Evasion attack against Multivariate Singular Spectrum Analysis based IDS

Vikas Maurya, Rachit Agarwal, and Sandeep Shukla

Department of Computer Science and Engineering
Indian Institute of Technology Kanpur
Kanpur India

September 14, 2023

Outline

- Industrial Control System
- Process-level IDS
- Attack Model
- MSSA based IDS
- Evasion attack
- Experiment and Result
- Conclusion

Industrial Control System (ICS)

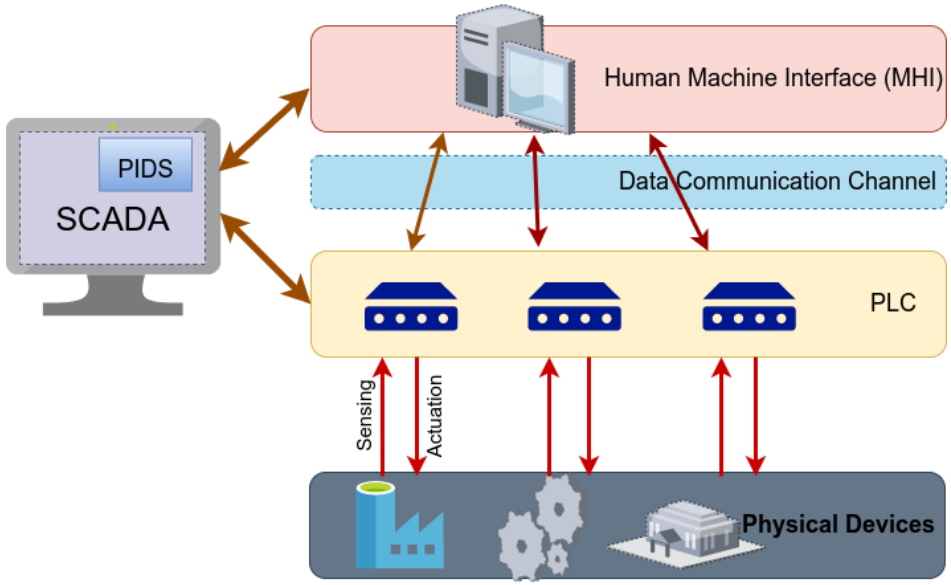


Figure 1: High level architecture of system model.

- CIs are mostly maintained by ICS
- Consists of various workstations
- Some past incidents are:
  - Iranian nuclear plant in 2009
  - German steel mill in 2014
  - Saudi petroleum refinery in 2017
  - Indian nuclear plant (NPCIL) in 2019
  - Israel water treatment plant in 2020

Process level IDS

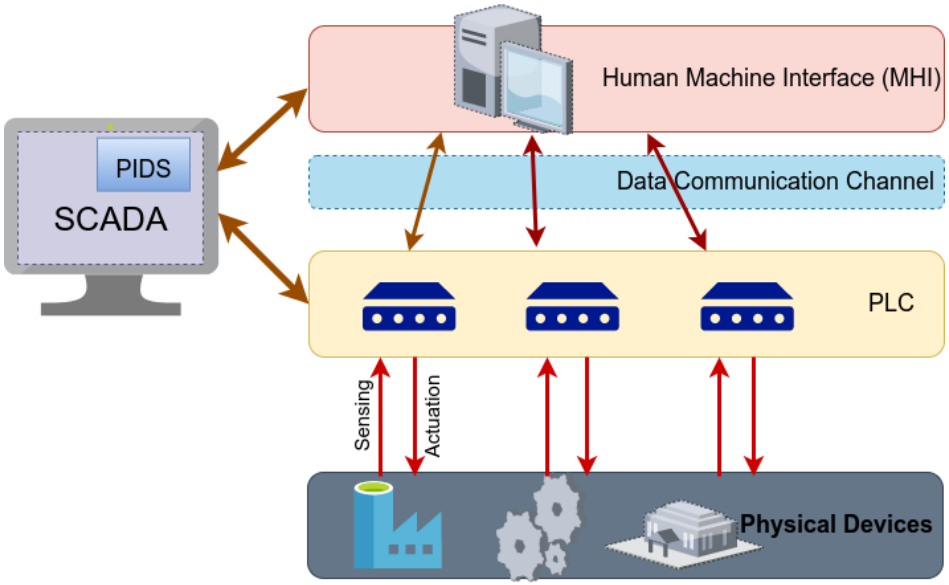


Figure 2: High level architecture of system model.

