Effectiveness of Association Rules Mining for Invariants Generation in Cyber-Physical Systems

Koyena Pal, Sridhar Adepu and Jonathan Goh
iTrust, Center for Research in Cyber Security
Singapore University of Technology and Design
Singapore
Email: adepu_sridhar@mymail.sutd.edu.sg, jonathan_goh@sutd.edu.sg

***Abstract*—Cyber-Physical Systems (CPS), which integrate controls, computing and physical processes are critical infrastructures of any country. They are becoming more vulnerable to cyber attacks due to an increase in computing and network facilities. The increase of monitoring network protocols increases the chances of being attacked. Once an attacker is able to cross the network intrusion detection mechanisms, he can affect the physical operations of the system which may lead to physical damages of components and/or a disaster. Some researchers used constraints of physical processes known as invariants to monitor the system in order to detect cyber attacks or failures. However, invariants generation is lacking in automation. This paper presents a novel method to identify invariants automatically using association rules mining. Through this technique, we show that it is possible to generate a number of invariants that are sometimes hidden from the design layout. Our preliminary study on a secure water treatment plant suggests that this approach is promising.**

***Index Terms*—Cyber Physical System, Association Rules Mining, Cyber Security, Attack Detection, Artificial Intelligence, cyber attacks, Secure Water Treatment testbed.**

I. Introduction

Cyber Physical Systems (CPS) are built from, and depend upon, the integration of computational algorithms and physical components. Such systems include large public infrastructure such as water treatment, oil and gas, and transportation. Rapid technology development of remote monitoring and controlling of the CPS opens vulnerabilities to attackers. Examples of cyber attacks [7], [12], [15] illustrate the importance of security in critical infrastructures. Given the potential and rising attempts of cyber attacks, it is necessary to design mechanisms for defending against such attacks.

Traditional intrusion detection systems use network traffic to monitor the CPS. However, once the attacker has breached the network layer, he is able to inject different attacks on the sensor and actuator communication channel [3], [2]. Monitoring physical behavior is necessary as it is the last layer of defence in critical infrastructures [4], [1]. In a water treatment plant, Adepu et al. [2] manually derived the physical invariants from the system design and used it for monitoring the plant. A process invariant, or simply an invariant, is a mathematical relationship among physical and/or chemical properties of the process controlled by the PLCs in a CPS [4]. In this paper, we derive these invariants using a data driven approach from a dataset [9] which is collected in an operational water treatment plant [14]. In essence, we use the concept of association rule mining to derive the invariants based on the sensor and actuator data of a Secure Water Treatment (SWaT) plant.

Related work: The literature related to attack detection in CPS can be divided into three categories: signature based, specification based, and behaviour based [13] techniques. Signature based approaches look for known patterns of the attacks [8]. However, this technique is not capable of detecting zero day attacks. Specification based approaches model the system behaviour using formal models which are mathematical models [11], [4]. It is hard to capture all system behaviour exactly using specification based approaches because, one needs human effort and his domain knowledge. In addition, hardware faults due to aging cannot be handled by specification based methods. In order to overcome these problems, data driven approaches like behavior based methods can be utilised. Behaviour based methods monitor the actual behaviour of the physical systems using the data obtained from the system [10]. This paper focuses on using a behaviour based approach to identify invariants.

Objective of the study: The study reported here was undertaken with the long term goal of developing robust defense mechanisms for CPS. As a short term goal, we automatically generated the invariants to detect cyber attacks [4] to support the work of Adepu et al. [4] . This study was performed on a SWaT testbed [14]. The results obtained and the methodology used in this study could serve as a basis for developing further investigation of approach to detect attacks and failures.

Research questions: Focus of the study described here is on the following questions where SWaT [14] is considered as a CPS:

RQ1: What is the procedure to automate the generation of invariants in a CPS?

RQ2: How do we verify the accuracy of the invariants generated?

This paper focuses on research questions RQ1 and RQ2 in detail.

