AORM: 위협 수준 탐지를 위한 다계층 구조 기반 위험 모델 제안

박준영
세종대학교
pjy010218
@sju.ac.kr

장우현
LIG넥스원
woohyun.jang2
@lignex1.com

김연재
LIG넥스원
yeonjae.kim
@lignex1.com

허재원
LIG넥스원
jaewon.huh
@lignex1.com

박기웅†
세종대학교
woongbak
@sejong.ac.kr

AORM: Proposal on a Multi-Layer Risk Model For Threat Level Detection

Junyeong Park Woo Hyun Jang Yeon Jae Kim Jae Won Heo Ki-Woong Park†

1. 서 론

현대의 드론에 대한 위협은 단순 물리적 파괴를 넘어, 적에게 탈취된 후 내부 핵심 기술이 역공학 되는 '탈취 후' 시나리오로 진화하고 있다. 하지만 기존의 안티-탬퍼링(Anti-Tampering) 연구는 주로 시스템 외부적인 침투를 탐지하는 데 중점을 두어, OS가 장악된 이후 시스템 내부에서 발생하는 논리적 공격에는 취약점을 가진다 [1].

이러한 한계를 극복하기 위해, 본 연구는 탬퍼링 상황을 탐지하고 공격을 특정하기 위한, 접근 통제 행렬을 재해석한 2차원 매트릭스 기반의 '행위-객체 위험 매트릭스(Action-Object Risk Matrix, AORM)'을 제안한다. 이 모델은 시스템의 기능적 중요도에 따라 계층화한 '행위' 축과 접근하고자 하는 데이터의 중요도에 따라 계층화한 '객체' 축으로 구성된다. 본 연구에서는 공격자가 낮은 위험도의 좌표에서 더 치명적인 높은 위험도의 좌표로 이동하려는 시도, 즉 전통적 정의의 위협 전이를 위협 수준 격상의 핵심 지표로 정의하면서도, 2차원 매트릭스를 활용하여 단순한 '허용/거부'의 보안 방식이 아닌 '맥락 인식' 위험 모델을 제안한다.

본 모델은 이를 통해 '위협 전이'를 탐지함으로써 단순한 무결성 훼손 여부를 넘어, 공격자의 침투 깊이를 상황 인식적으로 판단할 수 있는 새로운 시각을 제시하며, 드론에만 국한되지 않는 넓은 범용성을 지닌다. 또한, eBPF와 같은 커널 추적 기술을 통해 현대 운영체제에 효율적으로 구현 가능하며, 모든 임베디드 시스템의 위협 탐지에 이용 가능한 보안 패러다임으로 기능할 수 있다.

2. 관련 연구

2.1 이론적 기반

본 모델은 컴퓨터 보안 분야에서 검증된 고전적인 보안 모델들의 어깨 위에 서있다. 본 모델의 구조는 접근 통제 행렬(Access Control Matrix, ACM)에서 비롯된다. AORM이 채택한 객체와 행위는 ACM에서 정의 및 이용한 주제, 객체, 권한의 관계에서 착안했으며, AORM은 '행위'의 위험도를 정량화하는 '동적 매트릭스'로 확장했다는 점에서 차별성을 가진다. 또한, AORM의 '객체' 축은 벨-라파듈라 (Bell-LaPadula, BLP) 모델의 '데이터 등급화' 철학에서 착안했다. BLP 모델이 정보의 기밀성을 보장하기 위한 규칙을 적용한 것과 유사하게, AORM은 각 객체를 기밀성, 무결성, 가용성에 기반하여 중요도를 분류해 공격의 목표가 되는 객체와 그 영향을 체계적으로 평가하는 기반을 마련한다.

3. AORM 모델

AORM은 컴퓨터 시스템 내부의 모든 이벤트를 '행위'와 '객체'의 두 축으로 나누고, 이를 조합해 리스크를 정량화한다. 추상적 구조는 그림 1과 같다.

3.1 분류 체계

AORM은 시스템의 모든 요소를 '역할과 경계' 철학에 따라 범주화한다. 객체는 기밀성, 무결성, 가용성에 끼치는 영향을 기준으로, 행위는 시스템에 침투하는 깊이를 기준으로 각각 4개의 레벨로 분류된다. 객체는 L0(신뢰의 근원), L1(시스템 무결성), L2(운영 및 설정), L3(일반 및 감사), 행위는 L0(커널 조작), L1(특권 명령 실행), L2(일반 상호작용), L3(사용자 공간 실행)로 정의된다.

