Cyber Physical Systems under Sparse Adversarial Attacks

Chandrasekhar S

SPC Lab, Department of ECE

Indian Institute of Science

October 12, 2019

System Model

$$
\begin{aligned}
x(t+1) &= Ax(t) + B[u(t, y(0), y(1), \dots, y(t)) + w(t)] \\
y(t) &= Cx(t) + e(t)
\end{aligned}
$$

Here, $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^p$ and $u(t) \in \mathbb{R}^m$, The sparse vector $e(t) \in \mathbb{R}^p$ represents attack injected in different sensors, and $w(t) \in \mathbb{R}^m$ represents the attack on the actuators.
Assumption The set of attacked nodes do not change with time.

Only Sensor Attacks

Goal

- To estimate the initial state $x(0)$ in the presence of sensor attacks using observations $(y(t))_{t=0,1,\ldots,T-1}$.
- Decoder $D : (R^p)^T \to \mathbb{R}^n$.
  $\hat{x}(0) = D(y(0), y(1), \ldots, y(T-1))$
- q errors are correctable after T steps if $\forall\ x(0), \forall$ $K \subset \{1, 2, \ldots, p\}$ s.t. $|K| \leq q$ and $\forall\ e(0), e(1), \ldots, e(T-1)$ s.t $supp(e(t)) \subset K$, $\exists\ D$ s.t $D(y(0), \ldots, y(T-1)) = x(0)$

$$\begin{aligned} x(t+1) &= Ax(t) \\ y(t) &= Cx(t) + e(t) \end{aligned}$$

Number of Correctable Attacks

Proposition The following are equivalent

1. There is a decoder that can correct q errors after T steps.
2. $\forall\, z \in \mathbb{R}^n \setminus \{0\}$,
   $|\text{supp}(Cz) \cup \text{supp}(CAz) \cup \ldots \cup \text{supp}(CA^{T-1}z)| > 2q$.

We can write relation between observations and initial state as

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(T-1) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{T-1} \end{bmatrix} x(0) = \mathcal{O}x(0)$$

By Cayley-Hamilton theorem, one can also see that the number of correctable errors cannot increase beyond $T = n$

Proof.

(i) $\implies$ (ii) By contradiction, Take that vector z for which (ii) is false, then $\mathcal{O}z$ has less then $2q$ elements non-zero for each $y(i)$, an attack of size q which zeros out same q non-zero entries of $y(i)$ makes it indistinguishable from $x(0) = 0$ $\square$

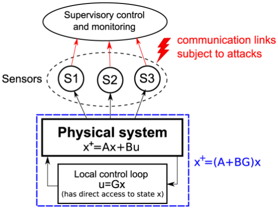
Proposition For almost all pairs $(A, C) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{p\times n}$ the number of correctable errors after $T = n$ steps is maximal and equal to $\frac{p}{2} - 1$

Proof.

Consider $\mathcal{O}_i = \begin{bmatrix} e_i^T C \\ e_i^T CA \\ \vdots \\ ei^T CA^{n-1} \end{bmatrix}$ consider $f_i(A, C) = det(\mathcal{O}_i)$

Note f_i is not identically 0, hence the zero set of f_i has measure 0 on $\mathbb{R}^{n\times n} \times \mathbb{R}^{p\times n}$ □

Question: Can we find a matrix G for feedback such that if we can add $u = Gx$ then number of correctable attacks $q = \lceil p/2 - 1 \rceil$



Lemma Assuming A has n eigen values of distinct magnitudes the following are equivalent

1. q errors are correctable after n steps.
2. $\forall$ eigen vector v of A $|\mathrm{supp}(Cv)| > 2q$

Proof Sketch Since any vector u can be written as linear combination of eigen vectors of A

$$CA^t u = \sum_{i=1}^{n} \alpha_i \lambda_i^t C v_i$$

$$\frac{CA^t u}{\lambda_1^t} = \alpha_1 C v_1 + \sum_{i=2}^{n} \alpha_i \frac{\lambda_i}{\lambda_1}^t C v_i$$

