Journal of Scientific & Industrial Research
Vol. 80, April 2021, pp. 347-353

Cybersecurity by Prediction of Time Synchronization using Bayesian Base Gradient Descent Approach

Amutha Arunachalam[1]*, K Seetharaman[2] and Ashish Agarwal[1]

[1]Time & Frequency Metrology, National Physical Laboratory, New Delhi, India

[2]Dept. of Computer & Information Sciences, Annamalai University, India

Received 26 July 2020; revised 13 March 2021; accepted 22 March 2021

Time Commerce tends to struggle, which necessities an improved time framework.Legal escalations for conflicts of time commerce in the digital economy demand a solution that helps to address technology, standards, and policies. To meet the demand, we have to build a system that can understand every domain essential for building an inter-organizational system. "Date" and "Timestamp" reflect the root of the current term "Date Trade" in the cyber world. The threat to these roots has been studied in-depth and proposed solutions specific to UTC NPLI. The electricity grid shifts to the energy network to improve operating efficiency and reliability by developing advanced information and communication technology. However, the Internet also provides a range of entry points dependent on the internet, which produce additional vulnerabilities due to malicious cyber-attacks, thereby threatening Nations' economic health. This paper proposes therefore a new mechanism to protect critical infrastructure against these malicious attacks, based on interval state predictors. This paper uses the prediction-based approach for reducing the impact of such attacks from cyberspace. In prediction, we have used a machine learning approach like Bayesian classifier by Bayesian approach to forecasting time synchronization concerning universal time clock (UTC). In our analysis, we have taken the basic UTC, UTC, and UTC likelihood proposed approach on basis of communication. This work has improved considerably the results to take care of CPS against such cybersecurity threats.

Keywords: Big data, Cyber-physical systems**,** Indian standard time, Time dissemination**,** Timestamp

Introduction

A cyber-physical system (CPS) is often a network comprised of a framework of Cyber elements (— for example, control, and computational software) and physical components (— for example, sensors). Instances of CPS also contain the smart grid, the gas distribution network, and ad-hoc vehicle network.[1] CPS cybersecurity has received much attention in recent years.

Time Synchronization Attacks in Cyber-Physical Systems

We believe it is possible to describe the cyber-physical structure as follows:[2]

$$z_t = h(x_t) + e_t \quad \ldots (1)$$

Here, $z_t \epsilon \mathbb{R}^m$, $x_t \epsilon \mathbb{R}^n$, $e_t \epsilon \mathbb{R}^m$. Here, the output z_t represents the output signal of (1) produced at time t from initial system state, $h(x_t)$ is indeed a set of m nonlinear x_t functions, and e_t represents thenoise signal affecting the plant. In practise, e_t is usually believed to be a Gaussian zero-mean vector with known covariance $cov\ (e_t) = R$, and $h(x_t)$, a linear function also approximates the nonlinear functions.

$$z_t = C\hat{x}_t + e_t \quad \ldots (2)$$

Here, $\hat{x}_t = C^T R^{-1} C)^{-1} C^T R^{-1} z_t$ is usually estimated system state (at t time) and $C \epsilon \mathbb{R}^{mXn}$ is known as the *Jacobian* matrix of the system's topology.

CPS Under Attacks Model

From an opponent's viewpoint, we believe that the opponent seems to have the capacity to modify the dynamics either through the wireless channel or output signals. The linear time-invariant system descriptor in attack can indeed be defined, neglecting process nonlinearities and the noise present on the measurements:[2,3]

$$z_t = C\hat{x}_t + Da_t \quad \ldots (3)$$

Here, $D \epsilon \mathbb{R}^{m(mXn)}$. The attack signal t↦ $a_t \epsilon \mathbb{R}^{m+n}$ reflects a particular attack, $\bar{x}$ and $\bar{z}$ seem to be the device and measurement states that may be under