3.1 리스크 점수 계산 공식

AORM은 분류한 이벤트를 정량적인 수치로 변환하기 위해 아래의 공식을 사용한다. 최종 수치는 리스크 점수로, 공격의 파급효과(Impact, I), 공격 용이성(Exploitability, E), 범위(Scope, S), 그리고 위협 맥락(Threat Context, T)을 종합적으로 고려해 산출된다.

$$AORM\,Score = Roundup[(Base\,Score) \times S_{factor} \times T_{factor}]$$

$$Base\,Score = (w_I \bullet I_{score}) + (w_E \bullet E_{score})$$

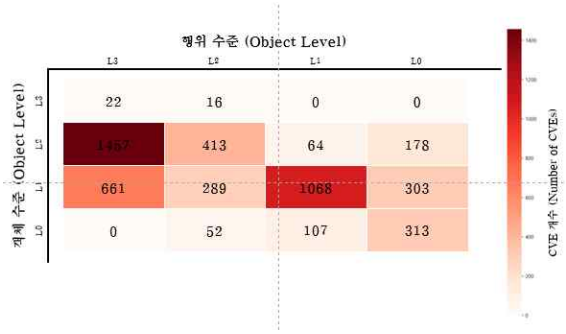
각 변수는 객체의 CIA 등급, 행위의 레벨과 복잡도, 공격 범위 변경 여부, 그리고 알려진 공격 패턴과의 일치 여부에 따라 결정되며, 이는 정성적 위협 평가를 재현 가능한 수치로 변환 가능하게 한다.

4. 실험 및 결과

본 연구에서 제안하는 AORM 모델의 논리적 타당성을

객체 \ 행위	L3 (사용자 공간 실행)	L2 (일반 상호작용)	L1 (특권 명령 실행)	L0 (커널 조작)
L3 (일반 및 감사)	Low	Low	Medium	High
L2 (운영 및 설정)	Low	Medium	High	Critical
L1 (시스템 무결성)	Medium	High	Critical	Critical
L0 (신뢰의 근원)	High	Criticai	Criticai	Critical

(그림 1) AORM 모델



(그림 2) AORM 리스크 점수 히트맵

검증하기 위해, 약 7,600개의 CVE에 대해 AORM 리스크 점수를 계산했다. 해당 실험의 결과는 그림 2에서 볼 수 있듯이, 실제 취약점들이 AORM 매트릭스 전반에 걸쳐 불균등하게 분포하며, 특히 [L1, L1], [L3, L2] 등 특정한 영역에 집중 분포되는 경향을 보인다. 이는 시스템 무결성을 해치는 행위와, 낮은 권한으로 시스템 설정을 변경하려는 시도가 현실 세계에서의 빈번한 공격 방식임을 의미하며, 실례로 CVE-2018-0830의 경우 [L1, L1]에 분포하고 CVSS 점수 4.7와 AORM 리스크 점수 91을 가지는데, CVE-2018-0830은 2019년 CWE에서 발표한 '가장 위험한 소프트웨어 약점 25개' 중 4위인 CWE-200, 정보 노출 (Information Exposure) 약점을 악용한 공격으로, 이는 AORM의 위험도 정의가 현실 세계의 위협 동향과 일치함을 증명하는 사례이다.

반면, [L3, L0]나 [L0, L3]와 같이 CVE가 존재하지 않는 셀도 존재하는데, 이는 오히려 모델의 타당성을 뒷받침하는 근거이다. [L3, L0]은 현대 운영체제의 구조상 거의 불가능한 공격 행위이며, [L0, L3]는 이미 최고 권한을 획득한 공격자가 일반 객체를 공격에 활용하는 비합리적인 공격임을 나타내어, AORM 모델이 시스템과 운영체제의 보안 구조적 제한사항과 공격자의 합리적 행위를 모두 반영하고 있음을 보여준다. 더 나아가, 각 CVE별로 산출한 AORM 점수와 기존 CVE의 CVSS 점수 간의 상관관계를 분석한 결과, 전반적인 경향성이 비슷하게 나타나 상관관계가 분명하게 존재하지만, 주목할 만한 불일치 사례들 또한 존재한다.

