순차 패턴 알고리즘을 사용한 SCADA 시스템의 통신 특성 반영에 관한 연구

정우석, Baraka D. Sija, 박지태, 김명섭

고려대학교

{hary5832, sijabarakajia25, pjj5846, tmskim}@korea.ac.kr,

A Study on Reflecting the Communication Characteristics of SCADA System by Using Frequent Sequential Pattern Mining Algorithm

Woo-Suk Jung, Baraka D. Sija, Jee-Tae Park, Myung-Sup Kim

Korea Univ.

요 약

SCADA 시스템은 정해진 기기들이 제한된 통신만을 반복적이고 주기적으로 사용하며 동작한다. 이러한 특징 때문에 SCADA 네트워크에서 화이트리스트기반 보안기법이 많이 사용되고있다. 하지만 현재의 화이트리스트 모델은 표현력이 너무 단순하여 SCADA 시스템의 다양성을 반영하지 못한다는 단점이 있다. 본 논문에서는 현재의 화이트리스트 기반 ACL(Access Control List)에 SCADA 시스템의 기계적이고 반복적인 통신특성을 반영하기 위하여 통신 간의 고정된 발생순서 정보를 확장한 화이트리스트 모델을 제안한다.

1. 서론

SCADA 시스템은 폐쇄망에서 비공개 프로토콜을 사용해 동작하였기 때문에 최근까지 보안과 관련한 위협이 크지 않았다. 하지만 최근 비즈니스 시스템과의 통합으로 인해 네트워크가 확산됨에 따라 폐쇄망과 업무망 그리고 인터넷망 간의 연결이 증가하였으며, 외부 협력업체와의 협업으로 인한 연결성과 다양한 매체의 이용 등으로 인한 보안상 취약점들이 증가하고 있는 추세이다. 이러한 환경에서 리소스가 적고 시스템 동작 패턴 또는 네트워크 통신 트래픽이 규칙적인 제어시스템 환경에서 보안을 담보할 수 있는 효율적 방안으로 주목 받고 있다[3].

SCADA 네트워크는 정해진 기기들이 제한된 통신만을 반복적이고 주기적으로 사용하며 동작한다. 하지만 현재 SCADA 네트워크에서의 ACL 자동 생성[2][3]에 관한 연구를 통해 생성된 ACL 은 단순한 5-tuple 정보만을 포함하므로 SCADA 시스템의 기계적이고 반복적인 통신특성을 반영하지 못한다는 한계점을 가진다[1]. 또한 ANY-ANY 같은 규칙이 많이 생성된다는 점과 한 번만 발생한 통신이라도 규칙으로 생성한다는 등의 단점이 존재한다

본 논문에서는 위에서 설명한 ACL 의 한계점을극복하고 SCADA 시스템의 기계적이고 반복적인 통신특성을 반영하기 위해서 ACL 들의 발생 순서를 추출하고, 추출된 발생 순서 정보를 ACL 에 추가하여 SCADA 시스템의 기계적 통신 특성이 반영된 ACL 를 생성한다.

2 장에서는 순차 패턴 알고리즘을 이용한 발생 순서 추출 방법에 대해 서술하고, 3 장에서는 결론 및 향후 연구에 대해 서술한다.

이 논문은 2015 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업(No.2015R1D1A3A01018057)과 2016 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 과학기술인문융합연구사업임(No.NRF-2016M3C1B6929228).

