EPFL

Thèse n° 8577

Time-Synchronization Attacks against Critical Infrastructures and their Mitigation

Présentée le 21 septembre 2021

Faculté informatique et communications
Laboratoire pour les communications informatiques et leurs applications 2
Programme doctoral en informatique et communications

pour l'obtention du grade de Docteur ès Sciences

par

Marguerite Marie Nathalie DELCOURT

Acceptée sur proposition du jury

Prof. A. Lenstra, président du jury
Prof. J.-Y. Le Boudec, directeur de thèse
Prof. A. Abur, rapporteur
Dr V. Lenders, rapporteur
Prof. M. Paolone, rapporteur

École polytechnique fédérale de Lausanne

2021

There are things known and there are things unknown,
and in between are the doors of perception.
— Aldous Huxley

To all my loved ones.

Acknowledgements

First and foremost, I would like to thank Prof. Le Boudec for his help and support during this PhD. It has been a real honour to learn from him. His unwavering logic and calmness have been essential in my remaining focused and centered.

Second, I would like to express my sincere gratitude to Dr. Lenders, Prof. Abur, Prof. Paolone, Prof Le Boudec and Prof. Lenstra for accepting to serve in my PhD thesis committee. Thank you for your time and support through this final step of the PhD.

Third, I would like to thank my colleagues from KTH, Ezzeldin Shereen and György Dàn for our fruitful collaborations. I really appreciated our brainstorming sessions. I am also grateful to Prof. Paolone from the DESL lab for ensuring the correctness of the electrical engineering aspects of this work. His knowledge has been very insightful and I thank him for taking the time to join our brainstorming sessions. I am thankful to my lab colleagues Eleni, Ludovic, Ehsan, Hossein, Wajeb, Jagdish, Maaz, Alaeddine, Stéphan and the others for our interesting conversations and for our collaboration in making both the TCPIP course and the smart grid course some of the best courses at EPFL.

Of course I am undeniably thankful to Prof. Lenstra with whom I started my academic life. I have always loved relying on his advice and exchanging unusual ideas around home-made cakes. This brings me to Benjamin Wesolowski with whom I shared an office during part of the PhD. Thank you Benjamin for being a really fun and good friend. I thank my former LACAL colleagues Dušan, Novak, Thorsten, Jens, Anja and Rob for their help and for our entertaining conversations.

My PhD life in Lausanne has also been very lively and exciting thanks to the Srpski-Italian-French alliance that is my relationship with PEL and the rest of the Serbian group. Hvala puno and dulce!

I thank my family for always providing the love and support that I needed. Our adventures in the mountains have been a real breath of fresh air.

Last but not least, I am most grateful to Marko for being my rock in this PhD journey. I thank him for helping me be the best version of myself, hvala ljubavi!

Lausanne, April 7, 2021 M. D.

Abstract

The security of critical infrastructures is of the utmost importance for the well-being of society and economy. In this thesis, we focus on their security against a certain type of malicious attack called *time-synchronization attack* (TSA). A TSA can impact any network that relies on the real-time analysis of data, by altering the time synchronization between its nodes. As a result, networks that rely on timely observations can start malfunctioning or even failing. Certain TSAs can be thwarted by the usual cyber-security tools such as authentication and confidentiality algorithms. However, such tools cannot counter TSAs that are implemented by physical attacks. Such non-typical attacks are undetectable themselves but because they affect the functionality of the system, they may be detected if they lead to non-plausible observations. The identification of TSAs requires in-depth knowledge of the system's operation and its normal dynamics.

This work focuses on TSAs in two different settings. We first consider the setting of smart grids. Their control and operation require the timely knowledge of the system state, which is inferred from an estimate computed from measurements. We consider phasor measurements taken from phasor measurement units (PMUs) because they are accurate, time and phase-aligned, and have fast streaming rates. However, they require a precise time synchronization, which is a weakness as existing time-synchronization techniques are known to be vulnerable to attacks. We aim to assess the vulnerability of synchrophasor-based applications, in particular the state estimation of a system, by exploring the feasibility and detectability of TSAs on PMUs.

A widespread technique for making the state estimation more robust against attacks is to couple it with a bad-data detection (BDD) scheme. However, it is known that false data injection (FDI) attacks and TSAs can impact the state estimation without being detected by the BDD algorithms. We present practical attack strategies for undetectable TSAs. We also present novel vulnerability conditions. One of our new conditions is a sufficient static condition that does not depend on the measurement values. We propose a security requirement that prevents it and we provide a greedy offline algorithm that enforces it. If this security requirement is satisfied, there is still a possibility that the grid can be attacked, although we reason that it is very unlikely. We identify two sufficient and necessary vulnerability conditions which depend on the measurement values. For each, we provide a metric that shows the distance between the observed and vulnerability conditions. We recommend their monitoring for security. Enforcing our static security requirement requires increasing the amount of measurement points in certain areas of the grid. In order to secure the grid by adding a minimal amount of measurement points, we investigate the potential of utilizing the three-phase model instead of the direct-sequence model. We show analytically that if the power system is unbalanced, then the use of the three-phase model as input to BDD algorithms enables to detect attacks that would be undetectable if only the direct-sequence model was used. Our results provide a new

argument for the adoption of three-phase models for BDD, as their use is a simple, yet effective measure for reducing the vulnerability of PMU measurements to TSAs. Numerical results obtained from simulations performed on the IEEE 39-bus benchmark with real load profiles taken from the Lausanne grid, confirm our findings.
The second setting we consider is that of sensor networks for passive-source localization. Localization methods rely on the timely analysis of measurements such as angles of arrival (AOA), time differences of arrival (TDOA) between sensors, frequency differences of arrival (FDOA) between sensors or a combination of them. Therefore, an accurate synchronization of the time reference between sensors is essential, which constitutes a weakness for localization systems. We focus specifically on the localization of a passive source from TDOA measurements. By nature, TDOA measurements are highly sensitive to time-synchronization offsets between sensors. We first illustrate that TSAs can severely affect the localization process and we show that residual analysis does not enable the detection and identification of TSAs. Second, we propose a two-step TDOA-localization technique that is robust against TSAs. It uses a known source to define a weight for each pair of sensors, reflecting the confidence in their time synchronization. Our solution then uses the weighted least-squares estimator with the newly created weights and the TDOA measurements received from the unknown source. As a result, our method either identifies the network as being too corrupt to localize, or gives a corrected estimate of the unknown position along with a confidence metric. Numerical results illustrate the performance of our technique.

Keywords: critical infrastructure, smart grid, phasor measurement unit (PMU), power system state estimation, synchrophasor, time-synchronization attack, bad data detection (BDD), clock servo, three phase system, delay attack, optimal PMU placement (OPP), communication system security, electronic warfare, localization, time-difference of arrival (TDOA), wireless sensor networks, outlier detection, non-line of sight (NLOS).

Résumé

La sécurité des infrastructures critiques est de la plus haute importance pour le bien-être de la société et de l'économie. Dans cette thèse, nous nous concentrons sur leur sécurité face à un certain type d'attaque, appelée *attaque de synchronisation du temps* (AST). Une AST peut avoir un impact sur n'importe quel réseau sensible au temps en modifiant la synchronisation entre différents points du réseau. En conséquence, les réseaux qui nécessitent une bonne synchronisation peuvent devenir obsolètes. Certaines ASTs peuvent être contrées par les outils usuels de cyber-sécurité tels que les algorithmes d'authentification et de confifentialité. Néanmoins, ces outils ne permettent pas de contrer les ASTs qui sont implémentées à travers des attaques physiques. Ces attaques inhabituelles ne sont pas détectables en elles-mêmes mais peuvent tout de même être détectées si elles entraînent des observations non-plausibles. L'identification d'ASTs nécessite une connaissance en profondeur du fonctionnement du système et de ses comportements habituels.

Ce travail porte sur les ASTs dans deux contextes différents. Le premier concerne les réseaux électriques intelligents. Leur contrôle et leur fonctionnement nécessite la connaissance en temps réel de l'état du système, qui est déduit d'une estimation calculée à partir de mesures. Nous considérons des mesures de phaseurs prises par une unité de mesure de phaseurs (PMU) car elles sont précises, alignées selon le temps et la phase, et ont une vitesse de diffusion rapide. Cependant, ces mesures requièrent une synchronisation précise du temps. Ceci est une faiblesse car nous savons que toutes les techniques de synchronisations qui existent sont vulnérables aux attaques. Nous avons pour but d'évaluer la vulnérabilité des applications basées sur les synchrophaseurs, en particulier l'estimation d'état du système, en explorant la faisabilité et la détectabilité des ASTs sur les PMUs.

Une technique répandue pour rendre l'estimation d'état plus robuste face aux attaques, est de la coupler avec un algorithme de détection de mauvaises données (DMD). Néanmoins, il est bien connu que les attaques d'injection de fausses données et les ASTs peuvent impacter l'estimation d'état sans être détecté par les algorithmes de DMD. Nous présentons des stratégies d'attaque pratiques qui permettent d'implémenter des ASTs indétectables. Nous présentons aussi de nouvelles conditions de vulnérabilité. Une de nos nouvelles conditions est une condition suffisante et statique, qui ne dépend pas des valeurs des mesures. Nous proposons une exigence de sécurité qui empêche cette condition d'être satisfaite et nous fournissons un algorithme hors ligne et glouton qui met l'exigence de sécurité en vigueur. Si cette exigence n'est pas satisfaite, il est encore possible que le système puisse être attaqué de façon indétectable, bien que nous raisonnons que ce soit peu probable. Nous identifions deux conditions de vulnérabilité qui sont nécessaires et suffisantes et qui dépendent des valeurs que prennent les mesures. Pour chacune, nous fournissons un indicateur de vulnérabilité qui montre la distance entre les conditions observées et celles

de vulnérabilité. Nous recommandons par sécurité de les surveiller. Mettre en vigueur notre exigence de sécurité statique nécessite l'augmentation du nombre de points de mesures dans certaines parties du réseau. Afin de sécuriser le réseau en ajoutant un nombre minimal de points de mesures, nous enquêtons sur le potentiel à utiliser un modèle triphasé plutôt que le modèle à séquence directe, qui est en une dimension. Nous montrons analytiquement que si le système n'est pas balancé, alors l'utilisation du modèle triphasé pour la DMD permet de détecter des attaques qui resteraient indétectables si seulement le modèle à séquence directe était utilisé. Nos résultats donnent un nouvel argument pour l'adoption de modèles triphasés pour la DMD car leur utilisation est une mesure simple et efficace qui réduit la vulnérabilité des mesures de PMU face aux ASTs. Nos résultats numériques obtenus à travers des simulations sur le système IEEE à 39 bus avec de vrais profils de charges pris sur le réseau électrique de Lausanne, confirment nos découvertes.

Le second contexte que nous considérons est celui des réseaux de senseurs pour la localisation de sources passives. Les méthodes de localisation se basent sur l'analyse en temps réel de mesures telles que l'angle d'arrivée du signal, la différence de temps d'arrivée (TDOA) entre senseurs, la différence de fréquence d'arrivée entre senseurs ou une combinaison de ces mesures. Par conséquent, une bonne synchronisation du temps entre les différents senseurs est essentielle, ce qui constitue une faiblesse pour les systèmes de localisation. Nous nous concentrons spécifiquement sur la localisation d'une source passive, à partir de mesures TDOA. De par leur nature, les mesures TDOA sont très sensibles aux décalages de synchronisation entre senseurs. Nous illustrons tout d'abord que les ASTs peuvent avoir un impact drastique sur le processus de localisation et nous montrons que l'analyse de résidus ne permet pas de détecter ou d'identifier les ASTs. Ensuite, nous proposons une technique de localisation en deux étapes, basée sur les mesures TDOA, qui est robuste face aux ASTs. Cette technique utilise une source connue afin de définir un poids pour chaque paire de senseurs, qui reflète le niveau de confiance que nous pouvons avoir en leur synchronisation du temps. Notre solution utilise ensuite l'estimateur des moindres carrés pondérés avec les mesures TDOA reçues depuis la source inconnue à localiser et nos nouveaux poids. Deux résultats peuvent découler de notre solution : elle peut soit détecter que le réseau est trop corrompu pour continuer à localiser, soit donner une estimation corrigée de la position recherchée avec un indicateur de confiance en l'estimation. Nos résultats numériques illustrent la performance de notre solution.

Mots clefs : infrastructure critique, réseau électrique intelligent, unité de mesure de phaseurs (PMU), estimation d'état, synchrophaseur, attaque de synchronization du temps, détection de mauvaises données, commande d'horloge, système triphasé, attaque de délai, placement optimal de PMU, sécurité des systèmes de communication, guerre électronique, localisation, différence de temps d'arrivée, réseau de senseurs sans fil, détection de cas aberrants, environnement sans visibilité directe.

Contents

Acronyms

AOA Angle of Arrival.

ARP Address Resolution Protocol.

BDD Bad Data Detection.

BF Brute Force.

CC Control Center.

CS Calibration Source.

EKF Extended Kalman Filter.

ERR Effective Rank Ratio.

FDI False Data Injection.

FDOA Frequency Difference of Arrival.

FIR Finite Impulse Response.

GPS Global Positionning System.

HTI Hypothesis Testing Identification.

IIR Infinite Impulse Response.

ILP Integer Linear Program.

IOS Index of Separation.

LAV Least Absolute Value.

LM Levenberg-Marquadt.

LNR Largest Normalized Residual.

LOS Line of Sight.

LS Least Squares.

LSE Linear State Estimation.

MITM Man in the Middle.

NLOS Non Line of Sight.

NTP Network Time Protocol.

OCP Output-Constrained Proportional Controller.

OCPI Output-Constrained Proportional Integral Controller.

PF Particle Filter.

PMU Phasor Measurement Unit.

PTP Precision Time Protocol.

PVT Position Velocity and Time.

RAIM Receiver Autonomous Integrity Monitoring.

RTT Round-Trip Time.

Rwgh Residual Weighting.

SCADA Supervisory Control and Data Acquisition.

SNR Signal to Noise Ratio.

TDOA Time Difference of Arrival.

TOA Time of Arrival.

TSA Time-Synchronization Attack.

UKF Unscented Kalman Filter.

WLS Weighted Least Squares.

Introduction

Critical infrastructures are defined as the resources and structures which are necessary for a well-functioning society and economy. For example, water-supply systems, electricity-supply systems, transportation networks, or communication systems, are essential for human life and economic organization. The failure or malfunction of such systems can have disastrous consequences ranging from an economic crisis to human deaths. As a result, their security is of the utmost importance.

Critical infrastructures can be described as cyber-physical systems composed of three main components: the physical system itself, the operating system and the information system. The first component includes the mechanical and electrical elements, the second component directly controls the physical system and the third component links the information from the operation systems throughout the infrastructure, which enables the command and control of the overall infrastructure. Different security tools and strategies can be used for the three different components. For example, cyber-security mechanisms are commonly used to secure the information system, while secure facilities robust against human intrusion, fire, or water floods, are used to secure parts of the physical system.

In this thesis, we focus on a certain type of malicious attack called *time-synchronization attack* (TSA), that can impact any time-sensitive network. The goal of such intentional attacks is to alter the time synchronization of nodes of a network that is possibly deployed over a large geographical area. As a result, networks that rely on timely observations can start malfunctioning or even failing. TSAs can be performed in two ways. The first consists of modifying the synchronization information that is received by the attacked node. This type of attack can be prevented by enforcing the authentication of the synchronization information. The second way to mount a TSA is to target the physical medium in which the synchronization information is sent, by introducing propagation delays. This is not a typical type of attack because it is a physical attack that is performed on parts of the information system of the infrastructure. As a result, they cannot be thwarted by the usual cyber-security tools.

Chapter 1 is an introductory chapter which explains that the time synchronization of network nodes can be achieved through a space-based protocol such as GPS, or through a network-based protocol such as PTP. In both cases, we describe how they work, how a delay attack on them can be performed, and what state-of-the-art countermeasures exist. We explain that these countermeasures are not enough to prevent TSAs.

TSAs themselves are undetectable but because they affect the functionality of the system,

they may be detected if they lead to non-plausible observations. The identification of an attack requires in-depth knowledge of the system's operation and its normal dynamics in order to differentiate observations with and without an attack.

This work focuses on TSAs in two different settings and as such, this thesis is divided into two parts. The first part of this thesis considers Smart Grids. A smart grid is an electricity network in which electricity and data flows can be bidirectional. It uses digital communications technology in order to dynamically control the system in real time. The control and operation of interconnected power grids often require the timely knowledge of the system state. Accurate information on the state enables or improves the performance of fundamental functions, such as security assessment, voltage control, and stability analysis. Legacy measurement technologies have low measurement and streaming rates which induces a relatively low refresh rate of the system state-estimation. Nonetheless, the emerging measurement technology of phasor measurement units (PMUs) makes it possible to acquire phasors that are accurate, time and phase-aligned (i.e., synchrophasors) with streaming rates of the order of tens of measurements per second [1, 2, 3]. PMUs require, however, precise time synchronization [4], which is a weakness as existing time-synchronization techniques are known to be vulnerable to attacks, as we show in Chapter 1. Therefore, in order to assess the vulnerability of synchrophasor-based applications, in particular the state estimation of a system, it is essential to explore the feasibility and detectability of TSAs on PMUs.

A significant advantage of measuring phasors with PMUs is that the state estimation problem becomes linear. This is because PMUs are capable of directly measuring voltage and current phasors, compared to traditional supervisory control and data acquisition (SCADA) measurements (e.g., the power-flow and power injections, which are nonlinear functions of the system state). A widespread technique for making the linear state estimation (LSE) more robust against attacks is to couple it with a bad-data detection (BDD) scheme (e.g., χ^2 and the largest normalized residual tests (LNR)). However, it was shown in [5] that false-data injection attacks can have a non-negligible impact on the state estimation, without triggering any reaction from the BDD algorithms. Subsequent works focused on undetectable attack strategies [6], on vulnerability identification [7, 8, 9] and on the mitigation of such attacks [10, 11, 12, 13]. A TSA is another type of attack that can impact the state estimation without triggering any reaction from the BDD algorithms. It consists of altering the time reference of PMUs. Unlike with false-data injection attacks, the data is never forged nor modified. The authors of [14] propose a strategy to compute undetectable TSAs against vulnerable pairs of PMUs, each measuring a single phasor. They also discuss how vulnerable pairs of PMUs can be targeted simultaneously in order to maximize an attack. Their techniques require that a specific attack-angle matrix is of rank equal to 1, we refer to such attacks as rank-1 TSAs.

Chapter 2 is a second introductory chapter in which general background on PMU-based state estimation and undetectable TSAs is provided. The system model is defined, the

most widespread state estimators and BDD schemes are described, and the results of [14] on rank-1 TSAs are presented.

In Chapter 3, we take into consideration the fact that the clock adjustments of a PMU are performed by a clock controller called clock servo, which prevents too large offsets from being implemented. If they are too large, the clock servo can modify the injected attack angles, which may render the attacks detectable. This cannot easily be addressed with the existing attack strategies, as the undetectable attack values form a discrete set and cannot be continuously adjusted as would be required to address the problems posed to the attacker by the clock servo. Going beyond prior work, this chapter first shows how to perform undetectable attacks against more than two PMUs, so that the set of undetectable attacks forms a continuum and supports small adjustments. Second, it shows how an attacker can anticipate the operation of the clock servo while achieving his/her attack goal and remaining undetectable. Third, the chapter shows how to identify vulnerable sets of PMUs, each measuring a single phasor. Numerical results on the 39-bus IEEE benchmark system illustrate the feasibility of the proposed attack strategies. Chapter 3 is based on our journal paper [15] published in IEEE Transactions on Instrumentation and Measurement.

In Chapter 4, we present novel vulnerability conditions in the general case where PMUs measure any number of phasors and can share the same time reference. This differs from prior work and the previous chapter because it was previously assumed that targeted PMUs measure a single phasor. Yet, PMUs are capable of measuring several quantities. One of our new conditions is a sufficient condition that does not depend on the measurement values. We propose a security requirement that prevents it and provide a greedy offline algorithm that enforces it. If this security requirement is satisfied, there is still a possibility that the grid can be attacked, although we reason that it is very unlikely. We identify two sufficient and necessary vulnerability conditions which depend on the measurement values. For each, we provide a metric that shows the distance between the observed and vulnerability conditions. We recommend their monitoring for security. Numerical results, on the IEEE-39 bus benchmark with real load profiles, show that the measurements of a grid satisfying our security requirement are far from vulnerable. Chapter 4 is based on our journal paper that is provisionally accepted for publication in IEEE Transactions on Control of Network Systems [16].

In Chapter 5, we investigate the potential of utilizing the three-phase model instead of the direct-sequence model for mitigating the vulnerability of state estimation to undetectable TSAs. Among the symmetrical components computed from the phasors in three-phase systems, the standard practice only uses the direct-sequence component for state estimation and BDD [17]. We show analytically that if the power system is unbalanced, then the use of the three-phase model as input to BDD algorithms enables to detect attacks that would be undetectable if only the direct-sequence model was used. Simulations performed on the IEEE 39-bus benchmark using real load profiles recorded

on the grid of the city of Lausanne, confirm our analytical results. Our results provide a new argument for the adoption of three-phase models for BDD, as their use is a simple, yet effective measure for reducing the vulnerability of PMU measurements to TSAs. Chapter 5 is based on our paper that is currently under review for publication in IEEE Transactions on Smart Grid.

The second part of this thesis considers sensor networks for passive-source localization. The problem of localizing uncooperative sources that emit radio frequency signals has been extensively studied in the field of electronic warfare [18], as well as for civil applications [19]. Solutions were proposed in various settings such as sensor networks, radar, sonar or wireless communication [20, 21, 22]. Localization methods rely on the timely analysis of measurements such as angles of arrival (AOA), time differences of arrival (TDOA) between sensors, frequency differences of arrival (FDOA) between sensors or a combination of them [18, 23, 24]. Therefore, an accurate synchronization of the time reference between sensors is essential, which constitutes a weakness for localization systems.

In Chapter 6, we focus on the localization of a passive source from TDOA measurements. TDOA values are computed with respect to pairs of fixed sensors that are required to be accurately time-synchronized. By nature, TDOA measurements are highly sensitive to time-synchronization offsets between sensors. We first illustrate that TSAs can severely affect the localization process. With a delay of a few microseconds injected on one sensor, the resulting estimate might be several kilometers away from the true location of the unknown source. We show that residual analysis does not enable the detection and identification of TSAs. The main contribution of this chapter is then to propose a two-step TDOA-localization technique that is robust against TSAs. It uses a known source to define a weight for each pair of sensors, reflecting the confidence in their time synchronization. Our solution then uses the weighted least-squares estimator with the newly created weights and the TDOA measurements received from the unknown source. As a result, our method either identifies the network as being too corrupt to localize, or gives a corrected estimate of the unknown position along with a confidence metric. Numerical results illustrate the performance of our technique. Chapter 6 is based on our journal paper [25] published in IEEE Transactions on Information Forensics and Security.

Chapter 7 concludes the thesis. Specifically, it combines the contributions of the previous chapters and explains how they can be used to either mount undetectable attacks or secure the system. This chapter also presents alternative ideas for TSA identification.

1

Delay Attacks

In critical infrastructures and in networks in general, the notion of time synchronization is essential: the inspection of time-aligned data from different geographical areas of a network enables the analysis of the link and causality between different log entries; the physicists at CERN rely on time-synchronized data acquired in their large particle accelerator for their research. In cyber-physical networks, the function of the system itself often relies on an accurate and precise time-synchronization. For example, the control and operation of a power grid relies on an estimation of the state of the system, that is itself estimated from the phase difference of complex measurements. An inaccurate time-synchronization of the measurements translates to an inaccurate phase difference and thus to a misestimation of the state of the system. Similarly, an inaccurate time-synchronization of sensors used for target localization translates to a position error. For example, we show in Chapter 6 that an error of $2.47\mu s$ in the synchronization between two sensors, leads to an error of approximately 1km in the location estimate. An inaccurate time-synchronization can not only lead to denial-of-service, but it can have devastating effects if it goes unnoticed. For example in power grids, a misestimation of the system state can lead to a blackout or to asset degradation, and in localization networks, it can lead to a vehicle entering on enemy territory.

The level of accuracy that is required is highly dependent on the system. For example, an accuracy of a few milliseconds is typically sufficient for the analysis of log files in a banking network. In contrast, due to the speed of light, an accuracy level of a few microseconds in sensing devices for localization, leads to location errors that are in the order of magnitude of a kilometre. The two systems that we study in this dissertation are highly sensitive to time-synchronization errors, they require an accuracy of the order of one to ten nanoseconds.

We now explain how time-synchronization can be achieved. The unit of a second is defined with respect to the stable oscillation frequency of the cesium-133 atom. In identical environments, each cesium-133 atom produces this identical frequency. As a result, atomic clocks are nowadays considered as the most accurate and precise time clocks in practice. Note however that current research towards optical clocks shows that

even more accurate clocks may be constructed in the future and may lead to a new definition of the second [26]. Atomic clocks are very expensive and it is not possible to simply equip every device with one. The common practice is to deploy less accurate but less expensive clocks, such as quartz clocks, and to synchronize them frequently to an external time-source that has an atomic clock.

Satellites are usually equipped with atomic clocks, they broadcast synchronization signals to receivers on earth [27]. Devices can also synchronize themselves by connecting to time servers over the internet. These servers themselves can have varying time sources: another time server, a radio clock or an atomic clock. In other words, time-synchronization can be achieved via a space-based protocol or via a network-based protocol. In any case, note that the best practice is to use several time sources in order to ensure that there is no single point of failure. For example, it is customary to use a network-based protocol to disseminate the time from one master node to the other nodes of a network and to further equip every node with a GPS receiver [28, 29]. As a result, network nodes can synchronize themselves by using a combination of their received GPS signals and of the synchronization packets that they received over the telecommunication network. We now describe these two families of time-synchronization methods and their vulnerabilities.

1.1 Space-based Synchronization

Space-based synchronization is achieved through Satellite communication between receivers on earth and a satellite constellation such as GPS, Glonass, Galileo or BeiDou-2, using the trilateration technique. The accuracy of the time synchronization that can be achieved with space-based solutions is approximately 10ns to 40ns. Such solutions are vulnerable to jamming, spoofing or replaying attacks, which can be achieved inexpensively and which can have a non-negligible negative impact on the studied network [30, 31, 32, 33].

1.1.1 The Trilateration Technique

As mentioned previously, satellites broadcast navigation signals containing their coordinates and local time: (x_i, y_i, z_i), t_i for satellite i. The receiver coordinates (x, y, z) and time offset Δ_{t_i} between its time clock and the satellite's time clock, are linked to the received satellite coordinates by the following equation

$$d_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} + c\Delta_{t_i},$$

where d_i is the distance between the receiver and satellite i and c is the speed of the signal. If the offset Δ_{t_i} is known, this equation defines a sphere around satellite i, on which the receiver is located. Observe that the receiver characteristics consist of 4 unknown

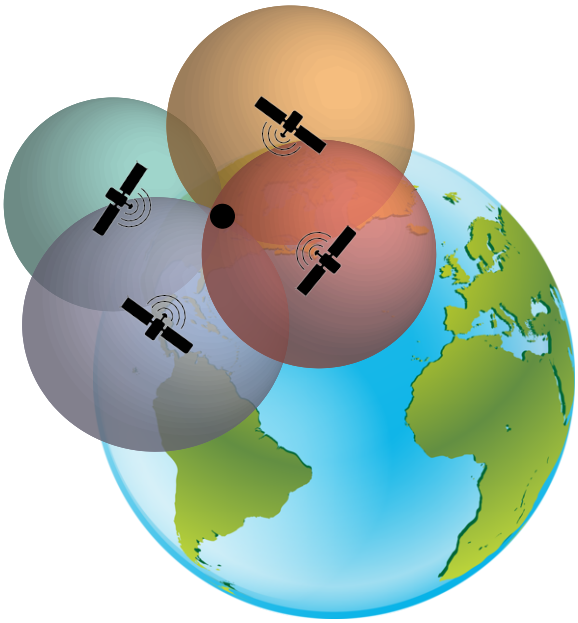


Figure 1.1 – The trilateration technique for space-based time synchronization and localization.

variables: its 3 coordinates and its time offset. Therefore, as depicted in Figure 1.1, the receiver is able to compute its characteristics by finding a unique intersection of the spheres of four satellites.

1.1.2 Delay Attack by GPS Spoofing

Both military and civilian GPS signals are vulnerable to delay attacks by spoofing or jamming with cheap devices. However, only the civilian GPS signals can be forged because the civilian infrastructure uses the legacy L1 signal which does not benefit from security enhancements found in its military counterparts. More specifically, it does not provide authentication or encryption.

For jamming attacks, the attacker superimposes a signal with a large signal-to-noise ratio over the genuine signal, which causes a denial of service. In spoofing attacks, the receiver receives a counterfeit signal from which it computes a wrong location and wrong offset. Spoofed signals are either genuine signals that are replayed with a delay (i.e. this is a threat to both civilian and military signals) or fake signals (i.e. this is a threat only to civilian signals). In both cases, the receiver is fooled. One seamless technique to perform a spoofing attack is to start sending accurate data at a low power at the same frequency as the true satellite. Then, the attacker increases the power of his signal until the receiver locks on the attacker signal (i.e. it seamlessly switches from the true signal to the attacked signal). Finally, the attacker can start delaying the arrival of data by injecting a delay of his choosing, as shown on Figure 1.2. Observe that spoofing attacks introduce always positive delays because the delayed signal takes a detour by the

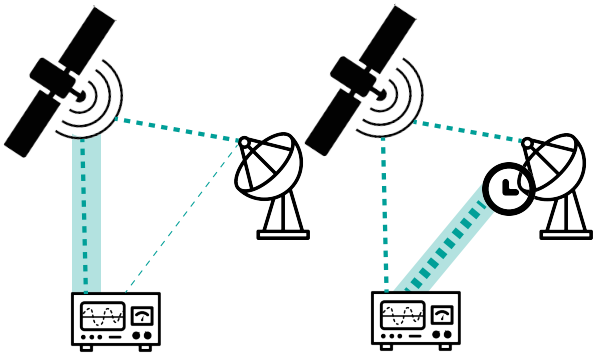


Figure 1.2 – Delay attack by GPS spoofing.

attacker's antenna before reaching the GPS receiver.

1.1.3 State-of-the-art Countermeasures

As mentioned previously, the lack of authentication allows the signals to be forged. An asymmetric cryptography signature scheme for navigation signals was proposed in [34] as a countermeasure. However, it requires that the satellite sends two signals, which takes time and is an issue for accurate time-synchronization. Another problem with this solution is that it requires modifying all of the GPS system: modification in the satellites and in the receivers, which is too expensive and not practical. Similarly, modifying the communication type from broadcast signals to bidirectional communication and enforcing a handshake protocol is not a feasible solution on such a legacy system.

In contrast, solutions that require only modifying the receivers exist, although they are not perfect. One such solution is to use a constellation of receivers. If an attacker spoofs a signal, then all of the receivers will be impacted in the same manner. Hence, they will all compute their location at the same point. By comparing their computations they can thus detect suspicious signals [35, 36, 37]. In order to remain undetected, the attacker would need to spoof each receiver individually and to inject different delays. Even though this solution is reliable, it is not always feasible to use it on small vehicles.

Another solution that does not require any modification to the satellites, is to use other signal characteristics that are unique with respect to the receiver position [38]. For example, the angle of arrival or the direction of arrival of the received signals can be leveraged to distinguish a genuine signal from an attacked signal. Once the receiver computes its location, it can verify that the signal characteristics are plausible given the satellites' and its own coordinates. The receiver can also measure the signal-to-noise ratio of the received signals and raise an alarm at a sudden increase. In order to accept a signal, GPS receivers check the correlation between a replica of a pseudorandom code corresponding to each satellites and the code contained in visible signals. The authors

of [39] proposed an efficient spoofing detection technique which consists of observing visible correlation auxiliary peaks instead of only the strongest. However, if the attacked signal is too powerful, then the genuine signal is covered by noise and this solution does not work. A similar technique can be performed by an additional GPS device that is placed between the satellite and the receiver and which acts similarly to a firewall [40].

Other solutions include the fusion of GPS receivers with inertial measurement units. In such solutions, extended Kalman filters are used to track and validate the device position, velocity and time (PVT) solutions [41]. However, the initial parameters of the filter are prone to error, which can be leveraged by attackers to mount an undetectable attack in which the parameters adapt to the spoofing coordinates [42]. As mentioned previously, the best practice is to use several sources simultaneously. Hence, the PVT solutions can be computed with respect to various satellite constellations and compared. However, all civilian satellite systems are vulnerable to spoofing attacks, hence this solution is not reliable because an attacker could simply spoof all of the signals.

It is common practice for GPS receivers used in aviation to perform receiver autonomous integrity monitoring (RAIM) [43, 44]. It detects outlier satellites whose signals are excluded during the PVT computation. Hence, it requires redundant signals, which means that the receiver must have more than 4 satellites within visible range at all times. Another technique called Crowd-GPS-Sec [45] enables the detection and localization of a GPS spoofer targeting aircrafts by comparing air traffic data sensed on the ground by a crowdsourced infrastructure, with the positions computed by aircrafts from signals that are possibly spoofed. Although this solution is efficient, it is highly dependent on the crowdsourced ground network which measures flight traffic. Hence, it can only be used in aerospace applications.

In spite of all of the existing countermeasures, delay attacks by spoofing satellite signals are still a threat in practice as they can be achieved undetectably and inexpensively with off-the-shelf devices. In this dissertation we explore countermeasures which consist of analysing the output of the system itself. For example in power grids, we aim to distinguish the dynamics of an attacked grid from those of an unattacked grid.

1.2 Network-based Synchronization

Although the use of GPS synchronization is widespread, some applications require the synchronization of devices which are located in areas where GPS signals are inaccessible. This can be the case in mountainous areas, for example. Also, the use of several time-synchronization mechanisms simultaneously is a requirement for security. Network-based synchronization can achieve high levels of accuracy by simply leveraging the usually already existing telecommunication network.

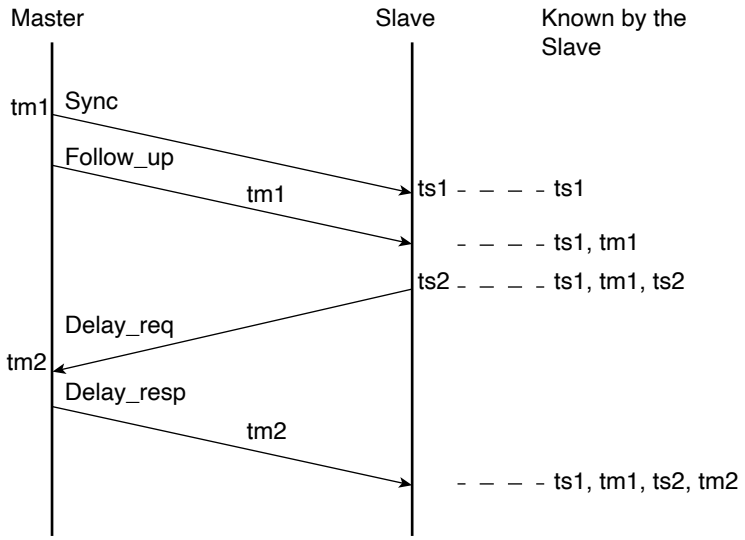


Figure 1.3 – The Precision Time Protocol

Network-based synchronization is achieved by the dissemination of time from master nodes to slave nodes. The two major families of protocols are the network time protocol (NTP) and the precision time protocol (PTP). In NTP, time-synchronization packets are sent hierarchically through layers called stratum. More specifically, an atomic clock is a stratum-0 device and a stratum-1 entity takes its time from a stratum-0. The stratum-1 can then act as a time source to stratum-2 entities and so on and so forth. Time servers provided by NIST can be reached online, they are usually stratum-2 servers. The NTP protocol can achieve an accuracy in the order of magnitude of milliseconds, making it impractical for some applications. In contrast, PTP aims to achieve an accuracy of a few microseconds. The superiority of PTP stems from hardware timestamping instead of software timestamping and accounts for device delay. Because the type of applications that we consider in this dissertation require a high time-synchronization accuracy, we focus on PTP and on White-Rabbit, which combines PTP with Synchronous Ethernet in order to achieve nanosecond accuracy [46].

