# Network-aware Mitigation of Undetectable PMU Time Synchronization Attacks

Ezzeldin Shereen and György Dán
Division of Network and Systems Engineering, School of Electrical Engineering and Computer Science
KTH Royal Institute of Technology, Stockholm, Sweden
E-mail: {eshereen, gyuri}@kth.se

***Abstract*—Time synchronization attacks are an emerging threat to many future smart grid applications, their mitigation is thus of utmost importance. In this paper we consider the problem of mitigating attacks that are undetectable by state-of-the-art power system state estimation, in precision time protocol networks. We formulate our problem as an integer linear program and show that it is NP-hard. We then provide a polynomial time approximation algorithm through a reduction from the group Steiner tree problem. We evaluate the performance of the proposed algorithm through extensive simulations compared to a greedy heuristic. Our results show that the approximation algorithm performs within a factor 1.8 of the optimal solution for synthetic topologies, while the greedy algorithm performs even better. On IEEE benchmark power systems the approximation algorithm performs within a factor 1.1 of the optimal solution, as good as the greedy heuristic.**

## I. Introduction

Phasor Measurement Units (PMUs) enable high-frequency timestamped phasor measurements for wide area monitoring and control (WAMPAC) applications in power systems. The high measurement frequency offers improved situational awareness compared to traditional measurement technology, and can be used for enhancing the performance of essential smart grid applications, such as islanding detection [1], oscillation monitoring [2], phase angle monitoring [3], event detection [4] and fault localization [5].

PMUs require precise time synchronization across large geographical areas. Time synchronization can be achieved using space-based or network-based synchronization. Space-based time synchronization relies on radio signals sent by satellites for delivering time references (e.g., GPS), whereas in network-based time synchronization the time references are transmitted by one or more *master* clocks through a packet switched network to each of the PMU clocks, which act as *slave* clocks. The state-of-the-art protocol for network-based time synchronization is the Precision Time Protocol (PTP) [6], and is expected to see wide-spread deployment for PMU time synchronization as a complement or as a replacement for space-based solutions.

Unfortunately, both space- and network-based time synchronization are vulnerable to time synchronization attacks (TSAs). Civilian GPS signals are unauthenticated and are thus vulnerable to GPS spoofing [7]. Besides, although message authentication is recommended in PTPv2.1 [8], it is not mandatory, rendering PTP also vulnerable to message spoofing [9]. A TSA against PMUs results in incorrect phase angle measurements, and can have a devastating effect on a variety of smart grid applications [10].

978-1-7281-6127-3/20/$31.00 ©2020 IEEE

Clearly, a trustworthy WAMPAC system should detect and mitigate TSAs against PMU measurements. An intuitive approach for detecting TSAs against PMUs would be to use bad data detection (BDD) techniques combined with linear state estimation (LSE) based on the PMU measurements [11]. Nevertheless, recent works have shown that it is possible to create TSAs that are undetectable by any BDD technique, both in theory [12] and under practical consideration of PMU clock servo constraints [13], and provided a computationally efficient method for identifying vulnerable sets of PMU measurements.

The existence of undetectable TSAs calls for mitigation in the form of securing time references. A cost-effective approach to mitigating TSAs would be to combine BDD with securing time synchronization to a subset of the PMUs so as to ensure that undetectable TSAs become infeasible. Securing time synchronization comes with significant investment and management cost, as network equipment need to be upgraded and key management has to be put in place. Ideally, the cost of securing time synchronization should be minimal, but identifying a least cost set of PMUs for mitigating undetectable TSAs is a non-trivial problem due to the coupling between the communication topology and the power system topology.

In this paper, we address this problem and make the following contributions. First, we formulate the problem of mitigating undetectable TSAs in a PTP network with minimum cost as an integer linear program. Second, we show that the problem is NP-hard and propose a polynomial time approximation scheme with bounded approximation ratio. Third, we evaluate the proposed solution using extensive simulations on synthetic network topologies and on IEEE benchmark power systems, compared to a greedy heuristic, and we show that significant cost savings are achievable by mitigating only practically undetectable TSAs.

The rest of the paper is organized as follows. Section II discusses the related work. Section III presents our model for time synchronization in power systems and for TSAs. We then formulate the problem of mitigating practically undetectable TSAs with minimum cost in Section IV, and propose an approximation algorithm for solving the problem in Section V. In Section VI we evaluate the performance of the proposed algorithm using extensive simulations. Finally, Section VII concludes the paper.

## II. Related work

Several recent works have considered the detection or mitigation of TSAs, for both space-based and network-based time

978-1-7281-6127-3/20/$31.00 ©2020 IEEE

synchronization. For detecting GPS spoofing against PMUs, [14] proposed utilizing the carrier-to-noise-ratio of the GPS signal. For PTP, [15] proposed introducing "guard clocks" in order to detect PTP delay attacks. Using a different approach, [16] proposed model-based and data-driven detectors based on the correlation between PMU phase angle changes and PMU clock offset adjustments. These works focus on detecting TSAs, but are typically prone to false negatives, i.e., they may not detect skillful adversaries.

Recent works on mitigating TSAs against GPS proposed to authenticate GPS messages [17]. Authors in [18] considered eight countermeasures for detecting and mitigating TSAs against GPS. Mitigation schemes proposed for PTP [19] focused strong message authentication schemes for PTP messages. However, these works do not consider the system level cost of TSA mitigation, and are thus cost-inefficient for power systems. In this paper we combine knowledge from the power system and the communication network to reduce the mitigation cost.

Similar in spirit to our work is [20], which considers network-aware mitigation of false data injection (FDI) attacks on power systems. Similar to previous works on FDI attacks against legacy SCADA measurements and PMUs [21][22], the mitigation schemes considered in [20] focus on protecting the integrity of measured data. On the contrary, in this work we focus on security controls for time synchronization, taking into account the impact of TSAs on measurement data integrity.