2. 순차 패턴 알고리즘

Frequent Pattern Mining 은 시간이나 다른 sequence 들과의 연관관계가 있는 빈번히 발생하는 패턴을 찾아내는 데이터 마이닝 분야이다. 순차 패턴 알고리즘은 크게 Apriori 기반 알고리즘과 Pattern-Growth 기반 알고리즘으로 나뉘는데 Pattern-Growth 알고리즘은 분할 정복 방법을 사용함으로써 Apriori 기반 알고리즘보다 데이터 세트를 줄여 효과적으로 탐색 공간을 줄임으로써 높은 효율성과 확장성을 보장하기때문에 본 논문에서는 Pattern-Growth 기반의 PrefixSpan 알고리즘을 사용한다. PrefixSpan 알고리즘의 적용을 통해 정해진 기기들이 제한된 통신만을 반복적이고 주기적으로 사용하는 SCADA 네트워크의 특성을 반영할 수 있는 ACL 들 간의 발생 순서 정보를 추출할 수 있다.

그림 1 은 ACL 들 간의 순서 정보 추출의 개념을 도식화한 것이다. 왼쪽의 여섯 개의 육각형의 색은 각각 하나의 ACL 을 의미하고, 육각형의 형태는 해당 ACL 의 사용 시간을 의미한다. 그림의 여섯 개의 ACL 에 순차 패턴 알고리즘 적용을 통해 주기성과 발생 순서 정보 그리고 서버 클라이언트 정보 등을 파악한다.

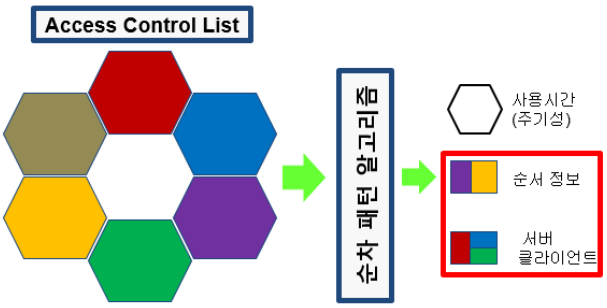
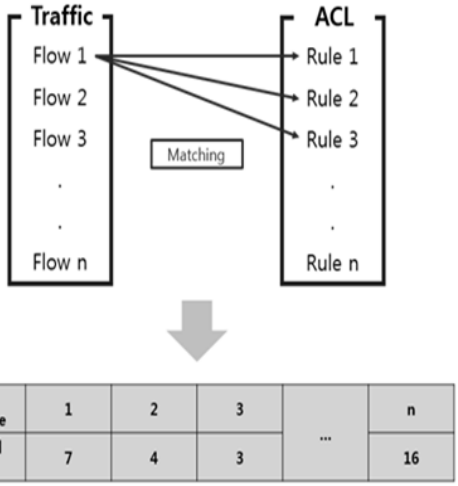


그림 1. ACL 들 간의 발생 순서 정보 추출 개념

3. PrefixSpan 알고리즘을 이용한 발생 순서 추출



Flow sequence	1	2	3	...	n
Matched Rule	7	4	3		16

그림 2. 트래픽에 ACL 을 매칭하는 과정

본 논문에서 제안하는 화이트리스트 모델에 확장될 ACL 들의 발생 순서 정보 추출은 세 단계로 이루어진다.

첫 번째 단계에서는 발생 순서정보를 추출할 ACL 에 각각 고유번호를 부여한다. 이는 이후 Transaction 을 생성 과정에서 ACL 을 트래픽에 매칭하여 매칭 정보를 추출하기 위해 사용된다.

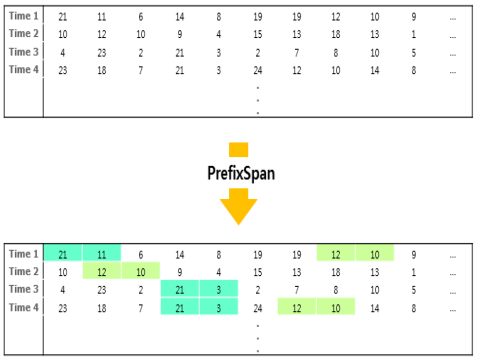


그림 3. PrefixSpan 알고리즘을 통한

두 번째 단계는 Transaction 생성 단계이다. 본 단계에서는 ACL 과 해당 ACL 을 생성하는데 사용한 트래픽을 사용하여 Transaction 을 생성한다. Transaction 은 그림 2 와 같이 트래픽을 플로우 단위로 ACL 에 매칭시키고, 매칭되는 ACL 의 고유 번호를 플로우의 시간 순서대로 나열한 것이다. 그림 2 는 연관관계 추출의 두 번째 단계를 도식화 한 것이다. 입력 트래픽의 각 플로우에 ACL 을 매칭하여 매칭된 ACL 의 고유번호를 시간 순서대로 나열한다.