- An ICS is protected by various layers of protection
- But an attacker can evade such protection
- The ultimate attacker's aim gets reflected in physical process
- A process level IDS monitors the sensor measurements to detect attack induced abnormalities

Attack Model

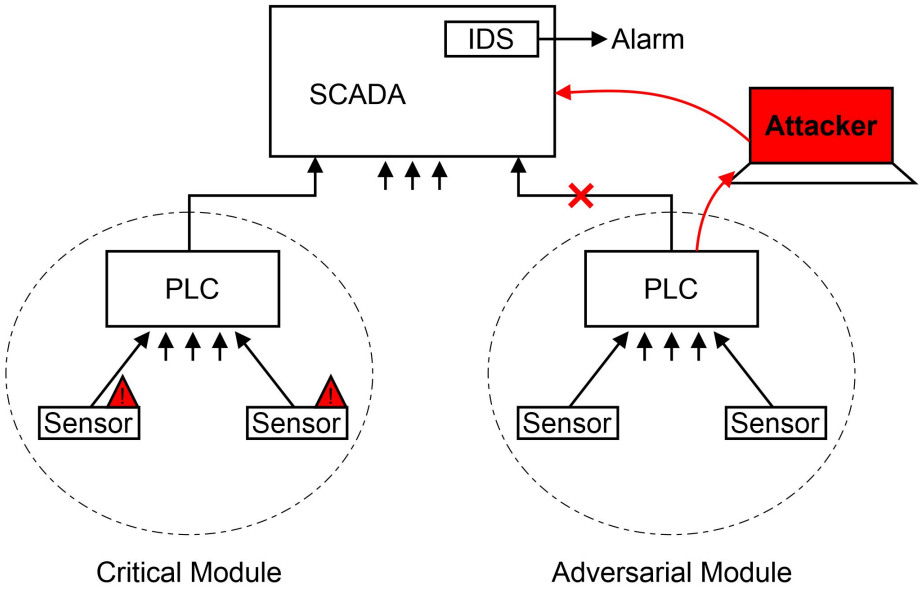


Figure 3: High level architecture of attack model.

- Attacker targets to the most critical component
- But the critical components are more protected
- Attacker finds some easy manipulable components
- A MITM attacker manipulate sensor measurements [1, 2, 3]
- Proposed perturbation method to craft adversarial measurements

MSSA based IDS

- A multivariate IDS
- MSSA based IDS [4] is:
  - Computationally efficient
  - Captures temporal information
  - Captures mutual correlation
  - Supports noise cancellation property
  - Capable to detect even a stealthy attack
  - Suitable for large scale ICS/IIoT networks

Working of MSSA based IDS

- It works by projecting the recent sensor measurements on noise free signal subspace
- Attack is detected based on departure of projected measurement