## III. System Model

### A. Power System and Network Model

We consider a WAMPAC system with $N$ buses and $M \geq N$ PMU measurements that ensure system observability. We denote by $z \in \mathbb{C}^M$ the vector of PMU measurements $z_m = |z_m|e^{j\theta_m}$, where $j = \sqrt{-1}$, and by $x \in \mathbb{C}^N$ the system state (voltage phasors at the buses). The measurement model is given by

$$z = Hx + e, \tag{1}$$

where $H \in \mathbb{C}^{M \times N}$ is the measurement matrix, and $e \in \mathbb{C}^M$ is white Gaussian measurement noise. We can define the verification matrix $F = H(H^\dagger H)^{-1}H^\dagger - I$, where $H^\dagger$ is the conjugate transpose of $H$, and use it for expressing the measurement residual $\hat{r} = Fz \in \mathbb{C}^M$, which is typically used for BDD, e.g., using the LNR test [23]. Bad measurement data identified by the BDD is typically discarded.

The measurements are taken by a set $\mathcal{T} = \{\tau_1, \ldots, \tau_T\}$ of PMUs, which rely on precise time synchronization. We denote by $M_i$ the number of measurements taken by PMU $\tau_i$. We consider delivering secure time references to a subset of these PMUs using PTP, and thus we model the communication infrastructure used for time synchronization in the WAMPAC by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{T} \cup \mathcal{N}$ is the set of vertices, and $\mathcal{N}$ is the set of network switches and routers. We consider that $\mathcal{G}$ is a tree that spans the PMUs $\mathcal{T}$. This assumption is motivated by that the active communication topology in power systems typically is a tree, and PTP networks use a tree topology as well for disseminating time references in a network [6]. Furthermore, we denote by $r \in \mathcal{V}$ the root vertex of the tree, which corresponds to where the PTP master clock is deployed.

### B. TSAs and Undetectablity

Consider now an attacker that is able to spoof PTP messages or the GPS signals, and can this way manipulate the time references of a subset $\mathcal{T}^a = \{\tau_1^a, \ldots, \tau_P^a\} \subseteq \mathcal{T}$ of PMUs, where $P$ is the number of manipulated time references (and PMUs). We define the attack-measurement matrix $\Psi \in \{0,1\}^{M \times P}$, which reflects the dependence of measurements on attacked time references, i.e., $\Psi_{m,p} = 1$ if measurement $m$ is measured by PMU $\tau_p^a$ and $\Psi_{m,p} = 0$ otherwise. Due to the attack the phasors measured by PMU $\tau_p^a$ will be rotated by $\alpha_p$. Thus, using the notation $u_p = e^{j\alpha_p}$ the $m^{th}$ measurement taken by PMU $\tau_p^a, p \in \{1, \ldots, P\}$ becomes $z_{p,m}^a = |z_{p,m}|e^{j(\theta_{p,m}+\alpha_p)} = z_{p,m}u_p$ , where $|z_{p,m}|$ is the measured magnitude, $\theta_{p,m}$ is the unattacked phase angle, and $\alpha_p$ is the angle shift introduced by the attack. Let $z^a$ be the measurement vector under the attack. If linear state estimation based on (1) is employed in the WAMPAC then a TSA should ideally be detected by BDD. It is thus natural to introduce the notion of undetectability as follows.

**Definition 1.** *A TSA against PMUs $\mathcal{T}^a$ is undetectable if it does not change the measurement residual, i.e., $Fz = Fz^a$.*

In what follows we recall fundamental results from [12; 13] about undetectable TSAs.

**Lemma 1.** *Consider a TSA against PMUs $\mathcal{T}^a$. The TSA is undetectable if and only if the vector $u \in \mathbb{C}^P$,, s.t., $u_p = e^{j\alpha_p}, p \in \{1, \ldots, P\}$ satisfies*

$$W_{\mathcal{T}^a}(u - 1) = 0, \tag{2}$$

*where $W_{\mathcal{T}^a} = \Psi^T diag(z)^\dagger F^\dagger F diag(z)\Psi$ is the complex attack angle matrix, $W_{\mathcal{T}^a} \in \mathbb{C}^{P \times P}$, and is Hermitian.*

**Lemma 2.** *Consider a TSA against PMUs $\mathcal{T}^a$, and $rank(W_{\mathcal{T}^a}) = 1$.*

- *If $P = 1$ then there is no undetectable TSA.*
- *If $P = 2$ then there is $1$ undetectable TSA.*
- *The pairs of time references that are vulnerable to undetectable TSAs for $P = 2$ form equivalence classes $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_C\}$, where $C = |\mathcal{C}|$ is the number of equivalence classes.*
- *For $P \geq 3$, let $\mathcal{C}' \subseteq \mathcal{C}_i \in \mathcal{C}$, $|\mathcal{C}'| = P$. Then there is a continuum of TSAs against time references $\mathcal{C}'$.*

The above result characterizes the set of vulnerable time references, and allows us to make two important observations. First, consider an equivalence class $\mathcal{C}_i$, $C_i = |\mathcal{C}_i|$. Then any subset of $3 \leq P \leq C_i$ time references in equivalence class $\mathcal{C}_i$ can be attacked in an undetectable manner. Furthermore, an undetectable attack against any subset of $P = 2$ time references can be constructed, but since there is only $1$ such attack, the attack can be detected by the BDD, as shown in [13].

## IV. Problem Formulation and Complexity Analysis

In this section we formulate the problem of mitigating undetectable TSAs at minimal cost.

### A. Minimum Cost TSA Mitigation Problem

We consider a power system operator that wants to upgrade its network infrastructure to mitigate TSAs. We assume that

LSE based on (1) is used with BDD, and hence the objective is to mitigate attacks against the collection $\{\mathcal{C}_1, \ldots, \mathcal{C}_C\}$ of equivalence classes of time references vulnerable to undetectable TSAs, as defined in the previous section.

