TTP 시나리오 예측 기반 멀티 클라우드 위협 대응을 위한 Genetic Algorithm 활용

손우영, 김세훈, 홍찬희, 김홍현, 최상훈, 박기웅

세종대학교 정보보호학과

wooyoung@pel.sejong.ac.kr, sehoon518@gmail.com, ghdcksgml220@gmail.com,
khh010203@gmail.com, csh0052@gmail.com, woongbak@sejong.ac.kr

Utilizing Genetic Algorithms for Proactive Threat Response in Multi-Cloud Environments Based on TTP Scenario Prediction

Wooyoung Son, Sehoon Kim, Chanhee Hong, Honghyeon Kim,

Sang-Hoon Choi, Ki-Woong Park

Dept. of Computer and Information Security, Sejong University

요 약

멀티 클라우드 환경의 수요가 증가됨에 따라, 해당 환경에서 발생하는 보안 위협을 분석하는 것을 넘어, 향후 발생할 수 있는 공격 시나리오를 예측하고, 이에 선제적으로 대응할 수 있는 통합 보안 시스템의 필요성이 증대되고 있다. 이에 본 논문에서는 로그 데이터를 기반으로 MITRE ATT&CK 프레임워크의 TTP (Tactics, Techniques and Procedures)와 매핑을 수행하고, TTP의 흐름을 분석하여 향후 발생 가능한 공격 시나리오를 예측한다. 또한, 이를 바탕으로 Genetic Algorithm을 활용함으로써 능동적인 선제적 대응을 수행할 수 있도록 하는 전략을 제공함과 동시에 반복적 진화를 통해 정교하고 다양한 전략을 도출하는 멀티 클라우드 보안 시스템을 제안한다.

Ⅰ. 서 론

클라우드 기술의 발전은 사용자가 필요로 하는 자원을 신속하고 유연하게 제공해줌과 동시에 언제 어디서든 작업을 수행할 수 있는 환경을 제공하였다. 이에 따라, 많은 조직들은 다양한 클라우드 플랫폼을 활용하여 운영 효율을 높이고 있는 실정이다. 하지만, 이러한 멀티 클라우드 환경으로의 전환은 공격 표면을 넓혀, 보안 위협이 발생할 가능성을 높이고 있다 [1].

이러한 환경에서는 서로 다른 체계를 가진 클라우드 서비스들을 포괄적으로 관리하고 보호할 수 있는 통합 보안 시스템의 구축이 필수적이다. 이와 더불어 대부분의 보안 관제 시스템은 클라우드 서비스를 모니터링하여 이상 행위에 대한 탐지에 집중하고 있으나, 이는 여전히 보안위협이 발생할 가능성을 열어두고 있어 모니터링된 정보에 기반하여 향후 발생 가능한 공격 시나리오를 예측하고, 선제적으로 대응할 수 있는 능력이 요구되고 있다.

이에 따라 본 논문에서는 TTP (Tactics, Techniques and Procedures) 시나리오 예측을 기반으로 멀티 클라우드 환경에서의 위협에 대응하기 위한 보안 시스템을 제안하며, 제안된 시스템의 경우, Genetic Algorithm (GA)을 활용함으로써 능동적인 선제적 대응을 수행할 수 있음을 보인다. 제안된 시스템은 멀티 클라우드 환경에서 수집된 로그들을 기반으로 MITRE ATT&CK 프레임워크의 TTP와 매핑을 수행하고, 시간 순으로 나열한 매핑된 TTP 흐름을 분석함으로써 향후 수행될 가능성이 높은 공격 시나리오를 예측한다. 또한, 예측된 시나리오를 바탕으로 선제적 대응을 수행할 수 있는 전략을 제공하며, GA를 활용함으로써 사전에 정의되지 않은 상황에서도 최적의 대응 전략을 도출하고, 반복적인 진화 과정을 통해 점차 정교하고 다양한 전략을 생성한다.

