ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
논문번호 22-03-08
j.inst.Korean.electr.electron.eng.Vol.26,No.3,389~395,September 2022
https://doi.org/10.7471/ikeee.2022.26.3.389
57

사이버 물리 전력 시스템에 대한 허위 데이터 주입 공격에 관한 고찰

An Overview of False Data Injection Attack Against Cyber Physical Power System

배 준 형*★

Junhyung Bae*★

Abstract

With the evolution of technology, cyber physical systems (CPSs) are being upgraded, and new types of cyber attacks are being discovered accordingly. There are many forms of cyber attack, and all cyber attacks are made to manipulate the target systems. A representative system among cyber physical systems is a cyber physical power system (CPPS), that is, a smart grid. Smart grid is a new type of power system that provides reliable, safe, and efficient energy transmission and distribution. In this paper, specific types of cyber attacks well known as false data injection attacks targeting state estimation and energy distribution of smart grid, and protection strategies for defense of these attacks and dynamic monitoring for detection are described.

요 약

기술의 진화와 함께, 사이버 물리 시스템(Cyber Physical System)은 향상되고 있고 이에 따라 새로운 유형의 사이버 공격도 발견되고 있다. 사이버 공격에는 여러 가지 형태가 있으며 모든 사이버 공격은 대상 시스템을 조작하기 위해 이루어진다. 사이버 물리 시스템 중 대표적인 시스템이 사이버 물리 전력 시스템, 즉 스마트 그리드이다. 스마트 그리드는 신뢰할 수 있고 안전하며 효율적인 에너지 전송 및 분배를 제공하는 새로운 유형의 전력망이다. 본 논문에서는 스마트 그리드의 상태 추정과 에너지 분배를 타깃으로 하는 허위 데이터 주입 공격(False Data Injection Attack)으로 잘 알려진 특정 유형의 사이버 공격 구성 방법과 이러한 공격의 방어를 위한 보호 전략과 탐지를 위한 동적 모니터링 기법을 소개한다.

Key words : cyber physical system, smart grid, state estimation, bad data detection, false data injection attack

Ⅰ. 서론

사이버 물리 시스템(Cyber Physical System, CPS)의 설계는 컴퓨팅 및 통신 기능을 물리적 세계의 실체에 대한 모니터링 및 제어와 통합하는 것을 말한다[1]. 기존의 임베디드 시스템과 달리, CPS는 지능형 컴퓨팅 코어에 의해 통합, 모니터링 및 제어되는 물리적 시스템이다. 스마트 그리드, 프로세스 제어 시스템 및 운송 시스템을 포함한 다수의 CPS는 첨단 컴퓨팅 및 통신 기술을 사용하여 개발될 것으로 예상된다. 스마트 그리드는 대표적인

* School of Electronic and Electrical Engineering, Daegu Catholic University
★ Corresponding author
E-mail : baejh80@cu.ac.kr, Tel : +82-53-850-2764
Manuscript received Aug. 23, 2022; revised Sep. 5, 2022; accepted Sep. 13, 2022.

전기에너지 기반 CPS로 물리적인 전력 송전 시스템과 컴퓨팅 및 통신 네트워크의 사이버 프로세스를 통합한다.

허위 데이터 주입 공격은 사이버 물리 시스템의 안전 및 보안에 대한 새롭고 강력한 공격 부류이다[2]. 공격자의 목표는 시스템 자체를 공격하지 않고 시스템이 잘못된 결정을 내리도록 유도하기 위해 시스템에 허위 입력 데이터를 주입하는 것이다. 공격자는 일반적으로 사이버 물리 시스템의 물리적 플랜트의 파라미터를 측정하는 센서를 공격하는 사이버 공격을 구현한다.

그림 1은 제어의 관점에서, 즉 플랜트를 제어하고 관리하는 제어 루프로서 사이버 물리 시스템 모델을 보여주고 있다. 루프의 모든 구성 요소와 연결에 대해 서로 다른 유형의 네트워크 공격을 개시할 수 있지만, 그림 1에 표시된 것처럼 센서에 대해 허위 데이터 주입 공격이 실행된다. 시스템에 허위 데이터를 측정값으로 삽입하면 시스템이 잘못된 결정을 내리고 오도된 조치를 취할 수 있다.