1.2.1 The Precision Time Protocol and White-Rabbit

The goal of PTP is to determine the offset between a master clock and a slave clock

$$\Delta_t = t_s - t_m,$$

where t_m and t_s are the master and slave times measured at a fixed time-instant. The protocol is depicted on Figure 1.3. First, the master sends a *Sync* packet at time $tm1$ to the slave, who receives the packet at time $ts1$. Then, the master sends a *Follow_up* packet containing the timestamp $tm1$ to the slave. Upon reception of this packet, the slave knows timestamps $tm1$ and $ts1$. Second, the slave sends a *delay_req* packet to

the master at time $ts2$. Then, upon reception of the delay_req packet at time $tm2$, the master responds with a *delay_resp* packet containing the timestamp $tm2$. At the end of the exchange, the slave knows timestamps $tm1$, $ts1$, $tm2$ and $ts2$. Observe that

- $ts1 - tm1 = \Delta_t + d_{ms}$, where d_{ms} is the propagation delay from the master to the slave and Δ_t is the constant offset that the slave wants to determine,

- and $ts2 - tm2 = -\Delta_t + d_{sm}$, where d_{sm} is the propagation delay from the slave to the master.

Then, assuming that $d_{ms} = d_{sm}$, the slave can compute his offset

$$\Delta_t = \frac{1}{2}(ts1 - tm1 - ts2 + tm2).$$

If $d_{ms} \neq d_{sm}$ but that the difference is known, then there is a known asymmetry in the propagation time and the offset can still be computed. It is proven that it is impossible for the protocol to measure asymmetries in the propagation delay [47]. In fact, a theorem from [48] states that in a network of n nodes and l unidirectional links, the number of delays that can be measured by timestamping protocols is at most $l - n + 1$. Because this bound, referred to as *cyclomatic number*, is always below l, it is not possible to measure all unidirectional delays. Consequently, all network-based protocols assume that the transmission time of a packet between a slave and its master clock is either symmetric or has a known asymmetry. This vulnerability is inherent to all network-based time-synchronization protocols. We explain in Section 1.2.2 how it can be exploited to mount delay attacks.

White Rabbit is a deterministic packet-based protocol that achieves the time-synchronization of slave clocks with respect to a master clock with nanosecond accuracy. Before using PTP for the clock offset compensation, White Rabbit uses synchronous ethernet to achieve the syntonization of the clocks (i.e. to synchronize the frequency of the clocks). This provides a better base ground for PTP, which explains the improved accuracy. Synchronous ethernet is expensive, hence it is used only when necessary. One of its initial use was for synchronizing the local time of sensors dispatched in the CERN accelerator to acquire data on experiments.

1.2.2 Delay Attack by Delay-Box Insertion

Delay attacks on a network can be achieved either by modifying the content of the synchronization packets or by simply delaying them (i.e. intercept, wait and forward). The authors of [49] presented a MITM attack that can be achieved through ARP poisoning. If the synchronization packets are not cryptographically secured, then the attacker can

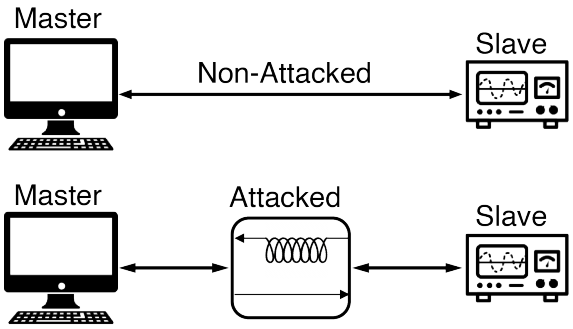


Figure 1.4 – Delay attack between a master and a slave node.

modify its content and introduce delays in the contained timestamps, thus performing a delay attack.

In contrast, the authors of [50] presented a physical delay-attack which bypasses all cryptographic protocols in place because it does not modify the packets but simply delays their arrival. This attack is based on the previously-mentioned vulnerability of network-based time-synchronization protocols. Specifically, it breaks the symmetry or known asymmetry between the master and slave nodes by injecting a one-way delay in the path between them. This is done by inserting a repeater-box on the fiber between the master and the slave. Within the repeater-box the bidirectional fiber is replaced by a short fiber in one direction and a long fiber coil in the reverse direction, as shown on Figure 1.4. The fiber lengths are chosen according to the delay that the attacker wishes to inject. This manipulation results in the slave clock to go ahead or behind the master clock. A longer fiber on the path from the slave to the master, injects a negative delay in the time reference of the slave clock, and the reverse injects a positive delay. Notice that with GPS spoofing only positive delays are injected but that with delay-box insertion the delays can be both positive or negative.

Such a physical attack is a realistic threat as it does not require a physical access to the often well-protected devices but only to the fiber cables that links well protected devices. Such devices are located at nodes which can be geographically separated by large distances, as is the case in power grids and in sensor networks for localization. The assumption that the cables of a large network are not all physically protected is thus realistic.

1.2.3 State-of-the-art Countermeasures

Although the security of PTP has been extensively studied [51, 52, 53, 54, 55], countermeasures against delay-attacks are scarce. Network security protocols such as IPsec [56], MACsec [57] and TLS [58] enable the encryption and signing of synchronization packets. Their use is recommended to counter network-based attacks, including delay attacks that

are achieved by modifying the content of the synchronization packets [59, 60]. However, as mentioned above, such protocols are not robust against delay-box insertion because the delay-box does not attempt to forge or modify synchronization packets.

The fact that delay attacks are a threat that could have a high impact on networks, including critical infrastructures, is very well acknowledged in the literature [61]. However, realistic and efficient countermeasures are rarely presented. Annex K of [62] provides a security extension to PTPv2 which aims to ensure message integrity and protection against replay attacks. However, these mechanisms are flawed [53, 63].

One line of research for countermeasures, focuses on monitoring the round-trip time (RTT) of synchronization packets between the master and slave clocks. This requires a precomputation phase in which the network is not attacked and RTT bounds are established [60, 64, 59]. A technique which bounds the clock offset by limiting the RTT is also proposed in [59, 64] and a hypothesis testing solution which analyzes the computed ratio between the master and slave clock rates is presented in [65]. However, such techniques are deemed unreliable. For example, if the attack delay is within the bounds, then it is undetected although it can already have a non-negligible negative impact [49]. The limits of techniques based on bounding the RTT as a countermeasure to delay attacks are also discussed in [60] and [66].

Another proposition from [60] is to use multiple redundant paths between master and slave nodes. The time information that comes through a path is accepted if its difference with the average time information received through the other paths is under a predefined threshold. This solution assumes that the paths are completely independent (i.e. no shared links) and that the attacker is only able to attack a minority of them. The authors also assume that all of the paths between two fixed nodes are symmetric and have a similar propagation time. Such strong assumptions are not scalable and are unrealistic. Moreover, using the same number of clocks and increasing the number of paths does not solve the fundamental problem of the cyclomatic number bound. In other words, this solution does not thwart delay attacks because it is still not possible to measure all unidirectional delays.

The authors of [64] propose to modify PTP in order to compensate for injected delays by averaging the measured delays or RTTs over a time interval. However, this technique does not detect any malicious activity but simply reduces the effect of the attack by averaging data that hopefully contains non-attacked data. If the attack is performed over the entire interval of time, then this solution does not mitigate anything. They also propose an alternative countermeasure which consists of inspecting the timestamps and flagging abnormal values. Although this technique could work against naive attackers, it is inefficient against smart strategies which inject slowly increasing delays [15].

A theoretical model that defines a set of necessary and sufficient conditions for secure

clock synchronization is established in [67]. In particular, they explain that two-way communication between the master and the slave is a necessity for secure clock synchronization. They show that the IEEE 1588 standard for PTP [68] does not satisfy all of the requirements for secure clock synchronization, although it is a two-way synchronization protocol. In order to secure against delay attacks, they propose a solution that relies on the knowledge of the true RTT that packets should take between the master and slave nodes. As discussed earlier, this type of mitigation technique is not sufficient for delay-attack detection.

As already mentioned several times in this chapter, it is recommended for security to use and compare time information received from multiple time sources. For example, the use of a redundant clock called guard clock is presented in [69]. In this solution, the clocks are equipped with a GPS receiver and the guard clock checks the plausibility of the received network-based and space-based time-synchronization information. The guard clock then notifies the impacted slaves. Note however that this solution is not immune to an attacker that performs both the GPS spoofing attack and the delay-box insertion attack.

In this dissertation, we suppose that an attacker is able to perform GPS spoofing attacks and delay-box insertions in order to alter the time-synchronization of network devices with delays of his choosing. Such attacks are referred to as time-synchronization attacks (TSAs). We investigate undetectable TSA strategies and propose countermeasures in the setting of smart grids and of sensor networks for localization.

PART I

SMART GRIDS

List of Variables of Part I

Indices	
m	Number of PMU measurements
n	Number of buses
q	Number of attacked time references
p	Number of attacked measurements
Parameters	
$z = (z_{\square}^{[1,m]} + jz_{\square}^{[m+1,2m]})^T \in \mathbb{C}^m$	Complex measurement vector
$z_{\square} = (\mathrm{Re}(z), \mathrm{Im}(z))^T \in \mathbb{R}^{2m}$	Measurement vector in rectangular coordinates
$x = (x_{\square}^{[1,n]} + jx_{\square}^{[n+1,2n]})^T \in \mathbb{C}^n$	Complex state vector
$x_{\square} = (\mathrm{Re}(x), \mathrm{Im}(x))^T \in \mathbb{R}^{2n}$	State vector in rectangular coordinates
$H \in \mathbb{C}^{m \times n}$	Complex measurement-to-state matrix
$H_{\square} \in \mathbb{R}^{2m \times 2n}$	Measurement-to-state matrix in rectangular coordinates
$e \in \mathbb{C}^m$	Complex error vector
$e_{\square} \in \mathbb{R}^{2m}$	Error vector in rectangular coordinates
$\hat{x} \in \mathbb{C}^n$	Estimated complex state vector
$\widehat{x_{\square}} \in \mathbb{R}^{2n}$	Estimated state vector in rectangular coordinates
$\hat{z} \in \mathbb{C}^m$	Estimated complex measurement vector
$F \in \mathbb{C}^{m \times m}$	Complex verification matrix to compute the complex LS residuals
Id	Identity matrix
S_i	Set of phasor indices at site i
$F^{S_i} \in \mathbb{C}^{m \times \lvert S_i \rvert}$	Submatrix of F with column indices in S_i
$z^{S_i} \in \mathbb{C}^{\lvert S_i \rvert}$	Complex measurements with indices in S_i
$C_{\square} \in \mathbb{R}^{2m \times 2m}$	Covariance of the measurement noise in rectangular coordinates
$r(z_{\square}) \in \mathbb{R}^{2m}$	Residual vector in rectangular coordinates
$G_{\square} = \begin{bmatrix} G_{\square,1} & G_{\square,2} \\ G_{\square,3} & G_{\square,4} \end{bmatrix} \in \mathbb{R}^{2m \times 2m}$	Verification matrix to compute the normalized WLS residuals in rectangular coordinates
$G_{\square,1},\ G_{\square,2},\ G_{\square,3},\ G_{\square,4} \in \mathbb{R}^{m \times m}$	Blocks of $G_{\square}$
$R = \frac{1}{2} \begin{bmatrix} G_{\square,1} - jG_{\square,2} \\ G_{\square,3} - jG_{\square,4} \end{bmatrix}$	Matrix useful to compute $r(z_{\square}) = Rz + \bar{R}\bar{z}$
δ	Attack angle in rad
z', $z'_{\square}$	Attacked measurement vector
Δz, $\Delta z_{\square}$	Difference between the attacked and unattacked measurement vectors

$u_i = e^{j\delta_i}$	Attack value
u_i^*	Specific computed u_i value
$\varphi \in \mathbb{N}^{m \times q}$	Attack indicator matrix
$W \in \mathbb{C}^{q \times q}$	LS attack angle matrix
$(n_i)_i$ and $(v_i)_i \in \mathbb{C}^p$	Vectors that span the null space of $F^{[S_1,...,S_q]}$
I^i, V^i and $S^i_{inj} \in \mathbb{C}^*$	Complex current, voltage and injected power values at time instant i
$l \in \mathbb{C}^*$	Colinearity coefficient between two vectors
$T = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{\frac{2j\pi}{3}} & e^{\frac{-2j\pi}{3}} \\ 1 & e^{\frac{-2j\pi}{3}} & e^{\frac{2j\pi}{3}} \end{bmatrix}$	Sequence transformation matrix
$T_Z \in \mathbb{C}^{3m \times 3m} = \text{block-diagonal}(T)$	Three-phase to complete-sequence measurement vector transformation matrix
$T_X^{-1} \in \mathbb{C}^{3n \times 3n} = \text{block-diagonal}(T)$	Three-phase to complete-sequence state vector transformation matrix
$D = \text{block-diagonal}([010])$	$m \times 3m$ matrix that selects the direct-sequence measurements from the complete-sequence model measurements
$T_{\square,Z} \in \mathbb{R}^{6m \times 6m}$	Three-phase to complete-sequence measurement vector transformation matrix in rectangular coordinates
$T^{-1}_{\square,X} \in \mathbb{R}^{6n \times 6n}$	Three-phase to complete-sequence state vector transformation matrix in rectangular coordinates
Models indicated by subscripts	
abc subscript	The variable is in the three-phase system model
012 subscript	The variable is in the complete-sequence model
1 subscript	The variable is in the direct-sequence model
Metrics	
$IoS_{i,t}(z_i, z_t) \in [0, 1]$	Index of Separation at sites i and t with respect to measurement values z_i and z_j
$IoS^*_{i,t} \in [0, 1]$	Infimum of $IoS_{i,t}(z_i, z_t)$ over all possible values of (z_i, z_t)
$ERR(F^{[S_1,..,S_q]}) \in [0, 1]$	Effective rank ratio of $F^{[S_1,..,S_q]}$

2

PMU-based State-Estimation and Undetectable Time-Synchronization Attacks

In this chapter, we present the background on PMU-based state estimation, bad-data detection (BDD) and time-synchronization attacks (TSAs), on which the rest of this part of the dissertation is built.

2.1 System Model

We consider a transmission network with a total of n buses. Some of the buses are equipped with phasor measurement units (PMUs) capable of measuring phase-to-ground voltage phasors, bus current-injection phasors and/or current-flow phasors, thus resulting in measurement vector $z \in \mathbb{C}^m$. Because PMUs are able to measure several phasors simultaneously, the total number of measurements can be higher than the total number of buses $m \geq n$. Phasors measured by PMUs are typically referred to as synchrophasors because they are time-synchronized. Measurements are typically noisy and therefore are not exactly equal to the true phasors. The true state vector of the grid $x \in \mathbb{C}^n$ consists of a vector with a voltage value for every bus, it fully specifies the power system and is linearly linked to the measurement vector through equation

$$z = Hx + e, \tag{2.1}$$

where $e \in \mathbb{C}^m$ is the complex measurement error and H is the $m \times n$ complex measurement-to-state matrix. The blocks of H with respect to the three types of phasor measurements, are constructed from binary values, admittance matrix values and line parameters. We refer the reader to Appendix A.1 of [70] for a detailed derivation of the blocks of H.

Definition 2.1. *A power network is observable if the set of measurements enables the unique computation of all the state variables.*

In other words, the system is observable if and only if H is full rank. Observe that if H is rank deficient, then there can be more than one plausible state vector x that satisfies Eq. (2.1), making it impossible to operate and control the grid. In principle, if the grid is non-observable, then the measurement points are either not well deployed on the grid or in insufficient number. By increasing the number of measurement points, we increase the measurement redundancy, thus reducing the number of plausible state vectors.

It may happen that a measured phasor suddenly becomes unavailable, this could be due to sensor malfunction for example. If removing this measurement renders the grid non-observable, we say that the measurement is critical. In general, grid operators prefer to ensure that no single measurement is critical, thus avoiding single points of failure.

Definition 2.2. *A set of measurement points is said to be critical if the submatrix of H obtained by removing the corresponding rows is not full column rank.*

Note that by full column rank, we mean that the columns are linearly independent vectors. Therefore, the entire set of measurement points is the largest critical set. However, a subset of a critical set can also be critical.

Definition 2.3. *A critical set of measurements is said to be critical minimal if none of its subsets is also critical.*

Finding minimal critical sets allows to anticipate weak points in the topology and to create systems that are more resilient to sensor failure. However, finding them is a computationally expensive combinatorial problem [71, 72].

2.2 State Estimation

The process of estimating a state vector $\hat{x} \in \mathbb{C}^n$ from a measurement vector $z \in \mathbb{C}^m$ and the measurement-to-state matrix H, is referred to as state estimation. The maximum likelihood estimation of x

- is the least squares (LS) estimation

$$\underset{\hat{x}}{\operatorname{argmin}}(H\hat{x} - z)^{\dagger}(H\hat{x} - z)$$

  if the rectangular coordinates of the measurement errors are drawn independently from the standard normal distribution,

- is the weighted least squares (WLS) estimation

$$\underset{\hat{x}_{\square}}{\operatorname{argmin}}(H_{\square}\hat{x}_{\square} - z_{\square})^{T}C_{\square}^{-1}(H_{\square}\hat{x}_{\square} - z_{\square}),$$

if the rectangular coordinates of the measurement errors are drawn from a normal (i.e. Gaussian) distribution centered in 0 and of covariance matrix $C_\square$, where $H_\square$, $\hat{x}_\square$ and $z_\square$ are the rectangular coordinates of H, $\hat{x}$ and z, respectively.

- is the least absolute value (LAV) estimation

$$\underset{\hat{x}_\square}{\operatorname{argmin}} \sum_{i=1}^{2m} |(H_\square \hat{x}_\square - z_\square)_i|$$

if the rectangular coordinates of the measurement errors are drawn independently from the standard Laplace distribution.

2.2.1 The LS Estimator

Due to the fact that Eq. (2.1) is linear, the LS estimate of the state vector has a closed-form solution

$$\hat{x} = (H^\dagger H)^{-1} H^\dagger z,$$

where $H^\dagger$ denotes the complex conjugate transpose of H. From the estimated LS state vector $\hat{x}$, we can construct the LS estimated measurement vector $\hat{z} = H\hat{x}$ and compute its difference with the observed measurement vector z. The result of this difference is called the LS residual vector

$$r = H\hat{x} - z.$$

Define the verification matrix

$$F = H(H^\dagger H)^{-1} H^\dagger - Id,$$

where Id refers to the identity matrix, then $r = Fz$. Observe that $r = 0$ if and only if there exists a state vector x such that $z = Hx$. Observe that the LS estimation is linear both in complex coordinates and in rectangular coordinates.

2.2.2 The WLS Estimator

As opposed to the LS estimator, the WLS estimator weighs the measurements by their noise variances. In a phasor measurement, noise of different standard deviation can appear in the phase and in the magnitude. In other words, the noise in polar coordinates is heteroscedastic, which means that the covariance of the noise cannot be expressed linearly over the complex numbers. As a result, the WLS estimator is not linear over $\mathbb{C}$ but linear over $\mathbb{R}$ (i.e. the WLS estimation is linear in rectangular coordinates but not in complex coordinates). To overcome this shortcoming, we use matrices and vectors in rectangular coordinates. Note that the transformation from polar to rectangular coordinates is non-linear and thus the fact that the measurement noise in polar coordinates follows a Gaussian distribution does not imply that the measurement noise in rectangular

coordinates also follows a Gaussian distribution. Nevertheless, according to [14] and [73], if the noise of the polar coordinates is Gaussian and small, as is assumed with the PMU class that we consider in this dissertation, then the noise of the rectangular coordinates can also be assumed small and Gaussian.

The linear measurement-to-state equation over the reals now becomes

$$z_\square = H_\square x_\square + e_\square,$$

where $z_\square = (\mathrm{Re}(z), \mathrm{Im}(z))^T \in \mathbb{R}^{2m}$ is the measurement vector in rectangular coordinates, $x_\square = (\mathrm{Re}(x), \mathrm{Im}(x))^T \in \mathbb{R}^{2n}$ is the state vector in rectangular coordinates, $H_\square \in \mathbb{R}^{2m\times 2n}$ is the rectangular-measurements-to-rectangular-state matrix and $e_\square \in \mathbb{R}^{2m}$ is the error vector in rectangular coordinates. The corresponding WLS state estimate is

$$\widehat{x}_\square = (H_\square^T C_\square^{-1} H_\square)^{-1} H_\square^T C_\square^{-1} z_\square,$$

where $C_\square \in \mathbb{R}^{2m\times 2m}$ is the covariance matrix of the measurement noise in rectangular coordinates. The computation of $C_\square$, from both measurement values and measurement noise standard deviation, is described in [73] as follows.

We denote the noisy measurement of the phasor $z^* = \rho^* e^{j\phi^*}$ by $z = \rho e^{j\phi}$ and we denote by $\mathrm{Re}(e)$ and $\mathrm{Im}(e)$ the real and imaginary parts of the complex noise, respectively. Then the mean and covariances of the noise in rectangular coordinates are given by

$$\begin{aligned}
\mathrm{Re}(e) &= \rho\cos(\phi) - \rho^*\cos(\phi^*),\\
E[\mathrm{Re}(e)] &= (e^{-\frac{1}{2}\sigma_{\phi^*}^2} - 1)\cos(\phi^*)\rho^*,\\
E[\mathrm{Re}(e)^2] &= \frac{1}{2}\left(1+\frac{\alpha^2}{9}\right)\rho^{*2}(1+e^{-2\sigma_{\phi^*}^2}\cos(2\phi^*)) + \rho^{*2}\cos^2(\phi^*)\left(1-2e^{-\frac{1}{2}\sigma_{\phi^*}^2}\right),\\
\mathrm{Im}(e) &= \rho\sin(\phi) - \rho^*\sin(\phi^*),\\
E[\mathrm{Im}(e)] &= (e^{-\frac{1}{2}\sigma_{\phi^*}^2} - 1)\sin(\phi^*)\rho^*,\\
E[\mathrm{Im}(e)^2] &= \frac{1}{2}\left(1+\frac{\alpha^2}{9}\right)\rho^{*2}(1+e^{-2\sigma_{\phi^*}^2}\cos(2\phi^*)) + \rho^{*2}\sin^2(\phi^*)\left(1-2e^{-\frac{1}{2}\sigma_{\phi^*}^2}\right),\\
E[\mathrm{Re}(e)\mathrm{Im}(e)] &= \frac{1}{2}\sin(2\phi^*)\left[\left(1+\frac{\alpha^2}{9}\right)\rho^{*2}e^{-2\sigma_{\phi^*}^2} - 2\rho^{*2}e^{-\frac{1}{2}\sigma_{\phi^*}^2} + \rho^{*2}\right],
\end{aligned}$$

where $E[\cdot]$ is the expectation value operator, $\alpha = 0.1\%$ is the maximum magnitude error and $\sigma_{\phi^*} = 10^{-4}/3$ is a function of the maximum phase error for class 0.1 PMUs. Unlike what is commonly assumed in the litterature, the real and imaginary parts of the noise are correlated. Observe that the noise covariances actually depend on the true phasor values, which are unknown in practice. These values were replaced by the measured

values in [73] and it was observed that this replacement has a negligible effect on the estimation.

The estimated measurement vector is given as follows

$$\widehat{z}_{\square} = H_{\square}\widehat{x}_{\square} = K_{\square}z_{\square},$$

where $K_{\square}$ is the well-known hat matrix

$$K_{\square} = H_{\square}(H_{\square}^{T}C_{\square}^{-1}H_{\square})^{-1}H_{\square}^{T}C_{\square}^{-1}.$$

This matrix is useful to determine the measurement redundancy. More specifically, if a diagonal term is larger than the off-diagonal terms of the row, then the estimated value of the corresponding measurement is largely based on its measurement value and not on a combination of redundant measurements.

The WLS residuals are obtained from the $2m \times 2m$ real verification matrix

$$G_{\square} = H_{\square}(H_{\square}^{T}C_{\square}^{-1}H_{\square})^{-1}H_{\square}^{T}C_{\square}^{-1} - Id = K_{\square} - Id$$

as $r_{\square}(z_{\square}) = G_{\square}z_{\square}$. Observe that the verification matrix with respect to the WLS estimation $G_{\square}$ depends on the measurement values, whereas the verification matrix with respect to the LS estimation F does not depend on the measurement values but only on the topology of the grid. The verification matrix $G_{\square}$ is also referred to as the residual sensitivity matrix in the literature because it represents the sensitivity of the measurement residuals to the measurement errors $e_{\square}$.

The hat matrix $K_{\square}$ and the verification matrix $G_{\square}$ are both useful for the identification of critical and leverage measurement points. A leverage point is a type of outlier that over-influences any state estimator that aims to minimize a function of the residuals by forcing their residuals to be close to 0. A measurement outlier can be caused either by a too large error term or by values of the corresponding row of H that are very large or small when compared to other rows of H. The first type of outlier corresponds to bad data that lies away in the sample space while the second type lies away in the factor space and is called a *leverage point*. Generally, points that are far from the center of the factor space or the input space, have an excessive influence on the estimation. Whatever the state estimator used, one way to detect such points is by using the closed form expressions inherent to the WLS method. As explained above, the diagonal elements of the hat matrix give a measure of the influence of the measurement value on the estimation. This is due to the fact that the hat matrix is in fact the projection matrix that projects the observations onto the estimated observations. If a diagonal-term is close to 1, then the estimate is almost entirely determined by this measurement, hence that point will very likely be a leverage point. A test to identify them is to check if these diagonal terms are higher than a threshold that is typically chosen to be twice their expected value.

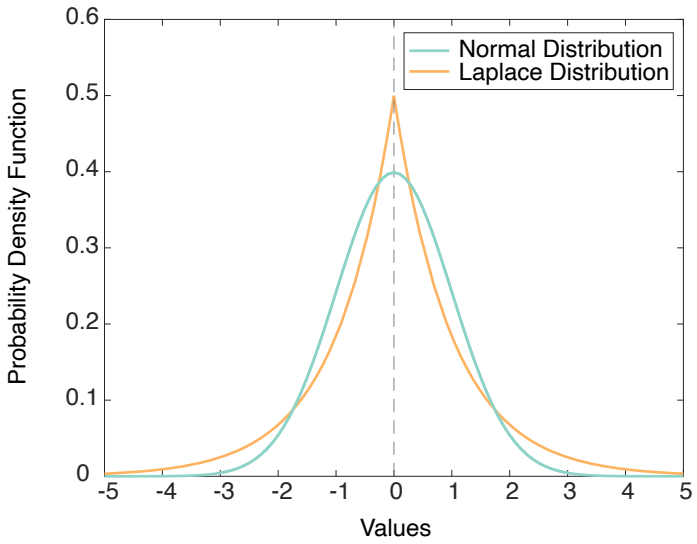


Figure 2.1 – Probability density functions of the Laplace and Normal distributions, both centered in 0 and of standard deviation 1.

One of the shortcomings of the WLS estimator is that it does not perform well in the presence of bad data or outliers. To overcome this, the common practice is to couple the state estimation process with a BDD algorithm. More details on BDD are given in Section 2.3.

2.2.3 The LAV Estimator

The LAV estimator is a robust state estimator. It has a larger breakdown point as it can tolerate more bad data than the WLS estimator before giving an inaccurate solution. The advantages of using the LAV estimator over the WLS estimator is that it is not as influenced by bad data and it has a similar time and memory complexity in our setting of PMU measurements. Recall that the LAV estimation of the system state is in fact the maximum likelihood estimation if the rectangular coordinates of the measurement errors are drawn from a Laplace distribution. As shown on Figure 2.1, the Laplace probability density function has longer tails than the normal probability density function, which means that it can tolerate a larger interval of error values. Nevertheless, it is also affected by leverage measurements. One countermeasure [74] is to use the WLS hat matrix $K_{\square}$ to identify and remove the leverage points, provided that the observability of the system is maintained. Another countermeasure is to use a strategic scaling of the measurement-to-state matrix $H_{\square}$.

Despite the similarity between the WLS and LAV estimators, there is no closed form solution for LAV. Nevertheless, the latter can be formulated as a single linear programming (LP) problem of dimension $2(m+n)$ which is easily solvable using techniques such as

the interior point or simplex methods.
The formulation of the LAV estimator as a single LP is explained in [75] and [74]. As previously mentioned, the goal is to minimize the sum of the absolute value of residuals which can be formulated as follows

$$\min_{\hat{x}_\square} c^T |r_\square|,$$
$$\text{such that } r_\square = z_\square - H_\square \hat{x}_\square,$$

where c is a vector of size $2m$ with all values equal to 1. Then, the authors of [74] reformulated the problem as follows

$$\min_{y_\square} \ c^T y_\square, \text{ such that}$$
$$M y_\square = z_\square$$
$$y_\square \geq 0,$$

where $M = \begin{bmatrix} H & -H & Id & -Id \end{bmatrix}$. This problem can be efficiently solved thanks to well-known and developed optimization tools on softwares like CPLEX or Gurobi.

It is argued in [74] that replacing the commonly used combination of the WLS estimator with a BDD algorithm by a single robust state estimator such as the LAV estimator is competitive and should be considered for the future of power networks. However, in an adversarial environment, attacked data can have unreasonably large values, thus negatively impacting a so-called robust state estimation. In the rest of this dissertation, we follow the common practice of assuming that the state estimator is the WLS estimator combined with a BDD algorithm.

2.3 Bad Data Detection

In power-system state estimation, it is common practice to couple the state estimation with a residual-based BDD algorithm. They typically flag unusual residual values obtained after the state estimation and remove the corresponding measurements before recomputing the state estimation. This process is repeated until no bad data is identified, in the worst case it is done $2m + 1$ times. This technique is typical in estimation theory because the residuals give insights on how well the estimate fits the measurements and therefore enables the assessment of the estimator accuracy.

Most BDD algorithms for power system state estimation in the literature are based on the analysis of residuals. In fact, a vast majority of the state-of-the-art techniques are variants either of the largest normalized residuals (LNR) test or of the χ^2 test. Article [76] proposes an efficient bad data detection scheme that scales well for large

power systems.Their technique is to cluster suspicious data into groups in which the LNR test is applied simultaneously. Paper [77] gives a modification of the χ^2 test in order to improve the bad data detection. Their modification is to tweak the threshold above which the data is flagged as bad, according to the residual covariance matrix. This paper also state that the χ^2 test is the most common bad data detection method used in several commercial state estimators. Finally, [78] uses the Adaptive Partitioning State Estimation method. The root idea of this method is similar to that of [76] as it is to apply a BDD test in partitions of the grid simultaneously and repeat the procedure until the faulty node is detected. The difference is that in [78], the authors use the χ^2 test while in [76], the authors use the LNR test. We now describe the LNR and χ^2 tests. We also provide an overview of the hypothesis testing identification method as it is known to perform better than the LNR test in the presence of multiple bad data whose residuals are correlated.

2.3.1 The LNR Test

As shown in [79] the WLS residuals are distributed according to a Gaussian distribution centered in 0 and of covariance matrix

$$\Omega = (Id - H_{\square}(H_{\square}^T C_{\square}^{-1} H_{\square})^{-1} H_{\square}^T C_{\square}^{-1}) C_{\square} = G_{\square} C_{\square}.$$

Recall that the WLS estimation assumes that the error is distributed according to a Gaussian distribution

$$e_{\square} \sim N(0, C_{\square}).$$

Because the residuals are linearly linked to the measurement errors, the WLS measurement residuals are also distributed according to a Gaussian distribution

$$r_{\square} \sim N(0, \Omega).$$

Then, the residuals can be normalized as follows

$$r_{\square,i}^N = \frac{|r_{\square,i}|}{\sqrt{\Omega_{ii}}},$$

and hence the normalized residual vector follows the Standard Normal distribution

$$r_{\square,i}^N \sim N(0, 1).$$

The core idea of the LNR test is to compare the largest normalized residual value against a predefined statistical threshold η. The choice of the threshold depends on the particular system. It is highly dependent on the number of samples and on the auto-correlation of residuals, even if they are normalized. The value of 3 is mentioned in [79] as an example but, in reality, it is important to set it according to a non-attacked control scenario on a

case-by-case basis, so as to avoid excessive false positives (false alarms). The steps of the LNR test are the following

1. perform the WLS estimation and compute the measurement residuals $r_\square$,
2. normalize the residuals $r^N_{\square,i} = \frac{|r_{\square,i}|}{\sqrt{\Omega_{ii}}}$ $\forall 1 \leq i \leq 2m$,
3. compare the largest normalized residual $\max(r^N_{\square,i})$ with the statistical threshold η,
4. if $\max(r^N_{\square,i}) > \eta$, then remove the corresponding measurement and go to step 1,
5. if $\max(r^N_{\square,i}) \leq \eta$, then stop because no bad data has been identified.

Note that the LNR test is only reliable either in the presence of a single bad-data or in the presence of multiple bad data whose residuals are uncorrelated. If there are several bad data whose residuals are correlated, then it was shown that the hypothesis testing identification (HTI) technique outperforms the LNR test in identifying them [79].

2.3.2 The Hypothesis Testing Identification Technique

Both the HTI method and the LNR test rely on the analysis of the normalised residuals to identify suspicious data. However, the core difference with the LNR test is that the HTI method further analyses the estimates of the suspicious-measurement errors. We now provide an overview of the HTI method for completeness but we refer the reader to Section 5.8 of [79] for a detailed explanation. The steps of the HTI method are the following

1. compute the normalized residuals obtained after the WLS estimation and select a set of suspicious measurements, whose normalized residuals have a high value, that are non-critical and that are linearly independent. The threshold for the selection of suspicious measurements is predefined by the user but it is assumed that all bad data is contained in the selected set. In other words, the remaining measurements are assumed to be free of bad data.
2. Compute the estimates of the errors of the measurements using the residuals and the inverse of the partition of the verification matrix that correspond to the selected set of suspicious measurements: $\hat{e}_s = G^{-1}_{\square,[s,s]} r_s$, where s is the index set of suspicious measurements.
3. Based on either a predefined fixed type I error probability (i.e. probability of false negatives) or on a predefined fixed type II error probability (i.e. probability of false positives), compute a decision threshold for each suspected measurement, according to the formulas given in Section 5.8.4 of [79].

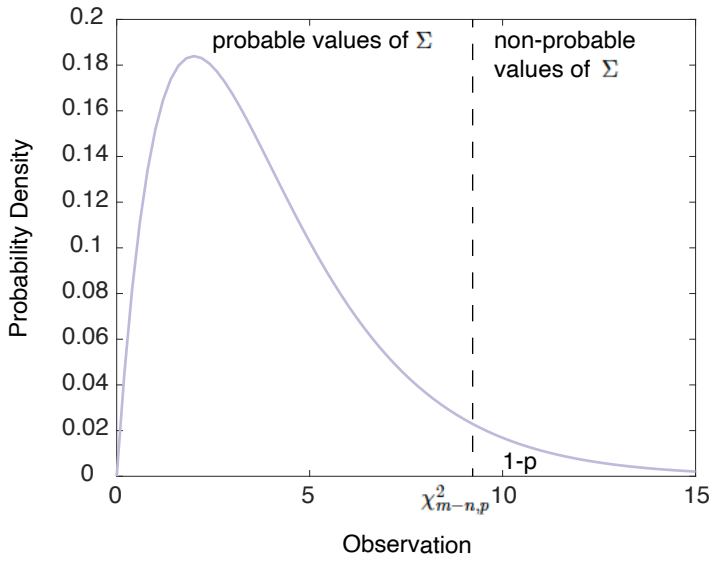


Figure 2.2 – Probability density function of the χ^2_4 distribution, the detection threshold is set for a probability value $p = 95\%$.

4. Select all measurements for whom the absolute value of the estimated measurement error is above its decision threshold computed at step 3.