We consider mitigation through authenticated time synchronization, e.g., using PTPv2.1. Securing the time reference of a PMU $\tau_t$ requires that a path in $\mathcal{G}$ from the root vertex $r \in \mathcal{V}$ to vertex $\tau_t$ has to be secured, including all intermediate vertices. Let $\mathcal{V}_{r\to\tau_t}$ be the set of vertices on the path from $r$ to $\tau_t$, including $r$ and $\tau_t$. In practice, the network equipment has to be upgraded to support authenticated PTP messages, and related key management. We thus define the cost of mitigation for a single time reference as the number of vertices $|\mathcal{V}_{r\to\tau_t}|$ on the path. For a set $\{\tau_1, \ldots, \tau_q\}$ of time references the cost is defined as the total number of vertices on the paths, i.e., $|\bigcup_{t=1}^{q} \mathcal{V}_{r\to\tau_t}|$. Before we present the problem formulation, we describe a naïve mitigation approach.

**Secure All (SA):** The straightforward way to mitigating TSAs would be to secure all time references $\bigcup_{i=1}^{C} \mathcal{C}_i$, which implies finding a tree in $\mathcal{G}$ rooted in $r$ that spans $\bigcup_{i=1}^{C} \mathcal{C}_i$. It is easy to see that securing these time references at minimal cost is equivalent to solving the minimum cost Steiner tree problem in graph $\mathcal{G}$ with the set $\bigcup_{i=1}^{C} \mathcal{C}_i \bigcup \{r\}$ as terminals.

**Minimum cost TSA Mitigation (MIN-TM):** While following the above approach mitigates TSAs, by Lemma 2 it is not necessary to secure the set $\bigcup_{i=1}^{C} \mathcal{C}_i$ of time references [13]. Instead, it is sufficient to secure $C_i - 2$ time references in each equivalence class $\mathcal{C}_i$ in order to mitigate practically undetectable TSAs, since a TSA against the remaining $P = 2$ time references would be detectable in practice. We can thus formulate the minimum cost TSA mitigation problem as follows.

**MIN-TM:** Consider the communication infrastructure graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a WAMPAC, root vertex $r \in \mathcal{V}$, and a collection of equivalence classes $\{\mathcal{C}_1, \ldots, \mathcal{C}_C\}$ s.t. $\mathcal{C}_i \subseteq \mathcal{V}, \forall i \in \{1, \ldots, C\}$, i.e., time references that are vulnerable to TSAs. Find a subtree $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*)$ of $\mathcal{G}$ with minimum $|\mathcal{V}^*|$ such that $r \in \mathcal{V}^*$ and $|\mathcal{V}^* \cap \mathcal{C}_i| \geq C_i - 2, \forall i \in \{1, \ldots, C\}$, where $C_i = |\mathcal{C}_i| \geq 3$.

Recall that an equivalence class $\mathcal{C}_i$ s.t. $C_i < 3$ does not allow to construct practically undetectable TSAs, hence we can assume $C_i \geq 3$.

### *B. Complexity Analysis*

Unfortunately, the MIN-TM problem formulated above is computationally hard, as we show next.

**Proposition 1.** *The MIN-TM problem is NP-hard.*

*Proof.* We prove the NP-hardness of the MIN-TM problem through reduction from the Group Steiner Tree (GST) Problem, which is known to be NP-hard [24].

**GST:** Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a root vertex $r \in \mathcal{V}$, edge costs $c_e \in \mathbb{Q}_{\geq 0}, \forall e \in \mathcal{E}$, and a collection of sets (groups) of terminal vertices $\mathcal{R} = \{\mathcal{R}_1, \ldots, \mathcal{R}_K\}$ s.t. $\mathcal{R}_i \subseteq \mathcal{V}, \forall i \in \{1, \ldots, K\}$. Find a subtree $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*)$ of $\mathcal{G}$ with minimum cost such that $r \in \mathcal{V}^*$ and $|\mathcal{V}^* \cap \mathcal{R}_i| \geq 1, \forall i \in \{1, \ldots, K\}$.

Given an instance of GST with a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, edge costs $c_e$, root vertex $r$, and terminal sets $\mathcal{R} = \{\mathcal{R}_1, \ldots, \mathcal{R}_K\}$, we construct an instance of MIN-TM with the graph $\mathcal{G}^m = (\mathcal{V}^m, \mathcal{E}^m)$, root vertex $r^m$, and equivalence classes $\mathcal{C}^m = \{\mathcal{C}_1^m, \ldots, \mathcal{C}_C^m\}$.

First, note that when $c_e = c, \forall e \in \mathcal{E}$, the objective of GST would be to minimize $|\mathcal{E}^*|$ which is equivalent to minimizing $|\mathcal{V}^*|$ (objective of MIN-TM) since in any tree $|\mathcal{V}^*| = |\mathcal{E}^*| + 1$. Second, observe that if GST is such that $|\mathcal{R}_i| = 3, \forall i \in \{1, \ldots, K\}$, and $c_e = 1, \forall e \in \mathcal{E}$ then we can set $\mathcal{G}^m = \mathcal{G}, r^m = r$, and $\mathcal{C}^m = \mathcal{R}$.

For the general case, we first set $\mathcal{G}^m = \mathcal{G}$ and $r^m = r$, and then for every $|\mathcal{R}_i| \neq 3$ we augment the graph $\mathcal{G}^m$ by adding three more vertices $\mathcal{V}_{it} = \{v_{i1}, v_{i2}, v_{i3}\}$, and adding an edge from each vertex in $\mathcal{R}_i$ to each vertex in $\mathcal{V}_{it}$. We denote the set of extra edges as $\mathcal{E}_{it}$. All $e \in \mathcal{E}_{it}$ are assigned the same cost, which is higher than the network diameter (the cost of the longest path in the network) in order to prevent introducing shortest paths that were not present in $\mathcal{G}$. Therefore, we set $\mathcal{G}^m = (\mathcal{V}^m, \mathcal{E}^m)$, where $\mathcal{V}^m = \mathcal{V} \cup \Big(\bigcup_{i:|\mathcal{R}_i|\neq 3} \mathcal{V}_{it}\Big)$ and $\mathcal{E}^m = \mathcal{E} \cup \Big(\bigcup_{i:|\mathcal{R}_i|\neq 3} \mathcal{E}_{it}\Big)$. Furthermore, we set $\mathcal{C}_i^m = \mathcal{R}_i, \forall i : |\mathcal{R}_i| = 3$ and $\mathcal{C}_i^m = \mathcal{V}_{it}, \forall i : |\mathcal{R}_i| \neq 3$ in the constructed MIN-TM, and hence all $\mathcal{C}_i^m$'s will be of cardinality 3.