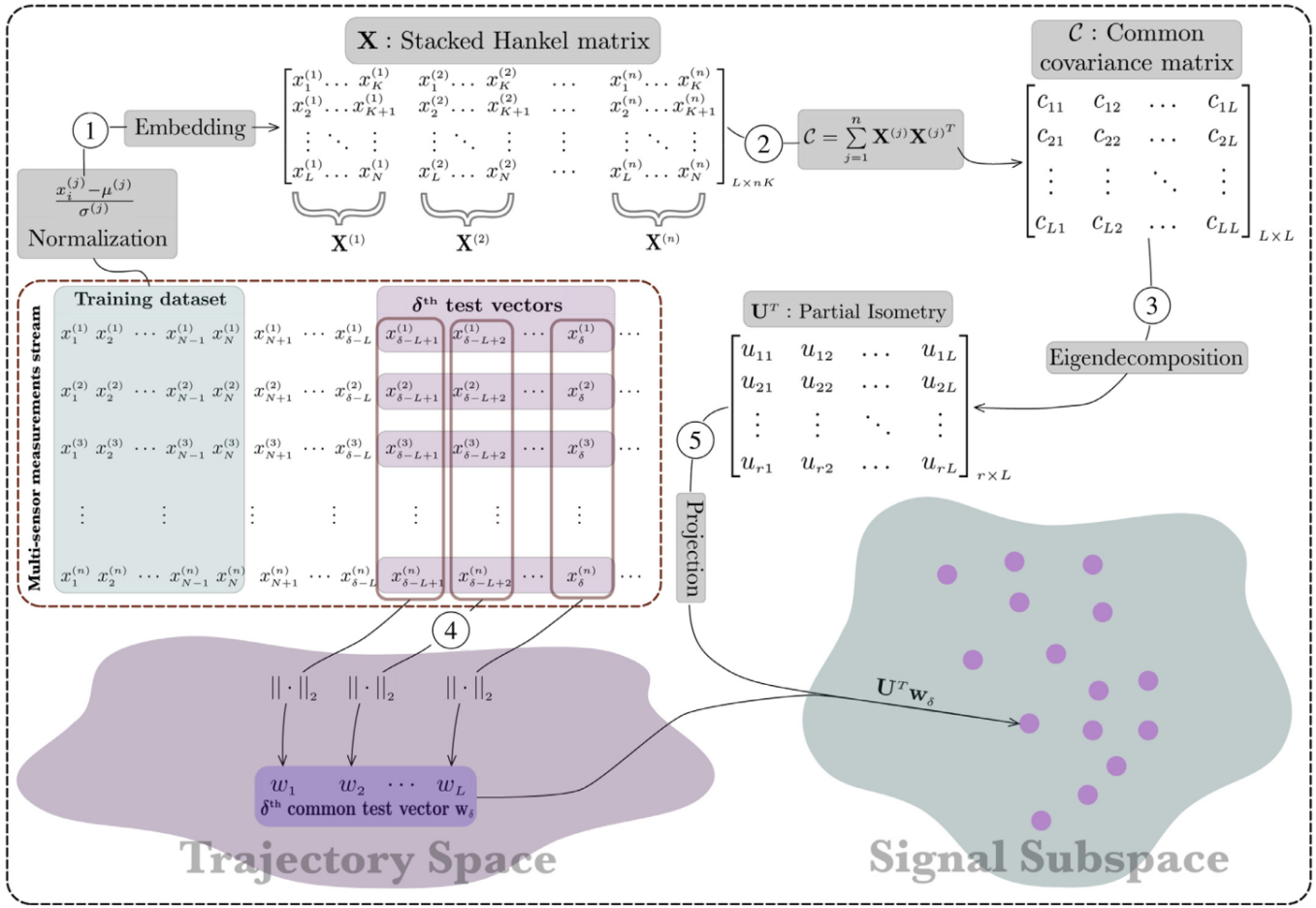


Figure 4: Working of MSSA Based IDS [4]

Model Parameters

- Parameters after training:
  - Projection Matrix: $U^T \in \mathbb{R}^{R \times L}$
  - Centroid Vector: $\hat{c} \in \mathbb{R}^R$
  - Classifier Threshold: θ
  - Lag Parameter L
  - $\mu^{(n)}$, $\sigma^{(n)}$

Online Detection

Consider an ICS consisting of N sensors where n^{th} sensor generates measurement $X_t^{(n)}$ at time t. The IDS performs following steps at timestamp t:

1. Normalization:
$$(X_t^{(n)} - \mu^{(n)})/\sigma^{(n)}$$
2. Compute aggregated measurement
$$m_t = ||X_t||$$
   - It forms an aggregated time series over time
$$\mathcal{T} = [\cdots, m_t, \cdots]$$
3. Compute L length lag vector:
$$w_t = [m_{t-L+1}, \cdots, m_{t-1}, m_t]^T$$
4. Compute the departure score:
$$D_t = ||\hat{c} - U^T w_t||^2$$
5. If $D_t > \theta$, then raise an attack alarm