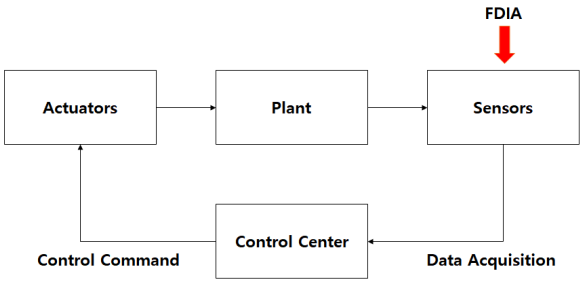

Fig. 1. Control Block Diagram of Cyber Physical System.
그림 1. 사이버 물리 시스템의 제어 블록도

센서가 일반적으로 현장에 있다는 점을 고려할 때 허위 데이터 주입 공격의 위력은 높다. 센서가 공격자의 제어 하에 들어갈 수 있으며, 온도, 압력, 화학 농도 측정 등과 같은 여러 경우 물리적 플랜트와의 인터페이스를 높은 정밀도로 조작할 수 있다. 간단한 예로 센서가 다양한 지점에서 온도를 측정하는 사이버 물리 시스템을 생각해 보자. 공격자는 임의기간 동안 측정값을 기록한 다음 기록된 값을 반복적으로 센서에 삽입할 수 있고, 일부 개입으로 인해 실제 온도가 크게 조작될 수 있다. 실제로, 물리적 시스템은 완전히 통제불능이 될 수도 있다. 잘 알려진 공격 중 하나인 스틱스넷 공격은 우라늄 농축용 원심분리기를 조작하여 회전 속도를 증가시켰으나 운영자는 회전 속도가 허용치 범위 안에 있는 것으로 모니터링하여, 결국 원심분리기가 파괴되어 장기간 플랜트 운영에 심각한 문제를 일으켰다[3, 4].

본 논문은 사이버 물리 전력 시스템인 스마트 그리드의 상태 추정에 대한 지능적인 사이버 공격의 일종인 허위 데이터 주입 공격에 초점을 맞춘다. 허위 데이터 주입 공격은 기존의 불량 데이터 감지 알고리즘에 의해 감지되지 않고, 선택된 측정을 협력적으로 수정하여 상태 추정 결과를 임의적이고 예측된 방식으로 조작한다. 시스템 토폴로지에 대한 지식을 통해 공격자는 몇 가지 측정만 수정하여 허위 데이터 주입 공격을 쉽게 구성할 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 전통적인 상태 추정 이론에 대하여 소개한다. III장에서 허위 데이터 주입 공격 구성 방법에 대해 설명한다. VI장에서는 허위 데이터 주입 공격 보호 전략과 탐지를 위한 동적 모니터링 기법을 설명한다. 마지막으로 결론과 향후 연구를 제시하였다.

II. 상태 추정

전력 시스템의 운전 조건은 매일 다양하기 때문에 지역 관제센터 소속 운영자는 계통을 정상적인 안전 상태로 유지해야 한다. 이 목표를 달성하려면 시스템 상태에 대한 지속적인 모니터링, 시스템 상태가 안전하지 않은 것으로 판명될 경우 필요한 예방 조치의 결정이 필요하다. 시스템의 현재 상태를 모니터링하는 것이 첫 번째이자 가장 중요한 단계이다. 현재 전력 시스템에서 SCADA 시스템이 구축되어 관제센터가 모든 종류의 아날로그 측정 및 회로 차단기 상태 정보를 수집할 수 있게 되었다. 그러나 SCADA 시스템이 제공하는 정보는 측정 오류, 원격 측정 오류, 통신 노이즈 등의 결과로 항상 신뢰할 수 있는 것은 아니다. 또한 수집된 측정으로 시스템의 해당 작동 상태를 직접 추출할 수 없다[5].