Next, we compute the greatest common factor for all edge costs $c_e, \forall e \in \mathcal{E}^m$, i.e., the largest $c_{gcf}$ such that $c_e = n_e c_{gcf}, \forall e \in \mathcal{E}^m$, where $n_e \in \mathbb{N}$. Then any edge $e \in \mathcal{E}^m$ can be replaced by a series of $n_e$ interconnected edges, each with the same weight $c_{gcf}$. Note that this procedure also introduces the addition of $n_e - 1$ intermediate vertices to $\mathcal{V}^m$.

After constructing the MIN-TM instance, we can reconstruct the solution tree for GST from the solution tree of MIN-TM by reversing the changes done to $\mathcal{G}$ in the construction process. We replace the series of edges with costs $c_{gcf}$ by the original edges with costs $n_e c_{gcf}$, and eliminate the added vertices $\mathcal{V}_{it}$ and edges $\mathcal{E}_{it}$ that appear in the MIN-TM solution tree. □

**Proposition 2.** *The MIN-TM problem is equivalent to GST.*

*Proof.* We have already proved that we can construct a MIN-TM instance for any instance of GST. To prove that the two problems are equivalent, we thus need to prove the converse, i.e., that we can construct a GST instance for any instance of MIN-TM.

Given an instance of MIN-TM with a graph $G = (\mathcal{V}, \mathcal{E})$, root vertex $r$, and equivalence classes $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_C\}$, we construct an instance of GST with the graph $\mathcal{G}^s = \mathcal{G} = (\mathcal{V}, \mathcal{E})$, root vertex $r^s = r$, and edge costs $c_e^s = 1, \forall e \in \mathcal{E}$.

First, observe that if MIN-TM is such that $C_i = 3, \forall i \in \{1, \ldots, C\}$ then we can set $\mathcal{R}^s = \{\mathcal{R}_1^1, \ldots, \mathcal{R}_{K^s}^s\} = \mathcal{C}$ s.t. $K^s = C$. Otherwise, for every $\mathcal{C}_i$, s.t., $C_i > 3$ we create the collection $\binom{\mathcal{C}_i}{3} = \{\mathcal{C}_i^* : \mathcal{C}_i^* \subset \mathcal{C}_i, |\mathcal{C}_i^*| = 3\}$ of all $\binom{C_i}{3}$ subsets of $\mathcal{C}_i$ of cardinality 3, and append $\binom{\mathcal{C}_i}{3}$ to the collection $\mathcal{R}^s$.

To show that this transformation does not change the solution, we need to prove that protecting at least $C_i - 2$ time references in $\mathcal{C}_i$ is equivalent to protecting at least one time reference in each $\mathcal{C}_i^* \in \binom{\mathcal{C}_i}{3}$. We do that by showing that protecting less than $C_i - 2$ time references in $\mathcal{C}_i$ is equivalent to the existence of $\mathcal{C}_i^* \subseteq \mathcal{C}_i, |\mathcal{C}_i^*| = 3$ for which no time reference is protected. First, protecting less than $C_i - 2$ ($C_i - 3$ or less) time references clearly implies the existence of a triplet for which no time reference is

protected. The converse follows as well, as the existence of such a triplet implies that the total number of protected references in $\mathcal{C}_i$ is at most $C_i - 3$.

Using the previous transformation, the collection $\mathcal{R}^s = \{\mathcal{R}^s_1, \ldots, \mathcal{R}^s_{K^s}\}$ will then contain a total of $K^s = \sum_{i=1}^{C} \binom{C_i}{3}$ terminal sets, and one vertex per $\mathcal{R}^s_i, i \in \{1, \ldots, K^s\}$ has to be included in the solution to MIN-TM. The transformation is polynomial, and the solution of the constructed GST problem is identical to the solution of the original MIN-TM problem. □

After showing that MIN-TM is equivalent to GST, we can formulate MIN-TM as the following integer linear programming (ILP) problem, using the notation of the constructed GST problem from Proposition 2.

$$\begin{aligned} \min_{y} \quad & \sum_{e \in \mathcal{E}} y_e \\ \text{s. t.} \quad & \sum_{e \in \partial\mathcal{S}} y_e \geq 1, \quad \forall \mathcal{S} \subseteq \mathcal{V}, s.t., r^s \in \mathcal{S}, \exists i, s.t. \mathcal{R}^s_i \cap \mathcal{S} = \phi \\ & y_e \in \{0,1\}, \qquad \forall e \in \mathcal{E}, \end{aligned} \tag{3}$$

where $y_e$ is a decision variable that indicates whether or not edge $e$ will be part of the solution, and $\partial\mathcal{S}$ denotes the set of edges with exactly one end-point in $\mathcal{S}$.

## V. TSA Mitigation using Randomized LP Rounding

In what follows we present an approximation algorithm for MIN-TM with bounded approximation ratio, based on linear relaxation followed by randomized rounding.