5. Repeat steps 1 to 4 until the selected set of suspicious measurements doesn't change from one iteration to the next.

2.3.3 The χ^2 Test

This test relies on the fact that the sum of m independent random variables distributed according to the Standard Normal distribution, will have a χ^2 distribution with m degrees of freedom χ^2_m. Recall that the normalized residuals are distributed according to the Standard Normal distribution, therefore the sum of the normalized residuals should have a χ^2 distribution with at most $m - n$ degrees of freedom

$$\sum_{i=1}^{m} |r^N_{\square,i}| \sim \chi^2_{m-n}.$$

Note that because there are n state values, there are only at most $m - n$ linearly independent residuals. In contrast to the LNR test which identifies bad data, the χ^2 test only detects the presence of bad data, it works as follows

1. perform the WLS estimation and compute the measurement residuals $r_\square$,
2. normalize the residuals $r^N_{\square,i} = \frac{|r_{\square,i}|}{\sqrt{\Omega_{ii}}}$ $\forall 1 \leq i \leq 2m$,
3. compute the sum of the normalized residuals $\Sigma = \sum_{i=1}^{m} r^N_{\square,i}$,

4. lookup the threshold value $\chi^2_{m-n,p}$ that corresponds to a detection confidence with probability p and $m-n$ degrees of freedom (i.e. any variable that comes from a distribution χ^2_{m-n} has a probability p of being smaller or equal to $\chi^2_{m-n,p}$),
5. compare the sum of the normalized residuals Σ with the identified detection threshold $\chi^2_{m-n,p}$,
6. if $\Sigma \geq \chi^2_{m-n,p}$, then there probably is bad data in the measurements,
7. if $\Sigma < \chi^2_{m-n,p}$, then the data is plausible and no bad data is suspected.

Figure 2.2 illustrates the link between the value of p, the detection threshold $\chi^2_{m-n,p}$ and the tested value Σ.

2.4 Undetectable Time-Synchronization Attacks

In spite of the efforts made towards a robust state-estimation process by using a BDD algorithm, the seminal work in [5] shows that the state estimation is vulnerable to bad-data injections that bypass the different BDD algorithms (both χ^2 and LNR). Subsequent articles focus on the characterization and mitigation of undetectable attacks, through either prevention or detection. The authors of [7, 10] focus on establishing an index of security that quantifies the vulnerability of sets of PMUs and on smart PMU allocations to mitigate it. Malicious undetectable attacks are proposed in [6] and techniques to identify vulnerable meters are given in [8]. Bad data injection attacks are made possible by the lack of a reliable cyber-infrastructure. They can be performed by corrupting measurement data at remote terminal units, tampering with the communication network or gaining access to the local area network of the control center. In other words, they could be prevented by securing the cyber-infrastructure with cryptographic protocols.

In contrast, time-synchronization attacks (TSAs) are physical attacks that are neither prevented nor detected by the cryptographic tools used by the synchronization protocol. In such attacks, no data is modified but simply delayed. Such an attack can alter the post-processing of measurements because the precision and accuracy of the estimation from phasor measurements depends on the time synchronization of the PMUs [4]. Hence TSAs can yield a false estimation of the system state as it was shown in [14, 15]. It was shown in [80] that GPS spoofing can be used to exclusively manipulate the time reference used by PMUs, thus shifting the phase of their measured synchrophasors. The attack aims at altering various smart-grid applications but does not tackle the issue of being undetected by BDD algorithms to impact durably the state estimation.

In contrast, the authors of [14] showed how TSAs can be performed without being flagged by the residual-based BDD algorithms. Because the rest of the first part of this dissertation builds on their work, we will now present it in detail. For simplicity

of presentation, complex matrices are used in this section. However, as described in Section 2.2, decoupled real matrices are used in practice because the WLS estimator is linear over $\mathbb{R}$ and not over $\mathbb{C}$.

Attack Model. The considered attacker is able to manipulate the time synchronization of $p \geq 2$ PMUs, via GPS spoofing or delay-box insertion on transmission-lines, such that the time reference of an attacked PMU is delayed or advanced. This is equivalent to introducing p attacking angles $\delta_i, i = 1 : p$, which correspond to the phase angle shifts of the synchrophasors measured by the attacked PMUs. The equivalence is given by $\Delta t_i = \frac{\delta_i}{2\pi * f * 10^{-6}} = g_\delta(\delta_i)$, where Δt_i is the offset caused by the attack given in μs, and $f \approx 50Hz$ is the instantaneous voltage signal frequency. It is assumed that a time reference affects only one synchrophasor location. Thus, for every $i \in \{1, \cdots, p\}$, an attack changes the measured phasor z_i to $z'_i = z_i\, u_i$, where $u_i = e^{j\delta_i} \in \mathbb{T}$, and $\mathbb{T}$ is the set of complex numbers of modulus 1, therefore the phasor magnitude is unchanged. Note that TSAs are multiplicative whereas traditional false-data injection attacks [5] are additive in nature. To identify targeted measurements, let Ψ be the $M \times p$ attack-measurement indicator matrix, defined by

$$\Psi_{m,i} = \begin{cases} 1 & \text{if } \delta_i \text{ targets } z_m, \\ 0 & \text{otherwise.} \end{cases}$$

It is supposed that the attacker knows H and can observe the synchrophasors z. These assumptions are realistic as standards enforce authentication but not encryption yet.

With such capabilities, the attacker's goal is to compromise the state estimation provoking wrong power attribution. For instance the objective could be to provoke a blackout by making the system over or under-estimate the power in a region of the grid.

2.4.1 Theoretically undetectable attacks

The condition for undetectability is that the attack must not modify the residuals. In other words, the residuals obtained from the state-estimation with the attacked and unattacked measurements must be identical. Hence, the resulting state-estimation while being false, remains plausible in the sense that it could be the result of an estimation after a natural trajectory of the grid. Therefore, it is expected that no BDD scheme based on residual analysis, is able to detect the proposed attacks.

The condition for attack undetectability translates to: $Fz = Fz'$, which is made more tractable by the authors of [14] by introducing the $p \times p$ attack-angle matrix W, as the Hermitian complex matrix given by

$$W \triangleq \Psi^T \operatorname{diag}(z)^\dagger F^\dagger F \operatorname{diag}(z) \Psi. \tag{2.2}$$

As shown in (Theorem 1, [14]), an attack $\delta = (\delta_1, \ldots, \delta_p)$ is undetectable if and only if

$$W(u-1) = 0, \tag{2.3}$$

where $u = (u_1, ..., u_p)^T$. Eq. (2.3) is called the undetectability condition for an attacking vector u. Note that Eq.(2.3) is independent of the noise model. Hence, it is valid whether PMU errors can be modelled by a Gaussian distribution or not [81].

Attacking a single measurement. In the case of a single delay, W and u correspond to single complex values. Indeed, assuming without loss of generality that the first measurement is affected by the delay, we obtain from Eq. (2.2) that

$$W = (F_{:,1}z_1)^{\dagger}(F_{:,1}z_1).$$

Observe that $W = 0$ if and only if $F_{:,1} = 0$ or $z_1 = 0$. Note that attacking a measurement equal to 0 does not make any sense because it won't have any impact: shifting the phase of a null measurement will still result in a null measurement. Shown in [82] and derived again from Lemma 3.2 of Chapter 3, the column $F_{:,1}$ is in fact equal to 0 if and only if the corresponding measurement (i.e. z_1 here) is critical. In other words, if the attacked measurement has a non-zero value and is not critical, then $W \neq 0$ and computing an undetectable attack for it results in $u = 1$ which means that the phase angle difference is $\delta = 0$, thus that there is no attack. However, if the targeted measurement is critical, then $W = 0$ and any attack angle δ will satisfy Eq. (2.3). Hence, if the measurement is critical, it can be attacked undetectably by any delay of the attacker's choosing. As it is customary for grid engineers to use a PMU allocation which prevents single critical measurements, no single PMUs can be attacked undetectably in practice. Hence, it is more interesting to consider attacks targeting multiple PMUs simultaneously.

Attacking a pair of measurements. In the case of two delays, W is a 2×2 complex matrix. As shown in (Theorem 2, [14]), for a pair of measurement points for which rank(W)= 1, there is one non-trivial undetectable attack vector $\delta = (\delta_1, \delta_2)$, given by

$$\begin{aligned}\delta_1 &= 2\arg(W_{1,1} + W_{1,2}) \mod 2\pi \\ \delta_2 &= -2\arg(W_{1,2}) + 2\arg(W_{1,1} + W_{1,2}) \mod 2\pi.\end{aligned}$$

Similarly to the case where there is only one attack delay, this attack vector applied to a pair of measurement points for which W is full rank (i.e. rank(W)= 2) gives only a trivial solution which corresponds to no attack (i.e. $u_1 = u_2 = 1$). However, if W is full rank, an attack vector computed from a rank-1 approximation of W could still be undetectable in practice.

2.4.2 Practically undetectable attacks

If for a pair of sites, that each measures a single non-critical phasor, the rank of W is not equal to 1, then the authors of [14] show that it is sometimes possible to use a rank-1 approximation of W for TSAs. Because W is hermitian, it is normal with real eigenvalues. Thus it is diagonalizable by a unitary similarity transformation

$$W = U\Lambda U^{\dagger},$$

where U is a complex unitary matrix (i.e. $UU^{\dagger} = U^{\dagger}U = I$) and Λ is a diagonal matrix consisting of real non-negative eigenvalues in descending order (i.e. $\Lambda = \mathrm{diag}(\lambda_{max}, \lambda_{min})$). Construct an approximation $\widetilde{\Lambda} = \mathrm{diag}(\lambda_{max}, 0)$ by setting the smallest eigenvalue to zero. Define the rank-1 approximation of W as

$$\widetilde{W} = U\widetilde{\Lambda}U^{\dagger}$$

and compute the attack vector from $\widetilde{W}$. This solution is effective if the index of separation (IOS), defined in [14] as the largest eigenvalue of W over the sum of both its eigenvalues, is close to 1

$$\mathrm{IoS} = \frac{\lambda_{max}}{\lambda_{max} + \lambda_{min}} \approx 1.$$

This occurs if the largest eigenvalue is significantly larger than the remaining one. Hence, pairs of PMUs for which W is either of rank equal to 1 or has an IoS approximately equal to 1 can be attacked undetectably. The IoS can be computed in closed form using the following formula

$$\mathrm{IoS} = \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\frac{\det(W)}{\mathrm{Trace}(W)^2}}. \tag{2.4}$$

Notice that W and thus the computation of the IoS depends on the measurement vector. In contrast, as shown in (Theorem 3, [14]), the minimum value of IoS taken over all measurement values can be computed independently from the measurements

$$\mathrm{IoS}^* = \min_z \mathrm{IoS} = \frac{1}{2} + \frac{|f_{12}|}{2(f_{11}f_{22})^{\frac{1}{2}}}, \tag{2.5}$$
$$\text{where } f_{ik} = \left(\sum_{l=1}^{m} \overline{F}_{l,i}F_{l,k}\right)\left(\sum_{l=1}^{m} \overline{F}_{l,k}F_{l,i}\right),$$

and where $\overline{F}_{l,k}$ is the complex conjugate of $F_{l,k}$. Hence, an attacker can find vulnerable pairs of measurements even if he does not have access to the measurement values. The authors of [14] propose to check the IoS^* value of each pair of measurements in order to find vulnerable pairs from knowledge on the topology and on the admittance matrix only (i.e. without checking the measurement values, which change at each time-instant). Observe that if the IoS^* is not close to 1, then there could still be a pair of measurement

values for which the IoS is close to 1, which happens if the ratio between the magnitude of the attacked measurements is either very small or very large.

Figure 2.3 summarises the different strategies that an attacker can use to perform an undetectable attack on a pair of PMUs which both measure a single synchrophasor only. Notice that without access to the measurement values an attacker can find valid attack locations but that in order to perform the attack he needs to be able to observe the measurement vector.

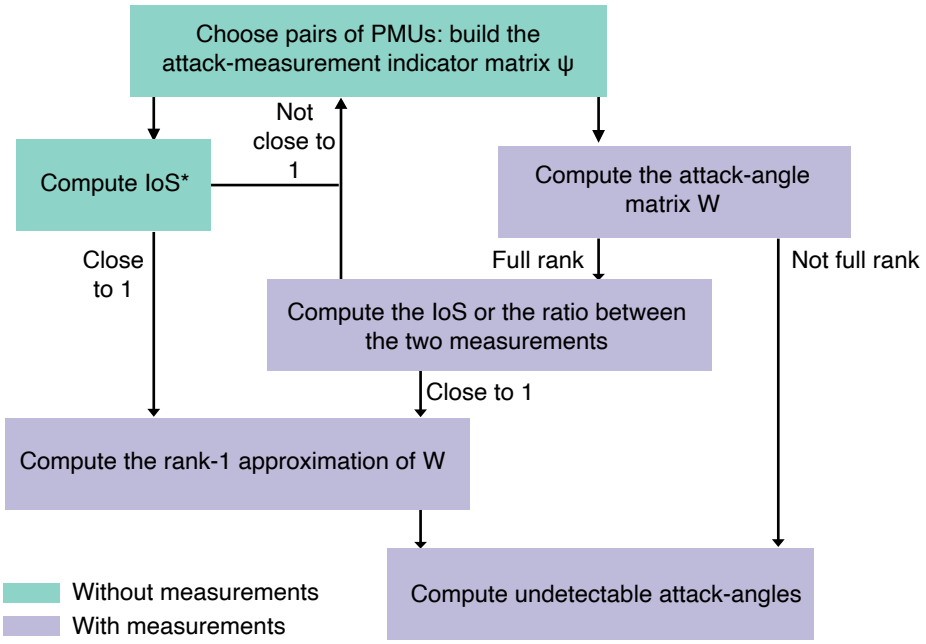


Figure 2.3 – Flow chart for finding valid attack locations and corresponding attack angles.

2.4.3 Extending the attack

Another important contribution of [14] is to demonstrate how attacks on disjoint pairs of PMUs can be performed simultaneously in a larger more damaging undetectable attack. This is due to the fact that attacking disjoint pairs of PMUs will impact disjoint pairs of measurements. Hence the construction of a pair's attack-angle matrix W is not altered by an attack on another pair. Therefore the attacks are independent of each other and can be performed simultaneously without influencing each other.

When it comes to overlapping pairs, the authors showed that an undetectable attack would have to be sequential instead of simultaneous. The attacker can use the IoS values computed prior to the attack to find valid locations as it was shown in (Theorem 7, [14]) that the IoS value is unchanged after an attack. In order to perform the overall attack simultaneously, he needs to compute the overall attack angle vector by searching

for the sequential attack-angles that he then combines. It was shown experimentally that different sequences of attacks on the same set of pairs of PMUs lead to different attack angles. It is thus in the attacker's best interest to optimize the attack in order to maximize his objective. For instance, the latter can be to create an overestimation of the power flow on a specified line. Finally, a greedy algorithm in which the pair that maximizes the attacker's objective is picked at each step, is proposed in [14].

3

Feasibility of Time-Synchronization Attacks against PMU-based State-Estimation

The results of [14] presented in Section 2.4 of Chapter 2 show how to compute an undetectable time-synchronization attack (TSA) against pairs of PMUs. It is done by solving a specific set of non-linear equations, yielding a discrete and finite set of attacks. In the case where these equations lead to no attack solution, an approximated set of equations that yields an attack solution is proposed. The authors also discuss how attacks against pairs of PMUs can be combined to maximize the attacker's objective, which could be to maximize the error on a particular line power-flow, for instance. However, the clock adjustment rate of a PMU is controlled in practice by a controller typically called the clock servo [83]. The latter ensures that clock adjustments always stay below a well-defined threshold. Hence, a delay above this threshold will not be implemented by the clock servo. Instead, it will transform the intended delay into a smaller one which may not be in the discrete and finite set of undetectable attacks. Therefore, if an attacker blindly performs an attack on pairs of PMUs according to the results of [14], without anticipating the actions of the clock servo, it is likely that the intended attack-angles will be modified into smaller detectable ones. Consequently, the impact on the measurements might not be the one intended by the attacker and the attack could become detectable by the BDD algorithms. In order to overcome the limitations of the results of [14], three major contributions are made in this chapter.

First, it is shown that the set of undetectable TSAs against three or more PMUs forms a connected compact set and that the number of valid attacks is uncountably infinite. This allows the attacker to anticipate the actions of the clock servo and to remain undetected by injecting small incremental delays over a period of time until the objective is maximized. It is proven that attacks can be tailored to remain undetected by reaching an optimal attack-angle in a continuous manner. It is also shown how to compute this set of valid attacks.

The second contribution of this chapter is to address the practical feasibility of performing TSAs by taking PMU clock constraints in consideration. Algorithms that can achieve practically undetectable attacks under realistic conditions are proposed. In order to find attack-angles against a specific set of PMUs, it is required to build an attack angle matrix using the system topology, the measurements and the choice of PMUs to attack. When this matrix is of rank approximately or exactly equal to 1, the results of this chapter show how to perform an attack. The vulnerability of a set of PMUs depends on how close the rank of the corresponding matrix is to 1. This measure is captured by the index of separation (IoS) introduced in [14] and presented in Chapter 2, which only considers pairs of PMUs.

As a third contribution, it is shown how vulnerable sets of an arbitrary number of PMUs can be found. The theory is extended in order to show that synchrophasors can be grouped in equivalence classes and that members of a class form a vulnerable set. This sufficient condition allows to efficiently find sets of PMUs that can be targeted in one simultaneous attack. By analysing the infimum of the IoS over all measurements, it is shown that vulnerable sets can be identified based on the system topology only, without having to read measurement values. Furthermore, the chapter provides evidence that under some topological conditions it is possible to attack PMUs that measure both voltage and current phasors (both sharing the same time reference). Finally, it also provides numerical evidence that attacks can be feasible against the time synchronization of PMUs whose measurements are not exactly critical.

This chapter is an extension of a preliminary conference version [84], where the practical feasibility of TSAs was not considered. Contrary to [84], this chapter shows that TSAs can be mounted undetectably by satisfying the constraints imposed by the PMU clock servo. In addition, it provides a sufficient condition for finding critical groups of measurements of arbitrary size, and it also shows that vulnerable sets without this condition exist. Finally, going beyond results in [84] it shows that it is possible to attack PMUs that measure two distinct synchrophasors simultaneously, and shows that the attacks bypass robust state estimation as well. The author of this dissertation joined the project after the publication of the preliminary conference version.

Attack Model. The system model used in this chapter is the same as the one presented in Section 2.4 of Chapter 2. The only feature added to the attack model of [14], also presented in Section 2.4 of Chapter 2, is that it is further supposes that the attacker is able to anticipate the actions of a regular PMU clock servo. This attack model is strong for two reasons. First, recent attacks on critical infrastructures, e.g., Stuxnet [85], had access to detailed system information. Second, such a strong model enables engineers to identify vulnerable data that require protection.

The rest of the chapter is organised as follows. Section 3.1 gives expressions to compute

the set of possible attack-angles. In Section 3.2 results on how to efficiently find sets of vulnerable PMUs are explained. In Section 3.3, the practical feasibility of deploying an undetectable attack is discussed by considering multiple strategies. Numerical results are presented in Section 3.4 to illustrate the effectiveness of the attack. Finally, Section 3.5 concludes the chapter.

3.1 Computing Undetectable Attack-Angles

This section presents a closed form expression to compute an undetectable TSA involving $p = 3$ time references. Then it shows how to extend this result for any $p \geq 2$. This contribution represents a great improvement compared to the case of $p = 2$ considered in [14], as it allows a continuum of non-trivial feasible attacks, which is needed to address the constraints required by the clock servo.

3.1.1 Computing Attack Angles for $p = 3$

The measurements taken by the three PMUs to be attacked are denoted by $[z_1, z_2, z_3]$, and the corresponding attack-angles by $\delta_1, \delta_2, \delta_3$. It has been shown in [14] that attacks are feasible when the effective rank of W is 1. In this case, (2.3) can be rewritten as

$$w_1(u_1 - 1) + w_2(u_2 - 1) = -w_3(u_3 - 1), \tag{3.1}$$

where $w = [w_1 w_2 w_3]$ is the row of the attack-angle matrix W that has the largest norm.

In what follows $C = (c, r)$ denotes a circle in the complex plane with center c and radius r. An algebraic approach is used to solve (3.1) and to provide a closed-form solution. Namely, the equation is interpreted as the intersection of the right-hand side with the left-hand side, which represent in the complex plane a circle $C_3 = (w_3, |w_3|)$, and an annular region defined by an inner circle $C_i = (-(w_1 + w_2), ||w_1| - |w_2||)$ and an outer circle $C_o = (-(w_1 + w_2), |w_1| + |w_2|)$. The following result characterises the set Θ_3 of feasible values for δ_3.

Proposition 3.1. *For $p = 3$ and rank$(W) = 1$, the set U_3 of feasible values of u_3, i.e., $U_3 = \{u_3 : u_3 = e^{i\delta_3} \ \forall \delta_3 \in \Theta_3\}$ is either a non-empty connected compact subset of $\mathbb{T}$ or the union of two non-empty connected compact subsets of $\mathbb{T}$, where $\mathbb{T}$ is the set of complex numbers of modulus 1. Furthermore, $u_3 = 1 \in U_3$.*

Proof. Let $\mathcal{I}_o$ and $\mathcal{I}_i$ be the set of intersection points of the circle C_3 with the outer and the inner circle, respectively. Four cases are distinguished.

1. $|\mathcal{I}_o| + |\mathcal{I}_i| = 1$, i.e., C_3 is tangent to one of the circles. This intersection point must be the one corresponding to $\delta_3 = 0$, because $\delta_1 = \delta_2 = \delta_3 = 0$ (no attack) is a

Algorithm 1 Compute-Feasible-Angles(w)

Input: w (row of largest norm of W)
$C_3 \leftarrow (w_3, |w_3|)$
$C_i \leftarrow (-(w_1 + w_2), ||w_1| - |w_2||)$
$C_o \leftarrow (-(w_1 + w_2), |w_1| + |w_2|)$
Compute $\mathcal{I}_i = C_3 \cap C_i = \{I_1, I_2\}$.
Compute $\mathcal{I}_o = C_3 \cap C_o = \{I_1, I_2\}$.
Compute Θ_3 using Proposition 3.1
Output: Θ_3

solution to (3.1). Thus $\Theta_3 = \{0\}$.

2. $2 \leq |\mathcal{I}_o| + |\mathcal{I}_i| < 4$, i.e., C_3 intersects with one of the circles at two points and could be tangent to the other circle. Let the two intersection points (not the tangent) correspond to angles δ_3^1 and δ_3^2. If $\{\delta_3^1, \delta_3^2\} \in [0, 2\pi]$ and $\delta_3^1 < \delta_3^2$ then we have two intervals, $[\delta_3^1, \delta_3^2]$ and $[\delta_3^2, \delta_3^1 + 2\pi]$, and the set of feasible values is the one including 0, since $\delta_3 = 0$ is a feasible solution. Hence, $\Theta_3 = [\delta_3^2, \delta_3^1 + 2\pi]$.

3. ($|\mathcal{I}_o| = |\mathcal{I}_i| = 2$), i.e., four intersection points. Let the corresponding angles in increasing order be $\{\delta_3^1, \delta_3^2, \delta_3^3, \delta_3^4\}$. Observe that due to the ordering, angles 1 and 2 correspond to intersection points with the same circle. The feasible set consists of the intervals between angles that correspond to intersection points with different circles. Thus, $\Theta_3 = [\delta_3^2, \delta_3^3] \cup [\delta_3^4, \delta_3^1 + 2\pi]$. Notice that the second interval includes $\delta_3 = 0$.

4. $|\mathcal{I}_o| + |\mathcal{I}_i| = 0$ or $|\mathcal{I}_o| + |\mathcal{I}_i| = \infty$. Since $\delta_3 = 0$ is a feasible solution, it is clear that $|\mathcal{I}_o| + |\mathcal{I}_i| = 0$ implies that C_3 is inside the annular region, while $|\mathcal{I}_o| + |\mathcal{I}_i| = \infty$ implies C_3 coincides with one of the circles. Thus, in both cases $\Theta_3 = [0, 2\pi[$

Note that Θ_3 always includes the intersection angles because they correspond to feasible solutions, hence the set of feasible solutions is closed. Furthermore, due to the structure of the circle group $\mathbb{T}$, an interval of feasible angles maps into a connected set. Moreover, in all four cases, $0 \in \Theta_3$. In other words, $1 \in U_3$. □

Algorithm 1 shows the pseudo-code for computing the set Θ_3 of feasible values for the attack angle δ_3. Let $C_x \cap C_y$ denote the intersection between two circles C_x and C_y. This can be efficiently computed by the following Lemma.

Lemma 3.1. *Consider two circles, $C_x = (c_x, r_x)$ and $C_y = (c_y, r_y)$, in the complex plane. Assume that $r_x > r_y$ and that the two circles intersect. Let $\mathcal{I} = \{I_1, I_2\}$ be the set of intersection points. I_1 and I_2 are given by*

$$I_1 = c_x + a + h, \quad I_2 = c_x + a - h$$

where

$$a = d \cdot \frac{d\bar{d} + r_x^2 - r_y^2}{2d\bar{d}}, \quad h = d \cdot i \cdot \sqrt{\frac{r_x^2 - a\bar{a}}{d\bar{d}}}, \quad d = c_y - c_x.$$

Proof. Figure 3.1 illustrates the problem of finding the intersection of the circles. Let p_f be the point of intersection of the line connecting c_x to c_y and the radical axis of the two circles. Let d be the vector directed from c_x to c_y, that is, $d = c_y - c_x$. Furthermore, let vector a be the vector directed from c_x towards p_f, and h be the vector directed from p_f to I_1. Note that a points in the same direction as d and h is perpendicular to both vectors. By inspecting the two triangles (c_x, p_f, I_1) and (c_y, p_f, I_1) the following two equalities hold

$$|a|^2 + |h|^2 = r_x^2, \quad (|d| - |a|)^2 + |h|^2 = r_y^2$$

solving the two equations for $|a|$ and using $|d|^2 = d\bar{d}$ leads to

$$|a| = \frac{d\bar{d} + r_x^2 - r_y^2}{2|d|}.$$

Since a is parallel to d, it is the case that $a = d \cdot \frac{|a|}{|d|}$, which yields the expression for a in the lemma. Because $|h| = \sqrt{r_x^2 - a\bar{a}}$, and since h is perpendicular to d, it must be that $h = d \cdot i \cdot \frac{|h|}{|d|}$, which in turn yields the expression for h in the lemma. The intersection points can be computed as

$$I_1 = c_x + a + h, \quad I_2 = c_x + a - h.$$

Note that if the two circles intersect only at one point, then $h = 0$, leading to $I_1 = I_2$. □

For each possible attack-angle δ_3, valid attack-angles δ_1 and δ_2 can be efficiently computed as follows. Substituting $s = -w_3(u_3 - 1)$ into (3.1) leads to

$$w_1(u_1 - 1) = s - w_2(u_2 - 1). \tag{3.2}$$

Proposition 3.2. *For each $\delta_3 \in \Theta_3$ there exist either one or two pairs of (δ_1, δ_2). A closed form expression of such a pair is given by*

$$u_1 = \frac{I}{w_1} + 1, \quad u_2 = \frac{-w_3(u_3 - 1) - I}{w_2} + 1 \tag{3.3}$$

$$\delta_1 = arg(u_1), \quad \delta_2 = arg(u_2), \tag{3.4}$$

where I corresponds to an intersection point between the left-hand side of (3.2) *denoted by $C_1 = (-w_1, |w_1|)$ and the right-hand side denoted by $C_2 = (w_2 - w_3(u_3 - 1), |w_2|)$.*

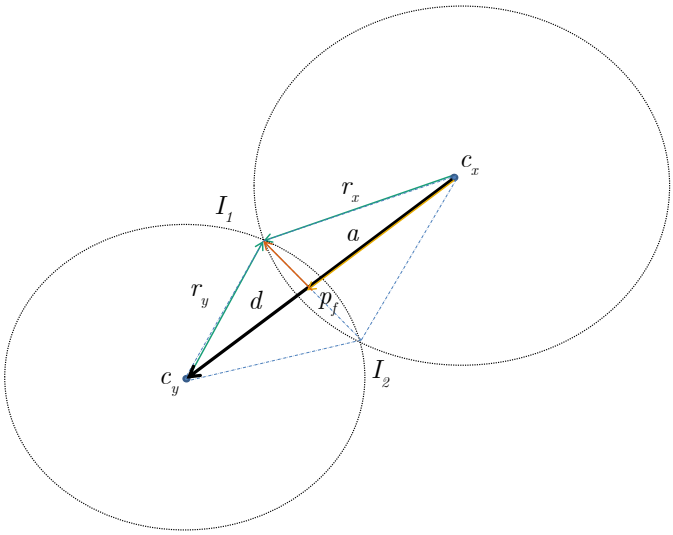


Figure 3.1 – Example illustrating the intersection between two circles

Algorithm 2 Compute-Angle-Pairs(δ_3, w)

Input: $\delta_3 \in \Theta_3$, and w (row of largest norm of W).
  $S \leftarrow \emptyset$
  $s \leftarrow w_3(e^{i\delta_3} - 1)$
  $C_1 \leftarrow (-w_1, |w_1|)$
  $C_2 \leftarrow (w_2 + s, |w_2|)$
  Compute $\mathcal{I}_{12} = C_1 \cap C_2 = \{I_1, I_2\}$.
  for all $I \in \mathcal{I}_{12}$ **do**
    Compute (δ_1, δ_2) by using I in equations (3.3) , (3.4)
    $S \leftarrow S \cup (\delta_1, \delta_2)$
  end for
Output: S

Proof. Both the left- and right-hand sides in (3.2) represent circles in the complex plane. They will be referred to as C_1 and C_2, respectively. C_1 is centered at $c_1 = -w_1$ with radius $r_1 = |w_1|$, and C_2 is centered at $c_2 = w_2 + s$ with radius $r_2 = |w_2|$. An intersection point of these circles corresponds to a solution to (3.2). The two circles intersect as $\delta_3 \in \Theta_3$. Again, Lemma 3.1 can be used to find the set of intersection point(s) I_{12}, and each intersection point corresponds to a pair (δ_1, δ_2). For each intersection point $I \in I_{12}$ the corresponding (u_1, u_2) and (δ_1, δ_2) can be computed by equating the left and right hand sides of (3.2) to I, yielding the expression in the proposition. □

The procedure of computing (δ_1, δ_2) is illustrated in Algorithm 2. To summarize, the set $\Theta(z) \subset \mathbb{R}^3$ of undetectable attacks for $p = 3$ is a two-dimensional manifold in $\mathbb{R}^3$, characterized by one degree of freedom, i.e. it is a curve.

3.1.2 Computing Attack Angles for any $p \geq 2$

The following describes an algorithm for computing undetectable attacks for the general case of $p \geq 2$. In this case, (3.1) becomes

$$\sum_{i=1}^{p-1} w_i(u_i - 1) = -w_p(u_p - 1), \tag{3.5}$$

where w_i is the entry in the row of the largest norm and the i^{th} column of W. In (3.5), the right hand side represents a circle $C_p = (w_p, |w_p|)$ in the complex plane, while the left hand side represents an annular region that is defined by an inner circle $C_i = (-\sum_{i=1}^{p-1} w_i, \max\{0, 2|w_{i^*}| - \sum_{i=1}^{p-1} |w_i|\})$ and an outer circle $C_o = (-\sum_{i=1}^{p-1} w_i, \sum_{i=1}^{p-1} |w_i|)$, where $i^* = \arg\max_{i \in \{1..p-1\}} |w_i|$. Similar to the procedure of the case when $p = 3$, the feasible set Θ_p of δ_p can be computed by Algorithm 1, given the parameters of the circles. For any choice of $\delta_p^* \in \Theta_p$ (and corresponding u_p^*) equation (3.5) can be rewritten as

$$\sum_{i=1}^{p-2} w_i(u_i - 1) = -w_{p-1}(u_{p-1} - 1) + s_p,$$

where $s_p = -w_p(u_p^* - 1)$. Again, Algorithm 1, with the appropriate parameters of the circles, can be used to compute the feasible range Θ_{p-1} of δ_{p-1}. Computing the feasible regions for $p - 2$ iterations results in

$$w_1(u_1 - 1) = \sum_{i=3}^{p} s_i - w_2(u_2 - 1). \tag{3.6}$$

Notice that (3.6) has the same form as (3.2). Therefore, δ_1 and δ_2 can be computed using Algorithm 2. Hence, it is expected that the set $\Theta(z) \subset \mathbb{R}^p$ of undetectable attacks is a $p - 1$ dimensional manifold in $\mathbb{R}^p$, characterized by $p - 2$ degrees of freedom.

3.2 Finding Sets of p Vulnerable Measurements

This section establishes a sufficient condition for finding vulnerable sets of p PMUs.

Recall that the attacks target sets of measurements whose corresponding W matrix is of rank approximately or exactly equal to 1. Therefore, the vulnerability of a set is given by a measure of how close to 1 the rank of W is. This quantity is introduced in [14] for $p = 2$ measurements as the index of separation (IoS), the authors also showed that its infimum (IoS*) can be computed from the topology only without access to the measurement values. For a pair of measurements (z_i, z_k), if the value of the IoS of the corresponding W matrix, denoted by $IoS_{(i,k)}(z_i, z_k)$, is close to 1, then W is well approximated by a matrix of rank 1 and the attack can be performed on the pair. If $IoS^*_{(i,k)} = \min_{z_i, z_k} IoS_{(i,k)}(z_i, z_k)$ is close to 1, then whatever the actual measurement value, $IoS_{(i,k)}$ will also be close to

1 hence the pair of PMUs is vulnerable. Note that a pair for which IoS^* is far from 1 might still be vulnerable at a particular time instant due to measurement values for which IoS gets close to 1. Both can be computed in closed form using Eqs.(2.4) and (2.5).
The next contribution of this chapter is to extend the theory by showing that measurements can be grouped in classes such that any combination of measurements within a class produces a W matrix of rank equal to 1. Our results rely on the following lemmas.

Lemma 3.2. *Assume that $m \geq n+1$ (otherwise every measurement is critical). Consider the W matrix associated with the set of measurements $\mathcal{P} = \{1, ..., p\}$. Let $i \in \mathcal{P}$ and assume that $z_i \neq 0$. The measurement i is critical if and only if $W_{i,i} = 0$.*

Proof. Observe that $W_{i,i} = |z_i|^2 \sum_{k=1}^{m} |F_{k,i}|^2$ and thus $W_{i,i} = 0$ if and only if the first column of the verification matrix is identically 0, which is equivalent to $Fe^{(i)} = 0$ where $e^{(i)}$ is the column vector with 1 in the i^{th} row and 0 elsewhere. This means that any attack against the i^{th} measurement alone is undetectable; by a reasoning similar to the proof of Theorem 1 in [6], this is equivalent to measurement i being critical. □

Lemma 3.3. *[84] Let $W^{(i_1,..,i_k)}$ be the $k \times k$ principal submatrix of W composed of the rows and columns or W of indices $\{i_1, .., i_k\}$, with $p \geq k \geq 2$. If W is hermitian positive semi-definite and $W^{(i_1,..,i_k)}$ is singular then W is singular.*

Proof. We will call a vector x isotropic with respect to a matrix W if $x^H W x = 0$. Null vectors (i.e. vectors such that $Wx = 0$) are obviously isotropic; the converse is not true in general, but is true when W is hermitian semi-definite (this can easily be seen by diagonalization). As $W^{(i_1,..,i_k)}$ is singular, there is a non-zero vector $x^{(i_1,..,i_k)} \in \mathbb{C}^k$ such that $W^{(i_1,..,i_k)} x^{(i_1,..,i_k)} = 0$. Subsequently, $x^{(i_1,..,i_k)}$ is isotropic with respect to $W^{(i_1,..,i_k)}$. We obtain the vector $\hat{x}^{(i_1,..,i_k)} \in \mathbb{C}^p$ by expanding $x^{(i_1,..,i_k)}$ to dimension p and by filling missing values with zeros (we assume W is a matrix of dimensions $p \times p$). $\hat{x}^{(i_1,..,i_k)}$ is isotropic with respect to W, hence is in the null-space of W. Since $\hat{x}^{(i_1,..,i_k)}$ is non-zero, W is singular. □

Theorem 3.1. *For a given value of the measurement vector z and for two measurements (i, k) we say that $i\mathcal{R}k$ if and only if $i = k$ or $IoS_{i,k}(z_i, z_k) = 1$.*

1. *The relation $\mathcal{R}$ defined in this way is an equivalence relation over the set of all possible measurements.*
2. *For any set $\mathcal{P}$ of $p \geq 2$ measurements, the corresponding W matrix has rank 1 if and only if all measurements in $\mathcal{P}$ are equivalent under $\mathcal{R}$.*

Proof. 1. Consider the 3×3 attack-angle matrix $W^{(i,j,k)}$ computed for a set of measurements with indices $\{i, j, k\}$. It is hermitian positive semi-definite because of (2.2). By

Lemma 3.2 the diagonal terms $W_{i,i}$ are non-zero because the considered measurements are non-critical and non-zero. The principal submatrix obtained by adding row and column j is singular because $IoS_{i,j} = 1$ by hypothesis. The same holds if we add row and column k instead of j. By Lemma 3.3, it follows that $W^{(j,k)}$ is singular, therefore $IoS_{(j,k)} = 1$. This shows that $\mathcal{R}$ is an equivalence relation.
2. Assume all measurements in $\mathcal{P}$ are in the same equivalence class. The rank of W is 1 by Lemma 3.2 and Theorem 15 in [86]. Conversely, if the rank of W is 1, then the rank of all principal 2×2 submatrices is ≤ 1 and, again by Lemma 3.2, is exactly 1, which shows that $IoS_{i,k}(z_i, z_k) = 1$ for all $i, k \in \mathcal{P}$. □

The first item implies that IoS has the transitivity property. In practice, Theorem 3.1 gives a sufficient condition for finding vulnerable sets of size p. An attacker will look for a set $\mathcal{P}$ of at least p measurements that mutually have $IoS = 1$. Any combination of at least 2 measurements within $\mathcal{P}$ has a W matrix of rank 1. Hence all or a subset of the measurements of this set can be the target of a powerful attack. This method is efficient, compared to a brute-force approach, which includes computing the rank of W for each combination of p PMUs, and thus has exponential complexity, which makes it intractable even for small size grids.