본 논문의 구성은 다음과 같다. 2장에서는 제안된 시스템에서의 핵심 기술로 사용되는 GA에 대해서 설명하며, 3장에서는 제안하는 멀티 클라우드 보안 시스템에 대하여 주요 프레임워크를 기반으로 설명한다. 4장에서는 GA를 활용함에 따라 사용자에게 전달할 대응 전략의 조합의 확장성에 대하여 수식을 기반으로 제시하며, 5장에서는 본 논문의 결론을 맺는다.

Ⅱ. Genetic Algorithm

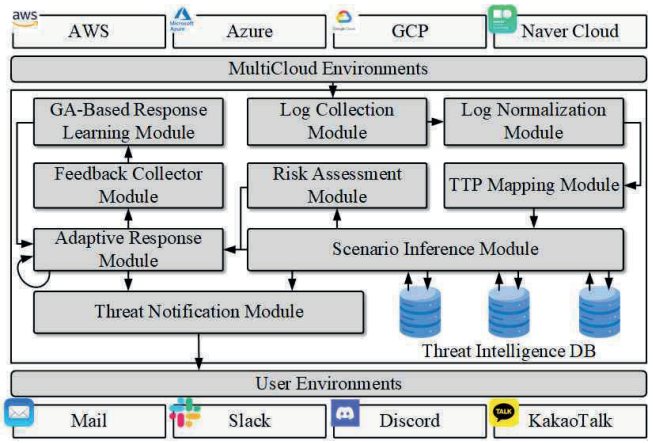
Genetic Algorithm은 John Holland에 의해 제안된 생물학적 진화 개념을 기반으로 하는 전역 최적화 기법이다 [2]. GA는 초기 해 집합인 Population을 구성하고, 각 해에 대한 Fitness를 계산하며, Selection, Crossover, Mutation의 유전 연산을 반복 적용함으로써 세대를 지날수록 더욱 적합한 해를 진화적으로 도출해낸다. Alhijawi은 GA가 다양한 문제에서 효과적으로 활용되고 있으며, 특히 탐지 및 식별 영역에서 사이버 보안에 폭넓게 적용되고 있다고 평가하였다 [3].

그러나 GA를 활용한 완화 및 대응 전략을 생성하고 학습하는 연구는 상대적으로 부족한 실정이다. 이에 본 논문에서는 GA를 활용하여 새로운 대응 방안을 설계하고 발전시키는 구조를 제안하며, 이를 통해 정적 Rule-Based 방식의 한계를 극복하고 멀티 클라우드 환경에서의 발생 가능한 보안 위협에 대한 능동적 대응 체계를 제안하고자 한다.

Ⅲ. 제안하는 멀티 클라우드 보안 시스템

기존 클라우드 보안 시스템의 경우, 단일 클라우드 서비스만을 타겟으로 하고 있거나, 이미 발생한 공격에 대한 사후 탐지에만 초점을 맞추고 있다. 따라서, 향후 발생 가능한 공격 시나리오를 예측하거나, 이에 대한 대응 방안은 제공하고 있지 않는다는 한계점을 지닌다. 이러한 한계점을 보완하기 위해, 본 장에서는 식별된 공격 시도 데이터를 기반으로 향후 공격 시나리오를 예측하고, 이에 대한 대응 방안을 제공하는 멀티 클라우드 보안 시스템을 제안한다. 제안된 시스템의 경우, (그림 1)을 통해 확인할 수 있다.

먼저, AWS, Azure 등 다양한 클라우드 환경으로부터 수집된 로그들의 경우, 로그 형식 및 로그 데이터 내 필드가 상이함에 따라, 로그 정규화 과정을 통해 통합된 형태로 변환된다. 정규화된 로그는 TTP Mapping Module에서 MITRE ATT&CK 프레임워크의 TTP 항목과 매핑되며, 이를 통해 도출된 결과를 시간 순서대로 나열한 후, Scenario Inference Module로 전송된다.