*Author for Correspondence
E-mail:amutha@gov.in

attack. TS attacks typically seek to alter the time measurement(s) for stamp(s) of byways of the insertion of a forged GPS signal. Observations with false time stamps are then sent to the control centre and eventually result in an inaccurate estimate of the state of the device. In the rest of this article, the time interval between the synchronized stamp and the existing timestamp is called d. Here, two traditional TS attack strategies are selected for analysis i.e. The DTS and the STS attack (Direct Time Synchronization and Stealth Time Synchronization), from the latest studies of TS attacks:

- ***DTS Attack:*** The attacker has the capacity to transfer a GPS spoofing signal to partially targeted TSM devices in this attack, which will impact certain time stamps on necessary measurements. The measurements index is impacted as shown in attacking set $K \subset \{1, \ldots., m\}$. For every i$\in K$, $(\bar{z}_t)_i = (\bar{z}_{t-d})_i$ and $(\bar{z}_t)_j = (\bar{z}_t)_j$ exists for all j$\notin$, where $0<d\leq t$.
- ***STS attack:*** With such an attack, we believe that the attacker will adjust the sample times of all measured data. $Z_t = (z_0, \ldots\ldots, z_{t-1})$ i.e. $\bar{z}=\bar{z}_{t-d}$ where $0 < d \leq t$.

Literature Survey

The literature brief shown in Table 1 is detailed at the end of this section. Humayed *et al*.[1] reported standardizing current CPS security research within a single structure. The structure comprises of 3 orthogonal coordinates: (1) we adopt some well-known taxonomy of risks, flaws, attacks, and controls from a security standpoint; (2) from viewpoint of the CPS elements, the analysis focuses on physical, hyperphysical, and cybercomponents; and (3) from the point of view of CPS frameworks, the researchers discuss specific CPS features and also descriptive systematic components. In a CPS framework, the model can be either generic to display specific component relationships, and detailed to collect any specifics when required. Zhang *et al*.[4] proposed a novel TSA: “Time Synchronization Attack” to attack smart grid timing information. As the synchronous measures are used by several smart grids but most measuring equipment is configured with GPS for accurate timing, it was extremely prone to attacks, the system of measurement by GPS spoofing.[5] For three smart grid PMU implementations, including monitoring the voltage stability, transmission line fault detection, event location, and efficacy of TSA is researched in the Paper. Through numerical simulations, the relevance of TSA is discussed.

Moussa *et al*.[6] checked the time synchronization frameworks for the power grid. The analysts evaluate their facets of security and study their drawbacks. The study addressed is the standardization initiatives to fulfil the timing requirements of the grid. Eventually, vulnerabilities identified by the study are identified and solutions for mitigation are suggested. Pasqualetti *et al*.[7] suggested a mathematical system for cyber-physical networks, monitors, and attacks; (ii) identified fundamental system-theoretical and graph-theoretical control constraints; and (ii) planned distributed and centralized control and detection of attacks. After this, through convincing illustrations, the investigators affirm the findings. Moussa[2] addresses the security implications of a main smart grid system enabling precise time synchronization. Synchronization of time is indeed an enormous necessity around the grid's realms, from its generation to the process of transmission, storage, and customer premises. To address the risk that is associated with PTP, the analyst focuses mostly on the substation, a fundamental part of the smart grid infrastructure, together with the recommended mechanism of time synchronization, the “Precision Time Protocol” (PTP) and recommends realistic and efficient identification, avoidance, mitigation strategies and techniques that harden and improve the protection and functionality of PTP in a substation. Fundamentals of stable clock synchronization theory were presented by Narula *et al*.[3] Precise clock synchronization is a cornerstone of power delivery control systems, database service activities, financial transactions, etc. Most of the clock synchronization (time transfer) models are centered around a single way of communication from master to slave clock, like in the satellite navigation systems. Wang *et al*.[8] designed a machine learning classifier denoted as "FDML i.e., first difference aware machine learning" to identify such attacks. Using the function of "first difference," adapted from statistics and economics, is the main principle underlying the classifier. IEEE 14-bus system simulations with real NYISO data have depicted that the FDML classifier can identify both TS attacks as well as other cyber-attacks efficiently. Zhang *et al*.[9] suggested a coordinated attack method known as ENFTA aimed at a time-based coordinated attack within a smart substation automation system network attack. Centered mostly around the system of time synchronization of smart sub-stations, this scheme