Theorem 3.2 establishes a similar result that holds for IoS^*. The proof is not given as it is similar to the proof of Theorem 3.1.

Theorem 3.2. *For two measurements (i, k) we say that $i\mathcal{R}^*k$ if and only if $i = k$ or $IoS^*_{i,k} = 1$.*

1. *The relation $\mathcal{R}^*$ defined in this way is an equivalence relation over the set of all possible measurements.*
2. *For any set $\mathcal{P}$ of $p \geq 2$ measurements, the corresponding W matrix has rank 1 if all measurements in $\mathcal{P}$ are equivalent under $\mathcal{R}^*$.*

Note that for any value of the measurement vector z, the relation $\mathcal{R}^*$ is a subset of $\mathcal{R}$.

Theorem 3.2 is thus more restrictive and enables to find fewer attackable measurements than Theorem 3.1 does. Nonetheless, the relation $\mathcal{R}^*$ can be computed a-priori without knowledge of the actual values of the measurements.

The next result shows that there is a link between the criticality of a set of measurements and its vulnerability to TSAs based on rank-1 approximations. First the independence of measurements is defined.

Definition 3.1. *Two measurements are said to be independent if the corresponding rows of the H matrix are linearly independent over $\mathbb{C}$. When two measurements are*

not independent, one is a known complex multiple of the other, i.e., they are essentially measuring the same complex quantity.

Using this definitions the following result can be formulated.

Theorem 3.3. *Assume that the number of measurements m and the number of states n satisfy $n \geq 3$ and $m \geq n+2$. Assume that every single measurement is non-critical. Then, a pair of measurements $\{i,k\}$ (with $i \neq k$) is a critical set if and only if i,k are independent and $IoS^*_{i,k} = 1$.*

Proof. Assume without loss of generality that the measurement pair is $\{1,2\}$. Let $H^T = (H_1^T, H_2^T)$ where H_1 is a $2 \times n$ complex matrix and H_2 is an $(m-2) \times n$ complex matrix. It follows that $H^\dagger H = H_1^\dagger H_1 + H_2^\dagger H_2$. Since the system is observable, H has full rank; define $A \stackrel{\text{def}}{=} (H^\dagger H)^{-1}$. The complex verification matrix $F = HAH^\dagger - I_m$ that corresponds to the pair $\{1,2\}$ can be put in the form

$$\begin{pmatrix} F_1 & F_2 \\ F_3 & F_4 \end{pmatrix} \stackrel{\text{def}}{=} \begin{pmatrix} H_1AH_1^\dagger - I_2 & H_1AH_2^\dagger \\ H_2AH_1^\dagger & H_2AH_2^\dagger - I_{m-2} \end{pmatrix}$$

Also define $Q \stackrel{\text{def}}{=} F_1^\dagger F_1 + F_3^\dagger F_3$ so that the W matrix is $W = \text{diag}(\bar{z}_{1:2})\, Q\, \text{diag}(z_{1:2})$.

1. Now the if part of the theorem can be proven. Assume that $IoS^*_{1,2} = 1$ and i,j are independent. By definition of IoS^*, W and therefore Q do not have full rank. Also, since 1 and 2 are independent, H_1 has full rank. The proof proceeds by contradiction: assume that $\{1,2\}$ is non critical, it then follows that H_2 has full rank, which is a contradiction by Lemma 5.

2. Conversely, assume that $\{1,2\}$ is a critical pair. The rank of H_2 is $\leq n-1$, thus nullity$(H_2) \geq 1$. By Lemma 3.5, nullity$(Q) \geq 1$ and thus Q does not have full rank; thus the same holds for W for any value of z, i.e. $IoS^*_{1,2} = 1$.

Furthermore, assume that measurements 1 and 2 are not independent. Since $\{1,2\}$ is a critical pair, one of the two measurements is also critical, which is impossible by hypothesis. □

Lemma 3.4. *For $v \in \mathbb{C}^n$: if $H_2 v = 0$ then $F_1 H_1 v = F_3 H_1 v = 0$.*

Proof. By definition of A, $AH^\dagger H = I_n$ hence $AH_1^\dagger H_1 + AH_2^\dagger H_2 = I_n$, thus

$$\begin{aligned} F_1H_1 &= H_1AH_1^\dagger H_1 - H_1 && (3.7)\\ &= H_1 - H_1AH_2^\dagger H_2 - H_1 = -H_1AH_2^\dagger H_2 && (3.8)\\ F_3H_1 &= H_2AH_1^\dagger H_1 = H_2 - H_2AH_2^\dagger H_2. && (3.9) \end{aligned}$$

The lemma follows immediately. □

Lemma 3.5. *If H has full rank then nullity$(H_2) \leq$ nullity(Q).*

Proof. By Lemma 3.4, for all v such that $H_2 v = 0$, $H_1 v$ is in the nullspace of Q.

Let $k = \text{nullity}(H_2)$ and let $(v_1, ...v_k)$ be k linearly independent vectors in the right nullspace of H_2. Let us show that $(H_1 v_1, ...H_1 v_k)$ are linearly independent. First observe that for $i = 1...k$ we have

$$Hv_i = \begin{pmatrix} H_1 v_i \\ H_2 v_i \end{pmatrix} = \begin{pmatrix} H_1 v_i \\ 0 \end{pmatrix}. \tag{3.10}$$

Now assume that for some complex numbers $\lambda_1, ..., \lambda_k$:

$$\lambda_1 H_1 v_1 + ... + \lambda_k H_1 v_k = 0 \tag{3.11}$$

it follows that

$$\lambda_1 H v_1 + ... + \lambda_k H v_k = 0 \tag{3.12}$$

$$H(\lambda_1 v_1 + ... + \lambda_k v_k) = 0 \tag{3.13}$$

Since H has full rank and $n < m$ it follows that the nullity of H is 0. The previous equation thus implies that

$$\lambda_1 v_1 + ... + \lambda_k v_k = 0 \tag{3.14}$$

which implies that $\lambda_i = 0$ for $i = 1 : k$. Thus the only null linear combination of $H_1 v_1, ...H_1 v_k$ is the trivial one, which means that $H_1 v_1, ...H_1 v_k$ are linearly independent. Since they are all in the nullspace of Q, it follows that $\text{nullity}(Q) \geq k = \text{nullity}(H_2)$. □

Lemma 3.6. *If H_1 and H_2 both have full rank then F_1, F_3 and Q also have full rank.*

Proof. 1. Since H_2 has full rank and $n \leq m-2$, the rank of $H_2 H_2^\dagger$ is n and $\text{range}(H_2 H_2^\dagger) = \mathbb{C}^n$. Similarly, A is invertible and thus $\text{range}(A H_2 H_2^\dagger) = \mathbb{C}^n$. Also, H_1 has full rank and $n > 2$ thus $\text{range}(H_1) = \mathbb{C}^2$. Thus $\text{range}(H_1 A H_2^\dagger H_2) = \mathbb{C}^2$.

Now using (3.8), it follows that $\text{range}(F_1 H_1) = \mathbb{C}^2$. Thus

$$\mathbb{C}^2 = \text{range}(F_1 H_1) \subseteq \text{range}(F_1) \subseteq \mathbb{C}^2, \tag{3.15}$$

hence $\text{range}(F_1) = \mathbb{C}^2$ and F_1 has full rank.

2. Let v be in the nullspace of F_3, i.e. $H_2 A H_1^\dagger v = 0$. Since H_2 has full rank and $n \leq m-2$, the nullspace of H_2 is reduced to 0, thus $A H_1^\dagger v = 0$. Now A is invertible thus $H_1^\dagger v = 0$. Furthermore, since H_1 (hence also $H_1^\dagger$) has full rank and $2 \leq n$, the nullspace

of $H_1^\dagger$ is reduced to 0 and finally $v = 0$. Thus the nullspace of F_3 is reduced to 0. Since $2 \leq m - 2$, F_3 has full rank.

3. Let v be in the nullspace of Q. It follows that $v^\dagger F_1^\dagger F_1 v + v^\dagger F_3^\dagger F_3 v = 0$, i.e. $\|F_1 v\|^2 + \|F_3 v\|^2 = 0$ thus $F_3 v = 0$ and $v = 0$ by item 2. Since Q is a square matrix, it follows that Q has full rank. □

In [14], large attacks targeted groups of pairs of PMUs such that their IoS^* values were equal to 1. In this chapter, Theorem 3.2 further establishes that such groups are in fact equivalence classes. Theorem 3.3 gives a novel condition for identifying vulnerable sets by finding critical pairs. This new technique is linked to the analysis of the rank of matrix H while studying IoS and IoS^* values correspond to analysing the rank of matrix W and values of matrix F respectively. In order to attack $p \geq 2$ PMUs, a target set can thus be found by identifying sets of measurements of cardinality p such that all combinations of two measurements are critical or have a corresponding IoS^* value equal to 1. Incidentally, Theorems 3.2 and 3.3 also establish that the criticality of a pair defines an equivalence relation on the set of non-critical measurements, a result of independent interest that can be used in other contexts than TSAs. Also note that vulnerability to TSAs with rank-1 approximations is not strictly equivalent to pairwise criticality: it is possible to find non-critical pairs of measurements that have $IoS \approx 1$ for some values of the measurement vector z (see Section 3.4).

3.3 Strategies for Implementing Undetectable Attacks

This section discusses how an attacker could use the presented methods to mount a TSA.

3.3.1 Computing an Optimal Undetectable Attack

For attacking $p = 3$ measurements, consider that the attacker has an objective function $\varphi(z, \delta_1, \delta_2, \delta_3)$ to maximize, e.g., the difference between the estimated and the actual power-flow on a transmission line. The attacker can observe the measurements taken at time instants $\{t^0, t^1, \cdots, t^k, \cdots\}$ and knows the instantaneous attack-angles $\delta^k = \{\delta_i^k, i \in \{1, \cdots, p\}\}$ that he already implemented. Therefore, given an observed measurement $z^{k'}$ taken at time t^k (possibly already attacked), the attacker can compute the non-attacked measurement $z^k = \{z_i^k : z_i^k = z_i^{k'} e^{-j\delta_i^k}, i \in \{1, \cdots, p\}\}$, and the angles $(\delta_1^{k*}, \delta_2^{k*}, \delta_3^{k*}) = \arg\max_{(\delta_1,\delta_2,\delta_3)\in\Theta(z^k)} \varphi(z^k, \delta_1, \delta_2, \delta_3)$ that would maximize the attack objective. Finding an approximately optimal solution is feasible, even if φ is non-convex, e.g., using a simple grid search over $\Theta_3(z^k)$ as shown in Algorithm 3. If $(\delta_1^{k*}, \delta_2^{k*}, \delta_3^{k*}) \neq (\delta_1^k, \delta_2^k, \delta_3^k)$, the attacker has to adjust the time references of the PMUs. Note that Algorithm 3 can be easily extended for any $p > 3$ by nesting an additional for loop for each additional attacked measurement $\{\delta_4, \delta_5, etc., \cdots\}$, and by updating the implementations of the

Algorithm 3 Optimize-Attack-Angles(z, w, L, φ)

Input: z (non-attacked measurement), w (row of largest norm of W), L (number of grid search points), and φ (attacker objective function)

$\mathcal{A} \leftarrow \emptyset$
$\Theta_3 \leftarrow$ Compute-Feasible-Angles(w)
$\eta^* \leftarrow \min\{\eta > 0 : |\{0, \eta, 2\eta, \cdots, 2\pi\} \cap \Theta_3| = L\}$
$A_3 \leftarrow \{0, \eta^*, 2\eta^*, \cdots, 2\pi\} \cap \Theta_3$
for $\delta_3 \in A_3$ **do**
  $S \leftarrow$ Compute-Angle-Pairs(δ_3, w)
  for $(\delta_1, \delta_2) \in S$ **do**
    $\mathcal{A} \leftarrow \mathcal{A} \cup (\delta_1, \delta_2, \delta_3)$
  end for
end for
$(\delta_1^*, \delta_2^*, \delta_3^*) \leftarrow \arg\max_{(\delta_1, \delta_2, \delta_3) \in \mathcal{A}} \varphi(z, \delta_1, \delta_2, \delta_3)$

Output: $(\delta_1^*, \delta_2^*, \delta_3^*)$

functions *Compute-Feasible-Angles()* and *Compute-Angle-Pairs()* with the appropriate circle definitions mentioned in Section 3.1.2. Alternatively a recursive procedure could be used to search through the grid.

Next, methods to implement the optimal angles in an undetectable manner are presented.

3.3.2 Clock Servo and Brute-Force Attack

Regardless of the employed time-synchronization mechanism, PMUs adjust their internal clock smoothly based on the external time reference. The component used to regulate this adjustment is typically called a clock servo. The clock servo can either be a hardware or a software component. Given a sequence of time instants $\{t^0, t^1, \cdots, t^k, \cdots\}$, the clock servo takes as input the observed clock offset Δt_{in}^k at time t^k, and outputs the target offset Δt_{out}^k that will be implemented by time t^{k+1}. In a TSA scenario, recall that the implemented target offset Δt_{out}^k in micro-seconds is related to the implemented attack-angle δ^k in radians by the relation $\Delta t_{out}^k = g_\delta(\delta^k)$. For now, assume that initially $\Delta t_{out}^0 = 0$ and $\delta^0 = 0$, i.e., there is no attack.

An attacker that compromises the time-synchronization mechanism can control Δt_{in}^k, and would want to set $\Delta t_{out}^k = \Delta t_{out}^{k*} = g_\delta(\delta^{k*})$, where δ^{k*} is the attack-angle computed for the PMU that maximizes the attack impact at time t^k computed as discussed above. The intended output offset by the attacker might not always be implemented by the clock servo. Therefore, $\widehat{\Delta t}_{out}^k$ and $\widehat{\delta}^k$ are respectively defined as the offset and the attack-angle intended by the attacker at time t^k; and Δt_{out}^k and δ^k denote the actual values implemented by the clock servo. Figure 3.2 summarises the procedure of computing and implementing a TSA.

Brute-Force Attack (BF): A naive attacker that is unaware of the clock servo would

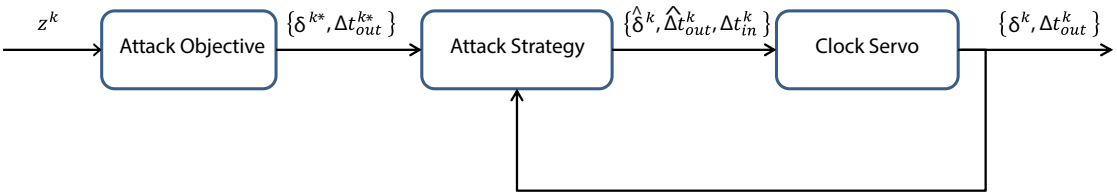


Figure 3.2 – Block diagram of clock servo aware timing attacks.

provide the servo with $\Delta t_{in}^k = \Delta t_{out}^{k*}$ at every time instant t^k. Nonetheless, due to the clock servo, the implemented Δt_{out}^k could be different from Δt_{in}^k and hence the attack might become detectable, as the adjustments of the clocks of the individual PMUs could result in a trajectory far from the set $\Theta(z^k)$ of undetectable TSAs. This attack corresponds to TSAs considered in previous work [14].

In what follows, it is shown how to implement attacks against two clock servo implementations: output constrained proportional-integral-controller (PI-controller), and output constrained P-controller. Since the latter is a special case of the former, the general case (the PI-controller) is considered first.

3.3.3 Output Constrained PI-Controller (OCPI) Clock Servo

The basic equation of a PI-controller is $y[t+1] = y[t] + K_p * e[t] + K_i \int_0^t e[\tau]d\tau$, where K_p and K_i are called the proportional and the integral gains of the controller, respectively. The measurement error at time t is $e[t] = y_{desired} - y[t]$. An output constrained PI-controller (OCPI(ρ)) clock servo adjusts the clock offset depending on the output of a PI-controller, and limits the change in the output during one time step by some threshold ρ. An example of such a clock servo is the one usually used in Precision Time Protocol version 2 (PTPv2). Therefore, the following provides a description of PTPv2 and its widely-used implementation, PTPd [87].

PTPv2 (IEEE1588-2008) is the latest standard protocol for network-based time synchronization [88]. It synchronizes the clock of one or more slave devices to that of a master clock by exchanging timestamps over a network. PTPd and LinuxPTP are widely used open-source implementations of PTPv2 for Unix based systems. The PTPd clock servo is used for adjusting the tick rate of the clock (the number of system clock ticks per second) as follows. First, the PMU clock estimates the master-to-slave d_{m2s} and the slave-to-master d_{s2m} delays from the exchanged timestamps, and uses them to estimate the one way propagation delay d_{prop} between the slave and the master (in μs) according to an infinite impulse response (IIR) low pass filter with the equation

$$s\, d_{prop}^k - (s-1) d_{prop}^{k-1} = (x^k + x^{k-1})/2,$$

where x^k is the filter input at time step k and s is the filter stiffness that controls the cut-off and the phase of the filter. The filter input is computed as $x^k = (d_{m2s} + d_{s2m})/2$. Furthermore, the estimated offset (clock error) from the master $\hat{o}^k$ is computed from d^k_{prop} using a finite impulse response (FIR) low pass filter (two sample average) according to

$$\hat{o}^k = ({\Delta t_{in}}^k + {\Delta t_{in}}^{k-1})/2, \tag{3.16}$$

where, ${\Delta t_{in}}^k$ is the observed offset computed as ${\Delta t_{in}}^k = d^k_{prop} - d_{m2s}$. Next, $\hat{o}^k$ is fed as the error signal to a discretized PI-controller that is used for computing the tick-rate adjustment

$$\Delta t^{k-1}_{out} - \Delta t^k_{out} = d^k_{i,b} + K_p\, \hat{o}^k. \tag{3.17}$$

To interpret (3.17), observe that $\Delta t^{k-1}_{out} - \Delta t^k_{out}$ represents the tick-rate adjustment that needs to be applied to the clock. The current controller error is $\hat{o}^k$, which is the estimated offset at time step k, and the bounded accumulated controller error (drift) at time step k is

$$d^k_{i,b} = \begin{cases} -\tau_d, & d^k_i < -\tau_d \\ d^k_i, & -\tau_d \le d^k_i \le \tau_d \\ \tau_d, & d^k_i > \tau_d \end{cases}, \tag{3.18}$$

where

$$d^k_i = d^{k-1}_{i,b} + K_i\, \hat{o}^k, \tag{3.19}$$

τ_d is a limit on the accumulated error, and $d^0_i = 0$. In real systems, typical values of the controller gains are $K_p = 0.1$ and $K_i = 0.001$. Furthermore, the servo makes sure that the adjustment magnitude is bounded, i.e., $|\Delta t^{k-1}_{out} - \Delta t^k_{out}| < \tau_d$ similar to (3.18). Finally, $\Delta t^{k-1}_{out} - \Delta t^k_{out}$ is passed to the Unix kernel function *adjtimex()* to implement the adjustment. Thus, the PTPd clock servo is an OCPI(τ_d) clock servo.

In the case of PTP, a TSA on a PMU can be implemented by changing the propagation times between the PTP master and the PMU (the slave), which causes a change in both the master-to-slave and the slave-to-master delays. In what follows it is shown that the attacker can manipulate Δt^k_{in} such that the attack angle follows a desired sequence, which it can use for performing an undetectable attack.

OCPI(ρ) Clock Servo Aware Attack: If the attacker is aware that a PMU uses an OCPI(ρ) clock servo, it provides the clock servo at every second with a computed Δt^k_{in} that results in a desired $\widehat{\Delta t}^k_{out} = g_\delta(\hat{\delta}^k)$, with the constraint $|\Delta t^{k-1}_{out} - \widehat{\Delta t}^k_{out}| < \varrho$, for some constant $\varrho > 0$. Note that, the notations $\widehat{\Delta t}^k_{out}$ and $\hat{\delta}^k$ are used because the values might not be implemented by the servo due to the constraint on d^k_i. The supplied Δt^k_{in} can be calculated by solving

$$\Delta t^k_{in} = \frac{d^{k-1}_{i,b} + \widehat{\Delta t}^k_{out} - \Delta t^{k-1}_{out}}{\frac{K_i}{2} + \frac{K_p}{2}} - \Delta t^{k-1}_{in}, \tag{3.20}$$

which is obtained by substituting (3.19) and (3.16) in (3.17). Note that at time t^k the values of $d_{i,b}^{k-1}$, Δt_{out}^{k-1} and Δt_{in}^{k-1} are already known, hence Δt_{in}^{k} is only a function of the desired $\widehat{\Delta t}_{out}^{k}$.

A simplified version of the OCPI-controller where $K_i = 0$ is also considered. It is referred to as an output constrained P-controller (OCP(ρ)) and the corresponding attack is referred to as the OCP attack. To the best of our knowledge, there are no PTP implementations where the servo is a P-controller. However, considering the OCP attack allows to evaluate the importance of the knowledge of K_i in performing an undetectable attack against an OCPI servo. Results obtained with the different attack strategies (Brute-Force, OCP and OCPI) are presented in the next section.

3.4 Numerical Results

In this section, simulation results based on the IEEE 39-bus system are provided. Section 3.4.1 describes the electrical model and the methodology considered to evaluate the proposed attacks. Section 3.4.2 considers an attack against $p = 5$ PMUs each measuring a single distinct synchrophasor. For this scenario, it is shown that using the Brute Force strategy as in [14], without taking the servo constraints into consideration, makes the attack detectable by BDD, and that using the attack strategies presented in Section 3.3 enables the attack to remain undetected. It is also shown that when attacking $p = 2$ PMUs among this set of 5, the discreteness property of the set of undetectable attack-angles prevents the implementation of servo-aware attacks, thus leading to attack detection. Furthermore, it is shown that implementing robust state-estimation techniques does not counter the presented undetectable attacks. Section 3.4.3 gives numerical evidence that attacks are also possible against PMUs that measure both voltage and current synchrophasors simultaneously, namely when both measurements at a bus share the same time reference. The theory that explains the conditions in which sets of measurement points that share the same time reference are attackable by the same delay, is developed in Chapter 4. Lastly, Section 3.4.4 shows that pairwise criticality or the shared equivalence class property of attacked measurements are not necessary conditions for vulnerability to TSAs, by demonstrating a practically undetectable attack, using rank-1 approximation of W, on a set of $p = 3$ PMUs where the pairwise IoS^* values of the measurements are strictly less than 1.

3.4.1 Electrical Model and Evaluation Methodology

In this simulation, it is assumed that the IEEE 39-bus system network has 13 PMUs that measure voltage phasors, 22 PMUs that measure injected-current phasors, and 12 zero-injection buses as illustrated in Figure 3.3. It is also assumed that Bus 31 is the connection point to the external grid. Note that a different topology would result in a

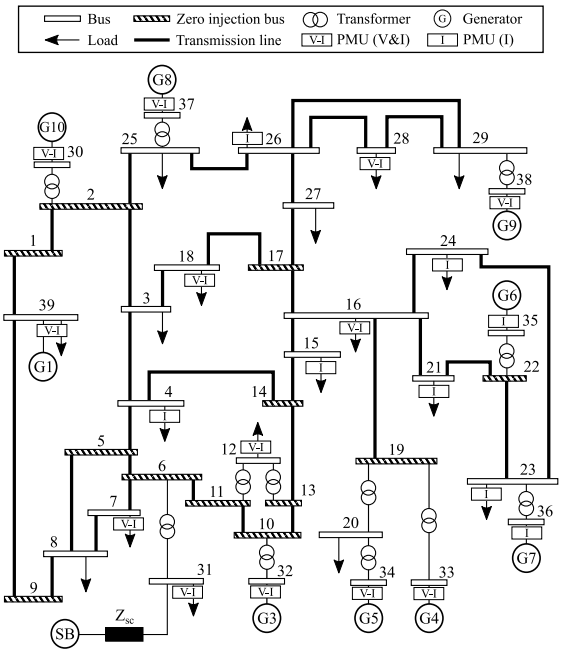


Figure 3.3 – Benchmark IEEE 39-bus transmission system and PMU locations.

different verification matrix and thus different attack-locations and attack-angles than the ones presented here. Also note that PMUs are considered to have an OCPI clock servo with threshold $\rho = 500\mu$s and that the OCP and OCPI attack strategies performed by an attacker are done with a threshold of $\rho = 20\mu$s. The used load profiles were obtained from real measurements taken at 50 frames-per-second by real PMUs installed in the 125-kV sub-transmission network of Lausanne, Switzerland. For this reason, the load profiles present time-domain behaviour typical of transmission networks. For a set of p attacked PMUs and for a target transmission line chosen by the attacker, a TSA is simulated using the following procedure:

- Creation of the non-attacked measurements:
  - At each time step k, a load flow is computed on the system to determine the true state of the power-grid, based on the load profiles.
  - The true state of the system is perturbed with randomly-generated Gaussian noise, depending on the accuracy of each PMU, to generate the measurement vector z^k (assuming class 0.1 voltage and current sensors).
- Attack computations:
  - Previous measurements are used to compute an estimate $\tilde{z^k}$ of z^k.
  - The estimate $\tilde{z^k}$ is used to compute the optimal intended attack-angles δ^{k*} by using grid search as in Algorithm 3.

  - The implemented attack-angle $\hat{\delta}^k$ is computed according to the chosen attack strategy: BF, OCP or OCPI; and applied to the PMU clock servo, resulting in the attacked measurement vector $z^{k'}$.

- State estimation and attack detection:
  - The WLS estimation is performed, both with the unattacked and attacked measurement vectors z^k and $z^{k'}$, and the measurement residuals $r^k = Fz^k$ and $r^{k'} = Fz^{k'}$ are computed.
  - The LNR, χ^2 and/or HTI tests are performed for the residuals r^k and $r^{k'}$.
  - The estimated power-flow is computed on the target line, with and without the attack.

At every new time-step, it is assumed that the attacker can use the previous measurement values to estimate the current measurement value and compute valid attack-angles with respect to this estimate. If the estimate is significantly inaccurate, the solution set of possible attack-angles will in fact be detectable. In order to determine the effect of sudden changes in the system state on the attack detectability, a sudden increase of factor 2 in the active power (referred to as an "inrush") is introduced in one of the buses after $t = 300$ seconds from the start of the simulation. Note that unless the clock-servo implements attack-angles that are different from the ones intended by the attacker, the presented strategies don't impact the residuals, hence detection methods based on the normality of the residuals are expected to fail in identifying the attack.

3.4.2 Practical Feasibility of Attacks

For the PMU allocation shown in Figure 3.3, the analysis presented in section 3.2 was applied to find equivalence classes under $\mathcal{R}^*$ by computing the pairwise IoS^* between measurements. The following equivalence classes were found:

- Class 1 contained 5 measurements: the voltage measurements at buses $\{28, 38\}$, and the current measurements at buses $\{26, 28, 38\}$.
- Class 2 contained 2 measurements: the voltage and current measurements at bus 34.
- Class 3 contained 2 measurements: the voltage and current measurements at bus 37.
- Class 4 contained 6 measurements: the current measurements at buses $\{16, 21, 23, 24, 35, 36\}$.

For the remaining measurements, each measurement constitutes a separate equivalence class.

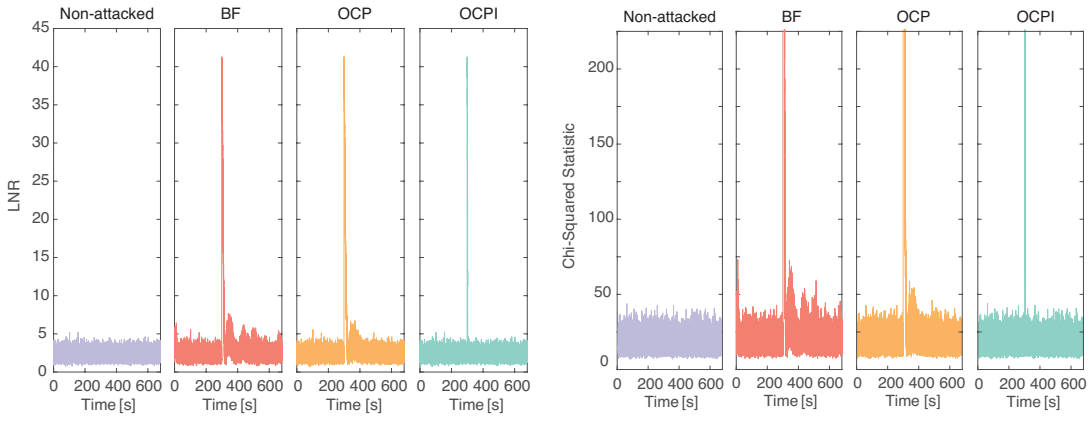


Figure 3.4 – Results for attacking the $p = 5$ equivalence class PMUs: (a) LNR Test results (b) χ^2Test results: the inrush spike goes above 2000, the figure was zoomed closer to zero for better comparison. In both cases the OCPI strategy is closest to the non-attacked scenario, except for the spike caused by the inrush, where all attacks are detectable.

This PMU allocation contained a total of 34 measurements, and thus the IoS^* of $\binom{34}{2} = 561$ pairs of PMUs had to be checked in order to construct the equivalence classes of attackable measurements. Note that without the knowledge on the equivalence classes, an attacker would have to compute the rank of the W matrix corresponding to $\binom{34}{3} = 5984$ combinations of measurements in order to find whether a $p = 3$ attack exists, or $\binom{34}{5} = 278256$ combinations in order to find whether a $p = 5$ attack exists.

A TSA was mounted on a subset of $p = 5$ PMUs from equivalence class 4, namely the current measuring PMUs at buses $\{21, 23, 24, 35, 36\}$. The goal of the attack was to minimize the apparent power-flow on the line between buses 16 and 24. In this scenario, the simulated inrush was located at bus 21, which is one of the attacked buses. To implement the attack strategies for $p = 5$, the $p = 5$ extension of Algorithm 3 was used to compute the optimal angles δ^{k*} at each time-step. First, the impact and detectability of this attack applied according to the different strategies described in Section 3.3 are discussed. Then, it is shown that a Robust State Estimation technique is just as vulnerable as a non-robust LSE. Finally, the number of PMUs that an attacker should target within this set of five is investigated.

3.4.2.1 Impact and Detectability of the Different Strategies

The results of applying the LNR and χ^2 tests on the obtained residual vectors are shown in Figures 3.4a and 3.4b respectively. The x-axis shows the simulation time, while the y-axis shows the value of the largest normalized residual and of the sum of squared normalized residuals at each time-instant, respectively. The figure shows that the two tests are equivalent in terms of attack detection. The lower the LNR and sum of

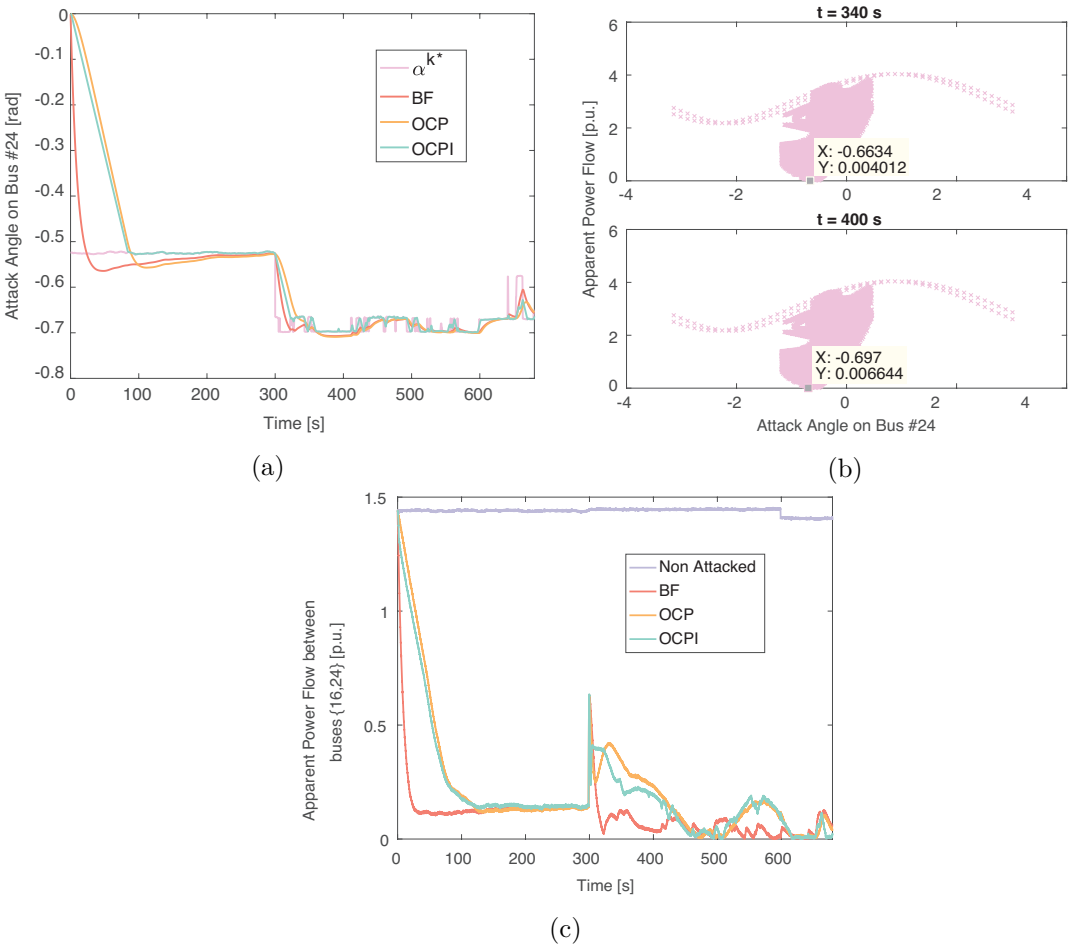


Figure 3.5 – Attack strategies and their impact: (a) Attack-angles for the optimal intended case δ^{k*} and the ones implemented by the different strategies on the current measuring PMU at bus 24. Notice the spikes in δ^{k*} after the inrush. (b) The power-flow on the transmission line between buses 16 and 24 as a function of the attack angle at bus 24 at two different time instants after the inrush. The optimal (minimal) angles changes significantly between the two instants explaining the spikes in the previous figure. (c) The estimated power-flow on the line between buses 16 and 24 is largely affected by the attack strategies.

squared normalized residuals, the stealthier the attack is. It is observed that the OCPI attack strategy yields values that are very close to the non-attacked measurements. The brute-force strategy and the OCP strategy, on the other hand, produce high values for long durations, especially after the inrush. The only time-instants when the OCPI attack produces high residuals are right after the inrush, where a spike is observed. This occurs because with the inrush the attacker mis-estimates the measurement value and does not compute the valid set of feasible attack-angles. As soon as the attacker is able to successfully estimate the new measurements, it is observed that the LNR and the sum of squared normalized residuals values decrease back to undetected values. However this decline is gradual as the clock-servo does not implement the new optimal attack-angles directly. This is the case for all strategies, nevertheless the OCPI strategy allows to

regain stable undetectable conditions faster. The figure suggests that sudden changes in the system state present a natural countermeasure to TSAs.