### A. Compact ILP Formulation

A serious drawback of formulation (3) is that the number of constraints can be exponential, since it involves iterating over all possible subsets of $\mathcal{V}$. In the following we present an alternative formulation of MIN-TM with a polynomial number of constraints, initially proposed for GST [25]. This formulation is based on converting $\mathcal{G}^s$ into a directed graph and solving a max-flow problem. The first step in the conversion is to replace each undirected edge $e \in \mathcal{E}$ with two directed edges, one in each direction. Moreover, we add a set $\mathcal{V}_t = \{v_{t,1}, \ldots, v_{t,K^s}\}$ of $K^s$ additional vertices to $\mathcal{G}^s$. Vertex $r^s$ will be the source node of the flow, while the vertices in $\mathcal{V}_t$ will be the sink nodes. Next, we add a directed edge from each of the three vertices in $\mathcal{R}^s_i$ to its corresponding sink vertex $v_{t,i}$. We let the augmented directed graph be $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ where $\mathcal{V}' = \mathcal{V} \cup \mathcal{V}_t$, and $|\mathcal{E}'| = 2|\mathcal{E}| + 3K^s$, yielding the ILP

$$\begin{aligned} \min_{y} \quad & \sum_{e \in \mathcal{E}} y_e \\ \text{s. t.} \quad & \sum_{(i,l) \in \delta^+(i)} f^k_{il} - \sum_{(l,i) \in \delta^-(i)} f^k_{li} = d_{i,k}, \quad \forall k \in \mathcal{V}_t, i \in \mathcal{V}' \\ & f^k_{il} \leq y_e, \qquad \forall e = \{i,l\} \in \mathcal{E}, k \in \mathcal{V}_t \\ & f^k_{il} \geq 0, \qquad \forall (i,l) \in \mathcal{E}', k \in \mathcal{V}_t \\ & y_e \in \{0,1\}, \qquad \forall e \in \mathcal{E}, \end{aligned} \tag{4}$$

**Algorithm 1** Randomized Rounding

**input**: $y_e, \forall e \in \mathcal{E}, 0 \leq y_e \leq 1$.
**output**: $y'_e, \forall e \in \mathcal{E}, y'_e \in \{0,1\}$.

```
M ← φ
for e ∈ E do
    if p(e) = φ and η ~ U(0,1) < y_e then
        M ← M ∪ e
    else if p(e) ≠ φ and η ~ U(0,1) < y_e / y_{p(e)} then
        M ← M ∪ e
    end if
end for
for e ∈ E do
    if p*(e) ∩ M = p*(e) then
        y'_e ← 1
    else
        y'_e ← 0
    end if
end for
```

where $f$ is an extra decision variable corresponding to the flow from $r^s$ to each terminal node in $\mathcal{V}_t$ on each directed edge in $\mathcal{E}'$, $\delta^+(i)$ is the set of directed edges $(i,l), \forall l \in \mathcal{V}'$ originating from vertex $i$, $\delta^-(i)$ is the set of directed edges $(l,i), \forall l \in \mathcal{V}'$ terminating at vertex $i$, and $d_{i,k}$ is defined as

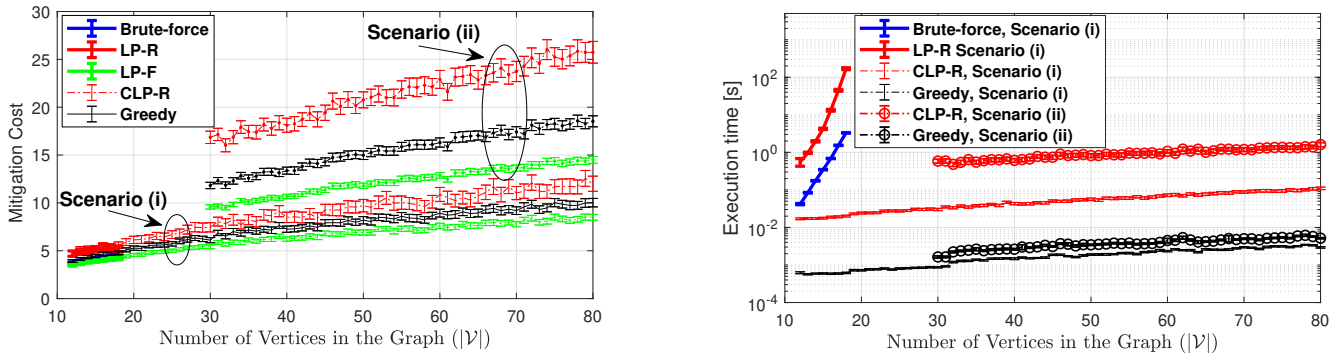
$$d_{i,k} = \begin{cases} 1, & i = r^s \\ -1, & i = k \\ 0, & i \in \mathcal{V}' \setminus \{r^s, k\}. \end{cases} \tag{5}$$

Although the resulting ILP has a polynomial number of constraints, it is still infeasible to optimally solve it for problem instances of practical interest. Nonetheless, it can serve as the basis for a polynomial time approximation, presented next.

### B. Approximation Algorithm

In what follows we present an approximation algorithm for MIN-TM based on a linear relaxation of (4) followed by randomized rounding of the fractional solution, originally proposed for GST [24]. The linear relaxation of (4) is obtained by replacing the constraint $y_e \in \{0,1\}, \forall e \in \mathcal{E}$ with the linear constraint $0 \leq y_e \leq 1, \forall e \in \mathcal{E}$. The LP can be solved in polynomial time.

Given the fractional solution, Algorithm 1 is executed $O(log|\mathcal{R}^s_{i^*}|\, ln\, 2K^s)$ times, taking the union of the resulting trees, where $i^* = argmax_{i \in \{1,\ldots,K^s\}} |\mathcal{R}^s_i|$. In Algorithm 1, $p(e)$ denotes the parent edge of an edge $e \in \mathcal{E}$ w.r.t. the root node $r^s$, $p^*(e)$ denotes the set of edges along the path from $e$ to $r^s$ with $e \in p^*(e)$, and $\eta \sim \mathcal{U}(0,1)$ is a standard uniform random variable. The algorithm marks an edge $e \in \mathcal{E}$ with a probability $\frac{y_e}{y_{p(e)}}$ if it has a parent edge and with probability $y_e$ otherwise, where $y$ is the optimal fractional solution of the LP. Next, $e$ is picked to be part of the solution ($y'_e = 1$) if $e$, as well as all its parent edges, were marked. Finally, for each set $\mathcal{R}^s_i$ that was not covered by Algorithm 1, we add the shortest path from a vertex $v \in \mathcal{R}^s_i$ to $r^s$ to the resulting tree.