높은 CVSS 점수에도 불구하고 낮은 AORM 점수를 보이는 CVE 사례는 원격에서 쉽게 악용 가능하나 시스템 파괴력이 제한적임을 의미하고, 낮은 CVSS 점수와 높은 AORM 점수를 가진 CVE는 악용 조건이 까다롭지만 시스템 파괴력이 충분한 CVE를 의미한다. CVE-2018-0830의 예시처럼, CVSS 점수는 낮게 산정되나 AORM 리스크 점수로 해당 공격의 파괴력을 산출할 수 있다. 결론적으로, AORM은 위협 분포를 정확히 모델링하는 논리적 타당성을 가질 뿐 아니라, CVSS와 같은 기존의 정적 평가 방식이 고려하지 못하는 '위협 전이성', '시스템 파괴력' 등의 동적 위험성을 정량적으로 평가하는 프레임워크임을 증명한다.

5. 고 찰

본 연구에서 제안하는 AORM 모델은 CVSS나 STRIDE와 같은 기존 보안 프레임워크를 대체하기 위함이 아니며, CVSS나 STRIDE에서 다루지 않는 '시스템 내 동적 위협'의 위험도를 정량적으로 평가하는 새로운 관점을 제시하는 모델이다. CVE 적용 실험의 분석 결과에서 나타난 CVSS와 AORM 점수 간 불일치는, '진입 용이성'만으로는 올바르게 평가할 수 없는 공격 행위의 '시스템 파괴력'이라는 위협을 평가할 수 있는 AORM 모델의 필요성을 증명한다.

하지만, 본 연구는 연구에 사용된 MITRE ATT&CK 매핑 데이터의 완성도에 의존하는 한계를 가지며, 리스크 점수 공식의 가중치를 운영 환경에 맞춰 최적화할 필요가 있을 수 있다. 그럼에도 불구하고, AORM은 산업 표준인 ATT&CK Matrix를 연결하고, CVE 데이터로 실증했다는 점에서 중요한 의의를 가진다.

(표 1) AORM 리스크 점수 산출 결과 일부

CVE_ID	CVSS Score	AORM Score	Critical_TTP
CVE-2018-0884	7.8	91	T1055.011
CVE-2018-0772	7.5	91	T1055.011
CVE-2018-0827	5.3	91	T1055.011
CVE-2018-0830	4.7	91	T1055.011
CVE-2019-1227	2.1	91	T1055.011
CVE-2020-15157	6.1	80	T1552.007
CVE-2020-14336	6.5	61	T1053.007
CVE-2020-14370	5.3	99	T1543.005

6. 결 론

본 연구는 드론 탈취 후 탬퍼링 상황을 효과적으로 탐지하기 위해 새로운 동적 위협 평가 프레임워크인 AORM을 제안했다. AORM은 시스템 공격 행위를 '행위'와 공격의 대상이 되는 '객체'의 2차원 매트릭스로 분석하고, 이를 정량적 수치인 리스크 점수로 산출함으로써, 기존 보안 모델과 다른 '시스템 파괴력'이라는 새로운 위험 차원을 측정한다. 본 연구에서는 MITRE ATT&CK 프레임워크와의 매핑 및 7천 여개의 CVE 데이터를 이용한 실험을 통해, AORM이 현실 세계의 위협 동향을 반영할 수 있고, CVSS와 상호 보완적인 가치를 가지고 있음을 증명했다. 향후 연구로 AORM을 기반으로 한 모니터링 에이전트를 개발해 실제 임베디드 시스템에서의 효고를 검증할 예정이다.

Acknowledgement

이 논문은 2022년 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(KRIT-CT-22-051)

참 고 문 헌

[1] Chae, H., Lee, C.S., & Kim, T.H. (2018). The Anti-tampering Process and Case Study by the Operating Mode of Various Unmanned Ground Vehicles. 2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 1432-1437.

[2] Strom, B.E., Applebaum, A., Miller, D., Nickels, K.C., Pennington, A.G., & Thomas, C. (2020). MITRE ATT&CK ® : Design and Philosophy.

[3] Lampson, Butler. (1974). Protection. ACM SIGOPS Operating Systems Review. 8. 18-24. 10.1145/775265.775268.