Figure 3.5a shows the optimal intended attack-angle δ^{k*} and the implemented attack-angles for the different strategies applied to the current measuring PMU on bus 24. Note how the BF strategy is the fastest to reach the intended angle, but at the expense of detectability. The other servo-aware strategies reach the intended attack-angle in a more gradual manner thus keeping the LNR values low to a lower normal-looking range. Furthermore, several spikes are observed in the intended optimal attack-angle δ^{k*} after the inrush in Figure 3.5a. These spikes can be explained by Figure 3.5b, where the x-axis represents the attack angle at Bus 24 and the y-axis represents the apparent power-flow on the line between buses 16 and 24. The sub-figures show all grid points considered by Algorithm 3 (its extension for $p = 5$) before choosing the optimal angles, at two different time-instants after the inrush ($t = 340$s and $t = 400$s). Note that multiple values of the apparent power-flow are obtained for every value of the attack-angle. These values correspond to different choices of the other attack-angles (on buses $\{21, 23, 35, 36\}$). By analysing Figure 3.5b it can be observed that the minimum power-flow is obtained by choosing substantially different values of the attack-angle ($\delta^{340*} \approx -0.66$ rad, $\delta^{400*} \approx -0.7$ rad), which is reflected by the spikes in Figure 3.5a. Note that the spikes of δ^{k*} values impacts the angles implemented by the BF strategy to a greater extent than the angles implemented by the OCP and OCPI strategies. This is due to the fact that the rate of change of the attack-angle for the OCP and OCPI strategies is limited by the threshold parameter ρ.

The impact of the different attack strategies on the target bus is illustrated in Figure 3.5c. The latter shows that the attacks are able to create a mis-estimation of the power-flow by one order of magnitude and that this mis-estimation is more gradual in servo-aware strategies.

3.4.2.2 Impact on Robust State-Estimation

In what follows, the LSE is assumed to be using the LNR test o be robust against bad-data. In order to simulate a robust LSE, the measurements with the highest normalized residual above threshold η_{BDD} are iteratively removed if removing the measurement does not impact the observability of the system. At the end of the process, if all measurements have normalized residuals below η_{BDD}, then the control center believes it has successfully removed bad-data and thus it will trust the computed state estimate. This robust LSE technique was performed after the attacks on the same $p = 5$ PMUs with $\eta_{BDD} = 5.5$. Figure 3.6a shows the apparent power-flow after the use of the bad-data removal scheme. The results are similar to those shown in Figure 3.5c, which shows that the scheme is not able to remove all of the attacked measurements from the data-set used for state-estimation and thus the impact of the attack is unaffected. This can be explained by the

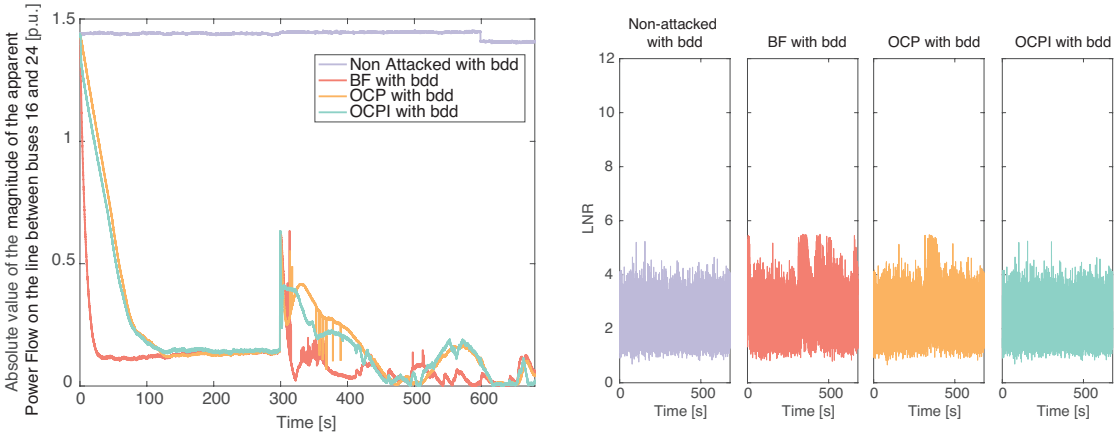


Figure 3.6 – Attack strategies impact with robust LSE: (a) Apparent power-flow obtained without attack and with attacks for $p = 5$ after bad-data removal shows that the mis-estimation is not countered by the robust technique employed. (b) LNR tests show that the LNR values are reduced, so the control center will be enclined to believe it managed to remove all bad-data.

fact that the robust state-estimation techniques rely on the analysis of the distribution of residuals, which is unchanged by an undetectable attack. Figure 3.6b shows that the control center has removed measurements with high LNR values but the mis-estimation shown in Figure 3.6a confirms that not all attacked measurements were removed.
We further performed the HTI method to attempt the identification of the attacked measurements. We performed it on the three attack strategies with three different techniques for the establishment of the initial set of suspicious measurements: 1) all measurements for which the absolute value of the normalised residual is above a threshold equal to 5 are selected, 2) all measurements for which the absolute value of the normalised residual is above a threshold equal to 4.5 are selected and 3) the ten measurements with highest measurement residual in absolute value are selected. We used the parameters and formulas given in Section 5.8 of [79] for a fixed type II error probability of 0.01. As a result, a set of unattacked data was suspected and at most one out of ten rectangular coordinates of attacked measurements was suspected. In other words, either all or a large majority of the attacked measurements are considered as true data. In the case of the OCPI attack strategy, only the imaginary part of the current phasor at bus 36 is suspected among the attacked data. This observation coincides with the fact that the robust state estimation combined with the LNR test does not remove all attacked data because the initial suspicious set of measurements depends on the values of the normalised residuals. Note that we performed the HTI method for all of the other attacks presented in this chapter and the same results were observed each time. In the rest of the dissertation, we analyse the detectability of an attack by only checking the LNR values because if they are small enough for undetectability, then the HTI method won't suspect any measurement either.

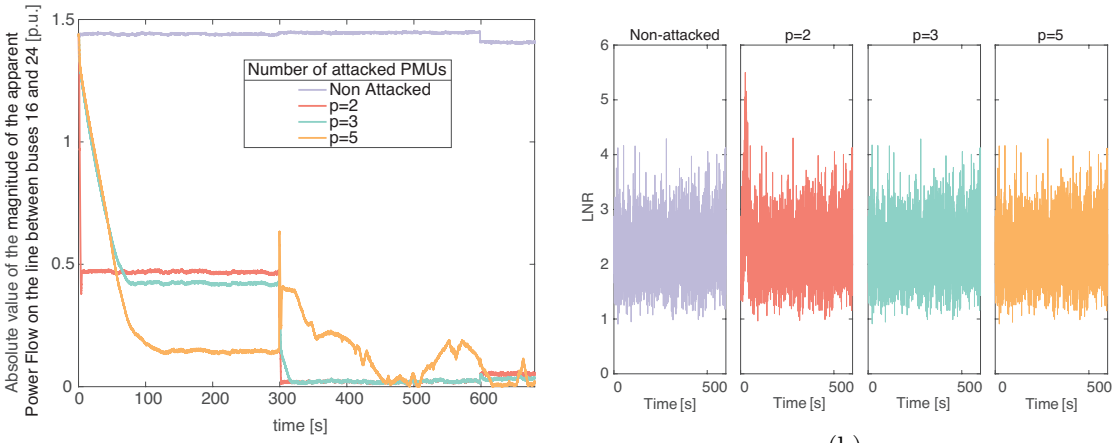


Figure 3.7 – Impact of the chosen number of PMUs to attack: (a) Apparent power-flow obtained without attack and with attacks with $p \in \{2, 3, 5\}$, notice that $p = 2$ is not gradual because the solution set of undetectable attacks is finite. (b) zoom on LNR tests at the beginning of the attacks shows that OCPI attacks for $p \geq 3$ are undetected but for $p = 2$ some residuals are too high which is explained by the drastic drop in the apparent power-flow in the previous picture.

3.4.2.3 Optimal Number of PMUs to Target for Impact and Undetectability

In Section 3.1 it was mentioned that when mounting an attack on $p = 2$ PMUs, the solution set of attack-angles that an attacker can choose to use while remaining undetected, is a singleton. Hence the attacker is not able to slowly change the attack angles in accordance with the PMU clock servo. However, by increasing the size of the set of targeted PMUs to $p \geq 3$, the set of undetectable attacks forms a continuum, which allows the attacker to slowly change the attack-angles so as to bypass the BDD. In order to illustrate the impact of choosing more or fewer PMUs to attack, the OCPI attacks were performed on subsets of the 5 PMUs attacked previously. A $p = 2$ case targeting PMUs 21 and 24, and a $p = 3$ case targeting PMUs 21, 23, and 24 are considered. For both cases, the attacker's objective is the same, namely to minimize the apparent power-flow on the line between buses 16 and 24.

Figure 3.7a shows the obtained apparent power-flow on the targeted line for the unattacked and attacked scenarios for different values of p with the best possible attack strategy. Notice that the apparent power-flow in the case of $p = 2$ drops abruptly at each change of optimal attack-angle, namely at the beginning of the attack and right after the inrush. In fact, in this case, the OCPI attack strategy corresponds to the brute-force strategy as the solution set of attack-angles is finite. Figure 3.7b shows the LNR test values for the different cases at the beginning of the attack. It can be noticed that the drastic change of apparent power-flow for $p = 2$ causes high, suspicious LNR values. Therefore, it is of the utmost importance that the attacker use a smart servo-aware strategy if he wishes to remain undetected, which requires targeting at least $p = 3$ PMUs. Furthermore, Figure 3.7a shows that in stable conditions, the apparent power-flow mis-estimation

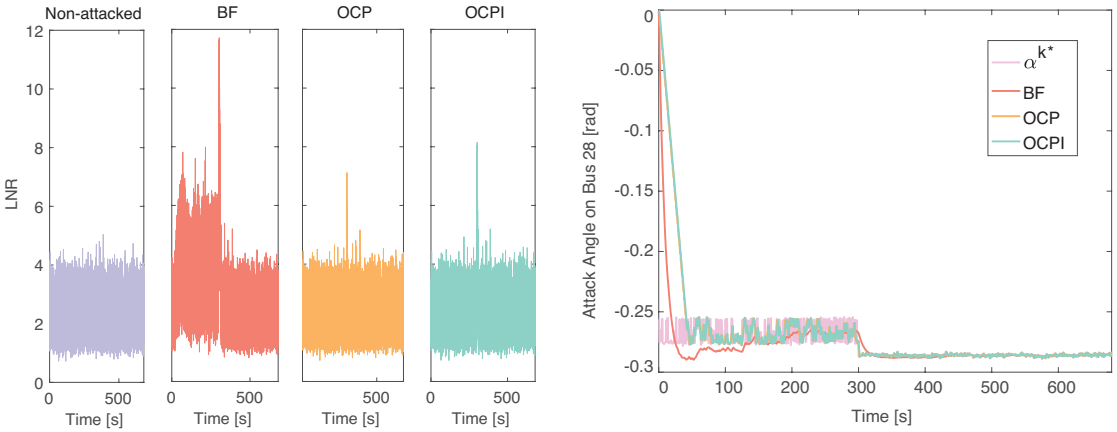


Figure 3.8 – Attacking PMUs measuring voltages and nodal injected-currents simultaneously: (a) The LNR test results in the non-attacked and OCPI attack scenarios cannot be differentiated, except at the inrush. (b) attack-angles implemented by the OCPI strategy follow the intended optimal angles closely. (c) The power-flow in per-unit on the transmission line between buses 28 and 29 is significantly decreased.

grows with p. However, one must be careful with choosing too high values of p as they also mean larger degrees of freedom which could lead to spikes in the optimal attack-angle and unstable-looking apparent power-flows as observed in Figures 3.5c and 3.7a, which could be used as a counter-measure for attack detectability. Also note that because of the increasing number of degrees of freedom, finding the optimal attack-angles takes an increasing amount of time. Therefore, an attacker would need to find a tradeoff between the cost of the attack and the level of mis-estimation he wishes to create. Experimentally, it was observed that $p = 3$ allowed for a large and stable mis-estimation of the power-flow, while remaining undetected.

3.4.3 Attacks on Voltage and Current Measuring PMUs

The previous scenario illustrated an attack against a set of $p = 5$ PMUs measuring current phasors only. One question that remains is whether it is possible to attack PMUs that measure both voltage and injected-current phasors simultaneously. Since both measurements are taken by one PMU (using one time reference), the implemented offset, and hence the attack-angles on both phasors will have to be the same, which poses a new constraint on the attacker.

Considering the same PMU allocation as before, another attackable set (pairwise $IoS^* = 1$) of 5 measurements taken by 3 PMUs was found. The attacked measurements are the injected-current phasor at bus 26, the voltage and injected-current phasors at bus 28, and the voltage and injected-current phasors at bus 38. Note that attacking these 5 measurements constitutes a $p = 3$ attack as the attack-angles applied to different measurements at the same bus will have to be the same. Since $p = 3$, there is still a continuum of undetectable attacks which enables the attacker to use the proposed attack strategies for gradually changing the attack-angles. In this scenario the attacker aims to minimize the apparent power-flow on the line between buses 28 and 29. The inrush location was also changed from bus 21 to bus 28 (closer to the attacked buses) to observe its impact on attack detectability. Figure 3.8a shows similar behaviour as for the previous scenario. Namely, the attack is not detectable with the OCPI strategy, except directly after the inrush when the optimal attack-angles change as shown in Figure 3.8b. Finally, Figure 3.8c shows a mis-estimation of the apparent power-flow on the target line by an order of magnitude. Hence, the presented attack strategies on PMUs measuring two synchrophasors simultaneously seem to be feasible. We present in Chapter 4 the exact undetectability conditions for attacks targeting sets of an arbitrary number of PMUs. We also provide a countermeasure against the most common vulnerability and Section 4.6 shows that if it is enforced, then the attacks presented in this chapter are no longer undetectable.

3.4.4 Attacking Non-critical Sets of PMUs

In the previous scenarios, the considered set of PMUs formed an equivalence class, which enabled the computation of undetectable attacks. To illustrate that practical TSAs can be implemented even if there are no critical pairs (when no equivalence classes under IoS^* exist), the feasibility of attacking sets of PMUs such that their pairwise IoS^* values are not equal to 1 is now considered. For this, a denser PMU allocation on the IEEE 39-bus system that does not allow for the existence of equivalence classes under IoS^* is considered. In this allocation, assume that a current measurement is installed on every bus in the system except for buses $\{16, 21, 23, 25\}$ as well as the zero-injection buses which are the same as in the previous setting. Moreover, a voltage measurement is installed on every bus in the system except for buses $\{28, 29\}$ as well as the zero-injection

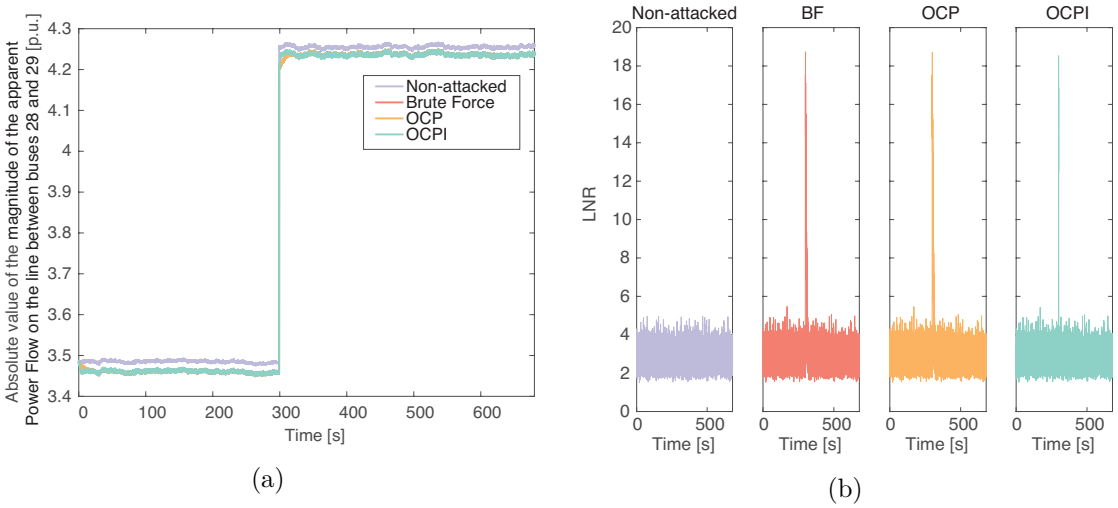


Figure 3.9 – Results when pairs of measurements are not critical: (a) Attacks create a mis-estimation of the apparent power-flow. (b) LNR Test for $IoS^* < 1$: all strategies are undetected except for the spike caused by the inrush.

buses. For this scenario, it is assumed that two synchrophasors at the same bus are measured by two distinct PMUs. An inrush at bus 28 is also considered in this scenario. There are now no pairs of PMUs with $IoS^* = 1$, i.e., no critical pairs. However, there are pairs of PMUs for which the IoS^* values are close to one, for example, the current measuring PMUs at buses $\{28, 29, 38\}$; $IoS^*(28, 29) = 0.9996$, $IoS^*(28, 38) = 0.9978$ and $IoS^*(29, 38) = 0.9993$. Hence, an attack against these $p = 3$ PMUs is considered, with the objective of minimizing the estimated power-flow on the transmission line between buses 28 and 29. At all time-steps, the corresponding W matrix has an IoS of maximum value $0.9989 < 1$ and minimum value 0.9975, hence the three measurements never form a critical set. Since the W matrix in this scenario is not of rank equal to 1, a rank-1 approximation is used to compute the attack-angles. The observed mis-estimation of the power-flow is shown in Figure 3.9a for the different attack strategies. The LNR test values from Figure 3.9b show that the attacks are undetectable by the LNR test, for all strategies. Note that the optimal attack-angles in this scenario are very small: they are, respectively, 0.023 and 0.015 rad before and after the inrush. As the optimal intended attack angles are small, they will not be significantly modified by the clock-servo and thus the residuals will not change, which explains the low LNR values for all strategies. Again, the only time instants when the residuals are high are right after the inrush because of the sudden large change of the set of feasible attack-angles.

3.5 Conclusion

In this chapter, it was shown that vulnerable sets of PMUs of arbitrary size can be found by grouping PMUs in equivalence classes with respect to the IoS. The practical feasibility

of attacks was studied and different clock servo-aware strategies for implementing an attack were proposed. Numerical results illustrate the importance of using a smart attack strategy in order for the attack to remain undetected. Using the proposed attack strategy, there is no mismatch between the intended attack and the one implemented by the PMU clock. The experiments also show that attacks can be detected upon the occurrence of an inrush at a nearby bus, yet when using robust state estimation with bad-data removal, the attack was successful and completely undetectable, even during the inrush. Thus the bad-data removal scheme could potentially work in favour of the attacker if the logs of removed measurements are not closely monitored. The effects of sudden changes to the grid could be investigated further to differentiate normal grid dynamics from the ones of an attacked grid. Furthermore, numerical evidence was provided for the feasibility of undetectable attacks when attacking PMUs that measure both voltage and injected current phasors simultaneously, and when attacking sets of PMUs that are not pairwise critical.

4

Security Measures for Grids against Rank-1 Undetectable Time-Synchronization Attacks

The attack strategies and vulnerability conditions proposed in [14] and [15] and presented in Chapters 2 and 3, require that each malicious offset alters a single PMU that measures a single phasor. Yet, PMUs are capable of measuring several phasors, and several PMUs can share the same time-reference. In order to understand and mitigate rank-1 time-synchronization attacks (TSAs) on real grids (i.e. TSAs where the rank of the attack-angle matrix W is of rank equal to 1), it is important to generalize the theory established in Chapter 3 to sets of PMUs that possibly share the same time reference and that measure an arbitrary number of phasors. Also, the techniques in Chapter 3 all rely on a complex verification matrix corresponding to the LS residuals when in fact the most commonly used state estimation technique is the WLS estimator. For exact attackability, the conditions are equivalent with the LS and the WLS matrices. However, for the establishment of vulnerability metrics, the use of the WLS matrices is more accurate because it is the WLS estimator that is used in practice and the WLS and LS residuals can be different. Due to the fact that the noise of phasors does not have circular symmetry, we cannot use the complex system model of Chapter 3 to compute the WLS estimates. Instead, we use the system model in rectangular coordinates introduced in Section 2.2.2 to compute vulnerability metrics based on the WLS residuals.

In this chapter, we establish the exact conditions in which a rank-1 TSA against sets of measurement points that possibly share the same time reference, is feasible. We call *site* a set of measurement points that share the same time reference. In other words, in the simulations of Chapter 3, all PMUs were in distinct sites. This does not have to be the case, for example in the simulations of this chapter we consider that all buses that are separated only by a converter belong to the same site. Hence, an offset on the time-reference of a site impacts all of the corresponding phasor measurements in the same manner. In order to establish the exact vulnerability conditions, we use the system model in complex form, i.e. the LS verification matrix, because it leads to a

structural vulnerability condition that does not depend on the measurement values. Such a structural condition cannot be derived using the WLS matrices. We then introduce vulnerability metrics that give insights on how far from vulnerable a group of sites is. For this purpose, we use the system model in rectangular coordinates because the vulnerability of a site depends on the detectability of an attack with respect to the analysis of the WLS residuals.

The first contribution of this chapter is to identify a sufficient and necessary vulnerability condition for rank-1 TSAs that target a single site that measures an arbitrary number of phasors: item (b) of Theorem 4.4. We also provide a metric that reflects how vulnerable a site is at a given time.

If two sites are not vulnerable by themselves, they could still be vulnerable as a pair. The second contribution of this chapter is to identify vulnerability conditions to rank-1 TSAs, for such pairs of sites. One is a sufficient condition that does not depend on the measurements: item (b) of Theorem 4.6. The other is a sufficient and necessary general condition that depends on the measurement values: item (a) of Theorem 4.6. We provide a second metric that reflects how vulnerable a pair of sites is at a given time.

The third contribution of this chapter is to show that rank-1 TSAs, on a set of more than two sites, are feasible if and only if they are feasible for any pair of sites within the set.

Finally, the fourth contribution of this chapter is to mitigate the feasibility of rank-1 TSAs on a grid by combining the first three results. The vulnerability of a grid can be first analysed statically. The sufficient structural vulnerability condition identified by item (b) of Theorem 4.6 does not depend on the measurement values. We establish a security requirement that prevents it and provide a greedy recursive algorithm that enforces it. The latter checks if the system is structurally vulnerable and if so it proposes a small set of measurement points that needs to be added in order to satisfy our requirement. However, the structural vulnerability condition is only a sufficient condition for vulnerability. Even if the grid is not structurally vulnerable, the measurement values may satisfy the general sufficient and necessary vulnerability conditions for single sites or for pairs of sites (i.e. item (b) of Theorem 4.4 or item (a) of Theorem 4.6, respectively). Although we reason in Sections 4.2.3 and 4.3.3 that this is unlikely to occur in practice, we cannot exclude it and hence we recommend the monitoring of the two provided metrics.

The chapter is structured as follows. Section 4.1 defines the system model, the attack model and explains the vulnerability conditions. Sections 4.2 and 4.3 give the exact vulnerability conditions for single sites and pairs of sites, respectively. They also provide vulnerability metrics, discuss the feasibility of the vulnerability conditions and compare the new conditions with the ones of the two previous chapters. Section 4.4 proves the third contribution of this chapter. Section 4.5 provides security measures to mitigate the feasibility of rank-1 TSAs on any observable grid. An algorithm to secure grids against

structural vulnerabilities is also provided in Section 4.5. Numerical results, on the IEEE-39 bus benchmark with real load profiles from the Lausanne grid, are given in Section 4.6. They show that the measurements of a grid satisfying our security requirements are far from satisfying the identified vulnerability conditions. Finally, Section 4.7 concludes the chapter.

4.1 System and Attack Models

In this chapter we use the same system model as in Chapters 2 and 3 with the difference that PMUs are able to measure an arbitrary number of phasors and that sets of PMUs may share the same time reference. More specifically, we say that measurement points that share the same time reference belong to the same site and that distinct sites have distinct time references. Note that by time reference, we refer to a site's clock and not to the time source, which could be space or network based. A same time source is used to synchronize the different time references (i.e. clocks) throughout the grid. We now provide new results on critical sets of measurements using the complex system model and introduce new notations for WLS residual computation using the system model in rectangular coordinates. We then present the attack model and the vulnerability conditions.

4.1.1 System Model in Complex Form: Results on Measurement Criticality

The following two theorems facilitate the understanding of the link between the criticality of a set of measurements and the rank of the verification matrix F.

Recall that Fz with $F = H(H^\dagger H)^{-1}H^\dagger - Id$, is the residual vector after the LS estimation. For ease of explanation, we introduce S_i as the set of phasor indices in site i, F^{S_i} as the $m \times |S_i|$ submatrix of F corresponding to the columns of F with indices in S_i and z^{S_i} as the vector of measurements with indices in S_i.

Theorem 4.1. *For a set S of measurement indices, F^S is full rank if and only if measurements with indices in S form a non-critical set.*

Proof. We show both directions.

- $S = \{1, \cdots, p\}$ *non-critical* $\Rightarrow$ F^S *full rank*: Assume that F^S is not full rank and that $|S| = p$, hence there exists a $y \neq 0 \in \mathbb{C}^p$ such that $F^S y = 0$. By construction $F \begin{pmatrix} y_1 & \dots & y_p & 0 & \dots & 0 \end{pmatrix}^T = 0$. Hence, by definition of F, vector $v = (H^\dagger H)^{-1}H^\dagger \begin{pmatrix} y \\ 0 \end{pmatrix} \in \mathbb{C}^N$ is such that $Hv = \begin{pmatrix} y \\ 0 \end{pmatrix}$ and $v \neq 0$ because $y \neq 0$. The

submatrix of H obtained by removing its first p rows is of dimension $(m-p)\times n$. Therefore, if $m-p\geq n$, then the submatrix of H obtained by removing its first p rows is not full rank, which implies that S is critical. Likewise, if $m-p\leq n$, then the number of available measurements is insufficient for observability. In other words, removing S renders the system unobservable which by definition means that S is critical. Hence, We have shown that if F^S is not full rank, then S is critical. This is equivalent to showing that if S is non-critical, then F^S is full rank.

- $S=\{1,\cdots,p\}$ *non-critical* $\Leftarrow$ F^S *full rank*: Recall that S critical means that the submatrix of H obtained by removing the rows corresponding to the measurements in S, is not full column rank. Hence, if S is critical, then there exists a $v\neq 0\in\mathbb{C}^N$ such that $Hv=\begin{pmatrix}h_1v & ... & h_pv & 0 & ... & 0\end{pmatrix}^T$, where h_i is the i^{th} row of H. Observe that $\begin{pmatrix}h_1v & ... & h_pv & 0 & ... & 0\end{pmatrix}^T$ is in the range of H and thus its orthogonal projection onto $Im(H)$ is equal to itself. Notice that $(F+Id)$ is the orthogonal projection matrix onto $Im(H)$. Hence, $(F+Id)\begin{pmatrix}h_1v & ... & h_pv & 0 & ... & 0\end{pmatrix}^T=\begin{pmatrix}h_1v & ... & h_pv & 0 & ... & 0\end{pmatrix}^T$ and $F\begin{pmatrix}h_1v & ... & h_pv & 0 & ... & 0\end{pmatrix}^T=0$. Removing the zero terms, we obtain $F^S\begin{pmatrix}h_1v & ... & h_pv\end{pmatrix}^T=0$. We assume that the system is observable, hence H is full rank. Notice that the number of rows of H (i.e. the number of measurements m) is larger or equal to its number of columns (i.e. the number of buses n). Hence, H full rank means that its kernel is equal to 0 and therefore that $\begin{pmatrix}h_1v & ... & h_pv\end{pmatrix}^T\neq 0$. Notice that F^S is of dimension $m\times p$ with $p\leq m$, thus $F^S\begin{pmatrix}h_1v & ... & h_pv\end{pmatrix}^T=0$ means that F^S is not full rank. We have shown that if S is critical, then F^S is not full rank. This is equivalent to showing that if F^S is full rank, then S is non-critical.

We conclude that F^S is full rank if and only if S is non-critical. □

Theorem 4.2. *For a set S of measurement indices, if rank$(F^S)=|S|-s$ with $1\leq s\leq|S|$, then all subsets of size $|S|-s+1$ are critical and there is at least one subset of size $|S|-s$ that is non critical.*

Proof. Recall that F^S is a submatrix of F of dimension $m\times|S|$, with $|S|\leq m$. Hence, if F^S or F^K with $k\subset S$ is full rank, then the columns of the matrix are independent. For a set S of measurement indices, if $\text{rank}(F^S)=|S|-s$ with $1\leq s\leq|S|$, then

- for all subset $K\subset S$ of size $|K|=|S|-s+1$, $\text{rank}\,(F^K)\leq|S|-s$, i.e. F^K is not full rank. By Theorem 4.1, the set of measurements with indices in K is critical. In other words, all subsets of size $|S|-s+1$ are critical.
- there exists at least one subset $K\subset S$ of size $|S|-s$ such that $\text{rank}\,(F^K)=|S|-s$, i.e. such that F^K is full rank. By Theorem 4.1, the set of measurements with

indices in K is non critical. In other words, there is at least one subset of size $|S| - s$ that is non critical.

□

Theorem 4.2 implies that if a set of measurements with indices in S is such that $\text{rank}(F^S) = 1$, then all pairs of measurement within the set are critical and that at least one measurement is non critical. Also observe that theorem 4.2 applied to sets of 1 measurement shows that a measurement is critical if and only if its corresponding column of F is the zero vector, which was also shown in [82].

4.1.2 System Model in Rectangular Coordinates

In Chapters 2 and 3, undetectable TSAs are proposed such that the LS residuals are unchanged. The formulation of the system model in complex form facilitates the derivation of the theory. However, in practice, it is the WLS estimator that is used for state estimation and because the LS and WLS residuals differ, the symptoms of an attack should be analysed with respect to the WLS residuals instead of with respect to the LS residuals. Recall that $r_\square(z_\square) = G_\square z_\square$ with $G_\square = H_\square(H_\square^T C_\square^{-1} H_\square)^{-1} H_\square^T C_\square^{-1} - Id$, is the residual vector in rectangular coordinates after the WLS estimation. We further express the rectangular coordinates of the WLS measurement residuals as a complex relation, using the complex conjugate of z as a variable

$$r_\square(z_\square) = Rz + \bar{R}\bar{z},$$

where $\bar{z}$ is the complex conjugate of z and where the complex matrix $R \in \mathbb{C}^{2m\times m}$ is computed from blocks of the real matrix $G_\square$

$$R = \frac{1}{2}\begin{bmatrix} G_{\square,1} - jG_{\square,2} \\ G_{\square,3} - jG_{\square,4} \end{bmatrix},$$

where $G_{\square,1},\ G_{\square,2},\ G_{\square,3},\ G_{\square,4} \in \mathbb{R}^{m\times m}$ are blocks of the $2m \times 2m$ real verification matrix

$$G_\square z_\square = \begin{bmatrix} G_{\square,1} & G_{\square,2} \\ G_{\square,3} & G_{\square,4} \end{bmatrix} \begin{pmatrix} \text{Re}(z) \\ \text{Im}(z) \end{pmatrix}.$$

4.1.3 Attack Model

As in the two previous chapters, we suppose that an attacker can observe the measurement vector z, that he knows the topology of the system (i.e. he knows H) and that he is able to manipulate the time reference of q sites via a GPS spoofing attack or a delay box insertion. An injected time offset d in the time reference of a site directly shifts

the phase of all measured phasors in this site by an angle of $\delta = 2\pi f d$ rad, where f in Hz is the instantaneous voltage frequency and the offset d is in seconds. Therefore, a measurement z_i affected by an attack angle δ is of the form $z'_i = z_i e^{j\delta}$: its phase is shifted by the attack angle and its magnitude is unchanged. Unlike in the previous chapters, the attack on q sites affects a total of $p \geq q$ measurements. The injection of an offset in the time reference of a site shifts the phase of all phasors measured in this site by the same attack angle.

4.1.4 Vulnerability Condition

As in the previous chapters, we say that an attack is undetected if and only if it does not modify the residual vector, i.e. if and only if $r(\Delta z_\square) = G_\square \Delta z_\square = 0$, where $\Delta z_\square = z'_\square - z_\square$ is the difference between the attacked and unattacked measurement vectors. The following lemma shows that for exact vulnerability, the use of F and R is equivalent because the complex LS residuals are equal to 0 if and only if the WLS residuals in rectangular coordinates are equal to 0, which happens if and only if the complex vector Rz is equal to 0.

Theorem 4.3. *For a measurement vector z,*

$$r_\square(z_\square) = 0 \iff Fz = 0 \iff Rz = 0.$$

Proof. $r_\square(z_\square) = 0 \iff \exists x_\square = (x_{\square,1}, x_{\square,2})^T \in \mathbb{R}^{2n}$ s.t $x_{\square,1}, x_{\square,2} \in \mathbb{R}^n$ and $z_\square = H_\square x_\square$ $\iff z = Hx$, with $x = (x_{\square,1} + j x_{\square,2})^T \in \mathbb{C}^n \iff Fz = 0$.

If $r_\square(z_\square) = 0$, then $Rz + \bar{R}\bar{z} = 0$, hence $Rz = -\bar{R}\bar{z}$. Furthermore, if $r_\square(z_\square) = 0$ then there exist a state vector $x \in \mathbb{C}^n$ such that $z = Hx$. Hence, for any $\lambda \in \mathbb{C}$, $\lambda z = H(\lambda x)$ and thus $r_\square(\lambda z_\square) = 0$. Taking $\lambda = j$, we get $r_\square(z_\square) = jRz - j\bar{R}\bar{z} = 0$, hence $Rz = \bar{R}\bar{z}$. Because $Rz = -\bar{R}\bar{z} = \bar{R}\bar{z}$, we conclude that $Rz = 0$.

If $Rz = 0$, then $r_\square(z_\square) = Rz + \bar{R}\bar{z} = 0$. □

We choose to use F to establish exact vulnerability conditions because R depends on the measurement values and F does not. This enables us to provide a structural vulnerability condition that we cannot derive from R. In contrast, we use R for vulnerability metrics because it is linked to the WLS residuals. Thus it enables us to measure how detectable an attack is, or how vulnerable to attacks a set of sites is.

To the best of our knowledge, all known techniques to compute undetectable attack offsets require that the rank of the attack-angle matrix W is equal to 1. We refer to such attacks as rank-1 TSAs. To be more precise, if the rank of W is larger than 1, then the attack vector is required to satisfy a system of more than one equations. As the rank of W increases, it becomes increasingly unlikely that such a vector u exists. For example, it

was shown in [14] and recalled in Section 2, that if W is full rank, then there are no such u. It is also easy to show that if the $p \times p$ matrix W, such that $p > 2$, is of rank $p-1$, then it is very unlikely to find a vector $u \in \mathbb{C}^p$ such that $W(u-1) = 0$ and $|u_i| = 1$ for all i ranging from 1 to p. Indeed in that case, the kernel of W has dimension 1 and for all $\lambda \in \mathbb{C}^p$ such that $W\lambda = 0$, there must be an $l \in \mathbb{C}^*$ such that $\lambda = lv$, $v \in \ker(W)$. Hence, the existence of an attack vector u such that $|u_i| = 1$ for all i requires that there exists an $l \in \mathbb{C}^*$ such that $u - 1 = lv$, which translates to $l = \frac{u_1-1}{v_1} = ... = \frac{u_p-1}{v_p}$, $|u_i| = 1$. Thus, we would need to find a second point of intersection between p circles each of radius $1/|v_i|$ passing through a common origin. So all circles would need to be equal to one of two circles meaning that all coordinates of vector v can take values from a set of cardinality 2, which is highly unlikely. Finding new techniques to compute a valid attack vector u when $\text{rank}(W) = p - k$ for different values of k, is out of the scope of this thesis. We base our theory on what is known, specifically on rank-1 TSAs.