위에서 언급한 우려는 전력 시스템의 상태 추정이라는 기술로 해결되었다. 상태 추정은 전력 시스템의 상태 변수에 대한 최상의 추정치를 계산하기 위한 수학적 절차이다. 상태 추정을 통해 불량 데이터의 영향을 제거하고 신뢰할 수 있는 상태 추정을 생성할 수 있다. 모든 버스 전압 크기 및 위상각으로부터 계산된 송전선 유효/무효 조류, 그리고 조류로부터 계산된 버스 부하 및 발전으로 구성된 상태 추정기 출력은 경제 급전 프로그램, 상정 해석 프로그램의 기반이 된다.

1. 가중 최소제곱 상태 추정

가중 최소 제곱(Weighted Least Squares, WLS) 상태 추정기는 단순한 모델과 낮은 계산량 때문에 가장 널리 사

용되는 방법이다. WLS 상태 추정기는 가중 최소 제곱 기준을 따르며, 목표는 제곱 측정 잔차의 가중 합계를 최소화하는 것이다. 비선형 측정 모델 $z=h(x)+e$를 고려한다. 여기서 $z=(z_1,z_2,...,z_m)^T$는 측정벡터, $x=(x_1,x_2,...,x_n)^T$는 상태벡터, 그리고 $h(\bullet)$는 측정과 상태를 관련짓는 비선형 함수, e는 측정오차 벡터이다. 측정 오차가 독립 0-평균 가우시안 분포, 즉 측정 i에 대한 $e_i \sim N(0,\sigma_i^2)$에 부합한다고 가정한다. m개의 측정과 n개의 상태가 존재하고 $n<m$이다. WLS 상태 추정은 최적화 문제 $\min_x J(x)=[z-h(x)]^T W[z-h(x)]$로 수학적으로 공식화할 수 있다. 여기서 $J(x)$는 측정 잔차의 가중 합, W는 측정 오차의 공분산 역행렬이다[5].

전력 시스템에서 교류 상태 추정을 위해 상태벡터 x는 기준 버스를 제외한 모든 버스 전압 크기 및 버스 전압 위상각을 포함한다. 상태 x와 측정 z 사이가 비선형인 관계로, $J(x)$를 최소화하기 위해 반복 기법이 채택된다. 일반적으로 사용되는 기법은 $J(x)$의 기울기를 계산한 다음 Newton 방법을 사용하여 $J(x)$을 0으로 강제하는 것이다.

DC 상태 추정을 위해 선형 측정 모델은 $z=Hx+e$로 표현할 수 있고, 여기서 H를 측정 Jacobian 행렬이라고 한다. H 행렬은 시스템 토폴로지 및 송전선로의 저항에 의해 결정된다. 최적 해 $\hat{x}=(H^TWH)^{-1}H^TWz$에서 $J(x)$의 기울기는 사라지고, 측정 잔차 $r=z-Hx$의 추정치는 $\hat{r}=[I-H(H^TWH)^{-1}H^TW]e$으로 주어진다. DC 조류 모델에서 전압 크기는 일정하고 모든 버스에서 1 p.u.라고 가정하며, 무효 전력은 완전히 무시된다. 따라서 상태 변수는 기준 버스를 제외한 모든 버스의 전압 위상각으로만 구성된다.

2. 불량 데이터 탐지 및 식별

상태 추정치 $\hat{x}$ 및 $\hat{z}$가 결정되면 이러한 추정치가 표준 편차와 정확하게 연관되는지를 확인하여 불량 데이터의 존재를 확인할 수 있다. $J(\hat{x})$ 검정은 일반적으로 불량 측정의 존재를 탐지하는 데 사용된다. 이 검정은 랜덤 변수 $J(x)$가 자유도 $k=m-n$인 카이제곱 분포를 따른다고 가정한다. $J(\hat{x})$가 결정된 유의 수준을 가진 탐지 임계값보다 크면 불량 측정의 존재를 의심할 만한 충분한 이유가 된다. 불량 데이터의 존재가 감지되면 r^n 검정을 사용하여 불량 데이터를 식별한다. 여기서 r^n은 정규화된 잔차의 벡터이다. 이 검정은 잘못된 측정이 정규화된 잔차를 가장 크게 발생시킨다는 사실을 기반으로 한다. 불량 데이터의 식별 및 폐기는 상태 추정의 정확도를 향상시킨다.