(a) (b)

Fig. 1: Mitigation cost (a) and execution time (b) for synthetic graphs with either $C = 3$ and $C = 5$ equivalence classes.

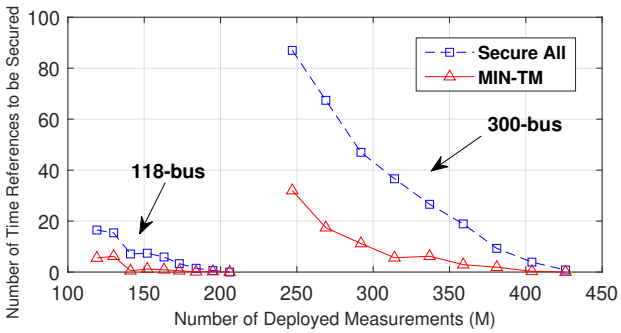



Fig. 2: The number of time references to be protected following *Secure All* and *MIN-TM*, computed for the IEEE 118-bus system and the IEEE 300-bus system.

**Proposition 3.** *The proposed approximation algorithm provides a* $O(logC + 3\,logC_{i'})$ *approximation for MIN-TM, where* $i' = argmax_{i\in\{1,\dots,C\}} C_i$.

*Proof.* LP relaxation of Problem (4), followed by randomized rounding as in Algorithm 1 provides an approximation ratio bound of $O(log^2|\mathcal{V}|\,logK^s)$ for GST, where $K^s$ is the number of groups [24]. If $\mathcal{G}$ is a tree then the bound is $O(log|\mathcal{R}^s_{i^*}|\,logK^s)$ [24]. In the case of MIN-TM there are $K^s = \sum_{i=1}^{C}\binom{C_i}{3}$ groups, each consisting of 3 time references. Thus the approximation ratio bound becomes $O\left(log\left(\sum_{i=1}^{C}\binom{C_i}{3}\right)\right)$. Furthermore, since $\binom{C_i}{3}$ is bounded by $C_i^3$, the approximation ratio bound becomes $O(log(CC_{i'}^3)) = O(logC + 3\,logC_{i'})$, where $i' = argmax_{i\in\{1,\dots,C\}} C_i$, which proves the result. □

## VI. Numerical Results

In the following we show results from extensive simulations on synthetic topologies and on IEEE benchmark power systems. All simulations were carried out on a notebook with Intel Core i7-8550 CPU @ 1.8 GHz with 16 GB of RAM.

Throughout the section we consider four algorithms for solving MIN-TM. The first one is a **brute-force** exponential search over all combinations of edges yielding the optimal solution. The second one is **LP Rounding (LP-R)**, in which we solve the linear relaxation of ILP (3), and then perform randomized rounding (Algorithm 1) on the solution. The third one is **Compact LP Rounding (CLP-R)**, in which we solve the linear relaxation of ILP (4), and perform the same randomized rounding on the solution. The fourth algorithm is a **Greedy** heuristic, in which we compute the shortest path from each vertex in each equivalence class $\mathcal{C}_i$ to $r$, and then include the shortest $C_i - 2$ paths per class in the resulting tree, as long as a class $\mathcal{C}_i$ was not already covered by an earlier shortest path, resulting in at most $\sum_{i=1}^{C} C_i - 2$ shortest paths.

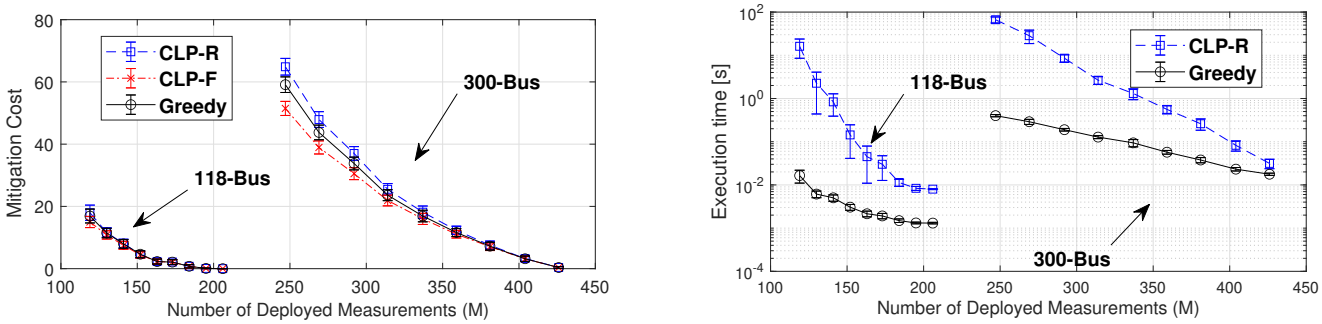
### A. Mitigation Cost for Synthetic Graphs

We start with showing results for synthetic graphs. To generate random tree graphs with $|\mathcal{V}|$ vertices, we choose $|\mathcal{V}| - 1$ edges randomly from all $\binom{|\mathcal{V}|}{2}$ edges, such that the chosen edges ensure the connectivity of the graph. We then choose the root vertex $r$ to be the vertex with the highest betweenness-centrality in $G$. Next, we randomly choose $C$ disjoint subsets of $\mathcal{V}$ as the vulnerable equivalence classes $\mathcal{C}$ for this graph. The equivalence class cardinalities $C_i$ are chosen uniform at random, as described below. We then solve the resulting MIN-TM problem using the four considered algorithms.