If for a pair of sites, that each measures a single phasor, the rank of W is not equal to 1, Chapters 2 and 3 show that it is sometimes possible to use a rank-1 approximation of W for rank-1 TSAs. This is possible when the index of separation (IoS), presented in Chapter 2 as the largest eigenvalue of W over the sum of both its eigenvalues, is close to 1. This occurs if the largest eigenvalue is significantly larger than the remaining one. The infimum of the IoS over all measurement values was introduced as IoS*. This quantity does not depend on the measurements. If it is equal to or close to 1, then the IoS is always equal to or close to 1. This condition is used, in the previous two chapters, to find vulnerable sets of PMUs, that each measures a single phasor, from the verification matrix only. For sites measuring an arbitrary number of phasors we are required, in Section 4.3, to study the effective rank of rectangular matrices with possibly more than two singular values. For this purpose, we introduce the effective rank ratio (ERR) of a matrix as its largest singular value over the sum of all its singular values. The ERR is close to 1 if the largest singular value is significantly larger than the others, in which case the effective rank of the matrix is close to 1. We use this condition to find vulnerable sets of sites that each measure an arbitrary number of phasors. We make the link between IoS and ERR in Section 4.3.4.

4.2 Vulnerability Conditions for a Single Site

We now provide a novel sufficient and necessary condition for rank-1 TSAs that target a single site. We then propose a vulnerability metric to measure how vulnerable a site is. Finally, we discuss the feasibility of the identified condition.

4.2.1 Vulnerability Condition

We assume that an attacker injects an offset in the time reference of $q = 1$ site; this affects p measurements. The remaining measurements are not affected by the attack. Note that an attack on a site with all p measurements equal to zero does not have any effect because the attacks create only phase shifts and a phase shift on a measurement equal to 0 does not modify the measurement. Recall that critical measurements can be attacked undetectably by any delay and that network operators typically ensure that no single-measurement is critical. Hence, we study the vulnerability of sites that measure at least one non-zero synchrophasor and where no measurement alone is critical. The following theorem gives necessary conditions in order to mount a rank-1 TSA.

Theorem 4.4. *Consider a rank-1 TSA on a single site measuring $p \geq 1$ phasors with indices in S_1, such that no measurement alone is critical and at least one measurement is not equal to* 0*.*

(a) If $p = 1$: such an attack is never feasible.

(b) If $p \geq 2$: such an attack is feasible if and only if z^{S_1} is in the kernel of F^{S_1}.

Proof. When $q = 1$, notice that by definition W corresponds to a single complex value: $W = \left(\sum_{i \in S_1} F_{:,i} z_i\right)^\dagger \left(\sum_{i \in S_1} F_{:,i} z_i\right)$. In this case, a rank-1 TSA corresponds to a single complex value u such that $W(u-1) = 0$. Observe that a non-trivial attack $u \neq 1$ exists if and only if $W = 0$. Thus, if and only if

$$\sum_{i \in S_1} F_{:,i} z_i = 0. \tag{4.1}$$

(a) If $p = 1$, then Eq. (4.1) is equivalent to $F_{:,1} z_1 = 0$. In other words, because z_1 is non-zero, a rank-1 TSA on a single site that measures a single phasor is feasible if and only if its corresponding column of F is equal to the null vector. Hence, such an attack is feasible if and only if the phasor is critical. As we assume that no single measurement is critical, we conclude that if $p = 1$, then no rank-1 TSA is feasible.

(b) If $p \geq 2$, observe that the left-hand-side of Eq. (4.1) is equal to $F^{S_1} z^{S_1}$. Therefore, Eq. (4.1) is satisfied if and only if the targeted measurement vector is in the kernel of F^{S_1}.

□

Observe that if a site is vulnerable to rank-1 TSAs, then any attack angle will be undetected: if site S_1 is vulnerable, then $F^{S_1} z^{S_1} = 0$. Hence, whatever the choice of attack vector u, $F^{S_1} \Delta z^{S_1} = F^{S_1}(z^{S_1} u - z^{S_1}) = F^{S_1} z^{S_1}(u-1) = 0$.

4.2.2 Vulnerability Metric: distance to item (b) of theorem 4.4

If a site is not vulnerable to rank-1 TSAs, it might still be close to it. In other words, if the residuals obtained after an attack are different but very close to the residuals obtained without an attack, the attack is undetectable in practice. The following theorem shows that $\|R^{S_1} z^{S_1}\|$ where $\|.\|$ is the l_2-norm, is a reliable vulnerability metric for single sites. The closer this metric is to 0, the more vulnerable the site is.

Theorem 4.5. *$\|r_\square(\Delta z_\square)\| \leq 4\|R^{S_1} z^{S_1}\| \ \forall \ \Delta z_\square = (Re(\Delta z), Im(\Delta z))^T$, where $\Delta z = z(u_1 - 1)$ and $u_1 \in \mathbb{C}$ such that $|u_1| = 1$, and where $\|.\|$ is the l_2-norm.*

Proof. Values of the complex vector $R^{S_1} z_\square^{S_1}$ can be written in polar form $\rho_i e^{j\phi_i}$ and any $u_1 \in \mathbb{C}$ such that $|u_1| = 1$ can also be written in polar form $u_1 = e^{j\delta}$. Hence,

$$\begin{aligned}
\|r_\square(\Delta z_\square)\| &= \|R\Delta z + \bar{R}\bar{\Delta}\bar{z}\| = \|2\mathrm{Re}(R\Delta z)\| \\
&= \|2\mathrm{Re}((u_1 - 1)R^{S_1} z^{S_1})\| \\
&= \sqrt{4\sum_{i=1}^{m} \rho_i^2 (cos(\phi_i + \theta) - cos(\phi_i))^2} \\
&\leq \sqrt{4\sum_{i=1}^{m} 4\rho_i^2} = 4\sqrt{\sum_{i=1}^{m} \rho_i^2} = 4\|R^{S_1} z^{S_1}\|.
\end{aligned}$$

□

4.2.3 Feasibility of the Vulnerability Condition: item (b) of theorem 4.4

If a rank-1 TSA is feasible, then there exists a relation among the measurement values with coefficients computed from values of vectors spanning the kernel of F^{S_1} (also referred to as its null-space). For example, in the case where $p = 2$, an attack is feasible if and only if the two involved measurements (z_1, z_2) are such that they satisfy the relation

$$\frac{z_1}{z_2} = \frac{n_1}{n_2}, \tag{4.2}$$

where $(n_1, n_2)^T$ is a fixed vector that spans $\ker(F^{S_1})$. This is because F^{S_1} is an m by 2 matrix, hence its rank is at most equal to 2. As none of the measurements are critical by themselves, its rank cannot be equal to 0 and as z^{S_1} is in the kernel of F^{S_1}, its rank cannot be equal to 2. Hence, its rank is equal to 1. By the rank theorem, $\ker(F^{S_1})$ is of dimension equal to 1. Hence, z^{S_1} must be a non-zero complex multiple of any vector that spans $\ker(F^{S_1})$, which leads to Eq. (4.2).

Also, such a relation seems unlikely to occur as measurement values depend on independent loads. For example, if z_1 is a voltage value V and z_2 a current value I, then at all time

instant i, a rank-1 TSA is feasible if and only if

$$\frac{z_1^i}{z_2^i} = \frac{V^i}{I^i} = \frac{V^i I^{\dagger i}}{|I^i|^2} = \frac{S_{inj}^i}{|I|^2} = \frac{n_1}{n_2}.$$

Hence, at all time instants the phase of the complex injected power S_{inj} must remain constant and equal to $\arg \frac{n_1}{n_2}$. Also, observe that the relation implies that $\frac{|V^i|}{|I^i|} = \frac{|n_1|}{|n_2|}$. As the magnitude of voltage values is always close to 1, the magnitude of the current values is approximately invariant and close to $\frac{|n_2|}{|n_1|}$. In practice, the injected current and power depend on the loads that vary in time due to external factors. Hence, it seems unlikely that they could be of constant magnitude or phase and even less likely that such constant values could be equal to specific values computed from the verification matrix only. Therefore, we conclude that the necessary conditions for a rank-1 TSA on a single site that measures two phasors are unlikely to occur on a realistic grid.

We present numerical values of the metric $\|R^{S_1} z^{S_1}\|$ obtained through realistic simulations in Section 4.6. In our simulations, we always encountered large values far from 0, which corroborates our analysis that necessary conditions to mount a rank-1 TSA on a single site seem unrealistic; but we still recommend to monitor this metric on a real grid.

4.3 Vulnerability Conditions for a Pair of Sites

If two sites are not vulnerable by themselves to rank-1 TSAs, they could still be vulnerable together. We identify two novel conditions to simultaneously mount rank-1 TSAs on such two sites. One of them is a sufficient vulnerability condition that is structural as it depends on the LS verification matrix only. The other is a general necessary and sufficient condition that depends on the measurement values. We then provide a vulnerability metric to measure the vulnerability of any pair of sites. We also discuss the feasibility of the identified general vulnerability condition. Finally, we establish the relation between our novel conditions and the vulnerability conditions presented in Chapters 2 and 3 for the special case where both sites measure only a single phasor (i.e. $p = 2$).

4.3.1 Vulnerability Conditions

We suppose that an attacker injects different offsets in the time reference of $q = 2$ sites, hence affecting a total of p phasor measurements with indices listed in S_1 and S_2 for the first and second sites, respectively. The following theorem gives necessary conditions in order to mount a rank-1 TSA.

Theorem 4.6. *Consider a rank-1 TSA on $q = 2$ sites measuring phasors with indices in S_1 and S_2, respectively, such that $|S_1| + |S_2| = p$, no measurement is critical by itself, neither site is vulnerable to rank-1 TSAs by itself and at least one measurement in each site is not equal to zero.*

(a) Such an attack is feasible if and only if $F^{S_1}z^{S_1}$ and $F^{S_2}z^{S_2}$ are colinear.

(b) If $\operatorname{rank}(F^{[S_1,S_2]}) = 1$, i.e. if all pairs of measurements in $S_1 \cup S_2$ are critical: such an attack is always feasible.

(c) If $\operatorname{rank}(F^{[S_1,S_2]}) = p$, i.e. if the combined set of measurements $S_1 \cup S_2$ is not critical: such an attack is never feasible.

Proof. Rank-1 TSAs are feasible if and only if $\operatorname{rank}(W) = 1$. By the rank properties of complex matrices, we have that $\operatorname{rank}(W) = \operatorname{rank}((F\operatorname{diag}(z)\varphi)^\dagger(F\operatorname{diag}(z)\varphi)) = \operatorname{rank}(F\operatorname{diag}(z)\varphi)$. Hence, rank-1 TSAs are feasible if and only if

$$\operatorname{rank}\left(\begin{bmatrix} \sum_{i\in S_1} F_{1,i}z_i & \sum_{i\in S_2} F_{1,i}z_i \\ \vdots & \vdots \end{bmatrix}\right) = 1. \tag{4.3}$$

(a) Eq. (4.3) is satisfied if and only if either

- one of the columns of the matrix is equal to the null vector. According to Theorem 4.4 this is equivalent to saying that a rank-1 TSA can be mounted directly on the corresponding site. However, this case is excluded because we suppose that no rank-1 TSA can be mounted on a site by itself.
- or the two columns of the matrix are colinear. Specifically $F^{S_1}z^{S_1}$ and $F^{S_2}z^{S_2}$ are colinear.

Since $F^{S_1}z^{S_1}$ and $F^{S_2}z^{S_2}$ are non-zero, they are colinear if and only if there exists an $l \in \mathbb{C}^*$ such that $F^{S_1}z^{S_1} + lF^{S_2}z^{S_2} = 0 \iff F^{[S_1,S_2]}(z^{S_1}, lz^{S_2})^T = 0 \iff (z^{S_1}, lz^{S_2})^T \in \ker(F^{[S_1,S_2]}) \iff \dim(E_Z \cap E_N) \geq 1$, with $E_Z = \operatorname{span}\{(z^{S_1}, 0)^T, (0, z^{S_2})^T\}$ and $E_N = \ker(F^{[S_1,S_2]})$. According to Grassmann's formula, this is equivalent to $\dim(E_Z) + \dim(E_N) - \operatorname{rank}([Z|N]) \geq 1$, where $[Z|N] = \begin{bmatrix} z^{S_1} & 0 & N \\ 0 & z^{S_2} & \vdots \end{bmatrix}$ and where N is a matrix with independent columns that span $\ker(F^{[S_1,S_2]})$. Therefore, rank-1 TSAs are feasible if and only if

$$\begin{aligned} &2 + p - \operatorname{rank}(F^{[S_1,S_2]}) - 1 \geq \operatorname{rank}([Z|N]) \\ \iff\ &1 + p - \operatorname{rank}(F^{[S_1,S_2]}) \geq \operatorname{rank}([Z|N]). \end{aligned} \tag{4.4}$$

(b) If $\operatorname{rank}(F^{[S_1,S_2]}) = 1$, then Eq. (4.4) $\iff p \geq \operatorname{rank}([Z|N])$. Observe that in this case, $[Z|N]$ is a p by $p+1$ matrix. Hence, it is always the case that the rank of $[Z|N]$ is at most equal to p. In other words, a rank-1 TSA is always feasible.

(c) If $\operatorname{rank}(F^{[S_1,S_2]}) = p$, then Eq. (4.4) $\iff 1 \geq \operatorname{rank}(Z)$, which is never satisfied as Z is a matrix of rank equal to 2. In other words, a rank-1 TSA is never feasible.

□

Interestingly, item (a) of Theorem 4.6 implies that if three different sites S_i, S_j and S_k, that are not vulnerable by themselves, are such that two pairs (S_i, S_j) and (S_i, S_k) are vulnerable, then (S_j, S_k) is also vulnerable. In other words, item (a) of theorem 4.6 defines an equivalence relation over the set of sites that are not vulnerable by themselves.

As mentioned previously, Theorem 4.6 establishes two novel vulnerability conditions:

- a *general vulnerability* condition defined by item (a) of Theorem 4.6: this condition is necessary and sufficient for vulnerability to rank-1 TSAs. It depends on the measurement values.
- a *structural vulnerability* condition defined by item (b) of Theorem 4.6: this sufficient vulnerability condition depends on the LS verification matrix only. In other words, if two sites satisfy this condition, then they are vulnerable but two sites that do not satisfy this condition might still be vulnerable.

If a pair of sites is not exactly *structurally* vulnerable to rank-1 TSAs, it can be practically so. Specifically, if the rank of $F^{[S_1,S_2]}$ is not equal to 1, its effective rank can be close to 1. This distance is captured by the ERR of $F^{[S_1,S_2]}$. Numerically, it can occur that a column of $F^{[S_1,S_2]}$ has significantly larger values than the remaining columns, in which case the effective rank is smaller than the computed rank. As a result, $\mathrm{ERR}(F^{[S_1,S_2]})$ can be very close to 1. In this case it is likely that a rank-1 TSA computed from a rank-1 approximation of $F^{[S_1,S_2]}$, is feasible in practice irrespective of the measurement values.

We present numerical results through realistic simulations in Section 4.6. Our simulations show that when the rank of $F^{[S_1,S_2]}$ is not equal to 1, its ERR is sometimes very close to 1, in which case the corresponding pair of sites is practically vulnerable.

4.3.2 General Vulnerability Metric: distance to item (a) of theorem 4.6

As in the single-site analysis, when a pair of sites is not exactly vulnerable to rank-1 TSAs, it can still be close to vulnerable. Theorem 4.3 implies that item (a) of theorem 4.6 can be equivalently written using the general WLS notations introduced in Section 4.1.2 as follows.

Theorem 4.7. *A pair of sites S_1 and S_2 with non-zero and non-critical measurements that are not vulnerable by themselves, are vulnerable together to undetectable TSAs if and only if there exist an $l \in \mathbb{C}^*$ such that $R^{S_1} z^{S_1} = l R^{S_2} z^{S_2}$.*

Proof. Vulnerability is equivalent to $r_\square(\Delta z_\square) = 0$ with $\Delta z = \begin{pmatrix}(u_1 - 1)z^{S_1} & (u_2 - 1)z^{S_2} & 0\end{pmatrix}^T$. Hence, by the last part of theorem 4.3, the vulnerability condition is equivalent to

$(u_1-1)R^{S_1}z^{S_1}+(u_2-1)R^{S_2}z^{S_2}=0$. In other words, the pair of sites is vulnerable if and only if there exists an $l=-\frac{u_2-1}{u_1-1}\in\mathbb{C}^*$ such that $|u_2|=|u_1|=1$ and $R^{S_1}z^{S_1}=lR^{S_2}z^{S_2}$. Note that we can divide by (u_1-1) because $u_1=1$ is the non-attack solution. □

In other words, an undetectable attack requires that the rank of the $m\times 2$ complex matrix $\left[R^{S_1}z^{S_1}|R^{S_2}z^{S_2}\right]$ is equal to 1. If it is not the case, the metric of vulnerability that we propose is $1-\mathrm{ERR}(\left[R^{S_1}z^{S_1}|R^{S_2}z^{S_2}\right])$. This metric shows how vulnerable a set of measurements corresponding to a pair of sites is. The closer this metric is to 0, the more vulnerable the pair of sites is.

4.3.3 Feasibility of the General Vulnerability Condition: item (a) of Theorem 4.6

If a pair of structurally non-vulnerable sites satisfy the general vulnerability condition, then there exists a relation between the measurements with coefficients computed from the verification matrix. Specifically, if $\mathrm{rank}\,(F^{[S_1,S_2]})\neq 1$ but z^{S_1} and z^{S_2} are such that $F^{S_1}z^{S_1}$ and $F^{S_2}z^{S_2}$ are colinear, then at least one of the measurements is directly determined by a combination of the other measurements and values of the vectors spanning $\ker(F^{[S_1,S_2]})$.

Relations for two sites each measuring two phasors. For example if phasors (z_1,z_2) and (z_3,z_4) are measured at the first and second sites, respectively, then a static analysis shows that $\mathrm{rank}\,(F^{[S_1,S_2]})\in\{2,3\}$. Notice that $F^{[S_1,S_2]}$ is a $m\times 4$ matrix with $m\geq 4$, hence $0\leq\mathrm{rank}\,(F^{[S_1,S_2]})\leq 4$. Because we assume that no single measurement is critical, the rank cannot be equal to 0, thus $1\leq\mathrm{rank}\,(F^{[S_1,S_2]})\leq 4$. Item (c) of Theorem 4.6 says that if $S_1\cup S_2$ forms a non-critical set, then no rank-1 TSAs are feasible on this pair of sites, therefore $F^{[S_1,S_2]}$ cannot be full rank: $1\leq\mathrm{rank}\,(F^{[S_1,S_2]})\leq 3$. From Theorem 4.2, we know that if the rank is equal to 1, then every pair of measurements in $S_1\cup S_2$ is critical, which contradicts the assumption that the two sites are not vulnerable to rank-1 TSAs on their own. Thus, $2\leq\mathrm{rank}\,(F^{[S_1,S_2]})\leq 3$.

- If $\mathrm{rank}\,(F^{[S_1,S_2]})=3$, then a rank-1 TSA on the two sites is feasible if and only if

$$\frac{z_1}{z_2}=\frac{n_1}{n_2}\text{ and }\frac{z_3}{z_4}=\frac{n_3}{n_4},\tag{4.5}$$

  where $(n_1,n_2,n_3,n_4)^T$ is a fixed vector that spans the kernel of $F^{[S_1,S_2]}$.

- If $\mathrm{rank}\,(F^{[S_1,S_2]})=2$, then a rank-1 TSA on the two sites is feasible if and only if

$$\frac{z_3}{z_4}=\frac{z_1(v_3n_2-v_2n_3)+z_2(v_1n_3-v_3n1)}{z_1(v_4n_2-v_2n_4)+z_2(v_1n_4-v_4n1)},\tag{4.6}$$

where $(n_1, n_2, n_3, n_4)^T$ and $(v_1, v_2, v_3, v_4)^T$ are fixed independent vectors that span the kernel of $F^{[S_1,S_2]}$.

As for the single-site vulnerability condition discussed in Section 4.2.3, we reason that such relations are unlikely to occur on real grids, as measurement values depend on independent loads. Numerical evidence provided in Section 4.6.4, shows that the relations given by Eq. (4.5) and by Eq. (4.6) are far from occurring during our simulations on the IEEE 39-bus benchmark system with real load profiles taken from the Lausanne grid.

Proof of Eq. (4.5) and of Eq. (4.6). Let us consider the two possibilities allowed by the non-zero and non-criticality assumptions.

- rank $(F^{[S_1,S_2]}) = 3$*:* by the rank theorem, the dimension of the kernel of $F^{[S_1,S_2]}$ is equal to 1. Let $(n_1, n_2, n_3, n_4)^T$ be a vector that spans the kernel of $F^{[S_1,S_2]}$. Recall that by Eq.(4.4), an attack is feasible if and only if $1+p-\text{rank}\,(F^{[S_1,S_2]}) = 2 \geq \text{rank}\,([Z|N])$ and that by construction $\text{rank}\,([Z|N]) \geq 2$. Therefore $\text{rank}\begin{pmatrix} z_1 & 0 & n_1 \\ z_2 & 0 & n_2 \\ 0 & z_3 & n_3 \\ 0 & z_4 & n_4 \end{pmatrix} = 2$, which leads to Eq. (4.5).

- rank $(F^{[S_1,S_2]}) = 2$*:* by the rank theorem, the dimension of the kernel of $F^{[S_1,S_2]}$ is equal to 2. Let $(n_1, n_2, n_3, n_4)^T$ and $(v_1, v_2, v_3, v_4)^T$ be independent vectors that span the kernel of $F^{[S_1,S_2]}$. By item (a) of Theorem 4.6, a rank-1 TSA is feasible if and only if there exist an $l \in \mathbb{C}^*$ such that $F^{[S_1,S_2]}(z_1, z_2, lz_3, lz_4)^T = 0$. Hence, (z_1, z_2, lz_3, lz_4) must be a linear combination of (n_1, n_2, n_3, n_4) and (v_1, v_2, v_3, v_4). By Theorem 4.2, because $\text{rank}\,(F^{[S_1,S_2]}) = 2$, all sets of three measurements are critical and thus all sets of three measurements are linearly dependant. As a result, all 3×3 submatrices of
$$\begin{bmatrix} z_1 & v_1 & n_1 \\ z_2 & v_2 & n_2 \\ lz_3 & v_3 & n_3 \\ lz_4 & v_4 & n_4 \end{bmatrix}$$
are not full rank, thus their determinant is equal to 0. More specifically, this means that
$$\det \begin{bmatrix} z_1 & v_1 & n_1 \\ z_2 & v_2 & n_2 \\ lz_3 & v_3 & n_3 \end{bmatrix} = 0, \text{ and that } \det \begin{bmatrix} z_1 & v_1 & n_1 \\ z_2 & v_2 & n_2 \\ lz_4 & v_4 & n_4 \end{bmatrix} = 0.$$
The computation of the first determinant leads to
$$\frac{z_1(v_3n_2 - v_2n_3) + z_2(v_1n_3 - v_3n1)}{z_3} = -l(v_1n_2 - v_2n_1).$$

Similarly, the computation of the second determinant leads to

$$\frac{z_1(v_4n_2 - v_2n_4) + z_2(v_1n_4 - v_4n1)}{z_4} = -l(v_1n_2 - v_2n_1).$$

Note that these equations are well defined because all measurements are non-zero. Observe that the left hand side of both equations are equal. By combining them and by rearranging terms, we obtain Eq. (4.6). Note that the denominator $z_1(v_4n_2 - v_2n_4) + z_2(v_1n_4 - v_4n1)$ cannot be equal to 0 because this would imply that all coefficients of the form $(v_in_k - v_kn_i)$ with $i \neq k$, are equal to 0. This would mean that all pairs of measurements are linearly dependant which is in contradiction with the fact that all sets of three measurements are critical minimal.

In our simulations presented in Section 4.6.4, the observed values of $1-\text{ERR}(\left[R^{S_1}z^{S_1}|R^{S_2}z^{S_2}\right])$ for pairs of sites that do not satisfy the structural vulnerability condition are always far from 0. This corroborates our intuition that the general vulnerability condition seems unlikely to occur for pairs of sites that are not already structurally vulnerable to rank-1 TSAs.

4.3.4 Relation with the Vulnerability Conditions of Chapters 2 and 3

We now show that if each site measures a single phasor (i.e. $p = 2$), then our vulnerability conditions are equivalent to the vulnerability conditions identified in Chapters 2 and 3. The following theorem gives the equivalence for the general vulnerability condition.

Theorem 4.8. *Consider a rank-1 TSA on $q = 2$ sites, each measuring a single phasor z_1 and z_2, respectively, such that no measurement is critical by itself or is equal to zero. Then $IoS_{1,2}(z_1, z_2) = 1$ if and only if $F^{S_1}z_1$ and $F^{S_2}z_2$ are colinear.*

Proof. As there are only 2 involved measurements, W is a 2 by 2 matrix. As no measurement is critical by itself, the rank of W is equal to either 1 or 2. By definition, the IoS of W is equal to 1 if and only if its smallest eigenvalue is equal to 0, which is equivalent to $\text{rank}\,(W) = 1$. Recall from the proof of Theorem 4.6 that the rank of W is equal to 1 if and only if Eq. (4.3) is satisfied. As no measurement is critical by itself, this is equivalent to the colinearity of $F^{S_1}z_1$ and $F^{S_2}z_2$. □

Similarly, the following theorem gives the equivalence for the structural vulnerability condition.

Theorem 4.9. *Consider a rank-1 TSA on $q = 2$ sites, each measuring a single phasor z_1 and z_2, respectively, such that no measurement is critical by itself or is equal to 0. Then $IoS^*_{1,2} = 1$ if and only if $\text{rank}\,(F^{[S_1,S_2]}) = 1$.*

Proof. We show both directions:

- $IoS^*_{1,2} = 1 \Rightarrow \text{rank}\,(F^{[S_1,S_2]}) = 1$: By definition of IoS^*, if it is equal to 1, then the IoS is equal to 1, whatever the values of z_1 and z_2. Therefore, Eq. (4.3) is satisfied even if $z_1 = z_2 = 1$, which is equivalent to $\text{rank}\,(F^{[S_1,S_2]}) = 1$.

- $IoS^*_{1,2} = 1 \Leftarrow \text{rank}\,(F^{[S_1,S_2]}) = 1$: recall from Eq. (2.5) that if $p = 2$, then

$$IoS^*_{1,2} = \frac{1}{2} + \frac{|f_{12}|}{2(f_{11}f_{22})^{1/2}}, \tag{4.7}$$

  with $f_{it} = \sum_{l,m}\sum_n \varphi_{l,i}\varphi_{m,t}\bar{F}_{n,l}F_{n,m}$. Notice that
  $|f_{12}|^2 = \left(\sum_{i=1}^m \bar{F}_{i,1}F_{i,2}\right)\left(\sum_{i=1}^m \bar{F}_{i,2}F_{i,1}\right)$,
  $f_{11} = \sum_{i=1}^m \bar{F}_{i,1}F_{i,1}$ and $f_{22} = \sum_{i=1}^m \bar{F}_{i,2}F_{i,2}$.
  If $\text{rank}\,F^{[S_1,S_2]} = 1$, then there exists $l \in \mathbb{C}^*$ such that $F_{:,1} = lF_{:,2}$ because no measurement is neither equal to 0 nor critical. Hence, $f_{11} = |l|^2 f_{22}$ and $|f_{12}|^2 = |l|^2 f_{22}^2$. By plugging them into Eq. (4.7), we get that $IoS^* = 1$.

□

Note that $IoS_{1,2}(z_1, z_2)$ is the IoS of the attack-angle matrix W computed from measurements (z_1, z_2). Also, $IoS^*_{1,2}$ is the infimum of $IoS_{1,2}(z_1, z_2)$ over all possible values of (z_1, z_2). Both $ERR(F^{[S_1,S_2]})$ and $IoS^*_{1,2}$ are independent of measurement values. Theorem 4.9 implies that one of them is equal to 1 if and only if the other is also equal to 1. Therefore, they are both equal to 1 for an exactly structurally vulnerable pair of sites measuring a single phasor. However, as we show next, $ERR(F^{[S_1,S_2]})$ may be close to 1 when $IoS^*_{1,2}$ is not. In other words, the structural vulnerability condition established in this chapter is better than the one introduced in Chapter 2 and used in Chapter 3, because it identifies more sets that are practically vulnerable.

Claim 4.1. *For any pair of sites,* $1 \geq ERR(F^{[S_1,S_2]}) \geq IoS^*_{1,2}$ *and* $ERR(F^{[S_1,S_2]})$ *can be close to* 1 *when* $IoS^*_{1,2}$ *is not if* F^{S_1} *has much larger values than* F^{S_2} *or vice-versa.*

Proof. Recall the structure of the attack-angle matrix:

$$W = \begin{bmatrix} |z_1|^2 f_{11} & \bar{z_1}z_2 f_{12} \\ z_1\bar{z_2} f_{21} & |z_2|^2 f_{22} \end{bmatrix},$$

where $f_{it} = (F^{S_i})^\dagger F^{S_t}$. Using this notation, recall from Eq. (2.5) that $IoS^*_{1,2} = \frac{1}{2} + \frac{|f_{12}|}{2\sqrt{f_{11}f_{22}}}$. Our metric $ERR(F^{[S_1,S_2]})$ is equal to $IoS(X)$ and $X = (F^{[S_1,S_2]})^\dagger F^{[S_1,S_2]} = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$. Recall from Eq. (2.4) that $IoS(X) = \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\frac{\det(X)}{\text{Trace}(X)^2}}$, where $\det(x) =$

$f_{11}f_{22} - |f_{12}|^2$ and $\text{Trace}(X) = f_{11} + f_{22}$ are the determinant and trace of X, respectively. Hence,

$$\begin{aligned} IoS(X) &= \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\frac{f_{11}f_{22} - |f_{12}|^2}{(f_{11}+f_{22})^2}} \\ &= \frac{1}{2} + \frac{1}{2}\frac{\sqrt{(f_{11}+f_{22})^2 - 4f_{11}f_{22} + 4|f_{12}|^2}}{f_{11}+f_{22}}. \end{aligned}$$

Denoting $P = f_{11}f_{22} \in \mathbb{R}^+$, $D = |f_{12}| \in \mathbb{R}^+$ and $S = f_{11} + f_{22}$, we have

$$\begin{aligned} IoS^*_{1,2} &= \frac{1}{2} + \frac{1}{2}\sqrt{\frac{D^2}{P}} \text{ and} \\ IoS(X) &= \frac{1}{2} + \frac{1}{2}\sqrt{\frac{S^2 - 4P + 4D^2}{S^2}}. \end{aligned}$$

Given P and S, values f_{11} and f_{22} exist if and only if $S^2 \geq 4P$. Define function $\phi(x) = \frac{x-4P+4D^2}{x} = 1 - 4\frac{P-D^2}{x} \leq 1$ because $P \geq D^2$. Because $S^2 \geq 4P$, the lowest value of the definition domain of $\phi(x)$ is $4P$: $\phi(4P) = \frac{D^2}{P}$. The derivative $\phi'(x) = \frac{P-D^2}{x^2}$ being positive implies that $\frac{S^2-4P+4D^2}{S^2} \geq \frac{D^2}{P}$, hence $1 \geq IoS(X) \geq IoS^*_{1,2}$. The inequalities are equalities if $D^2 = P$. If P is much larger than D^2, then $IoS^*_{1,2}$ is much smaller than 1 and if S is much larger than $2\sqrt{P}$, then $IoS(X)$ is close to 1. In that case, an attack is feasible, $IoS^*_{1,2}$ doest not enable the detection of the vulnerability but $IoS(X) = ERR(F^{[S_1,S_2]})$ does. This happens when $S >> 2\sqrt{P}$, thus when

$$\begin{aligned} f_{11} + f_{22} &>> 2\sqrt{f_{11}f_{22}} \\ (f_{11}+f_{22})^2 &>> 4f_{11}f_{22} \\ f_{11}^2 - 2f_{11}f_{22} + f_{22}^2 = (f_{11}-f_{22})^2 &>> 0 \\ f_{11} >> f_{22} &\text{ or } f_{22} >> f_{11} \end{aligned}$$

In other words, if the values of F^{S_1} are much larger than the values of F^{S_2} or vice-versa, then $ERR(F^{[S_1,S_2]})$ is close to 1 but $IoS^*_{1,2}$ is not. □

4.4 Vulnerability Conditions for an Arbitrary Number of sites

We now show that a rank-1 TSA targeting an arbitrary number of sites $q \geq 2$ that are not vulnerable by themselves, is feasible if and only if for every pair of sites among the targeted set of sites, a rank-1 TSA is feasible. As a result, the vulnerability analysis of a grid reduces to the vulnerability analysis of each site and each pair of sites. We then

establish the relation between the new result and those of Chapter 3 for the special case where all targeted sites measure only a single phasor.

4.4.1 Vulnerability Conditions

The following theorem establishes the general vulnerability condition for an arbitrary number of sites that are not vulnerable by themselves to rank-1 TSAs.

Theorem 4.10. *A set of $m \geq q \geq 2$ sites, each measuring an arbitrary number of phasors, such that none of the sites are vulnerable to rank-1 TSAs by themselves, are vulnerable together if and only if each pair of sites within the set is vulnerable to rank-1 TSAs.*

Proof. Define S_1, S_2, ..., S_q to be the set of measurement indices corresponding to phasors measured in the first, second, up to q^{th} targeted sites, respectively. Then, the rank of the attack-angle matrix W corresponds to the rank of the following matrix

$$T = \begin{bmatrix} \sum_{i \in S_1} F_{1,i} z_i & \cdots & \sum_{i \in S_q} F_{1,i} z_i \\ \vdots & & \vdots \end{bmatrix}.$$

- The set of q sites is vulnerable to rank-1 TSAs $\Rightarrow$ all pairs of sites in the set are vulnerable. The set of q sites is vulnerable if and only if $\text{rank}(T) = 1$. The rank of this m by q matrix is equal to 1 if and only if all columns of T are dependant. As none of the sites are vulnerable to rank-1 TSAs by themselves, no column of T is equal to the null vector. Therefore, all sub matrices consisting of two columns of T are of rank equal to 1. This means that the attack-angle matrix W that corresponds to the attack targeting the corresponding two sites is of rank equal to 1. In other words, if a set of q sites is vulnerable to rank-1 TSAs, then any two sites within the set of targeted sites are also vulnerable to rank-1 TSAs.

- The set of q sites is vulnerable to rank-1 TSAs $\Leftarrow$ all pairs of sites in the set are vulnerable. If there is a set of q sites such that all pairs of sites within the set are vulnerable, then all columns of T are dependant, which means that its rank is equal to 1. Therefore the large set of q sites is vulnerable to rank-1 TSAs.

□

Theorem 4.10 implies that a site is vulnerable to rank-1 TSAs either if it is vulnerable by itself or if its combination with at least one other site forms a vulnerable set where all pairs of sites are vulnerable together. Hence, mitigating the feasibility of rank-1 TSAs for each site and each pair of sites is sufficient to mitigate the attack feasibility of the

Vulnerability	Structural	General	Vulnerability metric
Single site	none	$F^{S_1} z_{S_1} = 0$	$\|R^{S_1} z_{S_1}\|$
Pair of sites	$\text{ERR}(F^{[1,2]}) \approx 1$	$F^{S_1} z^{S_1}$ and $F^{S_2} z^{S_2}$ colinear	$1 - \text{ERR}(\left[R^{S_1} z^{S_1} \middle\| R^{S_2} z^{S_2}\right])$

Table 4.1 – Vulnerability conditions and vulnerability metrics for each site and pair of sites of a grid with non-critical measurements.

grid. Table 4.1 recapitulates the vulnerability conditions and distance to vulnerability metrics for a site and for a pair of sites with non-critical measurements.

4.4.2 Relation with the Results of Chapter 3

Chapter 3 shows that measurements can be grouped in equivalence classes, when the IoS values of the attack-angle matrices of all pairs of measurements are equal to 1. It then shows that a set of measurements is vulnerable to rank-1 TSAs if and only if the set of measurements belong to the same equivalence class. In other words, a set of sites, each measuring a single phasor, is vulnerable to rank-1 TSAs if and only if all pairs of sites within the targeted set is vulnerable. Recall that item (a) of theorem 4.6 defines an equivalence relation over the set of sites that are not vulnerable by themselves. Hence, Theorem 4.10 shows that a set of non-vulnerable sites is vulnerable if and only if the sites belong to the same equivalence class. Therefore Theorem 4.10 applied to sites measuring a single phasor coincides with the result of Chapter 3.