The contributions of this paper is to present a methodology to generate invariants automatically. The rest of the paper is

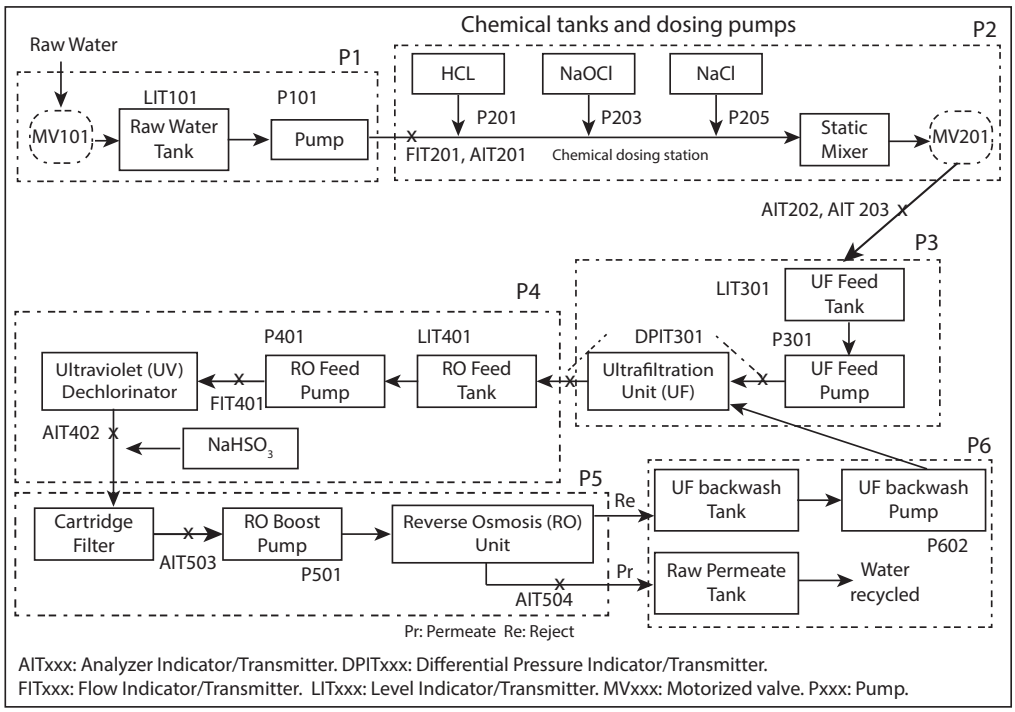

Fig. 1: Water treatment in SWaT: P1 though P6 indicate the six stages in the treatment process. Arrows denote the flow of water and of chemicals at the dosing station.

organised as follows: Section II discusses the SWaT dataset. The proposed method used for invariants generation is discussed in Section III. In section IV, we present the results of the experiments and conclude the paper in section V.

II. Architecture of the SWaT Testbed

SWaT [14] is a fully operational scaled down version of a water treatment plant. It is designed and built for research on the design of secure cyber physical systems. This testbed has a small footprint capable of producing 5 gallons/minute of doubly filtered water and mimics large modern plants for water treatment such as those found in modern cities.

Water treatment process: As illustrated in Figure 1, the treatment process used in SWaT consists of six distinct sub-processes termed P1 through P6. Each sub-process, referred to as a stage, is controlled by an independent Programmable Logic Controller (PLC). Thus, six PLCs work in concert to control the entire treatment process. Control actions are based on the system state estimated by the PLCs using data from sensors.

Stage P1 controls the inflow of water to be treated by the opening or closing of a motorized valve. State P2 is a chemical dosing station, while stage P3 is a Ultra Filtration (UF) process. A UF feed pump in P3 sends water via the UF unit to Reverse Osmosis (RO) feed water tank in stage P4. Here an RO feed pump sends water through an Ultraviolet dechlorination unit controlled by a PLC in stage P4. In stage P5, the dechlorinated water is passed through a 2-stage RO filtration unit. The filtered water from the RO unit is stored in the permeate tank and the rejected water is stored in the UF backwash tank. Stage P6 controls the cleaning of the membranes in the UF unit by turning on or off .

The communication infrastructure of the SWaT is presented in [3], it also describes what kind of attacks are possible in a water treatment system and the impact of the attacks.

A. SWaT Dataset

In this experiment, we used the SWaT Dataset [9]. The dataset was obtained by running SWaT non-stop from its empty state to a fully operational state for a total of 11-days. During this period, the first seven days consisted of normal operation, i.e. without any attacks. During the remaining days, 36 various cyber and physical attacks were launched on the testbed while data collection continued. The dataset [9] contains all the sensors and actuator values as well as the network traffic of the testbed over the said duration. The data which is collected from historian is almost 3 million lines which includes 53 attributes, we used complete data set in this paper. Till date, this is the most updated and complex open source dataset.

III. Methodology

To overcome the limitation of manually identifying invariants [4], we applied the association rules mining algorithm to the SWaT dataset. Essentially, this is a data mining process used to find rules that may govern associations and causal

TABLE I: Invariants based on design structure of the system

S.No	Description
SD1	MV101 is Open,FIT101 >delta
SD2	LIT101 is Low,MV101 is Open
SD3	LIT101 is High,MV101 is Close
SD4	LIT101 is LL (LowLow) P101 or P102 are Off
SD5	LIT301 is Low,P101 or P102 is On
SD6	LIT301 is High,P101 or P102 is Off

objects between sets of items [5]. Association rule mining works based on support and confidence.

