Ke LIU, Mufeng WANG, Rongkuan MA, Zhenyong ZHANG, Qiang WEI, 2022. Detection and localization of cyber attacks on water treatment systems: an entropy-based approach. *Frontiers of Information Technology & Electronic Engineering*, 23(4):587-603. https://doi.org/10.1631/FITEE.2000546

Detection and localization of cyber attacks on water treatment systems: an entropy-based approach

Key words: Industrial cyber-physical system; Water treatment system; Intrusion detection; Abnormal state; Detection and localization; Information theory

Corresponding author: Qiang WEI
E-mail: funnywei@163.com

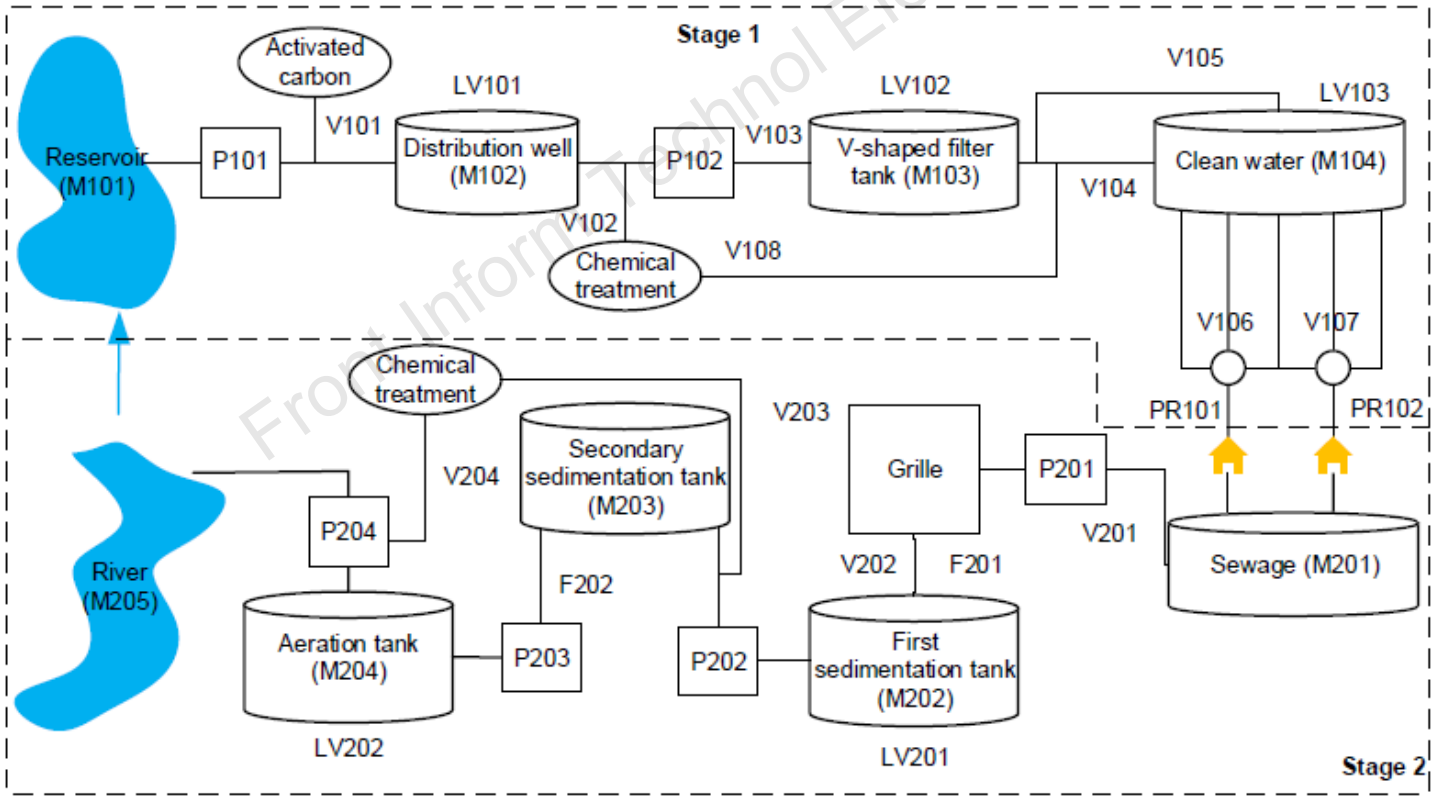
ORCID: https://orcid.org/0000-0002-0288-0086

Motivation

- In water treatment systems (WTSs), few types of metadata are used, which limits the performance of the intrusion detection system (IDS).
- Most existing IDSs are used to locate intrusion points rather than abnormal points. However, abnormal point localization is more helpful in recovering the system than intrusion point localization.