불량 데이터는 크게 단일 불량 데이터와 다중 불량 데이터로 분류할 수 있다. 잔차가 강하게 또는 약하게 상관된 측정에 불량 데이터가 여러 개 나타날 수 있다. $J(\hat{x})$ 검정과 r^n 검정 방법은 불량 데이터가 서로 모순이 있는 부적합(non-conforming) 불량 데이터일 경우, 단일 불량 데이터, 다수의 비 상호작용 불량 데이터, 그리고 다수의 상호작용 불량 데이터가 포함된 상황에서 매우 효과적이다. 그러나 다중 상호작용 및 불량 데이터가 서로 모순이 없는 적합(conforming) 불량 데이터의 경우, 올바른 측정은 정규화된 잔차를 가장 크게 나타낼 수 있는 반면, 잘못된 측정에서는 정규화된 잔차가 작거나 잔차가 전혀 없을 수도 있다. 우량 측정과 불량 측정을 구별하는 방법이 최적 조합 식별(Combinatorial Optimization Identification, COI) 방법이다[6]. 이 방법은 불량 데이터 집합에 해당하는 다중 정규화 잔차의 유클리드 놈이 최대라는 점에 기반한다. 모든 계량기의 신뢰도가 동일하다고 가정할 때, 불량 측정의 최소 수를 식별하는 것이 최적이다. 다수의 상호작용 및 적합 불량 데이터를 처리하는 또 다른 방법은 가설 검정 식별(Hypothesis Testing Identification, HTI) 방법이다[5]. 이 방법은 먼저 여분의 측정에 오류가 없다고 가정하고 정규화된 잔차에 따라 의심스러운 불량 측정 집합을 선택한다. 그런 다음 가설 검정을 사용하여 의심스러운 측정의 목록을 제거한다. 따라서 이 방법의 효과는 의심스러운 측정 집합의 초기 선택에 따라 달라진다.

Ⅲ. 허위 데이터 주입 공격

스마트 그리드 상태 추정은 전력 시스템의 안정적이고 경제적인 운영을 유지하는 데 중요한 역할을 한다. 기존의 상태 추정 접근법은 전통적으로 무작위 불량 측정을 탐지할 수 있다고 가정한다. 그러나 이들은 최근 고의적인 허위 데이터 주입 공격에 취약한 것으로 드러났다. 이러한 공격은 여러 계량기에서 취한 측정을 협력적으로 수정하여 감지되지 않고 상태 추정 결과를 왜곡한다. SCADA/EMS 시스템은 관제센터 LAN에 점점 더 많이 연결됨에 따라 잠재적으로 인터넷으로 액세스할 수 있다. 또한, 측정 데이터는 광섬유, 위성 및 마이크로파 연결 등으로 구성된 이기종 SCADA 통신 네트워크를 통해 암호화 없이 전송되는 경우가 많다. 위의 사실들로부터 상태 추정에 대한 허위 데이터 주입 공격으로 잠재적인

보안 위협을 야기하는 것을 알 수 있다.

[2]의 핵심 내용은 공격 벡터 a가 Jacobian 행렬 H의 열 벡터의 선형 조합인 경우, 즉 $a=Hc$이면 허위 데이터 주입 공격을 전혀 탐지할 수 없다는 것이다. 여기서 c는 0이 아닌 벡터일 수 있다. 주입된 공격을 측정 오차에 더해지는 부분으로 간주함으로써, 공격에 따른 추정된 측정 잔차는 $\hat{r}_{a=}\hat{r}+[I-H(H^TWH)^{-1}H^TW]Hc=\hat{r}$로 표현될 수 있으며, 이는 원래 측정과 정확히 동일하다. 기존의 불량 데이터 탐지 기법은 모두 측정 잔차를 기반으로 하기 때문에 허위 데이터 주입 공격은 전혀 탐지할 수 없다. 공격을 받는 상태 추정 해는 $\hat{x}_a=\hat{x}+(H^TWH)^{-1}H^TWHc=\hat{x}+c$이다. c는 0이 아닌 벡터일 수 있기 때문에 허위 데이터 주입 공격은 상태 추정 결과를 임의적이고 예측된 방식으로 조작할 수 있다. 또한 공격자가 전력 네트워크 구성 및 송전선로 파라미터 (즉, H 행렬)의 정보에 액세스할 수 있는 경우 허위 데이터 주입 공격을 쉽게 구성할 수 있다. 그리고 [7]에서 지적한 바와 같이 전력 시스템에서 H 행렬의 희소성 때문에 허위 데이터 주입 공격으로 몇 개의 계량기 데이터만 수정하면 된다.