Before explaining the general approach towards such data mining, there are two main terms that the reader needs to be associated with; *Minsup*,which indicates the minimum support i.e. the minimum number of times the items are found in the dataset and *Confidence value,* which shows the percentage of how many times a certain rule is found to be true.

A common strategy adopted by many association rules mining algorithms [6] to decompose the problem into two major sub tasks:

- Frequent Itemset Generation, whose objective is to find all the itemsets that satisfy the *minsup* threshold. These itemsets are called frequent itemsets.
- Rule Generation, whose objective is to extract all the high-confidence rules from the frequent itemsets found in the previous step.

Prior work [4] in identifying invariants were performed by applying the law of physics on the system design structure. However, if we applied an associative rules mining algorithm, we can identify various hidden relationships between different stages of a system that would have been tedious to find out if done manually. In order to see whether these rules are accurate enough to be set as conditions for a systems normal behaviour, there is a need to check if most of the constrains can be backed by Physics. If this succeeds, we can prove that a plausible way to detect cyber attack is through constraints generated by Association Rules Mining.

A. Procedure

Using the SWaT dataset[9], we generated constrains using the Apriori Algorithm. The Apriori Algorithm is an influential algorithm for mining frequent itemsets for boolean association rules. Figure **??** provides a brief visual representation of how the algorithm works. We refer the reader to [16] to understand Apriori more in detail.

As the the data consists of numerical values with some up to 9 decimal places, setting such values as itemsets are not only inaccurate but also difficult to understand as we did not know which number represents what sensor. In order to fit the data to the Apriori Algorithm accurately, we had to modify the dataset in a logical manner. We set ranges for sensor values such as High, Low, On and Off based on the current operation standard of the CPS in order to get more realistic constraints. i.e. to obtain a High range for sensor LIT101, the data must fall within 500 and 1000. We performed this procedure for all the sensor and actuators in SWaT.

Algorithm 1: Pseudo code for Apriori Algorithm[16]

Input:
D: transaction database;
Min_sup: the minimum support threshold
Output: frequent itemsets

Description:

1: L_1= find_frequent_1-itemsets(DB);
2: **for** (k=2; $L_{k-1} = \varphi; k++$) $\{$
3: C_k= Apriori_gen(L_{k-1});
4: **for each** transaction $t \in DB$ $\{$ //scan DB for counts
5: C_t = subset(C_k, t); //get the subsets of t that are candidates
6: **for each** candidate $c \in C_t$
7: $c.count++$;
8: $\}$
9: $L_k = \{c \in C_k | c.count \geq min\_sup\}$
10: $\}$
11: return $L = \bigcup_k L_k$;
12: Procedure Apriori gen(L_{k-1}: frequent$(k-1)$-itemsets)

IV. Results

After implementing the algorithm, we obtained about 11500 rules with all confidence values for 51 sensors. However, since we cannot discuss the effectiveness of each constrain due to space constraints, we are focusing on those sensors that affect P1 based on the invariants formed using the systems design structure. Tables I & II illustrates the invariants present in P1. We then compared the physical invariants with the automatically generated invariants (as shown in Table II) to see whether the generated constraints are valid.

In Table I, it is mentioned that when MV101 is Open, the value of FIT101 will be high. As illustrated in Table 2, the first row shows that this is 100% accurate. In order to check if the reverse is true, we wanted to see the confidence value when MV101 is Close and FIT101 is High. It came as 0.49%. This shows that this constraint is only seen in 0.49% of those itemsets when MV101 is Close. In Table I, this constraint is not shown. This shows that the confidence value for this rule is true and accurate as this value most probably came when the system was in a transition period.

TABLE II: Rules generated through apriori algorithm

Sensor 1	Sensor 2	Confidence Value (%)
MV101-Open	FIT101-H	100
MV101-Close	FTI101-H	0.49
MV101-Open	FIT101-L	0.012
MV101-Close	FIT101-L	99.5
LIT101-LL	P101-Off	100
LIT301-L	P101-On	95.7

Another instance can be taken where MV101 remains close. In the 4th row in Table II, it is shown that FIT101s value will most likely be Low. Similarly, if MV101 is Open, there is a negligible chance for FIT101 to be Low. This is evident based on the confidence value shown at 0.012%. The last row in Table II includes a sensor that does not primarily belong in P1 but affects the process as it takes in recycled water. According to the invariants set shown in Table I, LIT301-Low will result in P101-On. In Table II, this is shown to be true 95.7% of the time. Therefore, there is a match or coherence in the confidence values given for rules that have one sensor value controlled.