4.5 Mitigating Rank-1 TSAs

To minimize the feasibility of rank-1 TSAs, we now combine results from Sections 4.2, 4.3 and 4.4. Specifically, we propose as a security requirement to ensure that no pair of sites is structurally vulnerable to TSAs. Recall that a pair of sites is structurally vulnerable to TSAs if and only if $\text{ERR}(F^{[S_1,S_2]}) = 1$, in which case we say that the pair is exactly structurally vulnerable to TSAs. Recall further that in practice a pair of sites can be structurally vulnerable to TSAs if $\text{ERR}(F^{[S_1,S_2]})$ is close to 1, in which case we say that the pair is practically structurally vulnerable. We show that enforcing the exact structural security requirement with minimum changes, is an NP-hard problem that can be formulated as an integer linear program (ILP), which can be solved optimally. However, enforcing the exact structural security requirement is not sufficient for the practical security of the grid. We do not know how to formulate the practical structural security requirement as an ILP, but we propose a greedy offline algorithm that enforces it. This algorithm can also provide a heuristic solution that satisfies the exact security requirement. We show that our greedy algorithm applied to the grid of Section 4.6 for exact structural security, outputs the same optimal solution as the branch-and-bound

algorithm applied to the ILP formulation of the exact problem. Even if a grid is not structurally vulnerable, there is still an unlikely possibility that measurement values are such that some sites or pairs of sites are vulnerable. In order to check the non-vulnerability of the system, we recommend the monitoring of the vulnerability metrics.

4.5.1 Securing against the Structural Vulnerability Condition

Apart from vulnerability conditions, Theorem 4.6 also identifies a structural non-vulnerability condition. Item (c) of Theorem 4.6 states that if the combined set of measurements from the two sites does not form a critical set, then they are not vulnerable to rank-1 TSAs, irrespective of their measurement values. Hence, a natural idea to secure all pairs of sites is to ensure that none of them forms a critical set of measurements. However, from an engineering perspective, it is not realistic to impose this security measure, as it would either be impossible to enforce or require much redundancy in the measurements. This would require substantially more PMUs than what is required for observability of the system, which would be costly. For example on the grid used for simulations in Section 4.6, if we assume that sites are groups of buses separated by a converter only, then by placing PMUs measuring voltages and nodal-injection currents on every bus that is connected to a load or a generator, we obtain that 85 out of 253 pairs of sites are still critical. In this example, where the number of PMUs deployed on the grid is much higher than what is required for observability, approximatively 34% of the pairs of sites are still critical.

In contrast, it is possible to carefully increase the measurement redundancy of an observable grid's PMU allocation such that no pair of sites is structurally vulnerable. In what follows, we show that finding the minimum amount of PMUs to add in order to enforce the exact structural security requirement is a NP-hard problem which can be written as an ILP. This exact formulation of the problem can be solved heuristically but also optimally by algorithms such as the branch-and-bound algorithm. As mentioned previously, this is not sufficient to ensure the practical security of the grid. We propose a greedy heuristic which finds measurement points to add in order to enforce the practical security requirement.

4.5.1.1 Enforcing Exact Structural Security

Recall that theorem 4.2 implies that $ERR(F^{[S_1,S_2]}) = 1$ if and only if all pairs of measurements in $S_1 \cup S_2$ are critical. Hence, the exact structural security requirement is that there must be at least one non-critical pair of measurements per pair of sites. The optimal PMU placement problem, which aims to find the PMU allocation with the minimum number of PMUs such that the grid is observable, is known to be NP-hard [89]. Enforcing our security requirement with the minimum amount of PMUs corresponds to solving the optimal PMU placement problem with the added constraint that each pair of

sites has at least one non-critical pair of measurements. If we define the entire grid to be in a single site, then solving the problem with the added constraint actually solves the optimal PMU placement problem. Therefore, finding the minimum set of PMUs to add in order to enforce our security requirement, is also a NP-hard problem.

The optimal PMU placement problem can be formulated as an integer linear programm (ILP) [90] and our security requirement can be formulated as an additional linear constraint. The objective is to find the minimum amount of PMUs to add

$$\arg \min_{X_{new}, Y_{a,b,i,k} \ \forall(a,b,i,k)} c^T X_{new},$$

where

- c is a $n \times 1$ vector of ones, where n is still the number of buses,
- X_{new} is a $n \times 1$ vector of decision variables that can each take values in $\{0, 1, 2\}$, depending on the number of measurement points to add on the corresponding bus,
- $Y_{a,b,i,k}$ is a decision binary variable that is equal to 1 if measurement pair (a, b) in pair of sites (i, k), is non-critical.

The solution should satisfy the observability and security constraints

- network observability means that every bus is observed by at least one measurement, which means that it has a measurement point or it is adjacent to a bus that has a measurement point. The fact that there should be no single critical measurement requires that every bus should be observed by at least two measurements. These constraints can be combined as follows

  $$AX_{new} \geq B_2 - AX,$$

  where B_2 is the $n \times 1$ vector of twos, $X \in \{0, 1, 2\}^n$ shows the existing measurement points and A is the sum of the $n \times n$ identity matrix and the adjacency matrix of the network

  $$A_{st} = \begin{cases} 1 & \text{if buses } s \text{ and } t \text{ are connected,} \\ 1 & \text{if } s = t, \\ 0 & \text{otherwise.} \end{cases}$$

  Note that in the presence of zero-injection buses, i.e. buses that are not connected to a load or a generator, the notion of observability can be slightly modified. Indeed, the observability of a set of buses composed of a zero-injection bus and its adjacent buses, is maintained if at most one of the buses is not observed. This property of the zero-injection buses can be used to enforce the observability of the grid with

a decreased amount of PMUs. Other properties of zero-injection buses can also be leveraged to reduce the total number of required PMUs, other works [91] aim to model them as accurately as possible in order to take full advantage of their observability properties.

- the exact structural security requirement is that every pair of sites has at least one non-critical pair of measurements. This can be written as follows for every pair of sites (S_i, S_k)

$$AX_{new} + B_1(1 - Y_{a,b,i,k}) \geq B_1 - AX + AC_{a,b} \; \forall (a,b) \in S_i \cup S_k,$$
$$\sum_{(a,b)\in S_i \cup S_k} Y_{a,b,i,k} \geq 1,$$

where B_1 is the $n \times 1$ vector of ones and $C_{a,b}$ is a $n \times 1$ binary vector defined by

$$(C_{ab})_{s,1} = \begin{cases} 1 & \text{if } s = a, \\ 1 & \text{if } s = b, \\ 0 & \text{otherwise.} \end{cases}$$

The second inequality enforces that the first inequality is active for at least one pair of measurements of the pair of sites. The first inequality is active if decision variable $Y_{a,b,i,k}$ is equal to 1, in which case it enforces that the new set of measurement points without the pair of measurements (a, b), is such that every bus is observed by at least one measurement. In other words, it enforces that measurement pair (a, b) is not critical. In contrast, if $Y_{a,b,i,k} = 0$, then the first inequality is always true as it enforces that the new set of measurement points without the pair of measurements (a, b), is such that every bus is observed by at least zero measurements.

Known algorithms such as the branch-and-bound algorithm can be used to solve the ILP exactly. However, such algorithms are known to not scale well to large networks. Other strategies such as the extension of prior works on the optimal PMU placement problem [92, 93, 94], could be investigated further in order to find solutions that enforce our security requirement at a minimal cost. Alternatively, a solution can be found heuristically for both the exact and the practical structural security requirements. Algorithm 4, described in Section 4.5.1.2, implements a greedy strategy that adds measurement points to an observable grid that is not structurally secured.

In Section 4.6, Algorithm 4 applied for exact structural security, to a PMU allocation on the IEEE-39 bus benchmark, outputs the same solution as the matlab implementation of the branch-and-bound algorithm applied to the ILP formulation of the problem. In other words, the solution provided by Algorithm 4 is optimal in this specific scenario.

4.5.1.2 Enforcing Practical Structural Security

In an observable system, if a pair of sites is such that no measurement is critical and all pairs are critical, by adding one phasor measurement in one of the two sites, at least one pair of measurements will be non-critical. As a result, the ERR of $F^{[S_1,S_2]}$ will no longer be equal to 1. However, the practical security requirement is stronger and requires that the ERR is not close to 1. We do not know how to express such a requirement as a function of the level of observability of the network nodes, as is required for an ILP formulation of the problem. In contrast, in order to secure an observable but vulnerable grid, we propose to identify vulnerable pairs of sites such that the ERR of $F^{[S_1,S_2]}$ is close to 1 and to iteratively increase the number of phasors that they should measure until no critical pair of sites has a high $\text{ERR}(F^{[S_1,S_2]})$ value. A greedy strategy is to increase the number of measured phasors at sites that appear most frequently in the list of vulnerable pairs of sites. Algorithm 4 implements this greedy strategy by recursively building the secured set of measurement points. It takes as input the set of measurement points of the observable grid and it outputs a larger set of measurement points which includes the input set and the additional phasors required for structural security. Note that this algorithm only secures against structural vulnerabilities, it can thus be performed offline at each change of topology.

Recall the rank-1 TSA targeting $q = 5$ structurally vulnerable sites presented in Section 3.4.2. In Section 4.6, we secure the grid using Algorithm 4 and show that the attack is no longer feasible.

4.5.2 Monitoring the General Vulnerability Metrics

Once the grid is secured against structural vulnerabilities, the measurements of sites or of pairs of sites might still satisfy the general vulnerability conditions identified by item (b) of Theorem 4.4 and by item (a) of Theorem 4.6. As discussed in Sections 4.2.3 and 4.3.3, we conjecture that such conditions are unlikely to be satisfied in reality. By precaution, we propose to compute, at every estimation of the system's state, $\|R^{S_1} z^{S_1}\|$ for every site and $1 - \text{ERR}(\left[R^{S_1} z^{S_1} | R^{S_2} z^{S_2}\right])$ for every pair of sites. If over time it can be observed that a site or a pair of sites is frequently close to vulnerability, then the measurements satisfy either exactly or approximately a dangerous relation. In this case, we recommend breaking this relation by modifying the PMU allocation around the corresponding site or pair of sites.

4.6 Simulations

We validate our results with simulations on the IEEE 39 bus benchmark with real load profiles taken from the Lausanne grid at 50 Hz. Specifically, we show that the

Algorithm 4 Secure-Grid(M)

Input: M (set of measurement points for an observable grid)
Output: M_{new} (updated set of measurement points for an observable and structurally non-vulnerable grid)

$H \leftarrow$ Create topology matrix from admittance and M
$Vulnerable \leftarrow \emptyset$
$F \leftarrow H(H^\dagger H)^{-1}H^\dagger - Id$
for site i in grid **do** // We find pairs of sites that are vulnerable.
  $S_i = M(i)$
  for site $j \neq i$ in grid **do**
    $S_j = M(j)$
    $\Lambda \leftarrow$ Singular-Values($F^{[S_i,S_j]}$)
    if $\frac{\max(\Lambda)}{\sum\Lambda} \geq \eta$ **then** // If the ERR is larger than the predefined threshold, then the pair is considered as vulnerable. For exact structural security, set $\eta = 1$.
      $Vulnerable \leftarrow Vulnerable \cup (i, j)$
    end if
  end for
end for// Now $Vulnerable$ contains all vulnerable pairs of sites.
if $Vulnerable \neq \emptyset$ **then**
  while $Vulnerable \neq \emptyset$ **do**
    $freq \leftarrow$ Get the most frequent index in $Vulnerable$ // We find the site that is the most frequently found in vulnerable pairs.
    $M \leftarrow M \cup freq$ // We add a measurement point to it.
    for $tuple \in Vulnerable$ **do**
      if $freq \in tuple$ **then**
        $Vulnerable \leftarrow$ Remove $tuple$ from $Vulnerable$ // We remove all pairs containing this site from the list of vulnerable pairs.
      end if
    end for
  end while
  $M \leftarrow$ Secure-Grid(M)
end if// There are no more vulnerable pairs.
$M_{new} \leftarrow M$
return M_{new}

observable grid used in the previous chapters, is structurally vulnerable to rank-1 TSAs. After applying Secure-Grid, we show that all of the attacks presented previously in this thesis, are no longer feasible. Then, for each site and each pair of sites, we compute the vulnerability metrics and show that the structurally *non*-vulnerable grid is always far from vulnerable. In other words, we show that applying the security requirements established in the previous section, achieves the desired security.

4.6.1 Electrical Model

The PMU allocation we consider is depicted in Figure 4.1. Specifically, there are 12 zero-injection buses, PMUs measuring both voltages and currents are placed at

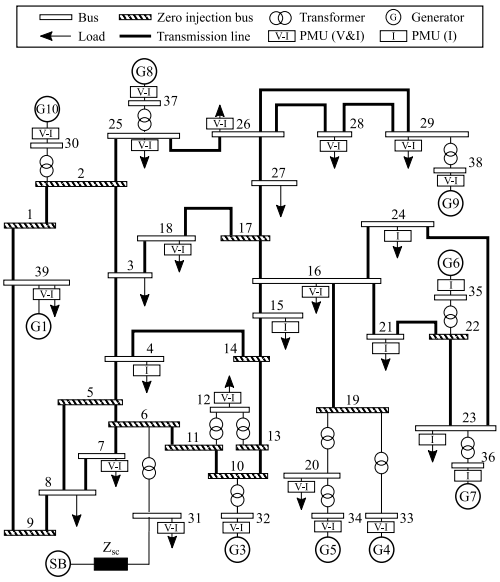


Figure 4.1 – PMU allocation on the IEEE-39 bus benchmark. Buses separated by a transformer are grouped in a site: PMUs in a site share the same time reference.

buses $\{30, 37, 28, 38, 18, 39, 12, 16, 7, 31, 32, 34, 33, 20, 25, 26, 29\}$ and PMUs measuring only currents are placed at buses $\{24, 35, 15, 21, 4, 23, 36\}$. We define sites to be groups of buses separated by transformers only: $\{2, 30\}$, $\{6, 31\}$, $\{10, 32\}$, $\{11, 12, 13\}$, $\{19, 20, 33, 34\}$, $\{22, 35\}$,$\{23, 36\}$, $\{25, 37\}$, $\{29, 38\}$, the other buses correspond to one-bus sites. With this allocation, no measurement is critical by itself, no PMU is critical by itself but some sites are critical.

Our simulations are done every 20ms over 700s, thus at $35'000$ different time instants. At each time instant, we create a measurement vector by computing the load flow. This results in the true state of the system. We then add randomly generated Gaussian noise to the true state, which results in the simulated measurement vector z. At 300s, we further introduce an inrush in the form of a sudden increase of factor 2 in the active power, at a nearby bus in order to observe the effect of sudden changes in the system state on the attack detectability.

4.6.2 Securing a Grid against Structural Vulnerabilities

The grid described above features pairs of sites such that all pairs of measurements are critical, i.e. such that $\text{rank}\,(F^{[S_1,S_2]}) = 1$. In other words, they are structurally vulnerable to rank-1 TSAs. It is the case for all pairs of sites among $\{21\}$, $\{22, 35\}$, $\{23, 36\}$ and $\{24\}$, which means that a rank-1 TSA can be mounted on any combination of at least two of these sites. In fact, Section 3.4.2 presents rank-1 TSAs on various combinations of the involved buses. Therefore, the grid needs to be secured from this structural vulnerability. Notice that the sites of vulnerable pairs are located close to each other, this means that

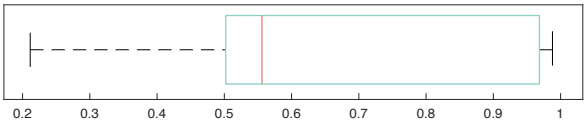


Figure 4.2 – Distribution of ERR values of $F^{[S_1,S_2]}$ matrices for all pairs of sites before applying Secure-Grid: some values are close to 1, the corresponding pairs of sites are structurally vulnerable in practice.

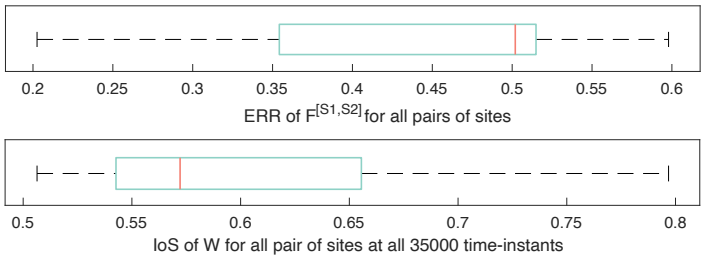


Figure 4.3 – Distribution of ERR values of $F^{[S_1,S_2]}$ and IoS values of W matrices for all pairs of sites after applying Secure-Grid: none are close to 1, no pairs of sites are structurally vulnerable in practice.

the corresponding region of the grid lacks measurement redundancy.

Secure-Grid applied on this grid for exact structural security, i.e. with $\eta = 1$, outputs the same result as the matlab implementation of the branch-and-bound algorithm. Specifically, they both require to add a measurement point on bus 35 exclusively. Therefore, our heuristic is in fact optimal in this scenario.

In order to gain insights on how vulnerable the other pairs of sites are in practice, we compute the ERR of their corresponding matrix $F^{[S_1,S_2]}$. The results given in Figure 4.2 reflect how close to 1 the ERR of $F^{[S_1,S_2]}$ matrices are. We observe that the median of ERR values is at 0.556 and that all values are between 0.2114 and 0.9939. The closer the ERR of $F^{[S_1,S_2]}$ is to 1, the more vulnerable the pair of sites is. Clearly, we can observe that several pairs of sites are vulnerable to rank-1 TSAs in practice, even though they have a non-critical pair of measurements.

Secure-Grid applied to this grid with $\eta = 0.8$, outputs a set of measurement points that includes additional phasor measurements. Specifically, buses 4, 15, 21, 24, 35 and 36 are required to measure an additional phasor. We observe on Figure 4.3 that there are no longer any pair of sites such that the ERR of $F^{[S_1,S_2]}$ is close to 1. Also on Figure 4.3, we observe that the IoS of the attack-angle matrix W, at all $35'000$ time instants is always much smaller than 1; i.e., there are no grid locations that are structurally vulnerable to rank-1 TSAs.

An attack on a set of three buses with different time references, which involves a total of 5 measurements, was also presented in Section 3.4.3. This attack is performed on a grid with a slightly different PMU allocation. Specifically, only one phasor is measured

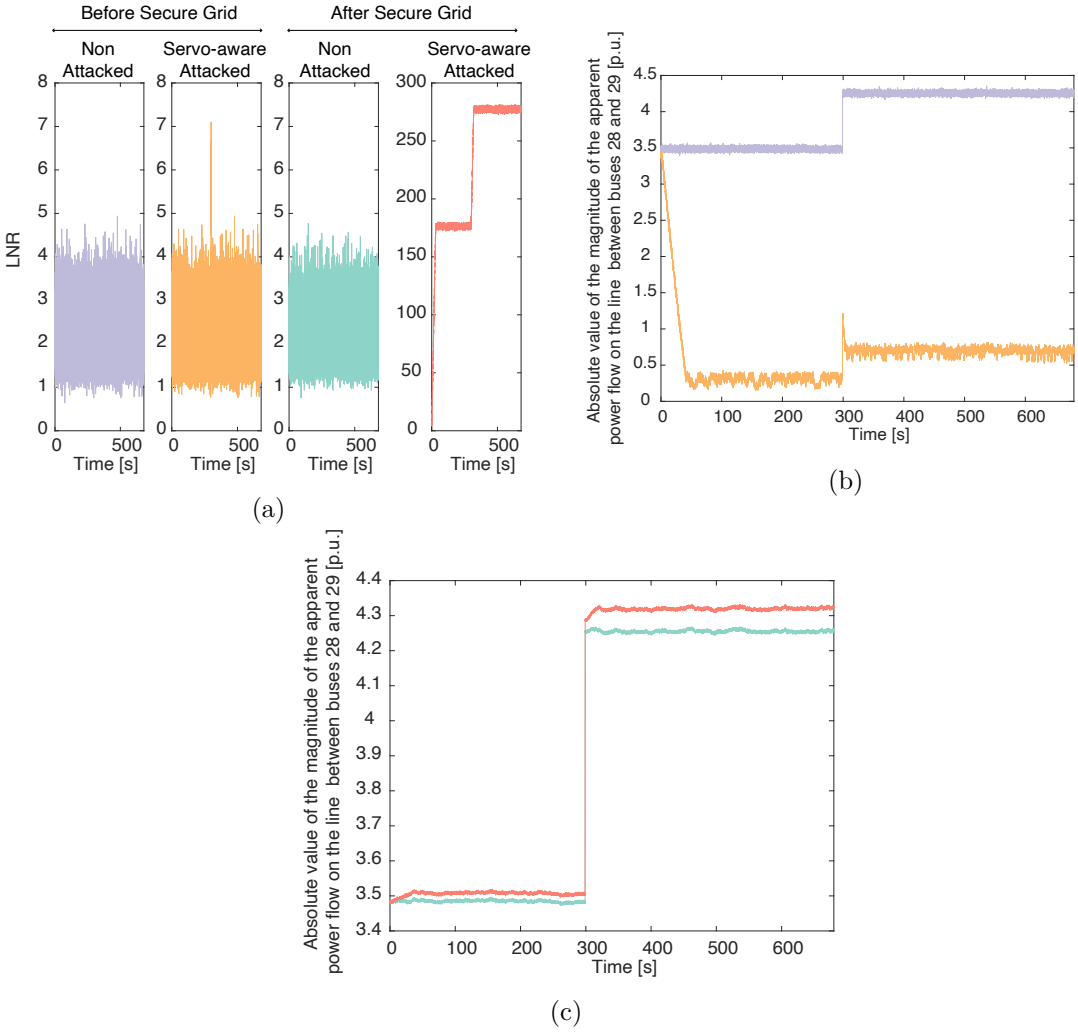


Figure 4.4 – Comparison of the impact and detectability of the servo-aware attack targeting buses 26, 28 and 38 before and after using Secure-Grid: (a) LNR values show that the servo-aware attack is undetectable before Secure-Grid and detectable after it; (b) The magnitude of the power-flow with and without the attack *before* Secure-Grid, the undetectable attack has a large impact; (c) The magnitude of the power-flow with and without the attack *after* Secure-Grid, the undetectable attack has an impact but is very detectable.

at bus 26 and no phasors are measured at bus 29. In this setting, the presented attack targeted buses 26, 28 and 38. Notice that buses 28 and 38 both measure two phasors simultaneously. This successful attack was found by trials and errors but was not explained by the theory of Chapter 3. We now understand that this set of buses was in fact structurally vulnerable because the rank of the corresponding sub matrix of F is equal to 1. We are now also able to secure the grid against this attack by adding phasors to measure at buses 26 and 29. Figure 4.4 compares the servo-aware attack impact and LNR values before using Secure-Grid (i.e. the results presented in Section 3.4.3) and after using Secure-Grid. It shows that the undetectable attack from Section 3.4.3 becomes clearly detectable once additional phasors are measured at the buses identified

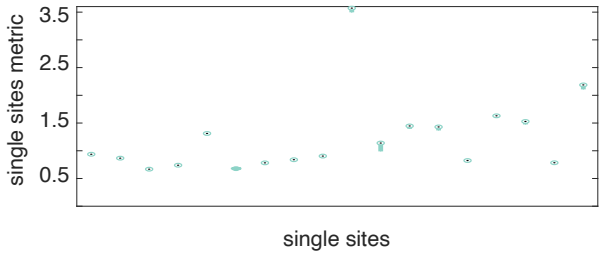


Figure 4.5 – After Secure-Grid: Distribution of the values of $\|R^{S_1} z^{S_1}\|$ for each site at 35′000 time instants. It is never equal to 0: the condition for a rank-1 TSA on a single site is never satisfied.

by Secure-Grid. The change of values in the middle of the simulation is due to a sudden increase of factor 2 in the active power introduced at a nearby bus. All attacks presented in [14] and Chapter 3 targeted sets of PMUs which were in fact structurally vulnerable to rank-1 TSAs. Our new security requirement therefore prevents all of them.

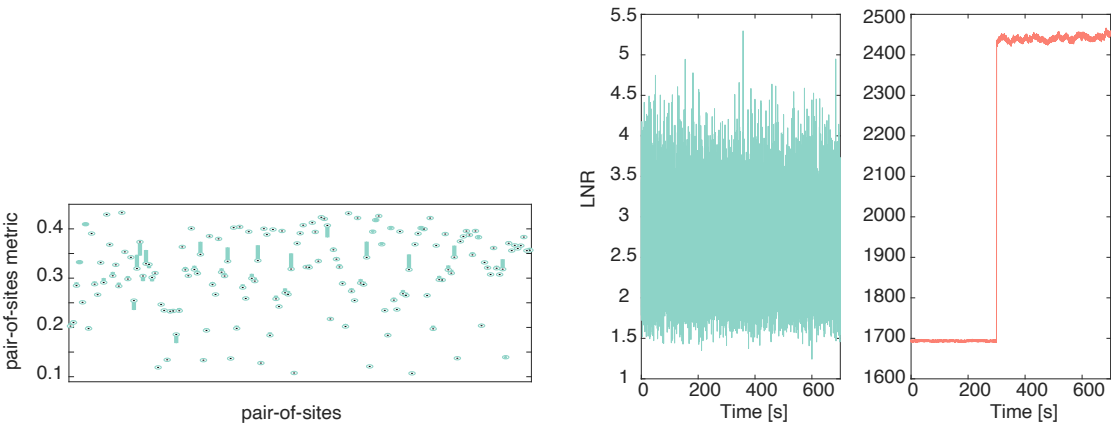
Even though the resulting grid is not structurally vulnerable, it is still possible that measurement values satisfy the general vulnerability conditions for a site or a pair of sites. We show that such specific conditions are far from appearing on our realistic grid.

4.6.3 General Vulnerability Condition for Single Sites

At each of the 35′000 time instants of the simulation, we compute the metric $\|R^{S_1} z^{S_1}\|$ introduced in Section 4.2.2. Recall that this metric is equal to 0 if and only if the site is vulnerable. The distribution of the obtained values of all sites are shown in Figure 4.5. We observe that the metric is never equal to 0, the most vulnerable site has a metric that is on average equal to 0.664. To illustrate that 0.664 reflects that the site is far from vulnerable in practice, we perform an attack on the corresponding site $\{31\}$ with a constant offset of $20\mu s$, which is the maximum offset allowed by the PMU clock controllers in Chapter 3. Recall that a vulnerable single site can be attacked undetectably with any attack-angle, including $20\mu s$. The obtained largest normalized residuals are approximately 6 times larger than those obtained without an attack. Such a difference is easily identified. In other words, the site that is the closest to satisfying the single-site vulnerability condition is far from vulnerable in practice. As a result, we observe that the measurement values of all sites of the grid are always far from satisfying the conditions necessary to mount a rank-1 TSA.

4.6.4 General Vulnerability Condition for Pairs of Sites

At each of the 35′000 time instants of the simulation we compute the vulnerability metric introduced in Section 4.3.2 for each pair of sites. Specifically, for each pair of sites we compute $1 - \text{ERR}(\left[R^{S_1} z^{S_1} | R^{S_2} z^{S_2}\right])$. Recall that the closer the metric is to 0, the more vulnerable the pair of sites is. The distribution of the obtained values for every pair



(a) (b)

Figure 4.6 – After Secure-Grid: (a) Distribution of the values of $1 - \mathrm{ERR}(\left[R^{S_1} z^{S_1} | R^{S_2} z^{S_2}\right])$ for each pair of sites at $35'000$ time instants. The most vulnerable pair has a mean metric value of 0.1561. (b) LNR values without an attack (left) and with an attack on the most vulnerable pair of sites (right); the values are clearly distinguishable, the attack is detected.

of sites at all time instants are given in Figure 4.6a. We observe that the minimum value is 0.1561, which is not close enough to 0 for undetectable attacks as it is shown on Figure 4.6b. The latter shows that an attack on the corresponding pair of sites (i.e. the one with the lowest vulnerability metric values) is detectable. All the other pairs have vulnerability metric values that are even further from satisfying the vulnerability conditions. As a result, we observe that at each time instant of the simulation, the measurement values of all pairs of sites of the secured grid are far from satisfying the conditions necessary to mount rank-1 TSAs.

Feasibility of Eq. (4.5) and of Eq. (4.6). Recall the special case considered in Section 4.3.3, where two sites that are not structurally vulnerable, measure two synchrophasors each. Then, a rank-1 TSA is feasible if and only if either Eq. (4.5) (if $\mathrm{rank}\,(F^{[S_1,S_2]}) = 3$) or Eq. (4.6) (if $\mathrm{rank}\,(F^{[S_1,S_2]}) = 2$) is satisfied.
On the secured grid, consider the two pairs of buses $(12, 32)$ and $(16, 20)$. Notice that bus 20 is in fact in a site with bus 34 but for this simulation we suppose that it forms a single-bus site. In both cases, the rank of the corresponding submatrix of F is equal to 3: $\mathrm{rank}\,(F^{[S_{12},S_{32}]}) = 3$ and $\mathrm{rank}\,(F^{[S_{16},S_{20}]}) = 3$. Therefore, a rank-1 TSA targeting these pairs is feasible if and only if Eq. (4.5) is satisfied. At the $35'000$ time instants t of our simulation, we compute the ratios of the left-hand-sides over the right-hand-sides of Eq. (4.5)

$$\frac{z_1^t / z_2^t}{n_1/n_2} \text{ and } \frac{z_3^t / z_4^t}{n_3/n_4} \tag{4.8}$$

and plot them in the complex plane. Note that both ratios must be equal to 1 for an attack to be feasible on the pair of buses. The obtained ratios are shown on Figures 4.7

and 4.8 for $(12, 32)$ and $(16, 20)$, respectively. Observe that in both cases they are far from 1. In other words, the conditions necessary to mount a rank-1 TSA are never satisfied. Observe that the values of the ratios for the pair $(12, 32)$ are grouped around different central values: this is due to two different sudden increases in the magnitude of the reactive power at a nearby bus during the simulation.

In order to investigate the feasibility of satisfying Eq. (4.6), we consider the pair of

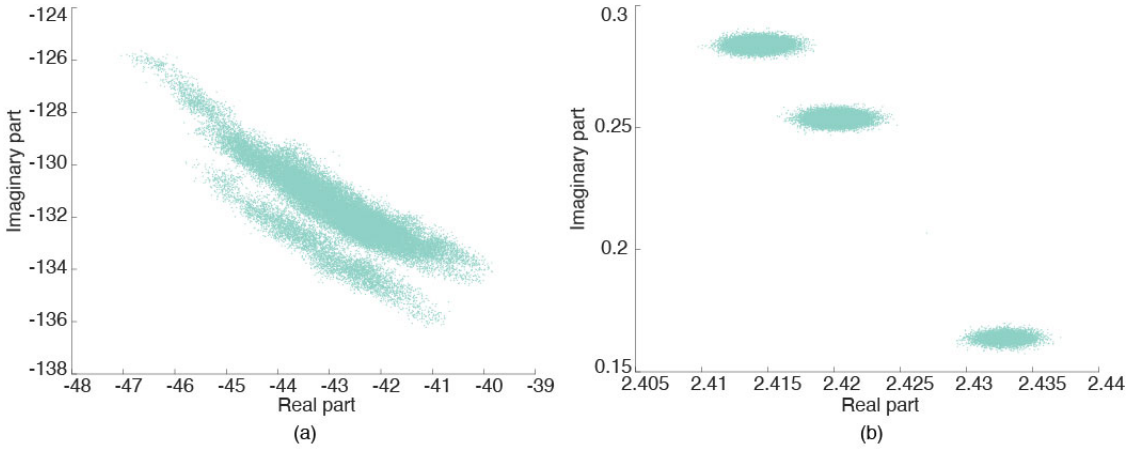


Figure 4.7 – After Secure-Grid: ratios (4.8) at bus 12 (a) and bus 32 (b) are always far from 1 at all $35'000$ time instants. Conditions for a rank-1 TSA are never satisfied.

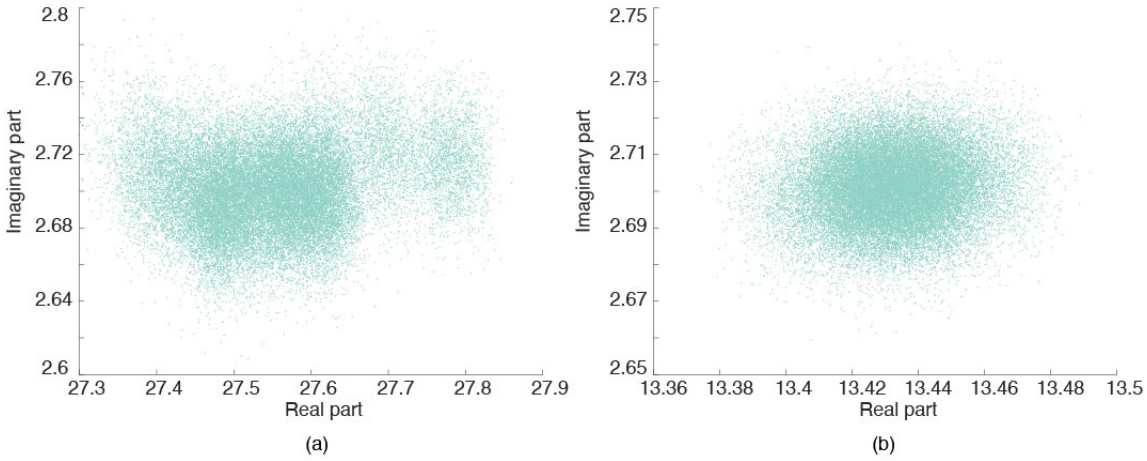


Figure 4.8 – After Secure-Grid: ratios (4.8) at bus 16 (a) and bus 20 (b) are always far from 1 at all $35'000$ time instants. Conditions for a rank-1 TSA are never satisfied.

buses $(20, 34)$ and suppose that they do not share the same time reference. Notice that they are separated only by a converter and thus were considered to be on the same site but for this simulation we suppose that they form two distinct sites. In this case, $\operatorname{rank}\left(F^{[S_{20}, S_{34}]}\right) = 2$, therefore a rank-1 TSA against this pair is feasible if and only if Eq. (4.6) is satisfied. As previously, we compute the ratio of the left-hand-side over the

right-hand-side of Eq. (4.6) at all 35′000 time instants t of our simulation

$$\frac{z_3^t/z_4^t}{(z_1^t(v_3n_2 - v_2n_3) + z_2^t(v_1n_3 - v_3n1))/(z_1^t(v_4n_2 - v_2n_4) + z_2^t(v_1n_4 - v_4n1))} \tag{4.9}$$

and plot it in the complex plane. As above, the ratio must be equal to 1 for an attack to be feasible on the pair of buses. Observe on Figure 4.9 that once again the ratios are far from 1. In other words, the conditions necessary to mount a rank-1 TSA are never satisfied.

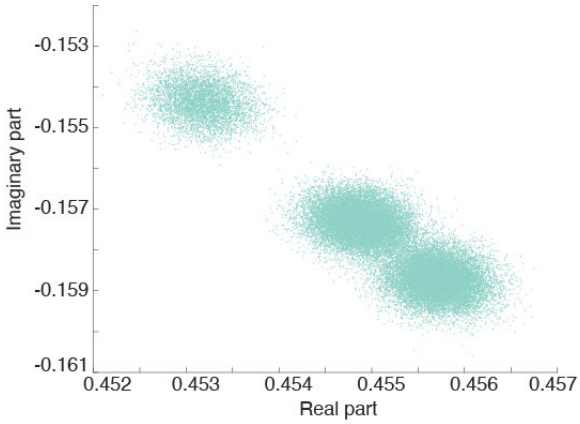


Figure 4.9 – After Secure-Grid: ratios (4.9) for pair $(20, 34)$ are always far from 1 at all 35′000 time instants. Conditions for a rank-1 TSA are never satisfied.

Note that we did not show in this chapter, how to compute a metric threshold value below which the metric shows vulnerability. We only showed that the smallest metric values obtained on our structurally *non*-vulnerable grid are not small enough for vulnerability. In order to establish a threshold, one would need to run tests on a vulnerable grid in order to find the threshold metric value from a predefined threshold LNR value. This is done in Section 5.3.2 of the next chapter.

