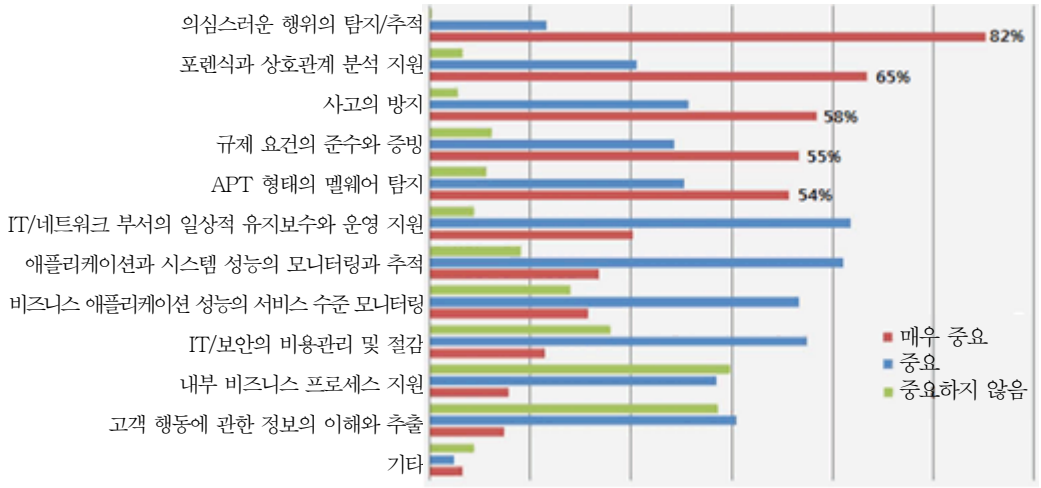
보안 분야의 빅데이터, 로그 분석 시간 부족

○ 빅데이터의 시대에는 축적한 방대한 양의 데이터로 무엇을 할 것인지를 아는 것 자체가 거대한 도전과제인데, 보안의 경우 로그 분석이 이에 해당

- 로그(log) 정보는 정보 시스템을 통과하는 데이터의 흐름을 추적하는 것으로, 다수의 사용자가 실시간으로 이용하는 시스템이 많을수록 방대한 로그 데이터가 기록됨
- 보안장비업체 SANS가 8번째로 발간한 '로그 및 이벤트 관리 서베이' 연간 보고서에 따르면, 기업의 IT 및 보안 부서들은 정상적인 로그 데이터로부터 위협을 초래할 수 있는 이벤트를 분리해 내는 것의 중요성을 인식하고 있음
- 그러나 상대적으로 소수의 기업만 공격자 식별 등의 목적으로 수집한 로그 데이터들을 잘 활용하고 있는 것으로 나타남
- 이번 서베이는 SANS가 전세계 600명 이상의 IT 전문가를 대상으로 실시

○ 로그 데이터를 축적하는 가장 중요한 이유는 내부 및 외부의 보안 이슈와 관련이 있는



<자료>: SANS, 2012. 4.

(그림 1) 로그 데이터를 수집하는 이유

* 본 내용과 관련된 사항은 정보서비스팀(☎ 042-710-1771)과 ㈜크로센트 박종훈 수석 아키텍트(☎ 02-2078-2088)에게 문의하시기 바랍니다.
** 본 내용은 필자의 주관적인 의견이며 NIPA의 공식적인 입장이 아님을 밝힙니다.