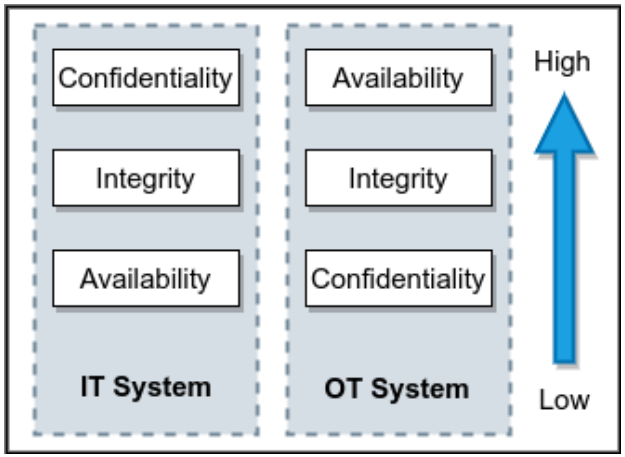
Related Works

- The current evasion attack methods [6, 7] are against deep-learning based IDS
- Do not consider time-series based models
- Deep-learning based IDS suffers from various limitations:
  - High computation cost
  - Noise cancellation property

The low computation cost and noise cancellation property of MSSA-based IDS make it one of the most suitable in large-scale ICS/IIoT networks, which motivates us to analyse it against evasion attacks.

Evasion Attack: Capabilities

- Accessibility



  - High availability supports non-encrypted communication [3, 5]
  - An attacker already present in the network can eavesdrop
- Manipulation capability
  - A rootkit can be deployed to PLC to manipulate the sensor measurements [1]
  - A victimized employee can be used for exploitation [2]
  - Various vulnerabilities are reported in OT networks, ≈83% violates communication [3]

Evasion Attack: Greedy Approach

- Objective Function:

$$\hat{m}_{t+1} = \arg\min_{m'_{t+1}} \left(D'_{t+1}\right)$$

- Simplify D'_{t+1} for unknown:

$$\begin{aligned} D'_{t+1} &= ||\hat{c} - U^T \cdot w'_{t+1}||^2 \\ &= ||\hat{c} - (U[1:L-1]^T \cdot w'_{t+1}[1:L-1] + U[L] * w'_{t+1}[L])||^2 \\ &= ||y - U[L] * m'_{t+1}||^2 \\ &= ||U[L]||^2 m'^2_{t+1} - 2(y^T \cdot U[L]) m'_{t+1} + ||y||^2 \end{aligned}$$

where,

$$y = \hat{c} - U[1:L-1]^T \cdot w'_{t+1}[1:L-1]$$

- Minima:

$$\hat{m}_{t+1} = \frac{y^T \cdot U[L]}{||U[L]||^2}$$

Evasion Attack: Reverse Engineering

- Estimate the adversarial measurements
  - Break $\hat{m}_{t+1} = ||X_{t+1}||$ into adversarial and non-adversarial:
    $$||X_{t+1}[\overline{adv}]||^2 + ||X_{t+1}[adv]||^2 = \hat{m}_{t+1}^2$$
  - Assumption $X_{t+1} \approx X_t$:
    $$||X_{t+1}[adv]||^2 = \hat{m}_{t+1}^2 - ||X_t[\overline{adv}]||^2$$
- Manipulation is performed only if departure score is above an estimated threshold

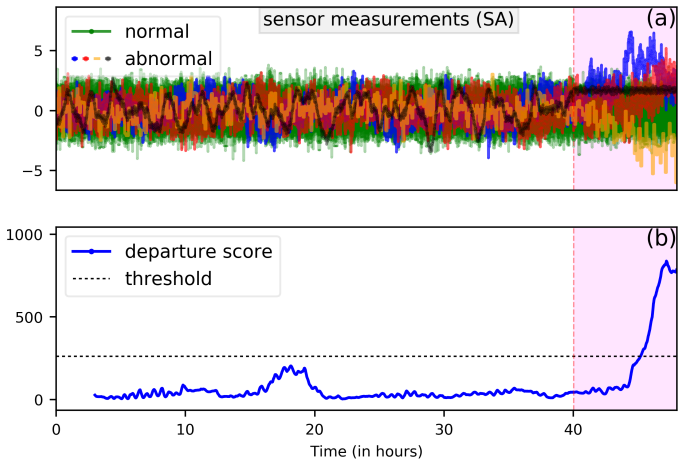
Evasion Attack: Constraints

- Accessibility constraints
  - White Box attacker: $(\mathcal{D}, \mathcal{X}, f, \phi)$
  - Gray Box attacker: $(\cancel{\mathcal{D}}, \mathcal{X}, f, \cancel{\phi})$
- Manipulation constraints
  - Manipulated measurement must be within the normal range