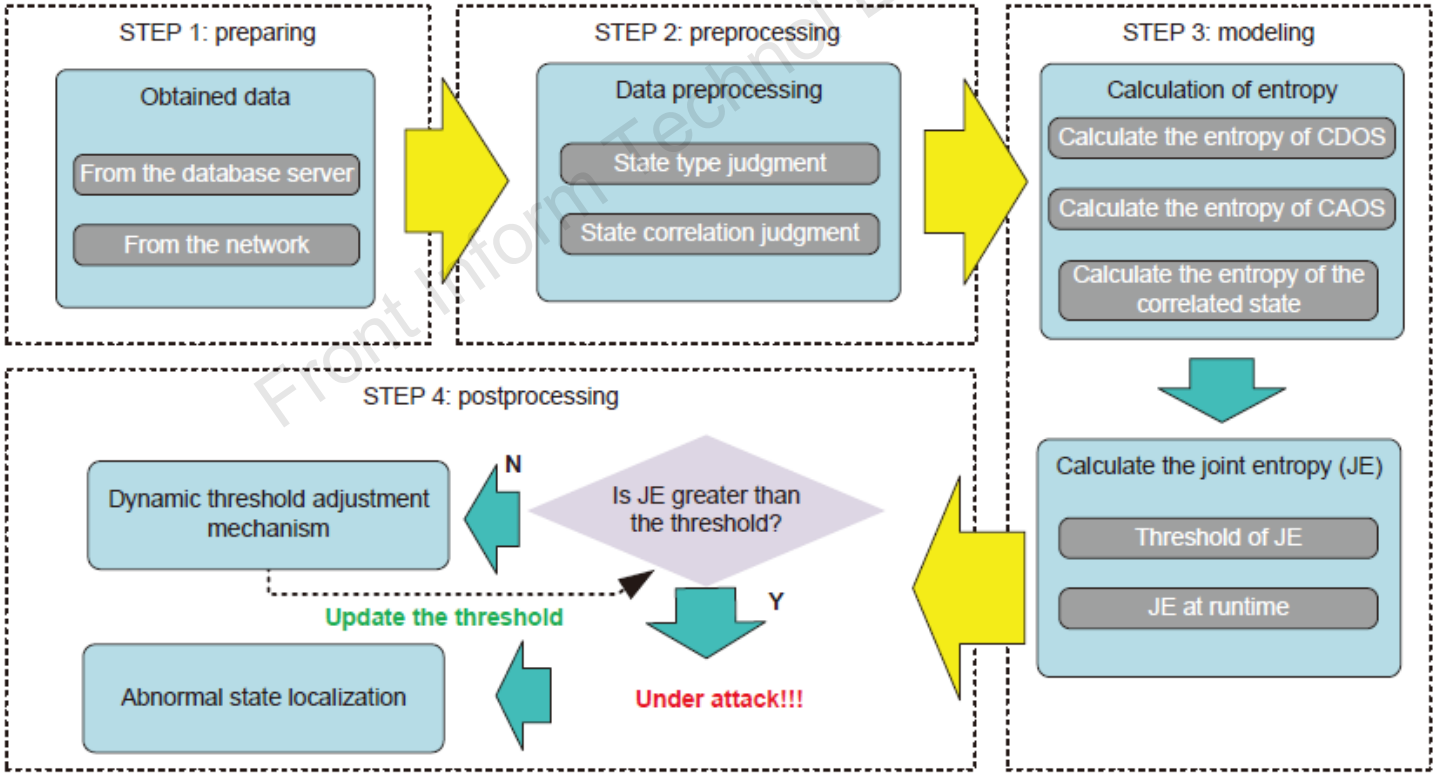
WTS testbed

- The two-stage WTS testbed includes two programmable logic controllers (PLCs) and more than 50 measurement points



Method overview

- Entropy-based intrusion detection method (EBID) includes mainly four steps: preparing, preprocessing, modeling, and postprocessing



Method

- Preparing
  - Obtain the controller's output state (COS) data in two ways: exported from human-machine interface (HMI) or database, captured from the network

- Preprocessing
  - Calculate entropy for COS
    - Discrete: $H_{\mathrm{D}}(X) = -\sum_{i=1}^{N} P(a_i)\lg(P(a_i))$
    - Analog: $H_{\mathrm{A}}(\phi_k) = -\int_S f(x)\lg(f(x))\mathrm{d}x$
  - Correlation between two states
    - Calculate the information gain ratio

$$R(\phi_i, \phi_j) = \frac{I(\phi_i, \phi_j)}{\min(H(\phi_i), H(\phi_j))} = \frac{H(\phi_i) - H(\phi_i|\phi_j)}{\min(H(\phi_i), H(\phi_j))}$$