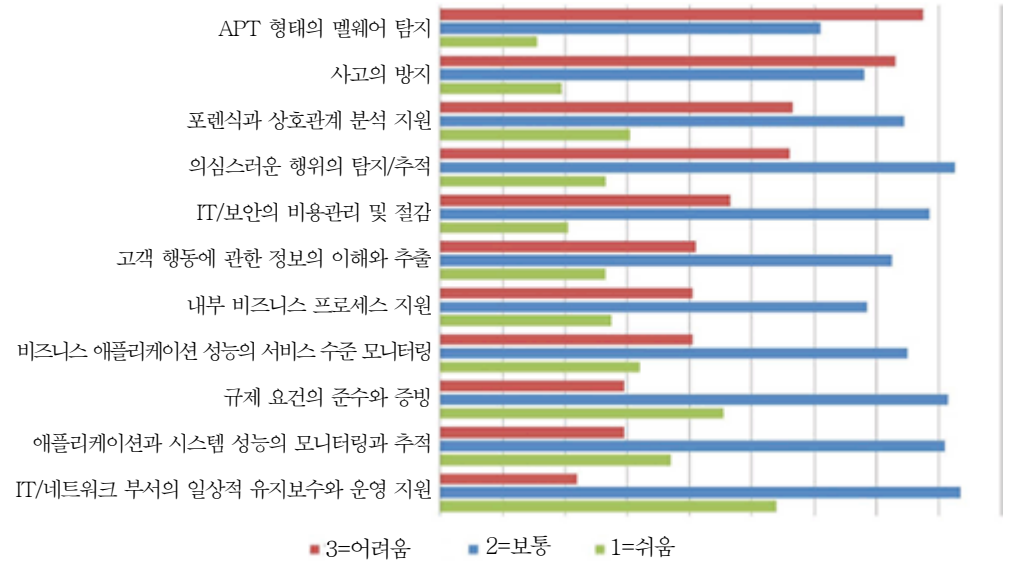
데, 82%가 의심스러운 행위의 탐지와 추적이라고 응답

- 의심스러운 행위란 인가받지 않은 접근과 내부의 데이터 오남용 등을 포함
- 두 번째 중요한 이유로는 포렌식 분석 지원이 꼽혔는데, 디지털 포렌식(forensic)이란 각종 디지털 데이터 및 통화기록, 이메일 접속기록 등의 정보를 수집·분석하여 DNA를 비롯하여 지문·핏자국 등 범행과 관련된 증거를 확보하는 수사기법을 의미

○ 로그 데이터를 수집하는 이유들 중 실제로 목적을 달성하기 가장 어려운 것은 APT 유형의 멀웨어를 탐지하는 것이라는 응답이 83%로 가장 높았음

- APT(Advanced Persistent Threat)는 단발성으로 끝나는 공격이 아니라 공격 대상 네트워크의 깊은 곳까지 침투하여 목적이 달성될 때까지 지속적으로 공격하는 캠페인 유형의 보안 위협을 의미하며, 최근 보안에서 가장 큰 이슈가 되고 있음
- APT 는 외부에 공개된 정보나 이전의 공격에서 얻은 데이터를 바탕으로 하기 때문에 공격 대상에 관해 더 자세히 알수록 공격은 더 정교한 형태로 이루어짐
- 많은 조직이 APT 공격과 싸우고 있으며, 네트워크의 백그라운드 노이즈로부터 위협을 초래할 수 잇는 데이터를 걸러내는 것은 어려운 일이 되고 있음

○ 로그 데이터를 수집하고 분석하는 방법으로는 크게 로그관리시스템을 이용하는 방식과



<자료>: SANS, 2012. 4.

(그림 2) 로그 데이터 수집 이유 중 달성하기 어려운 것

SIEM 을 이용하는 방식으로 구분

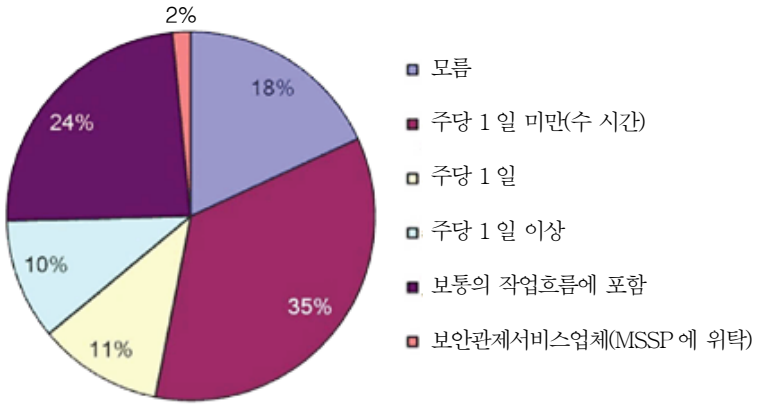
- 데이터를 호스트로부터 직접 수집하거나(19%), Syslog(UDP/TCP)로부터 로그를 수집(25%), 소스로부터 에이전트를 이용해서 데이터를 수집(14%)한 다음 로그 관리기로 로딩하는 등 로그관리시스템을 이용한다는 응답은 58%
- SIEM(Security Information Event Management)를 이용하여 로그 데이터를 수집, 분석하거나(22%), 로그서버 등 다른 수단에 의해 수집된 로그 데이터를 SIEM 을 이용해 분석하는 등 SIEM 을 이용한다는 응답은 37%
- 로그 관리 솔루션 분야가 발전함에 따라 로그 데이터를 관리하는 거의 모든 제품들은 데이터를 추출, 분석하고, 경보를 알려주는 하나 이상의 프로세스를 선탑재하여 출시
- 그럼에도 응답자의 22%는 로그 데이터 수집-분석-경보의 프로세스가 자동화되어 있지 않다고 밝혔으며, 이유는 시간과 비용 등 한정된 가용자원 때문