Fig. 1 shows the mitigation cost (Fig. 1a) and the average execution time (Fig. 1b) for each of the four algorithms. We also show the costs of the fractional solutions of linear relaxation of ILP (3) and ILP (4), denoted by LP-F and CLP-F, respectively. We show results for two equivalence class scenarios (i) $C = 3$ with $C_i \sim \mathcal{U}(3,4), i \in \{1,2,3\}$, and (ii) $C = 5$ with $C_i \sim \mathcal{U}(3,5), i \in \{1,\dots,5\}$. The network size $|\mathcal{V}|$ ranged from 12 to 80 vertices for scenario (i) and from 30 to 80 for setting (ii). Each point on the curves is the average of 200 simulations of different graphs with the same parameters, bars indicate the 95% confidence intervals.

Fig. 1a shows that both non-compact and the compact LP yield the same optimal cost, but their rounding doesnot necessarily yield the same solution. This manifests in that the curves for the fractional solution cost (LP-F and CLP-F) coincide, while those for the rounded solutions (LP-R and CLP-R) do not always coincide, due to randomized rounding.

The results show that the cost obtained by the approximation algorithms (LP-R and CLP-R) is within a factor of 1.5 of the fractional optimal solution for scenario (i), and within a factor of 1.8 for scenario (ii). As suggested by the approximation ratio bound, the approximation solution becomes worse as $C$ and $C_i$ increase. Also, interestingly, the greedy algorithm outperforms the approximation solutions for the synthetic graphs, despite not having an approximation ratio bound.




(a) (b)

Fig. 3: Mitigation cost (a) and execution time (b) for the IEEE 118-bus and the IEEE 300-bus systems.

Furthermore, we observe from Fig. 1b that the execution time of the brute-force algorithm and of LP-R increase exponentially with the network size, as expected, while the execution time of CLP-R and the greedy algorithm increase polynomially with the network size. The exponential increase in execution time for the brute-force algorithm and LP-R made it infeasible to execute them for scenario (ii) and larger networks in scenario (i). Among the polynomial time algorithms, the execution time of CLP-R was consistently one to two orders of magnitude higher than that of the greedy heuristic.

### *B. Securing IEEE Benchmark Power Systems*

We now turn to the evaluation on IEEE benchmark power systems. We used power system topology information in the MATPOWER package [26] in Matlab for computing the set of vulnerable time references and the equivalence classes $\mathcal{C}$ (and hence, $C$ and $C_i$) in the 118 bus and 300 bus IEEE benchmark power systems, by computing the index of separation (IoS*), introduced in [12] for identifying undetectable attacks.

**Vulnerability vs. Practical Undetectability:** As a first step, we assess the potential benefit of the proposed *MIN-TM* approach compared to *Secure All*. We do so by computing the number of time references to be secured following *Secure All* (i..e, $\sum_{i=1}^{C} C_i$), and following *MIN-TM* (i.e., $\sum_{i=1}^{C} C_i - 2$), as discussed in Section IV-A. Fig. 2 shows the results for the IEEE 118-bus system, and the IEEE-300 bus system. Each point on the curves is the average from 100 simulations, each corresponding to a different deployment of $M$ voltage or current injection measurements. The figure shows that the number of time references to be secured is significantly lower when mitigating only undetectable TSAs, except for very sparse PMU deployments, where the equivalence classes are large but few, and for very dense deployments, where TSAs are not feasible, i.e., $C = 0$. We can observe that the number of time references to be secured decreases with the number of phasor measurements, which is due to that the number of feasible TSAs decreases. Note that having many phasor measurements incurs a cost for the operator, and as such Fig. 2 illustrates a trade-off between the cost of deployment and the potential cost of securing PMU measurements against TSAs.

**Mitigation Cost on IEEE Benchmark Systems:** Next, we show results for solving the MIN-TM problem on the 118 bus and 300 bus IEEE benchmark power systems. For the computations we chose $\mathcal{V}$ to be the set of buses in the power system, and the set $\mathcal{E}$ to be a subset of connections between buses, such that $|\mathcal{E}| = |\mathcal{V}| - 1$, and $\mathcal{E}$ ensures the connectivity of the graph (i.e., $\mathcal{G}$ is a tree). We then chose $r$ to be the vertex with the highest betweenness centrality in $\mathcal{G}$. Furthermore, we set the set $\mathcal{T}$ of time references to be the buses (vertices) that have a PMU installed. In a practical deployment, each bus with a PMU may correspond to multiple vertices, e.g., one for the PTP switch in the corresponding substation, and one for the PMU itself, but this simplifcation does not change the solution to the corresponding MIN-TM problem.

Fig. 3a shows the mitigation cost achieved by the *CLP-R* and *Greedy* algorithms on the IEEE 118-bus and the IEEE 300-bus systems. Each point shows the average for 100 different deployments of $M$ voltage or current injection phasor measurements, along with corresponding 95% confidence intervals. We can observe that the proposed CLP-R algorithm performs close to optimal, as it achieves a mitigation cost that is within a factor of 1.04 to the optimal fractional solution (shown as CLP-F in the figure) for the IEEE 118-bus system, and within a factor of 1.1 to CLP-F for the IEEE 300-bus system. Moreover, the greedy algorithm performs very closely to CLP-R. One explanation for the better performance of CLP-R in real power systems can be that time references in an equivalence class are typically close to each other in the graph, which CLP-R can efficiently leverage.

Fig. 3b shows the execution time of the algorithms. The figure shows that the execution time decreases as $M$ increases, This is due to that the number and size of the equivalence classes (thus the number of triplets $K^s$) decreases with $M$, and both the size of ILP (4) and the number of paths established by the greedy algorithm depend on $K^s$ .

Our results show that the proposed approximation algorithm performs close to optimal. Although it is outperformed by the greedy algorithm on synthetic graphs, it achieves excellent performance on benchmark IEEE power systems on average, while providing a worst case performance guarantee.

