Data in Brief 14 (2017) 186–191

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset of anomalies and malicious acts in a cyber-physical subsystem

Pedro Merino Laso [a,*], David Brosset [a,b,**], John Puentes [a,c,**]

[a] *Chair of Naval Cyber Defense, École Navale - CC 600, F29240 Brest Cedex 9, France*
[b] *Naval Academy Research Institute, École Navale - CC 600, F29240 Brest Cedex 9, France*
[c] *Institut Mines-Télécom Atlantique, Lab-STICC CNRS UMR 6285, Equipe DECIDE, F-29238 Brest, France*

ARTICLE INFO

Article history:
Received 3 May 2017
Received in revised form
10 July 2017
Accepted 18 July 2017
Available online 20 July 2017

Keywords:
Anomaly
Cyber-physical system
Sensor data
Systems security

ABSTRACT

This article presents a dataset produced to investigate how data and information quality estimations enable to detect aNomalies and malicious acts in cyber-physical systems. Data were acquired making use of a cyber-physical subsystem consisting of liquid containers for fuel or water, along with its automated control and data acquisition infrastructure. Described data consist of temporal series representing five operational scenarios – Normal, aNomalies, breakdown, sabotages, and cyber-attacks – corresponding to 15 different real situations. The dataset is publicly available in the .zip file published with the article, to investigate and compare faulty operation detection and characterization methods for cyber-physical systems.

Specifications Table

Subject area	*Cyber-physical systems*
More specific subject area	*Anomaly detection and security*

[*] Corresponding author.
[**] Corresponding authors at: Chair of Naval Cyber Defense, École Navale - CC 600, F29240 Brest Cedex 9, France.
E-mail addresses: pedro.merino@ecole-navale.fr (P.M. Laso), david.brosset@ecole-navale.fr (D. Brosset), john.puentes@imt-atlantique.fr (J. Puentes).

http://dx.doi.org/10.1016/j.dib.2017.07.038

Type of data	*Raw signal measurements directly collected from a liquid storage and distribution cyber-physical subsystem, composed by one ultrasound depth sensor, four discrete sensors, two pumps, and a communication network*
How data was acquired	*A personal computer sounder recorded all the signals in synchroNous automatic mode, scanning every 0.1 s the system's programmable logic controller (PLC)*
Data format	*Comma separated values (CSV) files*
Experimental factors	*Fifteen situations were recorded separately including – Normal, blocked measures, floating objects on the liquid's surface, sensor failure, denial of service, spoofing, wrong connection, and hit of the tanks with different intensities – to illustrate five factual operational scenarios*
Experimental features	*Relations between dysfunctional components of the cyber-physical subsystem, operational scenario, and systemic effects are represented*
Data source location	
Data accessibility	*The data are available with this article. To access it open the index.html file included in the published .zip file*

Value of the data

- The dataset represents realistic sensors signals of a cyber-physical subsystem impacted by actual risks like aNomalies, sabotages, system breakdown, and cyber-attacks.
- The dataset can be used to validate detection and characterization algorithms for operational surveillance and security applications in cyber-physical systems.
- Included aNomalies and malicious acts can be studied to compare detection and characterization approaches for decision support.
- The dataset can be used to examine algorithms that assess data alteration and service degradation.

1. Data

The dataset contains 15 files of temporal series that represent 15 different situations related to 5 operational scenarios. Files' duration varies depending on the situation and dysfunctional component. Accordingly, affected components are two types of depth sensor, the underlying network, or the whole subsystem. These situations can be wrongly understood by a decision maker, or only identified for instance after the malicious act was accomplished. Since wrongly managed situations might have significant adverse operational costs, it is critical to detect and analyze in real time such events. Datasets covering such situations are currently rare, because of the complexity to acquire data from cyber-physical systems. In our case, the principle of reusable experimental platform [1] was applied, to collect diverse datasets for monitoring [2] and categorization of aNomalies [3].

2. Experimental design, materials and methods

Two tanks of different volumes that function as storage and distribution device for water or fuel, one ultrasound depth sensor, four discrete sensors, and two pumps, were used to acquire the dataset (Fig. 1). A computer controlled the system with a PLC connected to a monitoring network. The ultrasound depth sensor on the main tank (volume of 7 L) was calibrated relating the tank dimensions to 10,000 equidistant depth steps (0 corresponds to the full tank and 10,000 to the empty tank). Fig. 2 shows the tracked filling and emptying of the main tank. The four floating discrete sensors in the second tank (volume of 9 L), measured levels of liquid corresponding to four volumes: 1.25 L, 3.35 L, 8 L, and 9 L.

All signals – ultrasound depth sensor, pump 1, pump 2, and the four discrete level sensors – were acquired synchroNously for every situation described in Table 1, independently of the affected