○ 로그가 수집되는 주요 소스로는 윈도 서버와 방화벽 및 네트워크 장비가 주종을 이루며, 모바일 기기, 빌딩관리 장비, 클라우드 등은 아직 미약한 것으로 나타남

- 일반적으로 기업은 윈도와 유닉스 등의 서버, 보안 기기, 스위치와 라우터, 침입탐지 시스템, 안티바이러스 및 기타 보안 애플리케이션과 같은 네트워크 장비, 가상화된 서버와 하이퍼바이저뿐만 아니라 데스크탑과 노트북 등에서 로그 데이터를 수집
- 스마트폰과 태블릿 등 모바일 기기, HVAC(난방, 환기, 공조) 등 건물과 플랜트 운영을 위한 제어 시스템, 클라우드 기반 서비스와 애플리케이션 등의 소스로부터 로그 데이터를 수집한다는 응답은 10% 미만

○ 보안에 있어 로그 데이터 분석의 중요성을 인식하고 있고, 솔루션도 도입하고 있지만 기업은 아직 로그 분석에 충분한 시간을 투자하는 못하는 것으로 나타남

- 노이즈를 차단하고 필요로 하는 정보를 얻기 위해서는 이벤트 간 상관 관계 분석 기능을 향상시킬 필요가 있는데, 이는 IT 및 보안 관리자들이 먼저 로그에 익숙해지고, 무엇이 정상이고 비정상인지에 대한 기준을 설정할 것을 요구함
- 자동화된 로그 분석 도구들은 변칙적 트래픽과 보안에 관한 육감을 개발하고 있는 로그 분석가들을 완전히 대체할 수 없는 것으로, 역량있는 로그 분석가는 기본적으로 로그 데이터를 바라보는데 매일 상당한 시간을 투입해야 함
- 그러나 IT 전문 인력들이 통상 로그 데이터 분석에 얼마나 많은 시간을 할애하는가라



<자료>: SANS, 2012. 4.

(그림 3) 로그 분석에 투입하는 시간

는 질문에 "주당 몇 시간 안 된다"는 응답이 35%로 가장 높았음

- SANS 은 전반적인 로그 분석 시간이 작년 조사 때보다 낮아졌으며, 소규모 기업의 약 50%는 로그 분석에 시간을 쓰지 않거나 겨우 몇 시간만 투자한다는 사실에 주목

○ 보안 위협이 점차 정교해지고 빈번해짐에 따라 기업이 로그 분석에 사용하는 시간은 점차 증가할 것으로 예상되며, 또 그래야만 할 것임

- 기업이 빠르게 비정상적 행위를 감지하는데 가장 좋은 방법은 정기적으로 로그 데이터에 대한 검토와 분석을 통해 기업 자체적으로 기준점 혹은 정상적 행위에 대한 이해도를 높이는 것
- 빅데이터 시대에 데이터 트래픽이 급증하고 모바일 기기, 위치정보, 센서 정보 등 실시간 데이터 소스가 증가하면서 로그 데이터 분석은 기업이 지속적으로 수행해야 할 보안 활동의 핵심으로 부상하고 있음

(Network World, 5. 3 & Gov Info Security, 5. 5 & Market Watch, 5. 9.)