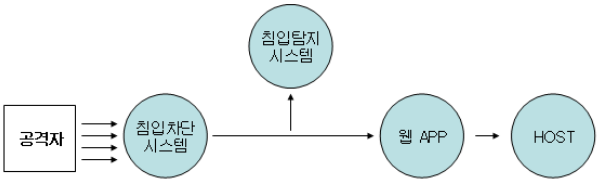

그림 2 공격시나리오 및 구성

4.2 실험 결과

표 3. 로그 & 경보 이벤트 결과

구분	로그	Alert 발생	Alert 축약	Alert 검증
침입차단시스템	12053	130	-	-
침입탐지시스템	150	-	10	4
WEB APP	12053	400	30	2
HOST	30	30	1	1
계	24,286	560	41	7

실험결과 표3와 같이 각각의 보안장비에서 발생하는 이벤트는 모델단계가 증가할수록 감소하고 있음을 알 수 있다.특히, 보안검증 단계에서는 침입탐지시스템의RAW DATA를 통해 웹공격의 POST 로 유입되는 파라미터를 확인할 수 있었으며 이를 통해 SQL Injection 공격을 확신할 수 있었다.

5. 결론 및 향후과제

이벤트 상관분석을 이용한 포렌식은 다양하고 대량의 이벤트와 경보가 발생하는 상황에서 필수적인 부분이고 이를 이용한 제품들도 다양하게 출시되고 있다 하지만, 대부분의 제품들이 침입탐지와 대응에만 초점을 맞추고 있고, 포렌식에 대해서는 미흡한 현실이다 이에 본 논문에서는 확장된 증거수집과 이벤트 상관분석 보안검증을 통해 탐지 및 대응, 분석시간을 최소화 할 수 있는 모델을 제시하였으며, 이를 통해 수집된 이벤트만으로도 신속하고 정확한 포렌식이 가능하였다 또한, 제안된 이벤트를 이용한 포렌식 모델은 전체 네트웍에서 일어나는 침해정보를 실시간으로 수집하고 분석하기 때문에 사건재구성이 용이하며 포렌식에 대한 신뢰성을 높일 수 있다.

향후에는 본 논문에서 제시한 확장된 증거수집에 대한 연구가 더욱 필요하다 해당 모델을 현재 존재하는 통합보안관제시스템(ESM)에 통합하는 방안이 필요하다 또한, 다른 포렌식분야보다 절차나 방법에서 더욱 발전한 시스템포렌식의 기술을 결합하여 분석하는 기술과절차에 대한 연구가 필요하다.

참 고 문 헌

[1] 박종성, 문종섭, "자동화된 침해사고대응시스템에서의 네트웍 포렌식 정보에 대한 정의" 정보처리학회논문지 14(4), pp. 115-126, 2004.

[2] Chen Kevin, Andrew Clark, Olivier De Vel and George Mohay "ECF - Event Correlation for Forensics" In Proceedings of 1st Australian Computer, Network & Information Forensics Conference Perth, Western Australia, 2003

[3] J Abbott, J Bell, A Clark, O De Vel, G Mohay "Automated recognition of event scenarios for digital forensics" Proceedings of the 2006 ACM symposium on Applied computing, 2006 .

[4] 이수형, 방효찬, 장범환, 나중찬 "효과적인 보안상황 분석을 위한 보안이벤트 처리" 전자통신동향분석 22(1), 2007.

[5] 최대수, 이용균 "ESM에서 보안이벤트 분석기술에 관한 연구" 한국컴퓨터종합학술대회 논문집34(1), 2007.

[6] Chr. Kruegel, W. Robertson, and Giovanni Vigna "Alert Verification Determining the Success of Intrusion Attempts" K.G. Saur Verlag, Muchen, 2004.

[7] 이형석, 김수형, 박혁로, 김민수, 노봉남 "공격흐름 파악을 위한 IP기반 상황분석 방법" 전남대학교 일반대학원 2005. 8

[8] Ali Reza Arasteh, Mourad Debbabi "Analyzing multiple logs for forensic evidence" Digital Investigation, 2007

[9] http://www.igloosec.com