adds to it an anti-correction and differentiated time control system of a dispatcher. Youssef *et al.*[10] extensively explores the standards and implementations of IEC 61 850 technology and highlights essential security flaws it presents in the framework of existing cyber-physical grid integration. Barreto *et al.*[11] posed a cyber-attack with high-accuracy specifications on "Packet-Based Time Synchronization Protocols" (PBTSP). From the viewpoint of PBTSP, a cyber-intrusion is impossible to detect and utilizes a flaw that would be in the essence of all PBTSPs. Irrespective of the cryptographic protocol under which PBTSP is secured, it can be performed successfully so it is undetectable inside the target slave clock by the clock-servo algorithm. Experimentally considers the effects of TSSA i.e. time synchronizationspoofing attacks based on WAMPAC i.e. synchro phasor-based wide-area protection, monitoring, as well as the control applications. Almas *et al.*[12] Phase angle control, damping applications, and anti-islanding safetyfor power oscillation are being examined. TSSA is developed using an IRIG-B real-time (RT) power system and signal generator models are performed as hardware-in-the-loop with the help of an RT simulator with commercially available PMUs coupled with it. NTP (Network Time Protocol) is a network protocol that is used to synchronize networked computing devices with clocks. So it is important to ensure proper functioning and safety of computing devices.[13] NTP a client-server model is a solution as peer–to–peer technology, multicasting, and broadcasting guarantees all computer devices to operate at the same time. If some systems are out of synchronization, environments may not only face ongoing problems but may also pose huge risks to cybersecurity. When several devices are in use, each resource would assume the appropriate time would be different without time synchronization. An in-depth understanding of the recent advances in Machine learning for Networked CPS is explained in the article by Felix *et al.*[14] is one of the latest references to this article.

Proposed Approach

Proposed Work

In this paper, the emphasis is done on a time synchronization system that creates an impact on the cybersecurity system. In other words, time calibration between nodes improves the prevention capabilities of cyber-attack and improves cybersecurity.[13] This paper involves the analysis of cybersecurity prevention. In the prevention of cyber-attack, time calibration of the UTC and the framework plays an important role.[11]

Research Questions

In this paper, the emphasis is done on two research questions as mentioned below:

- Do UTC frameworks improve time synchronization between different systems or nodes of national different nodes/units?
- Time synchronizations improve the prevention and detection of a cybersecurity attack.

Proposed methodology

Step1: Deploy network which synchronizes by UTC system

Step2: Data of UTC collect in a database with synchronization stamp

Step3: Learning by Bayesian network and define a threshold for prediction.

Step4: Optimize Bayesian network prediction by gradient descent

Step5: Analysis of different parameters between the proposed and existing framework.

Table 1 — Literature Survey

Author's Name	Year	Methodology Used	Proposed Work
Humayed *et al.*[1]	2017	The study of orthogonal coordinates	The analysis and systematization of current CPS protection research within a single structure.
Zhang *et al.*[4]	2013	Smart Grid time synchronization attack	A new TSA is suggested to attack smart grid timing data.
Narula *et al.*[3]	2018	Synchronization of precise clocks	Maintains a basic principle of safe synchronization of clocks
Moussa *et al.*[6]	2016	Protection evaluation of processes for time synchronization	Summary of the time synchronization processes for the power grid
Wang *et al.*[8]	2017	Techniques of Machine Learning	In cyber-physical systems, detecting time synchronization attacks
Zhang *et al.*[9]	2019	A coordinated model of attack	Research on the method of organized attack security based on the intelligent substation scheme of global time synchronization
Felix *et al.*[14]	2021	Techniques of Machine Learning	Resilient Machine Learning for Networked Cyber-Physical Systems

Proposed framework

The proposed framework Fig.1 follows the subsequent modules:

- In the framework, the first module shows the network connected with two nodes A and B. Node A and B are connected with the fibre of the network. The network architecture assumes cyber-attack is feasible.
- Node A displays several sensors in the systems, the measurement phase, time delay, and the learning unit. For an application, sensors collect data. Computer processes compute the task, time delay analysis used for applying time stamp, and sent to a common time base UTC framework. However, in this learning section, the sub-module learns the dynamic waiting time and adds the overhead of delay time using Bayesian prediction, and tries to calibrate the B node by learning module and UTC framework.