(그림 1) 제안하는 멀티 클라우드 보안 시스템

Scenario Inference Module은 TTP 간 전술 흐름, 기법 간 연관성, 과거 유사 사례 등을 Threat Intelligence DB로부터 수신하여 이를 기반으로 공격 시나리오를 구성한다. 구성된 시나리오는 Risk Assessment Module에서 위험도를 평가받고, 이 정보는 Adaptive Response Module로 전달된다. Adaptive Response Module은 Rule-Based 방식과 GA를 결합한 Hybrid 구조를 채택한다. 이에 따라, 정의되어 있는 공격 예측 시나리오가 도출된 경우, 사전에 정의된 대응 전략을 제공하고, 사전에 정의되지 않은 새로운 예측 시나리오에 대해서는 GA를 통해 최적의 대응 전략을 도출하여 제공한다.

대응 결과는 Feedback Collector Module에 의해 수집되며, 이때 수집되는 데이터로는 제공된 대응 전략 수행 시 발생된 대응 성공률, 대응 실행 지연 시간 등이 포함된다. 이렇게 수집된 피드백은 GA-Based Response Learing Module에게 전달되어 적합도 함수 계산 및 다음 세대 전략 생성에 활용된다. Adaptive Response Module에서 제공하는 대응 전략의 경우, Risk Assessment Module에서 계산된 위험도 점수가 임계값 이상일 때만 사용자에게 전달되며, 그렇지 않은 경우에는 Threat Notification Module을 통해 예측된 공격 시나리오만이 사용자에게 전달된다.

IV. Genetic Algorithm 기반 전략 조합 확장성 분석

3장에서 제시한 바와 같이, 제안된 멀티 클라우드 보안 시스템 내 Adaptive Response Module의 경우, 기존 Rule-Based 방식과 GA를 결합한 Hybrid 방식을 채택하고 있으며, 특히, GA 기반 대응의 경우, 반복적 진화를 통해 점차 정교하고 다양한 전략을 생성할 수 있다는 장점을 지닌다. 이에 따라 본 장에서는 제안된 Hybrid 방식의 위협 대응 구조가 제안된 시스템의 대응 확장성을 높임을 수식을 기반으로 제시한다.

기존 Rule-Based 방식의 경우, 사전에 정의한 대응 전략 집합을 기반으로 함에 따라, 대응 전략 공간이 고정되어 있다는 한계점을 지닌다. 도출된 공격 시나리오 내 향후 수행될 것으로 예측된 TTP 수를 n_r, 각 TTP i에 대한 Rule-Based 대응 전략 수를 r_i라고 하는 경우, Rule-Based 방식의 전체 대응 전략 조합 수 N_{Rule}은 수식 (1)과 같이 표현할 수 있다.

$$N_{Rule} = \prod_{i=1}^{n_r} r_i \quad (1)$$

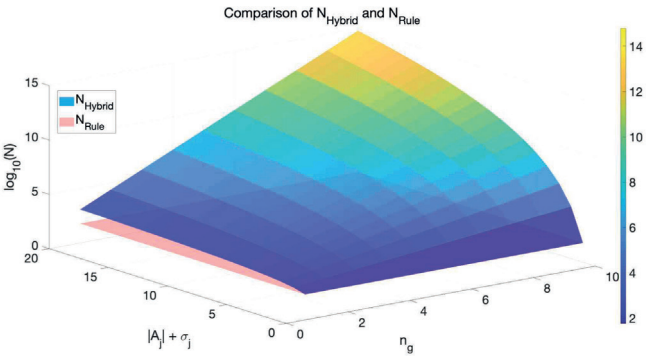
반면, 제안된 Hybrid 방식의 경우, TTP j에 대한 초기 대응 집합 A_j 외에도, Mutation 연산을 통해 생성되는 추가 후보군 σ_j를 포함함으로써 대응 전략 공간을 확장한다. n_g를 GA 기반 방식이 적용된 TTP의 수로 정의하면, 최종적으로 Hybrid 구조의 전체 대응 전략 조합 수 N_{Hybrid}는 수식 (2)와 같이 도출된다.