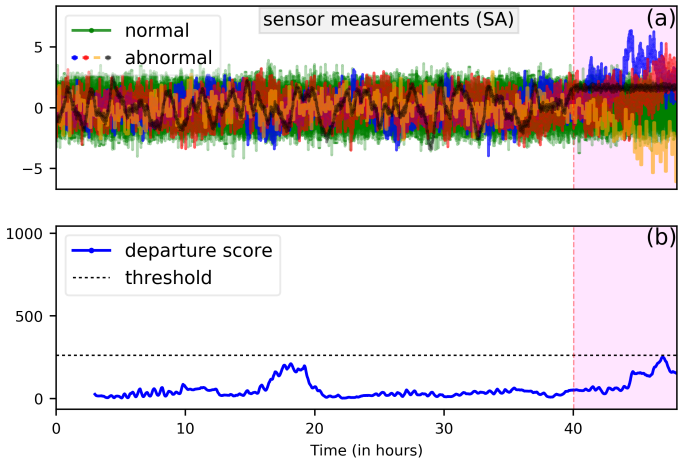
Data Preparation

- TE-process simulator [8] is used
- Generated normal measurements
- Generated stealthy attack dataset
- Generated Direct damage attack dataset
- Ensure main sensors impacted by attack

Stealthy attack scenarios

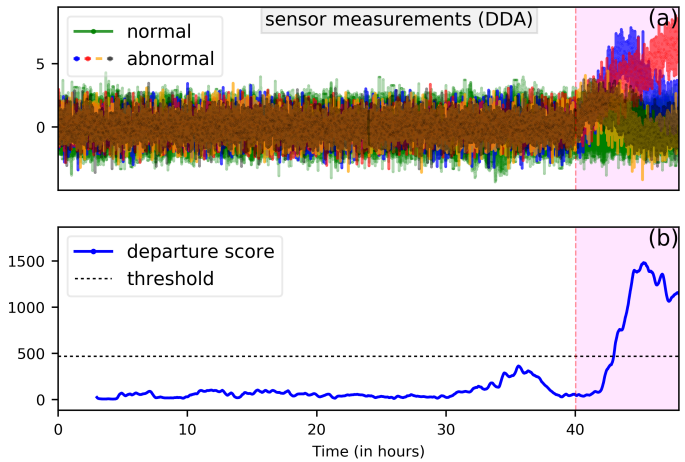


Before Manipulation

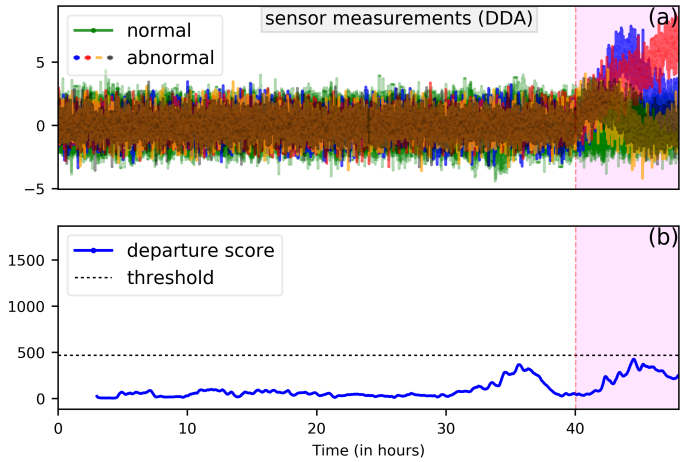


After Manipulation

Direct damage attack scenarios

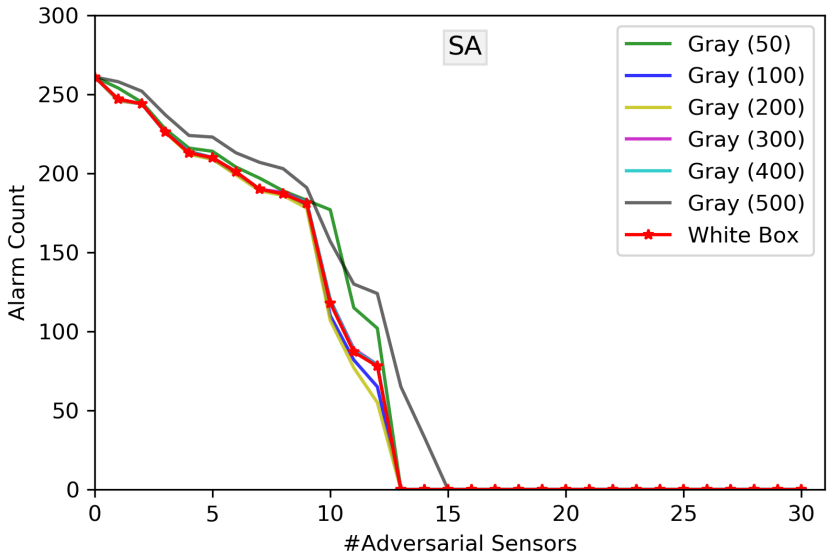


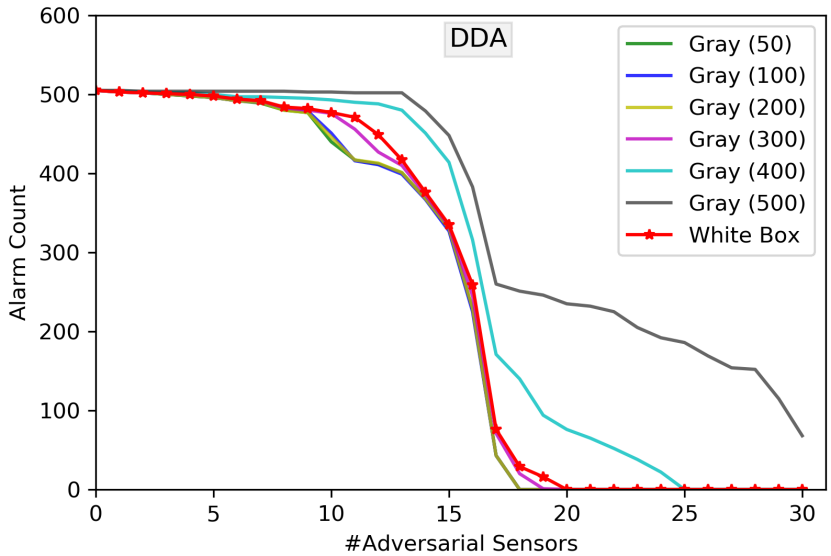
Before Manipulation



After Manipulation

White box and Gray box





Complexity Analysis

- Time Complexity: $O(RL)$
- Space Complexity: $O(RL)$
- Which is equal to the IDS
- Experimental: 53.7 μ-Sec

Conclusion

- We discussed vulnerability of MSSA based IDS
- A practicality of attack model
- Evasion attack against a time-series based IDS
- A novel perturbation method
- Demonstrated on SA and DDA attack scenarios

Garcia, L., Brasser, F., Cintuglu, M., et. al.

Hey, my malware knows physics! attacking PLCs with physical model aware rootkit.
In NDSS, San Diego (03, 2017).

Kovacevic, A., Nikolic, D.

Cyber attacks on critical infrastructure: Review and challenges
Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance (2015).

Yadav, G., Paul, K.

Assessment of scada system vulnerabilities
24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) (2019).

Aoudi, W., Almgren, M.

scalable specification-agnostic multi-sensor anomaly detection system for IIoT environments
International Journal of Critical Infrastructure Protection (2020).

Huitsing, P., Chandiaaa, R., Papa, M., Shenoi, S

Attack taxonomies for the modbus protocols
International Journal of Critical Infrastructure Protection (2008)

Erba, A., et al.

Constrained concealment attacks against reconstruction-based anomaly detectors in industrial control systems
Annual Computer Security Applications Conference, ACSAC (2020)

Li, J., Yang, Y., Sun, J.S., Tomsovic, K., Qi, H.

Conaml: Constrained adversarial machine learning for cyber-physical systems
ACM Asia Conference on Computer and Communications Security, AsiaCCS (2021)

Downs, J., Vogel, E

A plant-wide industrial process control problem
Computers & Chemical Engineering (1993)

Thank You!!