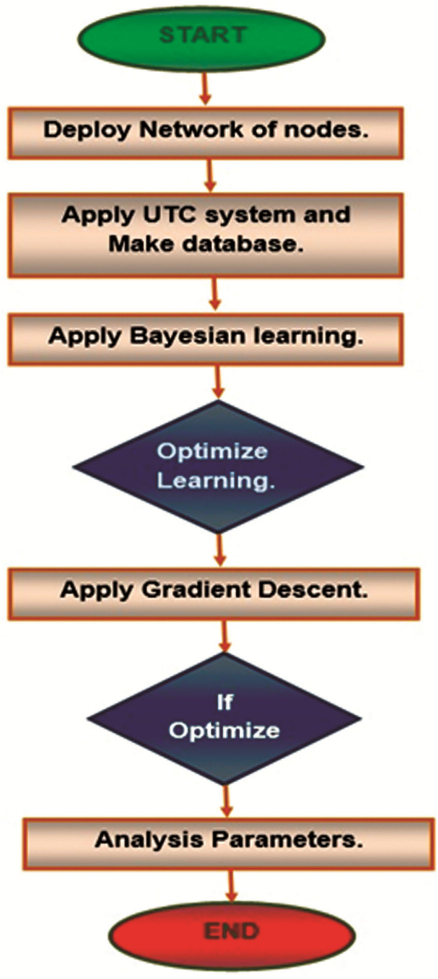

Fig.1 — Proposed Framework

Algorithm: Synchronize UTC by Optimization

Input: Plant Leaf dataset of Different number of Classes

Output: Optimize Network in Cyber Attack

Step0: Initialize Network parameters and Deployed Network.

Step1: Apply UTC Synchronize to every network node

Step2: N ⟵ Number of nodes

Step3: While (N>0)

Begin Apply Bayesian approach

P (X|Y) P (θ|D) =P (Y|X)P(X)P(Y)=P(D|θ)P(θ)P(D)

P (X|Y) P (θ|D) ⟵ after attack synchronize Prediction

Apply Gradient descent

For (i = 1 to P >0)

Begin

w = P – alpha * delta

if (W optimize)

goto step 4

else

goto step 2

END

END

Step4: Analysis Parameters

- Bayesian prediction finds the prior information by node A and calculates the maximum likelihood value. The maximum-likelihood value analysis or learn all previous values of time delay from different nodes and check dependency to UTC framework reference sources providing likelihood value based on the likelihood and maximum likelihood value. Besides, this forecasts a potential delay and synchronizes it accordingly.
- After the above process, the actual communication environment run and generate a graph explaining the relation between the systems and generate features of UTC and time value. After that, it implements the gradient descent, gradient descent approach analysis the value of the feature is a very small-time instance. If gradient descent converges quickly, it does not represent any simulated attack. In other cases, the maximum likelihood approach refers to identifying variance between time delay and estimating the deadline of time calibration. After the deadline time, if UTC synchronizes with the real-time, contact would be discarded accordingly.