$$N_{Hybrid} = N_{Rule} \times N_{GA} = \left(\prod_{i=1}^{n_r} r_i\right) \cdot \left(\prod_{j=1}^{n_g} (|A_j| + \sigma_j)\right) \quad (2)$$

N_{Hybrid}가 수식 (2)와 같이 도출되는 이유는, 하나의 시나리오 내에서 일부 TTP는 Rule-Based 방식을 기반으로, 나머지는 GA 방식을 기반으로 대응 전략을 선택하는 경우가 존재하기 때문이다.

예를 들어, 모든 TTP에 대하여 Rule-Based 대응 전략 수가 3개 (r_i=3), GA 기반 초기 전략이 3개($|A_j|$=3)이며, Mutation을 통해 추가로 3개(σ_j=3)가 생성되는 경우, N_{Rule}는 3^{n_r}, N_{Hybrid}는 $3^{n_r} \cdot 6^{n_g}$의 대응 전략 공간을 형성한다. 이는 $|A_j|, \sigma_j > 0$의 조건에서 항상 $N_{Hybrid} > N_{Rule}$이 성립함을 의미한다.

(그림 2)는 n_g와 $(|A_j| + \sigma_j)$ 변화에 따른 N_{Hybrid}와 정적인 값을 갖는 N_{Rule}을 비교한 로그 스케일 그래프이다. 특히 n_g가 증가할수록 N_{Hybrid}의 크기가 민감하게 기하급수적으로 커짐을 확인할 수 있다.



(그림 2) Hybrid vs. Rule-Based 전략 공간 비교 그래프

풍부한 대응 전략 공간이 존재함은, Hybrid 방식이 특히 알려지지 않은 위협 또는 복합적 시나리오에 대하여 더욱 적절한 대응 전략을 생성할 수 있다는 가능성을 보여준다. 또한, 이는 단순한 대응 전략의 공간 확장을 넘어, 최적 대응 전략의 선택 가능성을 높이고, 결과적으로 시스템의 보안성을 강화하는 기반이 될 것으로 사료된다.

V. 결론

본 논문에서는 TTP 시나리오 예측을 통해 보안 위협에 선제적으로 대응할 수 있는 멀티 클라우드 보안 시스템을 제안하였다. 제안한 시스템은 GA를 활용하여 사전에 정의되지 않은 상황에서도 반복적 진화 과정을 통해 다양하고 정교한 대응 전략이 도출될 수 있도록 하였다는 것에 의의가 있다. 하지만, GA는 복잡한 보안 위협 시나리오를 식별하고, 이를 예측할 시, path explosion 혹은 combination explosion 문제가 발생할 수 있다. 이에 향후 연구로는 GA와 concoliec execution을 기반으로 한 Hybrid 모델을 구성하여 복잡한 공격 시나리오에 대해서도 유기적으로 예측하여 대응할 수 있는 멀티 클라우드 보안 시스템을 구성하고자 한다.

ACKNOWLEDGMENT

한국연구재단(NRF) 중견후속연구사업(Project No. RS-2023-00208460, 100%)의 지원을 받아 수행된 연구임.

참 고 문 헌

[1] "멀티 클라우드 자원을 보호하는 방법", [Online]. Available: https://cloud.google.com/blog/ko/topics/threat-intelligence/protecting-multi-cloud-resources-modern-cyberattacks [Accessed: May. 5, 2025].

[2] S. Katoch et al., "A review on genetic algorithm: past, present, and future", Multimedia Tools Appl., vol. 80, pp. 8091 - 8126, 2021.

[3] B. Alhijawi and A. Awajan, "Genetic algorithms: Theory, genetic operators, solutions, and applications", Evolutionary Intelligence, vol. 17, no. 3, pp. 1245-1256, 2024.