실제로 협력적인 불량 데이터를 처리하는 상태 추정의 능력에 대한 근본적인 한계는 오래 전부터 인식되어 왔다. [8]에서 지적한 바와 같이, 허위 데이터 주입 공격은 상호작용하는 불량 데이터의 완전한 집합으로 볼 수 있으며, 이는 측정 잔차를 변경하지 않고 추정 상태를 $\hat{x}$에서 $\hat{x}+c$로 이끈다. [9]에서 허위 데이터 주입 공격의 성공에 대한 또 다른 설명을 제공하였다. 만약 $\hat{x}$가 실제 네트워크 상태이고 $\hat{x}$와 $\hat{x}+c$가 모두 유효한 네트워크 상태라면, 공격자의 주입 벡터 $a=Hc$는 관제센터가 실제 네트워크 상태를 $\hat{x}+c$라고 믿게 할 것이다. 어떤 검출기도 $\hat{x}$와 $\hat{x}+c$를 구별할 수 없기 때문에, 이러한 공격 벡터 a를 관측할 수 없는 공격이라고 한다. 허위 데이터 주입 공격을 구성하는 것은 네트워크에서 일부 계량기를 제거하는 것과 같으므로 네트워크를 관측할 수 없게 만든다.

[2]에서 공격자가 특정 계량기로 제한되거나 계량기를 손상시키는 데 필요한 리소스가 제한되는 두 가지 현실적인 공격 시나리오에서 상대가 어떻게 체계적이고 효율적으로 공격 벡터를 구성할 수 있는지에 대해 조사하였다. 앞서 설명한 허위 데이터 주입 공격 구성에서는 전력 시스템의 물리적 한계는 고려하지 않았다.

공격자가 공격 리소스 제한으로 인해 허위 데이터 주입 공격을 시작할 수 없는 경우 탐지 확률이 낮은 불완전한 허위 데이터 주입 공격을 구성할 수 있다. 이러한 공격은 약한 공격 체계로 분류된다.

허위 데이터 주입 공격에 대한 대부분의 연구는 DC 상태 추정에 기반한다. [10]에서는 선형 및 비선형 상태 추정기에 은밀한 속임수 공격을 시도하였다. 보다 현실적인 AC 상태 추정에 대한 허위 데이터 주입 공격 연구는 훨씬 어렵고 여전히 탐구 대상이다.

Ⅳ. 허위 데이터 주입 공격 보호 전략 및 동적 모니터링

이상적으로는 허위 데이터 주입 공격이 불가능하도록 전력 시스템이 완전하게 보호되어야 한다. 완전한 보호를 하기 위해서는 운영자가 n개의 측정을 보호하는 것이 필요한데, 이러한 측정값에 따른 H의 서브 행렬이 full rank가 되도록 선택한다. 수학적으로, Jacobian 행렬 H의 $n\times n$ 비특이 행렬 H_s에 대해 $c=0$인 경우에만 $H_sc=0$이 된다. 즉, 서브 행렬 H_s에 따른 측정이 보호되면 공격은 $H_sc=0$을 만족해야 하므로 공격 벡터를 구성할 수 없다. 이러한 측정을 기본 측정이라 하며, 이는 전력 시스템의 관측 가능성을 보장하는 데 필요한 최소 측정 집합이다. 그러나 시스템에서 상태 변수의 수가 일반적으로 크기 때문에 완전한 보호가 실제적으로는 불가능하다. 이 문제를 해결하기 위해서 효과적인 불완전한 보호 전략들이 제안되었다. [8]에서 전체 전력 시스템의 보안 수준을 높이기 위해 희소 공격 벡터와 관련된 보안 지수가 낮은 측정에 암호화 장치를 할당하였다. 또한 공격 및 보호 비용의 모델에서 가장 비용이 낮은 스텔스 공격을 찾는 알고리즘과 허위 데이터 주입 공격에 대한 불완전한 보호를 하기 위해 최대 최소 공격 비용과 최대 평균 공격 비용 모델을 기반으로 하는 두 가지 greedy 알고리즘을 제안하였다. [11]에서는 보호할 측정을 전략적으로 파악하는 greedy 알고리즘을 제안하였다. 이 전략은 허위 데이터 주입 공격을 시도하는 측정의 수만을 고려하고 전체 전력 시스템에 대한 공격의 영향을 고려하지는 않는다.