Bayesian Approach

Let X1, Xn represent sampled 'n' observations from a probability density p(x | ✓). In this segment, we

compose p (x | ✓) if we see a random variable, and p (x | ✓) presents the conditional density of X dependent on ✓. In comparison, if we interpret it as a deterministic value, we write p✓(x). In Bayesian inference, the rule of Bayes becomes extremely strong. The relation happens when we begin to view Bayes' rule variables as parameters (θθ) of a system and observed data (DD):

P (X|Y) P (θ|D) =P (Y|X)P(X)P(Y)=P (D|θ)P(θ)P(D)

Gradient Descent

Gradient Descent is indeed the method of decreasing one function by adopting the cost function gradients. It includes understanding both the cost form and the derivative so that you can recognize the gradient from a given point and travel in that path, – for example downhill to the minimal value. We may use a common technology known asstochastic gradient descent in machine learning to reduce a model's error in the training samples Table 2. The way it would work is just by presenting each training case to a model onceat a timeperiod.The method made a prediction for such a training sample, the erroneous part is calculated and the model is revised to minimize the error during the next prediction. The whole procedure is used to discover the series of coefficients in such a model that also results in the lowest error mostly on thetraineddata for the model. Using the equation, each iteration of the coefficients, labelled weights (w) in the language of machine learning is updated:

w = w – alpha * delta

In which w is indeed the optimized weight or coefficient, alpha represents the rate of learning thatyou need to customize (– for example, 0.1) and gradient seems to be the model's error on the weight assigned training data.

Results Analysis

The work analyses the above given framework. For experiment Deploy Network with UTC base synchronization. For synchronization using Time stamp function an simulate attack on different number of nodes by stealth time synchronization attack. After attack simulate use prediction base approach using Bayesian approach and optimize prediction by gradient descent. These experiments done on 10 to 100 nodes and experiment run 1-10 time and cumulative results show in the graph. By such a framework, the diverse communication and the actual time should be calibrated between the sender and the receiver. This also analyses the difference between the actual time and UTC time, UTC with ML and UTC with Bayesian prediction.

In Fig. 2 analysis of the actual calibration time and time for the UTC is given. The disparity is just so much that the risk of cyber-attack increases. The UTC performance is based on the following reasons:

- UTC is one such framework that will increase the delay as it represents an intermediate communicational node. It increases the overhead and is not able to calibrate as per the actual time.
- Do not store delay preceding time or know because delay does not accommodate.

Analysis of the actual time and the UTC +ML calibration time is given in Fig. 3. Its difference is reduced as compared to only the UTC but still the cyber-attack probability increases. The UTC performance is based on the following reasons:

- UTC does not have any information of network domain but the maximum likelihood approach such as gradient descent approach does not change according to network performance, so it can reduce time gap between the actual and the predicted value at an initial time.

In Fig. 4 analysis of the actual time and the UTC + Bayesian approach calibration time is given. Its

Table 2 — Results on time synchronization between two nodes in UTC enabled network

Number of Communication	Actual time	UTC Framework	UTC+ maximum Likelihood	UTC+ Bayesian
10'	20	27.41666667	21.23	20
20'	30.13	32.9	29	29
30'	32.12	35.37666667	33.23	31.23
40'	36.45	37.37666667	37.34	35.23
50'	37.56	38.6	40	36
60'	38.12	39.82333333	38	37.12
70'	40.12	43.86	45.23	39.23
80'	41.23	50.48666667	52.12	40.12
90'	50.23	55.115	54	50
100'	60	60	60	60

difference is reduced as compared to only the UTC but it reduces the cyber-attack probability. The UTC performance is based on the following reasons:

- UTC does not have any information about the network domain but the Bayesian approach learns with the help of maximum likelihood and reduces the delay between the actual and predicted time value.