Method (Cont'd)

- Modeling
  - Calculate the joint entropy of the system

$$H_U(\phi_1, \phi_2, \ldots, \phi_M) = \sum_{k=1}^{M} H(\phi_k) - \sum I(\phi_i, \phi_j)$$

  - Calculate the threshold

$$\bar{H}_U = H_U(1+\alpha),\ \alpha \in [0,1]$$

- Postprocessing
  - Dynamic threshold adjustment mechanism (DTAM)

$$\bar{H}_{U_n} = \begin{cases} \bar{H}_{U_0}, & \hat{\sigma} > V \\ (1 - \varepsilon\frac{\hat{\sigma}}{V})\bar{H}_{U_0}, & \hat{\sigma} \leq V \end{cases}$$

  - Abnormal state localization

Algorithm 1 Abnormal state localization method

Input: R, $\gamma(\phi)$, I, U
Output: Set of abnormal states Ω

```
for each state in γ(φ) do
    Flag ← 1
    for each substate in γ(φ) do
        if R(state, substate) > δ then
            Flag ← 0
        end if
    end for
    if Flag == 1 and state not in I then
        Add state to Ω
    end if
end for
for each state1 in γ(φ) do
    for each state2 in γ(φ) do
        if R(state1, state2) > δ and (state1, state2) not in U then
            Add state1 and state2 to Ω
        end if
    end for
end for
```

Major results

- Detection rate (DR) and false positive rate (FPR)

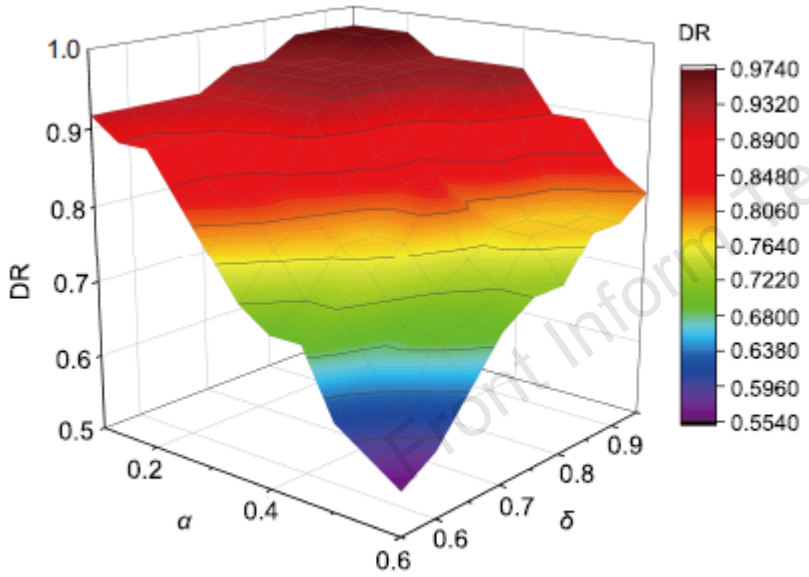


Fig. 7 Detection rate (DR) with different δ's and α's

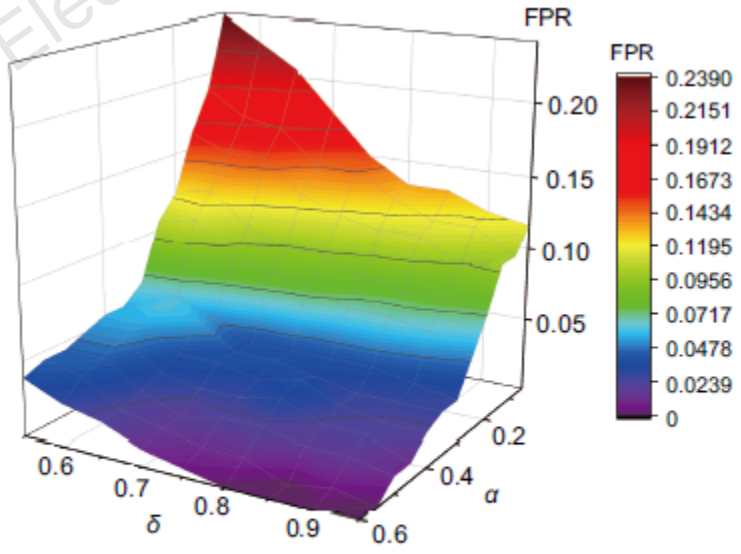


Fig. 8 False positive rate (FPR) with different δ's and α's

Major results (Cont'd)