세 번째 단계는 순서 정보 추출 단계이다. 본 단계에서는 순차 패턴 알고리즘인 PrefixSpan 알고리즘을 사용하여 Transaction 의 ACL 매칭 발생 순서 정보를 추출한다. 그림 3 은 PrefixSpan 알고리즘을 사용하여 Transaction 으로부터 ACL 매칭 순서 정보를 추출하는 과정을 도식화한 것이다. 21 번 ACL 다음에는 11 번 혹은 3 번 ACL 이 매칭되고, 12 번 ACL 다음에는 10 번 ACL 이 매칭되는 것을 확인 할 수 있다. 그림 4 는 12 번 ACL 다음에 10 번 ACL 이 매칭된다는 발생 순서 정보 추가하여 확장한 화이트리스트 모델 예시이다. 해당 모델은 독립적인 12 번과 10 번 ACL 을 발생 순서 정보 추출을 통하여 하나의 계층적 구조를 가지는 화이트리스트 모델로 구성되었다. 그림 4 의 모델은 계층적 구조를 통하여 12 번 ACL 이 매칭된 경우에만 10 번 ACL 을 매칭하므로 방화벽의 불필요한 탐색 공간을 줄일 수 있다. 또한 각 ACL 의 Max duration 정보를 포함하여 해당 ACL 이 매칭 된 경우에 Max duration 시간만 해당 규칙을 열어주면 되므로 모든 규칙들을 항시 열어야 하는 위험을 피할 수 있다.

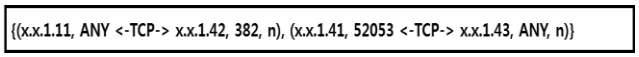


그림 4. 제안하는 화이트리스트 모델 예시

3. 결론 및 향후연구

본 논문에서는 SCADA 네트워크의 화이트리스트에 ACL 들 간의 발생 순서 정보를 추가한 화이트리스트 모델을 제안한다. ACL 들 간의 발생 순서 정보는 순차 패턴 알고리즘의 하나인 PrefixSpan 알고리즘 적용을 통해 추출하였으며, 발생 순서 정보를 추가함으로써 기존 화이트리스트가 가지는 표현력의 한계를 확장하는 효과를 기대할 수 있다. 또한 발생 순서 정보를 통해 SCADA 시스템의 통신 특성을 반영하였다.

향후 본 논문에서 제안한 화이트리스트 모델을 확장하여 계층적 구조를 가지는 화이트리스트 모델에 관한 연구를 진행할 예정이다. 또한, 화이트리스트를 자동으로 생성, 유지 및 보수하여 화이트리스트를 항상 최신의 것으로 유지할 수 있는 방법에 대한 연구를 할 계획이다.

참 고 문 헌

[1] Jung, Woo-suk, et al. "Whitelist representation for FTP service in SCADA system by using structured ACL model." Network Operations and Management Symposium (APNOMS), 2016 18th Asia-Pacific. IEEE, 2016.

[2] 유형욱, 윤정한, 손태식. "제어시스템 보안을 위한 whitelist 기반 이상징후 탐지 기법." 한국통신학회논문지 38.8 (2013): 641-653

[3] Barbosa, Rafael Ramos Regis, Ramin Sadre, and Aiko Pras. "Flow whitelisting in SCADA networks." International journal of critical infrastructure protection 6.3 (2013): 150-158.