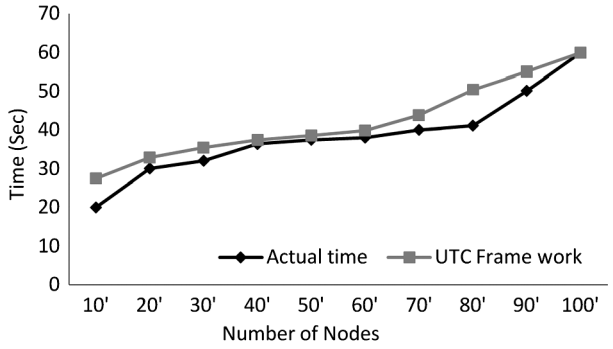

Fig. 2 — Comparison between Actual and UTC framework

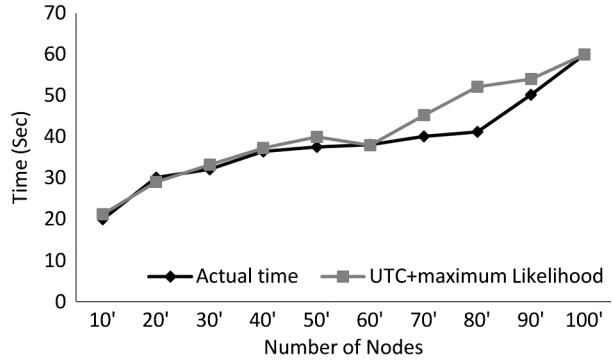

Fig. 3 — Comparison between Actual and UTC framework with ML

By result analysis following points conclude:

- Analysis The actual synchronization with three approaches first is physical resource UTC based its improve but not completely mainly its reduce performance in less number of nodes and high number of nodes
- In second approach ML or machine learning approach whichsignificantlyimprove in less number of nodes but not in high number of nodes
- In third Bayesian base proposed approach maximum improvement in high and low number of nodes

National Pyramid for UTC NPLI

For time can successfully resolve the demands of National Timing Echo System and provide the "Synchronised Traceable Time" and "Trusted Time Stamping" requirements of the country. The National Pyramid UTC NPLI model at Fig. 5 also known as IST identifies the traceability for national level timing infrastructure having Time and Timestamp functions and the organisations involved in timing activities.

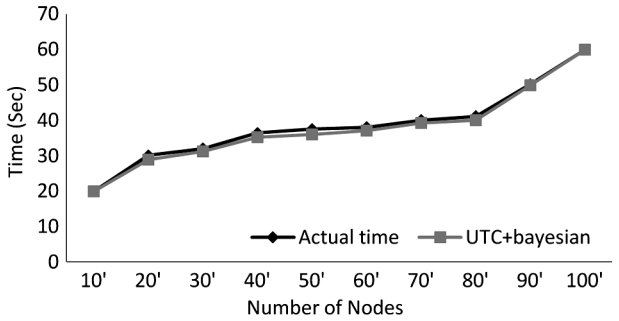

Fig. 4 — Comparison between Actual and UTC framework with a Bayesian approach

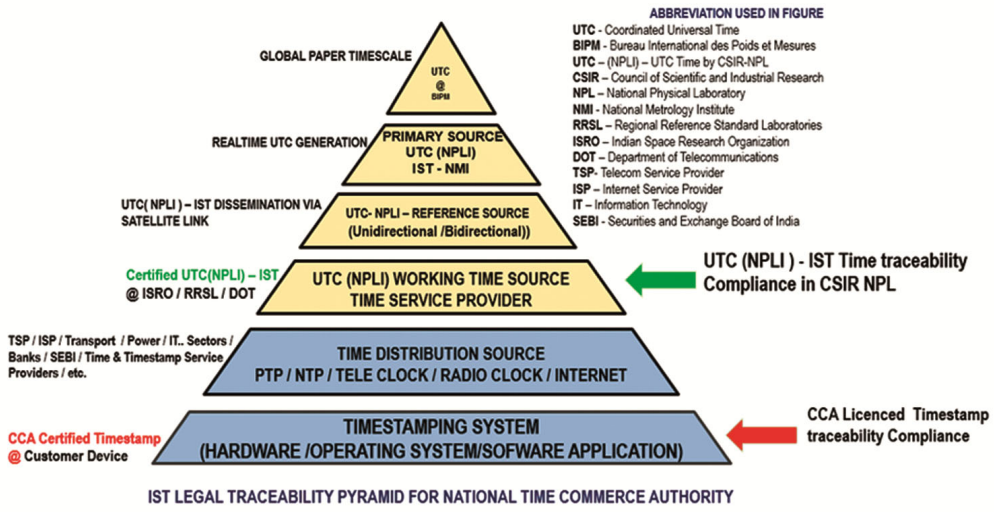

Fig. 5 — UTC NPLI also known as IST Legal Traceability Pyramid Model

Further the testing, certification and accreditation shall be done with reference to the proposed National pyramid. Under the proposed collaborative approach Timing Infrastructure that is being built shall provide trusted "Time Commerce" to government and the Citizens.