강건한 전력 시스템 설계에도 불구하고 현장에서의 운영은 예상치 못한 조건으로 인하여 고장과 실패가 발생한다. 문제의 조기 발견과 신속한 복구를 통해 지속적으로 안전한 운영을 보장하기 위해서는 현장에서 실시간으로 모니터링할 필요가 있다. 설계된 모니터는 모든 유형의 공격과 실패를 커버하기 위하여 모든 유형의 파라미

터 변경 및 결함을 탐지하여야 한다. 특히 허위 데이터 주입 공격의 경우 모니터는 입력되는 파라미터 측정을 관측하고 정상 작동으로부터의 편차를 탐지해야 한다. 센서 및 플랜트 결함으로 인한 광범위한 파라미터 변화와 실패는 강건한 계산 도구와 방법을 개발할 필요성을 제시한다. 그러한 파라미터 변화에 대한 플랜트 모니터링의 일반적인 기술인 상태 모니터링을 위한 기존의 방법은 제한적이며 모든 유형의 고장을 조기에 감지하고 진단하는 것은 사실상 어렵다. 강건하고 안전한 사이버 물리 전력 시스템을 위해 실시간 모니터를 개발하는 양호한 방법은 기존의 상태 모니터를 확장하여 사이버 공격, 특히 허위 데이터 주입 및 초기 고장과 같은 유형의 고장까지 탐지하는 것이다.

중요 시설의 상태 모니터링을 위한 계산 도구와 방법은 모델 기반 및 무모델 기반의 두 가지 접근법으로 분류할 수 있다[12]. 모델 기반 접근 방식은 결함이 없는 조건에서 모니터링되는 시스템의 역학에 대한 사전 정보를 활용한다. 이 정보는 상태 공간 표현 또는 등가 수식과 같은 모델에 포함된다. 무모델 기반 접근 방식은 원 데이터를 처리하여 이를 신경망과 같은 비모수적 근사 형태로 나타낼 수 있다.

사이버 공격 또는 고장을 나타내는 파라미터 변경을 감지하려면 무공격/무고장 시스템 작동과 문제가 있을 때의 작동 간의 편차를 감지해야 한다. 이 차이는 임계값을 초과하는 편차를 통해 탐지된다. 이 임계값은 효과적인 조기 감지뿐만 아니라 허위 경보를 방지하기 위해 설정값이 매우 중요하다. 공격과 고장을 감지하고 격리하는 방법의 신뢰성은 모니터링 시스템의 무공격/무고장 작동을 나타내는 모델의 정확도, 무공격/무고장 조건에서 모니터링 시스템의 출력을 위한 추정 방법의 정확도, 통계 의사 결정 절차의 정확성과 결함의 존재를 추론하는 데 사용되는 고장 한계값의 세 가지 주요 요인에 의해 결정된다. 대부분의 모델 기반 고장 진단 접근법인 선형 및 비선형 관측기 또는 필터가 사용되는 모델의 정확도는 신뢰성 측면에서 중요하다. 무모델 기반의 고장 진단 방법에서는 원 데이터에서 추출된 모델의 정확도가 신뢰성을 나타낸다. 이 정확도는 모델 검증 방법을 사용하여 측정할 수 있다. 시스템에는 결함이 없을 수 있지만 파라미터의 값은 모델에서 사용되는 것과 다를 수 있다.

모델이 유효한지 여부를 결정하기 위해 새로운 작동 조건에서 고장이 없는 시스템의 출력을 모델에 의해 제공되는 예상 출력과 비교한다. 이 작업은 잔차를 계산하고 통계적으로 분석하여 수행된다. 즉, 모델 유효성 평가를 하여 설정된 임계값을 벗어난 것을 보고 모델이 업데이트되어야 하는지 판단할 수 있다.