A. Limitations

Currently, this technique is limited to only pair wise sensors and actuators. In a CPS, all the sensors and actuators across multiple processes work in tandem. For example, P3 relies on P1 for water. This means that the actuators (P101 or P102) in P1 will be turned on. This action causes MV301 in P3 to be turned on. All these behavior is determined by the water level sensor LIT301 in P3. In this scenario, at least 3 sensors and actuators are working together to perform a request. Hence, for more accurate invariants generation, the technique adopted must be capable of deriving valid constrains across multiple sensors and actuators.

Another limitation of approach is to facing few false positives, false negatives. It is happening because at one state the invariant is true and same invariant is not true in other state. This paper is not checking the invariant based on state of system. We are interested to consider this aspect in the future work.

V. Conclusion

This paper presents invariants generation in cyber physical systems that is helpful to detect attacks and failures in CPS. We used association rule mining to generate invariants from a CPS dataset. As future works, we will overcome the limitations by focusing on identifying invariants across multiple processes. We will also include time as a bias to enhance the invariants generated.

VI. Acknowledgment

The authors wish to express their gratitude to Prof. Aditya Mathur and Kaung Myat Aung for invaluable assistance during this study. This work was supported by the National Research Foundation (NRF), Prime Ministers Office, Singapore, under its National Cybersecurity R&D Programme (Award No. NRF2014NCR-NCR001-40) and administered by the National Cybersecurity R&D Directorate

References

[1] S. Adepu and A. Mathur. Argus: An orthogonal defense framework to protect public infrastructure against cyber-physical attacks. In *IEEE Internet Computing Magazine, Cyber Physical Systems Security and Privacy (in press)*, 2016.

[2] S. Adepu and A. Mathur. Generalized attacker and attack models for cyber physical systems. In *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)*, volume 1, pages 283–292, June 2016.

[3] S. Adepu and A. Mathur. An investigation into the response of a water treatment system to cyber attacks. In *2016 IEEE 17th International Symposium on High Assurance Systems Engineering (HASE)*, pages 141–148. IEEE, 2016.

[4] S. Adepu and A. Mathur. Distributed detection of single-stage multipoint cyber attacks in a water treatment plant. In *the 11th ACM Asia Conference on Computer and Communications Security (in Press),*, May, 2016.

[5] R. Agrawal, T. Imieliński, and A. Swami. Mining association rules between sets of items in large databases. In *Acm sigmod record*, volume 22, pages 207–216. ACM, 1993.

[6] R. Agrawal, R. Srikant, et al. Fast algorithms for mining association rules. In *Proc. 20th int. conf. very large data bases, VLDB*, volume 1215, pages 487–499, 1994.

[7] P. Cobb. German steel mill meltdown: Rising stakes in the internet of things, 2015.

[8] W. Gao and T. H. Morris. On cyber attacks and signature based intrusion detection for modbus based industrial control systems. *The Journal of Digital Forensics, Security and Law: JDFSL*, 9(1):37, 2014.

[9] J. Goh, S. Adepu, K. Junejo, and A. Mathur. A dataset to support research in the design of secure water treatment plants. In *the 11th International Conference on Critical Information Infrastructures Security, (in Press),*, Oct, 2016.

[10] K. N. Junejo and J. Goh. Behaviour-based attack detection and classification in cyber physical systems using machine learning. In *Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security*, pages 34–43. ACM, 2016.

[11] E. Kang, S. Adepu, D. Jackson, and A. P. Mathur. Model-based security analysis of a water treatment system. In *In Proceedings of 2nd International Workshop on Software Engineering for Smart Cyber-Physical Systems (in press; SEsCPS'16)*, May 2016.

[12] R. Lipovsky. New wave of cyberattacks against Ukrainian power industry, January 2016. http://www.welivesecurity.com/2016/01/11.

[13] R. Mitchell and I.-R. Chen. A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys (CSUR)*, 46(4):55, 2014.

[14] SWaT: Secure Water Treatment Testbed, 2015. http://itrust.sutd.edu.sg/research/testbeds/.

[15] S. Weinberger. Computer security: Is this the start of cyberwarfare? *Nature*, 174:142–145, June 2011.

[16] M. J. Zaki, S. Parthasarathy, W. Li, and M. Ogihara. Evaluation of sampling for data mining of association rules. In *Research Issues in Data Engineering, 1997. Proceedings. Seventh International Workshop on*, pages 42–50. IEEE, 1997.