Conclusions

This paper has conceptualised Model to establish "National Time Commerce Authority" in consideration with the present legislation having the National Pyramid UTC NPLI also identified as Indian Standard Time IST. The requirements of inter-ministerial coordination were addressed when setting the national authority. The Time and Timestamp Services with proper certification program will help in attaining the trust to the customers. In this paper efforts have been attempted to enhance the performance of UTC by machine learning and AI based optimization approach. Machine learning Bayesian approach is used because of dependent synchronization relation between attacker and UTC synchronize value. In future studies, a Cyber Secure Architecture based Time Infrastructure is planned to be built for National Distribution of Time as per International Standards Quality and Cyber security.

References

1 Humayed A, Lin J, Li F & Luo B, Cyber-physical systems security—A survey, *IEEE Internet Things J*, **4(6)** (2017) 1802–1831.

2 Moussa B, Debbabi M & Assi C, Security assessment of time synchronization mechanisms for the smart grid—*IEEECommun Surv Tutor*, **18(3)** (2016) 1952–1973.

3 Pasqualetti F, Dorfler F & Bullo F, Attack detection and identification in cyber-physical systems —*IEEET AUTOMAT CONTR*, **58(11)** (2013) 2715 – 2729.

4 Zhang Z, Gong S, Dimitrovski A D & Li H, Time synchronization attack in smart grid: Impact and analysis —*IEEE T SMART GRID*, **4(1)** (2013) 87 – 98.

5 Psiaki ML, Humphreys T E & Stauffer B, Attackers can spoof navigation signals without our knowledge. Here's how to fight back GPS lies—*IEEE Spectr,* **53(8)** (2016) 26 – 53.

6 Moussa B, Debbabi M & Assi C, Security assessment of time synchronization mechanisms for the smart grid— *IEEE Commun Surv Tutor*, **18(3)** (2016) 1952 – 1973.

7 Pasqualetti F, Dorfler F & Bullo F, Attack detection and identification in cyber-physical systems —*IEEE Trans Automat Contr*, **58(11)** (2013) 2715 – 2729.

8 Wang J, Tu W, Hui L C, Yiu S M & Wang E K, Detecting time synchronization attacks in cyber-physical systems with machine learning techniques — *IEEE 37th ICDCS*, 2017, 2246 – 2251.

9 Zhang S, Cheng P, Wang B & Zhou X, Research on coordinated attack protection method based on global time synchronization system of an intelligent substation, *J Netw Comput Appl*, **4(1)** (2019) 14–20.

10 Youssef T A, El Hariri M, Bugay N & Mohammed O A, IEC 61850: Technology standards and cyber-threats —*EEEIC IEEE,* (2016) 1 – 6.

11 Barret S, Suresh A & Le Boudec J Y, Cyber-attack on packet-based time synchronization protocols: The undetectable delay box —*IEEET Instrum Meas* (2016) 1 – 6.

12 Almas M S, Vanfretti, Singh, R S & Jonsdottir G M, Vulnerability of synchrophasor-based WAMPAC applications to time synchronization spoofing — *IEEE Trans Smart Grid*, **9(5)** (2018) 4601 – 461.

13 The Implications of Network Time Protocol (NTP) for Cybersecurity —*https://www.cyber.nj.gov/alerts-advisories/ntp-and-its-implications-on-cybersecurity* (28-03-2021).

14 F O Olowononi, D B Rawat & C Liu, Resilient Machine Learning for Networked Cyber Physical Systems: A Survey for Machine Learning Security to Securing Machine Learning for CPS —*IEEE Commun SurvTutorials*, **23(1)** (2021) 524–552.