추정 방법의 정확도는 공격/고장 진단 프로세스에서 중요하다. 무공격/무고장 시스템의 추정기는 최소의 분산을 가져야 측정 노이즈의 영향이 제거되어 추정된 시스템 출력이 실제값에 근사할 수 있다. 다양한 상태 관측기 또는 필터가 있지만, 이 중에서 칼만 필터(Kalman Filter)가 최소 분산 추정치를 제공하는 것으로 알려져 있고 실제로 널리 사용되고 있다[13, 14].

칼만 필터는 다른 상태 관측기 또는 필터에 비해 계산 속도 측면에서 성능이 더 우수하고 빠른 수렴을 달성하므로 동적 시스템의 실시간 고장 진단에 적용이 가능하다. 더욱이 칼만 필터는 측정 노이즈 및 모델 오차에 대해 강건성을 보장하는 형태로 재설계될 수 있다. 따라서 칼만 필터는 최소 분산 추정치의 최적성을 보장하고 다른 상태 관측기 또는 필터 성능을 능가하기 때문에 사이버 물리 시스템에 적합한 추정 방법으로 사용될 수 있다.

공격과 고장을 탐지하기 위한 임계값의 최적 선택은 초기 고장을 포함한 선제 공격과 고장 진단 및 허위 경보율에 있어서 중요하다. 모델 검증과 유사하게 잔차 시퀀스는 공격 또는 고장을 탐지하는 통계 테스트의 확률적 변수를 결정하는 데 사용된다. 잔차 시퀀스의 요소는 0 평균 가우시안 분포를 따르고 공분산 행렬의 역수가 가중된 잔차 벡터의 제곱합이 카이제곱 분포를 따른다는 것을 알 수 있다. 이 분포의 신뢰 구간을 사용하여 무공격/무고장 모델과 모니터링되는 시스템 작동 간의 편차를 탐지할 수 있다.

칼만 필터 기반 상태 모니터를 사용하는 간단하고 효과적인 방법은 칼만 필터를 가상 센서로 사용하여 무고장 모드에서 플랜트의 센서 작동을 모방한 가상 센서의 결과와 실제 센서 측정 사이의 편차를 식별하는 것이다. 임계값을 벗어나는 편차로 공격 또는 실패 여부를 따진다. 통계적 의사 결정 기준과 결합된 이 방법은 스마트 그리드 센서에 대한 공격을 탐지하는 데 사용된다. 칼만 필터는 무고장 모드에서 그리드 센서 작동을 모방한 가상 센서로 사용되며, 그 출력을 실제 센서의 출력과 비교하여 잔차 시퀀스 벡터를 생성한다. 공분산 행렬의 역에 의해 가중된 이 잔차 벡터의 제곱은 카이제곱 분포를 따르는 랜덤 변수로 구성된다. 따라서 이 변수는 카이제곱 분포의 특성을 활용하고 신뢰 구간 접근 방식을 사용하여 이 통계 테스트의 임계값을 정의할 수 있다. 통계 테스트의 출력값이 임계값을 초과하면 허용가능한 범위를 벗어나 센서 작동이 비정상적임을 나타내므로 경보가 발

생한다. 중요한 것은 통계 테스트가 센서 클러스터에 적용될 수 있기 때문에 공격에 노출된 스마트 그리드의 부분을 식별할 수 있다는 것이다. 또한 각 센서마다 통계 테스트를 적용하여 손상된 센서를 분리할 수도 있다.

Ⅴ. 결론

본 논문에서는 대표적인 에너지 기반 사이버 물리 시스템인 스마트 그리드에 대한 허위 데이터 주입 공격에 대하여 서술하였다. 허위 데이터 주입 공격이 사이버 물리 시스템 영역에서 주요한 공격 중 하나이기 때문에 현재 이것에 대한 연구가 많이 이루어지고 있다. 그러나 대부분의 연구가 주로 센서 네트워크와 스마트 그리드 영역에서만 다루어지고 있는 것이 현실이다. 향후 다양한 영역의 안전필수 제어시스템에서도 허위 데이터 주입 공격 연구가 필요하다.