Optimization Formulation of the Decoder

$$\begin{aligned} \min_{\hat{x}\in\mathbb{R}^n, \hat{K}\subset\{1,\ldots,p\}} \quad & |\hat{K}| \\ \text{subject to} \quad & \operatorname{supp}(y(t) - CA^t\hat{x}) \subset \hat{K} \\ & \quad \text{for } t \in \{0, \ldots, T-1\} \end{aligned}$$

But the above optimization problem is NP-hard in general.

$$\begin{aligned} \Phi^{(T)} : \mathbb{R}^n &\to \mathbb{R}^{p\times T} \\ x &\to \begin{bmatrix} Cx & CAx & \ldots & CA^{T-1}x \end{bmatrix} \\ Y(T) &= \begin{bmatrix} y(0) & y(1) & \ldots & y(T-1) \end{bmatrix} \end{aligned}$$

Then the above optimization problem is

$$\arg \min_{\hat{x}\in\mathbb{R}^n} \|Y(T) - \Phi^{(T)}\hat{x}\|_{\ell_0}$$

Consider a ℓ_1 decoder for $r \geq 1$ that solves

$$D_{1,r}(y(0), y(1), \ldots, y(T-1)) = \arg \min_{\hat{x} \in \mathbb{R}^n} \| Y(T) - \Phi^{(T)} \hat{x} \|_{\ell_1/\ell_r}$$

$$\|M\|_{\ell_1/\ell_r} = \sum_{i=1}^{p} \|M_i\|_{\ell_r}$$

Proposition The following are equivalent

1. The decoder $D_{1,r}$ can correct q errors after T steps.
2. $\forall\ K \subset \{1, 2, \ldots, p\}$ with $|k| = q$ and $\forall\ z \in \mathbb{R}^n \setminus \{0\}$

$$\sum_{i \in K} \left\| \left(\Phi^{(T)} z \right)_i \right\|_{\ell_r} < \sum_{i \in K^c} \left\| \left(\Phi^{(T)} z \right)_i \right\|_{\ell_r}$$

Proof.

Prove (i) $\implies$ (ii) through contradiction choose $x(0) = 0$ and let K and z be such that (ii) is false and choose attack nodes as set K, then

$$\|Y(T) - \Phi^{(T)}z\|_{\ell_1/\ell_r} \geq \|Y(T)\|_{\ell_1/\ell_r}$$

$$\sum_{i\in K} \|(Y(T) - \Phi^{(T)}z)_i\|_{\ell_r} + \sum_{i\in K^c} \|(\Phi^{(T)}z)_i\|_{\ell_r} \geq \sum_{i\in K} \|(Y(T))_i\|_{\ell_r}$$

Choosing $(Y(T))_i = (\Phi^{(T)}z)_i$ for $i \in K \implies$ contradiction. □

Proof.

Prove (ii) $\implies$ (i) through contradiction, $\exists\ x(0)$, $z = x(0) + e$ and set of attacked nodes K such that

$$\|Y(T) - \Phi^{(T)}z\|_{\ell_1/\ell_r} < \|Y(T) - \Phi^{(T)}x(0)\|_{\ell_1/\ell_r}$$

$$\sum_{i\in K} \|(Y(T) - \Phi^{(T)}z)_i\|_{\ell_r} + \sum_{i\in K^c} \|(\Phi^{(T)}e)_i\|_{\ell_r} < \sum_{i\in K} \|(Y(T) - \Phi^{(T)}x(0))_i\|_{\ell_r}$$

$$\sum_{i\in K^c} \|(\Phi^{(T)}e)_i\|_{\ell_r} < \sum_{i\in K}[\|(Y(T) - \Phi^{(T)}x(0))_i\|_{\ell_r} - \|(Y(T) - \Phi^{(T)}z)_i\|_{\ell_r}]$$

$$\implies \sum_{i\in K} \left\|\left(\Phi^{(T)}z\right)_i\right\|_{\ell_r} \geq \sum_{i\in K^c} \left\|\left(\Phi^{(T)}z\right)_i\right\|_{\ell_r}$$

$\square$