- Impact of the calculation cycle

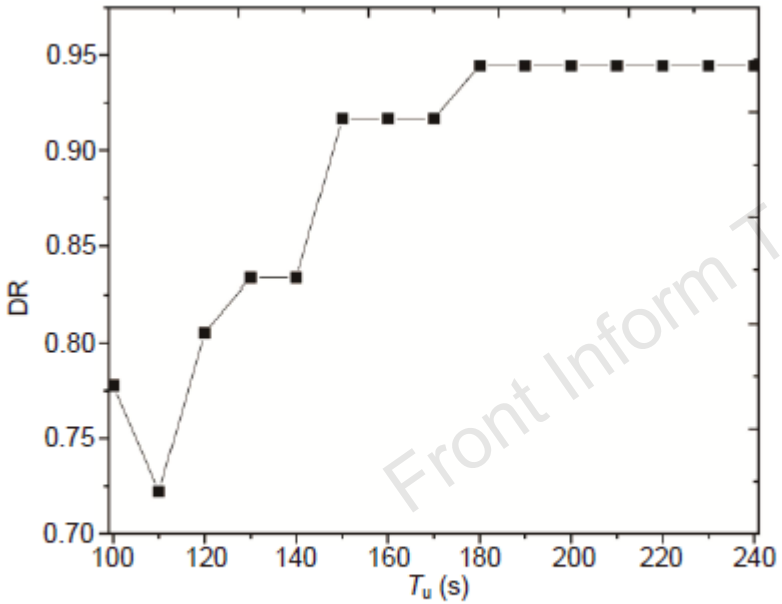


Fig. 10 The varying of the detection rate (DR) with different T_u's

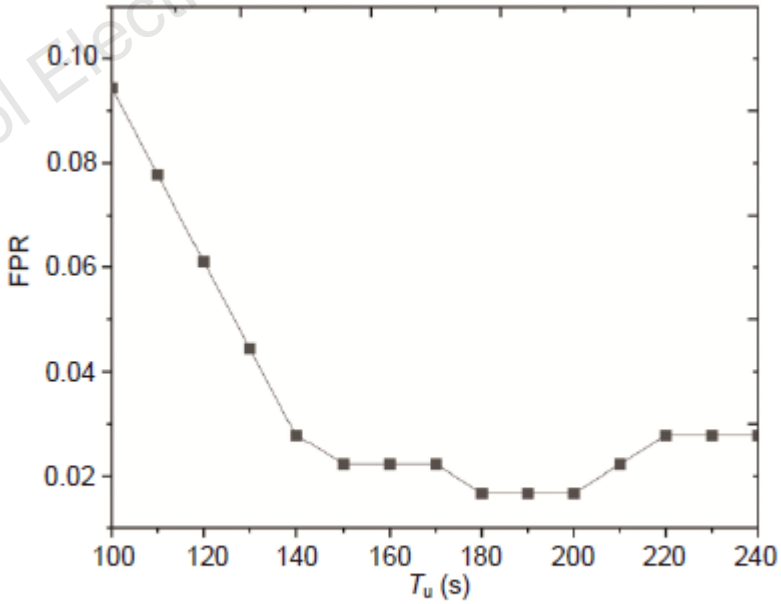


Fig. 11 The varying of the false positive rate (FPR) with different T_u's

Major results (Cont'd)

- Impact of the DTAM

Table 1 The varying of the detection rate (DR) with increasing ε

ε	DR without DTAM	DR with DTAM
0.04	94.44%	94.44%
0.08	94.44%	97.22%
0.12	94.44%	97.22%
0.16	94.44%	97.22%
0.20	94.44%	97.22%

DTAM: dynamic threshold adjustment mechanism

Table 2 The varying of the false positive rate (FPR) with increasing ε

ε	FPR without DTAM	FPR with DTAM
0.04	1.67%	1.67%
0.08	1.67%	1.67%
0.12	1.67%	2.22%
0.16	1.67%	3.33%
0.20	1.67%	6.11%

DTAM: dynamic threshold adjustment mechanism

Conclusions

- In this paper, we presented a novel EBID for intrusion detection by using COSs as input from the WTS. To improve the performance of EBID, we also proposed DTAM and the abnormal state localization method.

- EBID achieved a 97.22% detection rate and 1.67% false alarm rate.

- EBID can be extended to other industrial control systems (ICSs) by adding parsers and encoders of HMI data and network protocols.

Ke LIU is a fourth-year PhD candidate in the State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, China, and he received his BS degree in computer science from the State Key Laboratory of Mathematical Engineering and Advanced Computing in 2017. His research interests include ICS security, web security, and program analysis of binary code.

Qiang WEI received the PhD degree in computer science and technology from China National Digital Switching System Engineering and Technological Research Center, Zhengzhou, China. He is currently a professor with the State Key Laboratory of Mathematical Engineering and Advanced Computing. His research interests include network security, industrial internet security, and vulnerability discovery.