References

[1] S. Son, T. Park, and M. Won, "An Overview of Cyber Physical Systems," *Telecommunications Review*, Vol.24, No.4, pp.450-459, 2014.
DOI: 10.1007/978-3-030-43494-6_1

[2] Y. Liu, P. Ning, and M.K. Reiter, "False Data Injection Attacks against State Estimation in Electric Power Grids," *in Proc. of the 16th ACM Conf. Computer and Communications Security*, Chicago, IL, 2009. DOI: 10.1145/1952982.1952995

[3] G. Liang, J. Zhao, F. Luo, and S. R. Weller, "A Review of False Data Injection Attacks against Modern Power Systems," *IEEE Trans. Smart Grid*, Vol.8, No.4, pp.1630-1638, 2017.
DOI: 10.1109/TSG.2015.2495133

[4] Q. Wang, W. Tai, Y. Tang, and M. Ni, "Review of the false data injection attack against the cyber-physical power system," *IET Cyber-Physical Systems: Theory & Applications*, Vol.4, No.2, pp. 101-107, 2018. DOI: 10.1109/TSG.2015.2495133

[5] A. Abur, A.G. Exposito, *Power System State Estimation: Theory and Implementation*, Marcel Dekker, 2004.

[6] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*, Kluwer Academic, 1999.

[7] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Limiting False Data Attacks on Power System State Estimation," *44th Annual Conf. Information Sciences and Systems*, pp.1-6, 2010.
DOI: 10.1109/CISS.2010.5464816

[8] G. Dan, and H. Sandberg, "Stealth Attacks and Protection Schemes for State Estimators in Power Systems," *IEEE Conf. Smart Grid Comm.*, pp.214-219, 2010.
DOI: 10.1109/SMARTGRID.2010.5622046

[9] O. Kosut, L. Jia, R. Thomas, and L. Tong, "Malicious Data Attacks on the Smart Grid," *IEEE Trans. Smart Grid*, Vol.2, No.4, pp.645-658, 2011.
DOI: 10.1109/TSG.2011.2163807

[10] A. Teixeira, S. Amin, H. Sandberg, K. H. Johansson, and S.S. Sastry, "Cyber Security Analysis of State Estimators in Electric Power Systems," *Proc. IEEE Conf. Decision and Control*, 2010.
DOI: 10.1109/CDC.2010.5717318

[11] T. T. Kim, and H. V. Poor, "Strategic Protection Against Data Injection Attacks on Power Grids," *IEEE Trans. Smart Grid*, Vol.2, No.2, pp.326-333, 2011. DOI: 10.1109/TSG.2011.2119336

[12] G. Rigatos, D. Serpanos, and N. Zervos, "Detection of Attacks Against Power Grid Sensors Using Kalman Filter and Statistical Decision Making," *IEEE Sensors Journal*, Vol.17, No.23, pp.7641-7648, 2017.
DOI: 10.1109/JSEN.2017.2661247

[13] D. B. Rawat, and C. Bajracharya, "Detection of False Data Injection Attacks in Smart Grid Communication Systems," *IEEE Signal Processing Letters*, Vol.22, No.10, pp.1652-1656, 2015.
DOI: 10.1109/MCOM.2015.7045410

[14] A. Sargolzaei, K. Yazdani, A. Abbaspour, et al., "Detection and Mitigation of False Data Injection Attacks in Networked Control Systems," *IEEE Trans. Industrial Informatics*, Vol.16, No.6, pp. 4281-4292, 2020.
DOI: 10.1109/IISR.2018.8535978

BIOGRAPHY

Junhyung Bae (Member)

2004 : BS degree in Electronic Engineering and Avionics, Korea Aerospace University.
2006 : MS degree in Electrical Engineering, Hanyang University.
2017 : PhD degree in Information and Communication Engineering, DGIST.
2006~2010 : Researcher, DGIST.
2011 : Senior Researcher, Samsung Thales.
2017 : Postdoc., DGIST
2017~2020 : Senior Researcher, Korea Aerospace Industries.
2020~Current : Assistant Professor, School of Electronic and Electrical Engineering, Daegu Catholic University.