## VII. Conclusion

In this paper we considered the problem of mitigating time synchronization attacks against PMU-based state estimation in power systems using PTP. We formulated the problem of upgrading the minimum number of network equipment for mitigating the attacks, and we showed that the problem is NP-hard by reduction from the group Steiner tree problem. We presented an approximation algorithm with bounded approximation ratio,

and compared it to a greedy heuristic. Our results show that the greedy algorithm performs better for synthetic graphs, but the approximation algorithm performs equally good for IEEE benchmark power systems. The results show that the joint use of secure PTP and linear power system state estimation using PMU measurements can be promising in cost-efficient mitigation of time synchronization attacks in future power systems.

## ACKNOWLEDGEMENT

This work was partly funded by the Swedish Civil Contingencies Agency (MSB) through the CERCES project.

## REFERENCES

[1] Y. Guo, K. Li, D. M. Laverty, and Y. Xue, "Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches," *IEEE Trans. on Power Delivery*, vol. 30, no. 6, pp. 2544–2552, 2015.

[2] G. Liu, J. Quintero, and V. M. Venkatasubramanian, "Oscillation monitoring system based on wide area synchrophasors in power systems," in *iREP Symposium - Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability*, 2007, pp. 1–13.

[3] A. Xue, S. Leng, Y. Li, F. Xu, K. E. Martin, and J. Xu, "A novel method for screening the PMU phase angle difference data based on hyperplane clustering," *IEEE Access*, vol. 7, pp. 97 177–97 186, 2019.

[4] T. Xu and T. Overbye, "Real-time event detection and feature extraction using PMU measurement data," in *Proc. of IEEE SmartGridComm*, 2015, pp. 265–270.

[5] M. Jamei, A. Scaglione, and S. Peisert, "Low-resolution fault localization using phasor measurement units with community detection," in *Proc. of IEEE SmartGridComm*, 2018, pp. 1–6.

[6] *1588-2019 - IEEE Approved Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, 2019 (accessed June 1, 2020). [Online]. Available: https://standards.ieee.org/content/ieee-standards/en/standard/1588-2019.html

[7] S. Gong, Z. Zhang, M. Trinkle, A. D. Dimitrovski, and H. Li, "GPS spoofing based time stamp attack on real time wide area monitoring in smart grid," in *Proc. of IEEE SmartGridComm*, 2012, pp. 300–305.

[8] E. Shereen, F. Bitard, G. Dán, T. Sel, and S. Fries, "Next steps in security for time synchronization: Experiences from implementing IEEE 1588 v2.1," in *Proc. of IEEE ISPCS*, 2019.

[9] M. Han and P. Crossley, "Vulnerability of IEEE 1588 under time synchronization attacks," in *IEEE PES General Meeting (PESGM)*, 2019, pp. 1–5.

[10] M. S. Almas, L. Vanfretti, R. S. Singh, and G. M. Jonsdottir, "Vulnerability of synchrophasor-based WAMPAC applications' to time synchronization spoofing," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 4601–4612, 2018.

[11] H. Zhao, "A new state estimation model of utilizing PMU measurements," in *Intl. Conf. on Power System Technology*, 2006, pp. 1–5.

[12] S. Barreto, M. Pignati, G. Dán, J. Le Boudec, and M. Paolone, "Undetectable PMU timing-attack on linear state-estimation by using rank-1 approximation," *IEEE Trans. on Smart Grid*, 2016.

[13] E. Shereen, M. Delcourt, S. Barreto, G. Dán, J. Le Boudec, and M. Paolone, "Feasibility of time-synchronization attacks against PMU-based state estimation," *IEEE Trans. on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3412–3427, 2020.

[14] Y. Fan, Z. Zhang, M. Trinkle, A. D. Dimitrovski, J. B. Song, and H. Li, "A cross-layer defense mechanism against GPS spoofing attacks on PMUs in smart grids," *IEEE Trans. on Smart Grid*, vol. 6, no. 6, 2015.

[15] B. Moussa, M. Debbabi, and C. Assi, "A detection and mitigation model for PTP delay attack in an IEC 61850 substation," *IEEE Trans. on Smart Grid*, vol. 9, no. 5, pp. 3954–3965, 2018.

[16] E. Shereen and G. Dán, "Model-based and data-driven detectors for time synchronization attacks against PMUs," *EEE J. Sel. Areas Commun. (JSAC)*, vol. 38, no. 1, pp. 169–179, 2020.

[17] S. Bhamidipati, T. Y. Mina, and G. X. Gao, "GPS time authentication against spoofing via a network of receivers for power systems," in *IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2018, pp. 1485–1491.

[18] L. Heng, J. J. Makela, A. D. Domínguez-García, R. B. Bobba, W. H. Sanders, and G. X. Gao, "Reliable GPS-based timing for power systems: A multi-layered multi-receiver architecture," in *Power and Energy Conf. at Illinois (PECI)*, 2014, pp. 1–7.

[19] E. Itkin and A. Wool, "A security analysis and revised security extension for the precision time protocol," *IEEE Trans. on Dependable and Secure Computing*, vol. 17, no. 1, pp. 22–34, 2020.

[20] O. Vukovic, K. C. Sou, G. Dán, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *IEEE J. on Sel. Areas Commun. (JSAC)*, vol. 30, no. 6, pp. 1108–1118, 2012.

[21] G. Dán and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *IEEE SmartGridComm*, 2010, pp. 214–219.

[22] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. on Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[23] A. Abur and A. G. Expósito, "Power system state estimation : Theory and implementation." Marcel Dekker, 2004.

[24] N. Garg, G. Konjevod, and R. Ravi, "A polylogarithmic approximation algorithm for the group steiner tree problem," in *Proc. of ACM Symposium on Discrete Algorithms (SODA)*, 1998, p. 253–259.

[25] M. X. Goemans and Y. Myung, "A catalog of steiner tree formulations," *Networks*, vol. 23, pp. 19–28, 1993.

[26] R. D. Zimmerman and C. E. Murillo-Sanchez, "(2019). MATPOWER (version 7.0) [Software]," available: https://matpower.org.