4.7 Conclusion

We showed in this chapter that the analysis of the vulnerability of a grid to rank-1 TSAs reduces to the vulnerability analysis for every site and every pair of sites. We identified a sufficient vulnerability condition for pairs of sites that measure an arbitrary number of phasors. This condition does not depend on the measurement values. We established a security requirement to prevent this vulnerability. We also provided an offline greedy algorithm that enforces our security requirement. A limitation of our security requirement is that even if satisfied, it is still possible that measurement values are such that attacks are feasible, although we conjecture that it is unlikely. We identified sufficient and necessary vulnerability conditions for single sites and for pairs of sites,

each measuring an arbitrary number of phasors. We recommend the monitoring of two metrics associated to these conditions in order to check the non-vulnerability of the grid. Numerical results, on the IEEE-39 bus benchmark with real load profiles from the Lausanne grid, show that the measurements of a grid satisfying our security requirement are far from vulnerable to rank-1 TSAs. Finally, our results, applied to sites measuring a single phasor, coincide with the results of Chapter 3. The new theory established in this chapter, enables us to better understand and thus prevent attacks presented in Chapter 3.

5

Time-Synchronization Attack Detection in Unbalanced Three-Phase Systems

The theory of Chapter 4 is applicable to both single-phase systems and three-phase systems. On a three-phase system, although measurements are readily available in the three-phase model, it is customary to perform the state estimation and the bad data detection (BDD) in the direct-sequence model. This is due to the fact that it reduces the problem from three dimensions to one dimension, and thus decreases the computation complexity. In this chapter, we analyse the security benefits with respect to TSAs, of using the three-phase model instead of the direct-sequence model for state estimation and for BDD.

We make two important contributions. First, we provide an analytical characterization of the vulnerability of three-phase state estimation compared to state estimation based on a direct-sequence model. We show that in a balanced three-phase system the vulnerability conditions in the three-phase and in the direct-sequence representations are equivalent. However, if the system is unbalanced, we show that vulnerability in the direct-sequence representation does not imply vulnerability in the three-phase representation. These results indicate that three-phase state estimation is more resilient to time-synchronization attacks (TSAs). Second, we provide empirical evidence that confirms the superiority of three-phase state estimation in detecting TSAs using extensive simulations on the IEEE 39-bus benchmark using real load profiles measured by PMUs installed on the 125 KV sub-transmission power grid of the city of Lausanne [95, 96].

State of the art. As mentionned above, state estimation can be performed either using the full three-phase model, or using the simpler direct-sequence model of the power system [17]. However, due to unbalanced loads and untransposed transmission lines, a direct-sequence equivalent of the three-phase system will often be inaccurate, especially in distribution grids [97]. Therefore, a large body of research work has focused on developing accurate and efficient three-phase state estimators. Authors in [98] demonstrated the accuracy gain achieved through considering the three-phase model of a power system,

using a hybrid state-estimator that utilizes both SCADA measurements and PMU measurements. Authors in [97] evaluated the robustness of three-phase state estimators to different sources of uncertainty in distribution grids. More recent works focus on creating a three-phase state estimator that relies only on PMU measurements. Authors in [99] consider the problem of finding an optimal placement of PMUs in a power system to achieve full observability in three-phase distribution grids. Their solution aims at minimizing the number of deployed PMUs, while maximizing state estimation accuracy. Authors in [100] focus on the computational efficiency of three-phase state estimators, by using modal transformation to leverage the linearity of the state estimation process when only PMUs are used. Lately, motivated by advances in machine learning, [101] proposes using artificial neural networks for three-phase state estimation on sparse PMU measurements, instead of the traditional WLS estimation, and finds that the estimator achieves high estimation accuracy while meeting real-time requirements. Yet, the majority of power-system operators today perform state estimation based on the direct-sequence model despite the advantages of three-phase state estimation, due to the lack of reliable three-phase grid component models, and due to the higher computational burden.

The rest of the chapter is organized as follows. Section 5.1 presents the models considered for three-phase state estimation and for TSAs. Section 5.2 shows analytically that when the system has unbalances, the three-phase state estimator is more resilient to TSAs than the direct-sequence state estimator. Section 5.3 presents extensive simulations confirming our analytical results. Finally, Section 5.4 concludes the chapter.

5.1 System and Attack Models

We adapt the system model from Chapter 4 to three-phase systems, both in complex form and in rectangular coordinates. As explained in Section 5.2, the former is used for the analysis of exact vulnerability conditions and the latter is used to measure the distance to vulnerability of sites that are not exactly vulnerable. We then introduce the attack model.

5.1.1 System Model in Complex Form

We consider a three-phase system with a total of n buses and $3m$ phasor measurements measured by PMUs. The complex measurement vector $z_{abc} \in \mathbb{C}^{3m}$ is linearly linked to the state vector $x_{abc} \in \mathbb{C}^{3n}$ via the complex measurement-to-state matrix $H_{abc} \in \mathbb{C}^{3m \times 3n}$ as

$$z_{abc} = H_{abc} x_{abc} + e_{abc}, \tag{5.1}$$

where $e_{abc} \in \mathbb{C}^{3m}$ is the complex measurement error. We define the LS $3m \times 3m$ complex verification matrix $F_{abc} = H_{abc}((H_{abc})^{\dagger} H_{abc})^{-1}(H_{abc})^{\dagger} - Id$, where $\dagger$ denotes

the conjugate transpose and Id is the $3m \times 3m$ identity matrix. Then, we can compute the LS residuals in the three-phase model as $F_{abc}z_{abc}$. The three-phase measurement vector can be transformed into its complete-sequence model counterpart as $z_{012} = T_Z z_{abc}$, where T_Z is the $3m \times 3m$ block diagonal matrix that has along its diagonal the 3×3 sequence transformation matrix

$$T = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & e^{\frac{2j\pi}{3}} & e^{\frac{-2j\pi}{3}} \\ 1 & e^{\frac{-2j\pi}{3}} & e^{\frac{2j\pi}{3}} \end{bmatrix}.$$

Similarly, the three-phase state vector can be transformed into its complete-sequence model counterpart as $x_{012} = T_X^{-1}x_{abc}$, where T_X is the $3n \times 3n$ block diagonal matrix that has along its diagonal the 3×3 inverse sequence transformation matrix $T^{-1} = 3T^{\dagger}$, note that T is non-singular and therefore invertible. The measurement-to-state matrix in the complete-sequence model is obtained as $H_{012} = T_Z H_{abc} T_X$ and Eq.(5.1) in the complete-sequence model becomes

$$z_{012} = H_{012}x_{012} + e_{012},$$

where $e_{012} = T_X^{-1}e_{abc}$ is the complete-sequence measurement error. The verification matrix in the complete-sequence model is $F_{012} = H_{012}((H_{012})^{\dagger}H_{012})^{-1}(H_{012})^{\dagger} - Id$. The following lemma shows the relation between the verification matrix in the three-phase model and in the complete-sequence model.

Lemma 5.1. $F_{012} = T_Z F_{abc} T_Z^{-1}$.

Proof. We use the fact that $3T_Z^{\dagger} = T_Z^{-1}$.

$$\begin{aligned}
F_{012} &= H_{012}(H_{012}^{\dagger}H_{012})^{-1}H_{012}^{\dagger} - Id \\
&= T_Z H_{abc} T_X (T_X^{\dagger} H_{abc}^{\dagger} T_Z^{\dagger} T_Z H_{abc} T_X)^{-1} T_X^{\dagger} H_{abc}^{\dagger} T_Z^{\dagger} - Id \\
&= 3T_Z H_{abc} T_X (T_X^{\dagger} H_{abc}^{\dagger} H_{abc} T_X)^{-1} T_X^{\dagger} H_{abc}^{\dagger} T_Z^{\dagger} - Id \\
&= 3T_Z H_{abc} T_X T_X^{-1} (H_{abc}^{\dagger} H_{abc})^{-1} (T_X^{\dagger})^{-1} T_X^{\dagger} H_{abc}^{\dagger} T_Z^{\dagger} - Id \\
&= 3T_Z H_{abc} (H_{abc}^{\dagger} H_{abc})^{-1} H_{abc}^{\dagger} T_Z^{\dagger} - Id \\
&= T_Z H_{abc} (H_{abc}^{\dagger} H_{abc})^{-1} H_{abc}^{\dagger} T_Z^{-1} - T_Z T_Z^{-1} \\
&= T_Z F_{abc} T_Z^{-1}.
\end{aligned}$$

□

As mentioned previously, the standard practice only uses the direct-sequence component for state estimation and BDD. Let D be the $m \times 3m$ matrix that selects only the direct-sequence elements from the complete-sequence model. Specifically, D is a block

diagonal matrix with $\begin{bmatrix}0 & 1 & 0\end{bmatrix}$ along its diagonal. Then the direct-sequence residuals are $F_1 z_1 = DF_{012} z_{012}$.

5.1.2 System Model in Rectangular Coordinates

The system model in complex form is a theoretical model that enables to identify closed-form vulnerability conditions. However, in practice the system model that is used in measurement systems is in rectangular coordinates. This is due to the fact that the use of the WLS estimator as state estimator is widespread and that it is linear over $\mathbb{R}$ and not over $\mathbb{C}$, because the covariance of the error is not linear over $\mathbb{C}$.

We now present an adaptation to three-phase systems of the system model in rectangular coordinates introduced in Chapter 4. The $6m \times 1$ real three-phase measurement vector and the $6m \times 1$ real complete-sequence measurement vector in rectangular coordinates are denoted by $z_{\square,abc} = \left(\mathrm{Re}(z_{abc}^{[1]}), \mathrm{Im}(z_{abc}^{[1]}), \cdots, \mathrm{Re}(z_{abc}^{[3m]}), \mathrm{Im}(z_{abc}^{[3m]})\right)^T$ and $z_{\square,012}$, respectively; $z_{abc}^{[1]}$ denotes the measurement of index 1 of measurement vector z_{abc}. The measurement-to-state equation becomes $z_{\square,abc} = H_{\square,abc} x_{\square,abc} + e_{\square,abc}$, where $H_{\square,abc} \in \mathbb{R}^{6m \times 6n}$ is the three-phase measurement-to-state matrix in rectangular coordinates, $x_{\square,abc} \in \mathbb{R}^{6n}$ is the three-phase state vector in rectangular coordinates and $e_{\square,abc} \in \mathbb{R}^{6m}$ is the three-phase measurement error in rectangular coordinates. The $6m \times 6m$ real WLS verification matrix is $G_{\square,abc} = H_{\square,abc}(H_{\square,abc}^T C_{\square,abc}^{-1} H_{\square,abc})^{-1} H_{\square,abc}^T C_{\square,abc}^{-1} - Id$, where $C_{\square,abc}$ is the covariance matrix of the three-phase measurement error in rectangular coordinates. Therefore, the three-phase WLS residuals in rectangular coordinates are computed as $G_{\square,abc} z_{\square,abc}$.

The three-phase measurement, state and error vectors and the three-phase measurement-to-state and verification matrices in rectangular coordinates can be transformed into their complete-sequence model counterparts using the block-diagonal transformation matrices $T_{\square,Z} \in \mathbb{R}^{6m \times 6m}$ and $T_{\square,X} \in \mathbb{R}^{6n \times 6n}$. $T_{\square,Z}$ has the following 6×6 matrix along its diagonal

$$\frac{1}{3}\begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & \cos(\frac{2\pi}{3}) & -\sin(\frac{2\pi}{3}) & \cos(-\frac{2\pi}{3}) & -\sin(-\frac{2\pi}{3}) \\
0 & 1 & \sin(\frac{2\pi}{3}) & \cos(\frac{2\pi}{3}) & \sin(-\frac{2\pi}{3}) & \cos(-\frac{2\pi}{3}) \\
1 & 0 & \cos(-\frac{2\pi}{3}) & -\sin(-\frac{2\pi}{3}) & \cos(\frac{2\pi}{3}) & -\sin(\frac{2\pi}{3}) \\
0 & 1 & \sin(-\frac{2\pi}{3}) & \cos(-\frac{2\pi}{3}) & \sin(\frac{2\pi}{3}) & \cos(\frac{2\pi}{3})
\end{bmatrix}$$

and $T_{\square,X}$ has its inverse along its diagonal.

The relation between the matrices in the three-phase and complete-sequence models in rectangular coordinates is $H_{\square,012} = T_{\square,Z} H_{\square,abc} T_{\square,X}$ and $C_{\square,012} = T_{\square,Z} C_{\square,abc} T_{\square,Z}^{-1}$. The following lemma shows the relation between the WLS verification matrices in the three-phase and complete-sequence models.

Lemma 5.2. $G_{\square,012} = T_{\square,Z} G_{\square,abc} T_{\square,Z}^{-1}$.

Proof. The proof is very similar to the proof of lemma 5.1. By replacing

- $T_{\square,Z}^T = 3T_{\square,Z}^{-1}$,
- $T_{\square,X}^T = 3T_{\square,X}^{-1}$
- and $C_{\square,012}^{-1} = 3T_{\square,Z} C_{\square,abc}^{-1} T_{\square,Z}^T$

in $G_{\square,012} = H_{\square,012}(H_{\square,012}^T C_{\square,012}^{-1} H_{\square,012})^{-1} H_{\square,012}^T C_{\square,012}^{-1} - Id$, we obtain Lemma 5.2. □

As for the complex system model, we denote the direct-sequence verification matrix and measurement vector in rectangular coordinates by $G_{\square,1}$ and $z_{\square,1}$, respectively.

5.1.3 Attack Model

We consider an attacker that is able to manipulate the time reference of q sites, where each site is a group of buses where PMUs share the same time reference, as defined in Chapter 4. A manipulation of the time reference of a site by d seconds induces a phase shift $\delta = 2\pi f d$ rad in the phasors measured by the PMUs of the attacked site, where f is the voltage frequency. Let S_k be the set of phasor indices measured at site number k. If site k is attacked, then the attacked measurement vector at site k is $z_{abc}^{S_k\prime} = e^{j\delta} z_{abc}^{S_k}$: all the phases are shifted by δ and the magnitudes are unchanged. The values for other sites are unchanged:

$$(z'_{abc})_i = \begin{cases} e^{j\delta} z_{abc}^i & \text{if } i \in S_k \\ z_{abc}^i & \text{otherwise.} \end{cases}$$

Observe that the attacked measurement vector at site k in the complete-sequence model and in the direct-sequence model are subject to the same phase shift $z_{012}^{S_k\prime} = T_Z z_{abc}^{S_k\prime} = e^{j\delta} T_Z z_{abc}^{S_k} = e^{j\delta} z_{012}^{S_k}$ and $z_1^{S_k\prime} = D z_{012}^{S_k\prime} = e^{j\delta} D z_{012}^{S_k} = e^{j\delta} z_1^{S_k}$.

In order to identify and possibly remove anomalous measurements, state estimation in power grids is generally combined with BDD algorithms. As seen in Chapter 2, the most widely used techniques for BDD rely on the analysis of the measurement residuals obtained after state estimation, and are variants of either the largest normalized residual (LNR) test or the χ^2-test. The LNR-test checks the presence of unusually large residual values and the χ^2-test checks that the distribution of the sum of the residuals is plausible. Therefore, a TSA that does not modify the residuals is undetected by such BDD algorithms, which motivates the following definition.

Definition 5.1. *(Undetectable TSA) A TSA against a group of sites is undetectable under WLS if the residuals are not changed by the attack, i.e.,*

$$\begin{aligned} G_{\square,abc}\Delta z_{\square,abc} = 0, \text{ where } \Delta z_{\square,abc} = z'_{\square,abc} - z_{\square,abc}, \\ G_{\square,012}\Delta z_{\square,012} = 0, \text{ where } \Delta z_{\square,012} = z'_{\square,012} - z_{\square,012}, \\ G_{\square,1}\Delta z_{\square,1} = 0, \text{ where } \Delta z_{\square,1} = z'_{\square,1} - z_{\square,1}, \end{aligned}$$

depending on the chosen model.

Unfortunately, the computation of the WLS residuals is not linear over $\mathbb{C}$ but is linear over $\mathbb{R}$, whereas the computation of the LS residuals is linear over $\mathbb{C}$. The following lemma shows that despite this important difference, the vulnerability conditions are equivalent.

Lemma 5.3. *Consider the residuals under LS and under WLS, then*

$$\begin{aligned} G_{\square,abc}z_{\square,abc} = 0 &\iff F_{abc}z_{abc} = 0, \\ G_{\square,012}z_{\square,012} = 0 &\iff F_{012}z_{012} = 0, \\ G_{\square,1}z_{\square,1} = 0 &\iff F_1 z_1 = 0. \end{aligned}$$

Proof. The proof for the direct-sequence model is in the proof of Theorem 4.3. The proofs for the three-phase and complete-sequence models are analogous to that of the direct-sequence model. □

As a consequence, we can perform the vulnerability analysis of three-phase state estimation using the complex LS verification matrix.

5.2 Vulnerability Analysis of Three-Phase Systems

In this section, we first show that the vulnerability conditions of a grid in the three-phase and complete-sequence models are equivalent. We then show that in the presence of unbalances, the three-phase model of the system is less vulnerable to undetectable attacks than the direct-sequence model of the system, and thus a three-phase state estimator is more recommended for LSE. Finally, we consider the case of attacks that only slightly modify the residuals, thus potentially remaining undetected by the BDD mechanisms, and we show that if the system is balanced then WLS using the direct-sequence model is at least as vulnerable as WLS using the complete-sequence model.

5.2.1 Exact Vulnerability Analysis

We first show that the vulnerability conditions for a group of sites are equivalent in the two considered three-dimensional models, i.e., the three-phase and complete-sequence models. Second, we show that if a TSA is feasible on a three-dimensional model of the system, then necessarily the one-dimensional model (i.e. only the direct sequence) can also be attacked. These results rely on the following characterization of the vulnerability conditions in the three-phase and in the complete-sequence models, for a group of sites $S = \cup_k S_k$. As a shorthand, for an index set S let us denote by F^S a submatrix of F formed by the columns indexed by S.

Lemma 5.4. *Consider the set of measurement indices S, then*

$$F_{012}^S z_{012}^S = T_Z F_{abc}^S z_{abc}^S.$$

Proof. Because T_Z and T_X are block diagonal matrices with T and T^{-1} along their diagonal, respectively, the complete-sequence model verification matrix has the following structure

$$F_{012} = \begin{bmatrix} TF_{abc}^{[1:3,1:3]}T^{-1} & \cdots & TF_{abc}^{[1:3,3m-3:3m]}T^{-1} \\ \vdots & & \vdots \\ TF_{abc}^{[3m-3:3m,1:3]}T^{-1} & \cdots & TF_{abc}^{[3m-3:3m,3m-3:3m]}T^{-1} \end{bmatrix}.$$

Hence, $F_{012}^S = T_Z F_{abc}^S (T_Z^{-1})^S$. Similarly, we can write $z_{012}^S = T_Z^S z_{abc}^S$ and we obtain

$$F_{012}^S z_{012}^S = T_Z F_{abc}^S (T_Z^{-1})^S T_Z^S z_{abc}^S = T_Z F_{abc}^S z_{abc}^S.$$

□

Theorem 5.1. *A group of sites is vulnerable to undetectable TSAs in the three-phase model if and only if it is vulnerable to undetectable TSAs in the complete-sequence model.*

Proof. Theorem 4.10 shows that the analysis of the system vulnerability to TSAs reduces to the vulnerability analysis of every site and every pair of sites. We first show the equivalence for single sites. Theorem 4.4 states that a site measuring $p \geq 1$ phasors with indices in S_k, such that no measurement alone is critical and at least one measurement is not equal to 0, is vulnerable to undetectale TSAs if and only if $p \geq 2$ and z^{S_k} is in the null space of F^{S_k}. Hence, a site, with measurement indices in S_k, is vulnerable to TSAs in the three-phase model if and only if $F_{abc}^{S_k} z_{abc}^{S_k} = 0$, hence $T_Z F_{abc}^{S_k} z_{abc}^{S_k} = 0$ and thus by Lemma 5.4, $F_{012}^{S_k} z_{012}^{S_k} = 0$. Similarly, if the site is vulnerable to TSAs in the complete-sequence model, then $F_{012}^{S_k} z_{012}^{S_k} = 0$, hence $T_Z^{-1} F_{012}^{S_k} z_{012}^{S_k} = 0$ and thus by lemma 5.4, $F_{abc}^{S_k} z_{abc}^{S_k} = 0$.

We now show the equivalence for pairs of sites. Theorem 4.6 states that two sites measuring phasors with indices in S_i and S_k, respectively, such that $|S_i| + |S_k| = p$, no

measurement is critical by itself, neither site is vulnerable to TSAs by itself and at least one measurement in each site is not equal to zero, are vulnerable to TSAs if and only if $F^{S_i}z^{S_i}$ and $F^{S_k}z^{S_k}$ are colinear. Hence, a pair of sites, with measurement indices in S_i and S_k, is vulnerable to TSAs in the three-phase model if and only if $F_{abc}^{S_i}z_{abc}^{S_i}$ and $F_{abc}^{S_k}z_{abc}^{S_k}$ are colinear $\iff \exists l \in \mathbb{C}^*$ such that $F_{abc}^{S_i}z_{abc}^{S_i} + lF_{abc}^{S_k}z_{abc}^{S_k} = 0$. By Lemma 5.4, this is equivalent to $T_Z F_{012}^{S_i}z_{012}^{S_i} + lT_Z F_{012}^{S_k}z_{012}^{S_k} = T_Z(F_{012}^{S_i}z_{012}^{S_i} + lF_{012}^{S_k}z_{012}^{S_k}) = 0$. Because T_Z is invertible, this is equivalent to $F_{012}^{S_i}z_{012}^{S_i} + lF_{012}^{S_k}z_{012}^{S_k} = 0$, i.e., $F_{012}^{S_i}z_{012}^{S_i}$ and $F_{012}^{S_k}z_{012}^{S_k}$ are colinear. □

Theorem 5.2. *If a group of sites is vulnerable to undetectable TSAs in the three-phase or complete-sequence model, then its direct-sequence representation is also vulnerable to undetectable TSAs. The converse is not always true.*

Proof. As for Theorem 5.1, we use Theorem 4.10 which states that the analysis of the system vulnerability to TSAs reduces to the vulnerability analysis of every site and every pair of sites. We first show the relation for single sites. Recall that a site with measurement indices in S_k is vulnerable to TSAs in a three-dimensional model if and only if $F_{012}^{S_k}z_{012}^{S_k} = 0$. Hence, by definition, the direct-sequence residuals are $F_1^{S_k}z_1^{S_k} = DF_{012}^{S_k}z_{012}^{S_k} = 0$, where D is the matrix, defined in Section 5.1.1, that selects only the direct-sequence elements from the complete-sequence model measurement vector.

Similarly, we show the relation for pairs of sites. Recall that a pair of sites, with measurement indices in S_i and S_k, is vulnerable to TSAs if and only if $F_{012}^{S_i}z_{012}^{S_i}$ and $F_{012}^{S_k}z_{012}^{S_k}$ are colinear vectors, which is equivalent to stating that there exists an $l \in \mathbb{C}^*$ such that $F_{012}^{S_i\cup S_k}\begin{bmatrix} z_{012}^{S_i} \\ lz_{012}^{S_k}\end{bmatrix} = 0$. Then, $F_1^{S_i\cup S_k}\begin{bmatrix} z_1^{S_i} \\ lz_1^{S_k}\end{bmatrix} = DF_{012}^{S_i\cup S_k}\begin{bmatrix} z_{012}^{S_i} \\ lz_{012}^{S_k}\end{bmatrix} = 0$. Therefore, if $F_{012}^{S_i}z_{012}^{S_i}$ and $F_{012}^{S_k}z_{012}^{S_k}$ are colinear, then $F_1^{S_i}z_1^{S_i}$ and $F_1^{S_k}z_1^{S_k}$ are also colinear. Observe that D is not an invertible matrix, thus the converse is not always true. □

Theorem 5.2 shows that the set of vulnerabilities for the three-phase model is a subset of those for the direct-sequence model. In principle, the inclusion need not be strict, as shown by the next result.

Theorem 5.3. *For a balanced three-phase system, a group of sites is vulnerable to undetectable TSAs in the three-phase or complete-sequence model if and only if its direct-sequence representation is vulnerable to undetectable TSAs.*

Proof. Recall that by definition, if the system is balanced, then $z_b = \alpha^2 z_a$ and $z_c = \alpha z_a$, with $\alpha = e^{2j\pi/3}$. Hence, if the system is balanced, then $z_0 = z_2 = 0$ and $z_1 = z_a$. Also, if the system is balanced, then the verification matrix in the three-phase model is of the

form

$$F_{abc} = \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & & \vdots \\ P_{m1} & \cdots & P_{mm} \end{bmatrix}, \text{ where } P_{xy} = \begin{bmatrix} a_{xy} & b_{xy} & b_{xy} \\ b_{xy} & a_{xy} & b_{xy} \\ b_{xy} & b_{xy} & a_{xy} \end{bmatrix},$$

with $a_{xy},\ b_{xy} \in \mathbb{C}\ \forall 1 \leq x, y \leq m$. After transformation using T_Z, we obtain that the verification matrix in the complete-sequence model F_{012} is a block matrix with 3×3 blocks

$$Q_{xy} = \text{diag}(a_{xy} + 2b_{xy}, a_{xy} - b_{xy}, a_{xy} - b_{xy}).$$

Therefore, if the system is balanced, then $F_{012}^{S_k} z_{012}^{S_k} = F_1^{S_k} z_1^{S_k}$ because the other values are equal to 0. Hence, $F_{012}^{S_k} z_{012}^{S_k} = 0$ if and only if $F_1^{S_k} z_1^{S_k} = 0$, where S_k is the set of measurement indices at bus k. Similarly for a pair of sites with measurement indices in S_i and S_k, we obtain $\text{ERR}\left[F_{012}^{S_i} z_{012}^{S_i} | F_{012}^{S_k} z_{012}^{S_k}\right] = \text{ERR}\left[F_1^{S_i} z_1^{S_i} | F_1^{S_k} z_1^{S_k}\right]$. Hence, when the system is balanced, $F_{012}^{S_i} z_{012}^{S_i}$ and $F_{012}^{S_k} z_{012}^{S_k}$ are colinear if and only if $F_1^{S_i} z_1^{S_i}$ and $F_1^{S_k} z_1^{S_k}$ are colinear. Recall that the analysis of the system vulnerability to TSAs reduces to the vulnerability analysis of every site and every pair of sites. □

Theorem 5.3 implies that in a balanced three-phase system, the three-phase and direct-sequence models are equally vulnerable to TSAs. However, in an unbalanced three-phase system, three-phase state estimation is less vulnerable to undetectable TSAs than state estimation based on the direct-sequence model. In other words, a site or a pair of sites whose direct-sequence representation is vulnerable to TSAs may not be vulnerable using a three-phase representation. Intuitively, if the three-phase measurements are all taken into account, an attack at a site shifts three times more phasors than in the direct-sequence representation, making it harder to remain undetected.

5.2.2 Approximate Vulnerability Analysis

A group of sites that is not vulnerable to undetectable TSAs could potentially be vulnerable to attacks that only slightly change the residuals. In what follows we analyze the potential vulnerability of sites and pairs of sites to such attacks. Unlike for undetectable TSAs for which Lemma 5.3 allowed us to focus on the LS verification matrix, due to the change of the residuals we have to perform the potential vulnerability analysis of three-phase state estimation using the WLS verification matrix in rectangular coordinates.

For the purpose of measuring the distance to vulnerability we rely on and extend vulnerability metrics introduced in Chapter 4 for a single site and for a pair of sites. It is sufficient to consider these two metrics as the analysis of the system vulnerability to TSAs can be reduced to the vulnerability analysis of every site and every pair of sites, as shown in Theorem 4.10. Let us first recall the metrics for the direct-sequence model.

Definition 5.2. *The direct-sequence vulnerability metric for*

- *a site k is defined as* $\|R_1^{S_k} z_1^{S_k}\|$,
- *a pair of sites* (i,k) *is defined as* $1 - ERR(\left[R_1^{S_i} z_1^{S_i} | R_1^{S_k} z_1^{S_k}\right])$,

where $R_1 \in \mathbb{C}^{2m\times m}$ *and* $G_{\square,1} \in \mathbb{R}^{2m\times 2m}$ *are given by*

$$R_1 = \frac{1}{2}\begin{bmatrix} G^1_{\square,1} - jG^2_{\square,1} \\ G^3_{\square,1} - jG^4_{\square,1} \end{bmatrix} \text{ and } G_{\square,1} z_{\square,1} = \begin{bmatrix} G^1_{\square,1} & G^2_{\square,1} \\ G^3_{\square,1} & G^4_{\square,1} \end{bmatrix} \begin{pmatrix} Re(z_1) \\ Im(z_1) \end{pmatrix},$$

and ERR *is the effective rank ratio.*

The effective rank ratio (ERR) of a matrix is the ratio of the largest singular value in absolute value to the sum of all singular values in absolute value. The closer the metrics are to 0, the more vulnerable is the site or the pair of sites. If the metric is equal to 0 then the site or the pair of sites is in fact vulnerable to undetectable TSAs, while if the metric is close to 0 then an attacker may be able to compute an attack that could potentially remain undetected by the BDD algorithms. For three-phase analysis we propose to adapt Definition 5.2 as follows.

Definition 5.3. *The three-dimensional vulnerability metric in the three-phase model (resp. complete-sequence model) for*

- *a site k is defined as* $\|R_{abc}^{S_k} z_{abc}^{S_k}\|$ *(resp.* $\|R_{012}^{S_k} z_{012}^{S_k}\|$*),*
- *a pair of sites* (i,k) *is defined as* $1 - ERR(\left[R_{abc}^{S_i} z_{abc}^{S_i} | R_{abc}^{S_k} z_{abc}^{S_k}\right])$ *(resp.* $1 - ERR(\left[R_{012}^{S_i} z_{012}^{S_i} | R_{012}^{S_k} z_{012}^{S_k}\right])$*),*

where $R_{abc} \in \mathbb{C}^{6m\times 3m}$ *(resp.* $R_{012} \in \mathbb{C}^{6m\times 3m}$*) is defined from blocks of* $G_{\square,abc}$ *(resp.* $G_{\square,012}$*).*

Since the covariance matrix affects the structure of the WLS verification matrix, it is not possible to establish an equivalence result similar to that of Theorem 5.3 for the approximate vulnerability metrics, not even if the system is balanced. Nevertheless, in what follows we show that if the system is balanced then WLS using the direct-sequence model is at least as vulnerable as WLS using the complete-sequence model.

Theorem 5.4. *For a balanced three-phase system,*

- *the direct-sequence vulnerability metric of a site k is no more than its three-dimensional counterpart:* $\|R_{012}^{S_k} z_{012}^{S_k}\| \geq \|R_1^{S_k} z_1^{S_k}\|$,

- *the direct-sequence vulnerability metric of a pair of sites i and k is no more than its three-dimensional counterpart:*

$$1 - ERR([R_{012}^{S_i} z_{012}^{S_i} | R_{012}^{S_k} z_{012}^{S_k}]) \geq 1 - ERR([R_1^{S_i} z_1^{S_i} | R_1^{S_k} z_1^{S_k}]).$$

Proof. As previously, $z_0 = z_2 = 0$ if the system is balanced, therefore, $R_{012}^{S_k} z_{012}^{S_k}$ is equal to the product of the submatrix of $R_{012}^{S_k}$ consisting of only the direct-sequence columns, with the direct-sequence measurement vector $z_1^{S_k}$. Hence, $R_1^{S_k} z_1^{S_k}$ is a sub-vector of vector $R_{012}^{S_k} z_{012}^{S_k}$, i.e., $R_{012}^{S_k} z_{012}^{S_k}$ is equal to $R_1^{S_k} z_1^{S_k}$ concatenated with more values. As a result, $\|R_{012}^{S_k} z_{012}^{S_k}\| \geq \|R_1^{S_k} z_1^{S_k}\|$. With a similar reasoning, we obtain that matrix $[R_{012}^{S_i} z_{012}^{S_i} | R_{012}^{S_k} z_{012}^{S_k}]$ corresponds to matrix $[R_1^{S_i} z_1^{S_i} | R_1^{S_k} z_1^{S_k}]$ with added rows, thus $\text{rank}([R_{012}^{S_i} z_{012}^{S_i} | R_{012}^{S_k} z_{012}^{S_k}]) \geq \text{rank}([R_1^{S_i} z_1^{S_i} | R_1^{S_k} z_1^{S_k}])$. Observe that the ERR is equal to 1 if and only if the matrix is of rank equal to 1 and decreases as the rank increases. Therefore, we obtain that $\text{ERR}([R_{012}^{S_i} z_{012}^{S_i} | R_{012}^{S_k} z_{012}^{S_k}]) \leq \text{ERR}([R_1^{S_i} z_1^{S_i} | R_1^{S_k} z_1^{S_k}])$. □

5.3 Numerical Results

We use simulations based on measurements from the Lausanne power grid to illustrate our results, thus demonstrating the superiority of the three-phase model in detecting TSAs.

5.3.1 Electrical Model

We perform the evaluation on the IEEE 39-bus benchmark system. We consider a PMU allocation in which 21 PMUs are deployed in the system, as shown in Figure 5.1. Among the 21 PMUs, 13 PMUs measure both voltage phasors and injected current phasors at buses {7, 12, 16, 18, 28, 30, 31, 32, 33, 34, 37, 38, 39}, and 8 PMUs measure only injected current phasors at buses {4, 15, 21, 23, 24, 26, 35, 36}. We assume that the PMUs are able to measure the three phases of several phasors simultaneously. Moreover, we consider that the set of buses {1, 2, 5, 6, 9, 10, 11, 13, 14, 17, 19, 22} are zero-injection buses. We mentioned previously that groups of buses that share the same time clock are called sites. Throughout our simulations, we consider that the time reference of every bus can be compromised individually, i.e., sites consist of single buses.

We obtained active and reactive three-phase power measurements measured by 15 PMUs installed in the 125 kV grid of the city of Lausanne, recorded in 2016. Some of the PMUs monitor 1 power line and others monitor 2 power lines, resulting in a total of 22 available measurement points. We assigned one set of three-phase PMU measurements to each of the 19 loads of the IEEE 39-bus benchmark. After running the load flow, we

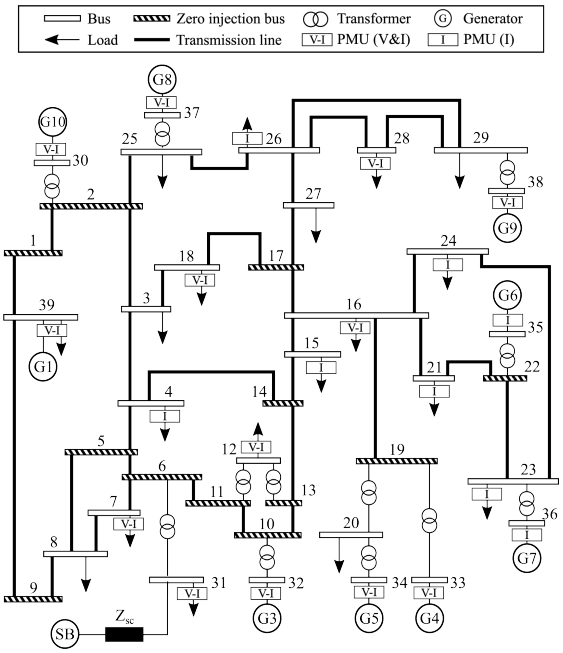


Figure 5.1 – PMU deployment on the IEEE 39-bus benchmark system.

obtained three-phase voltage phasors for each bus. These voltages are considered as state values. We then created the measurements by adding noise in their phase and magnitude according to the level of noise of 0.1-class PMUs. We consider that the frequency with which we obtain measurements is $50Hz$. All presented attacks are computed using the output constrained PI-controller clock servo aware (OCPI) attack strategy presented in Chapter 4. This attack strategy takes into account the presence of a clock servo in each PMU, used for controlling the clock adjustment rate. The attack strategy targets a minimum of three sites simultaneously, hence the attacks that we present always target a triplet of buses. The BDD algorithm used to illustrate the detectability of an attack is the LNR test on the WLS residuals as in the previous chapters.

5.3.2 Practically Balanced Measurements and Increasing Unbalances

In the real data obtained from the Lausanne grid, we found some periods of time in which the three phases of the loads are very close to being balanced. Figure 5.2 shows the phase and magnitude of the state voltage values at each bus at one time-instant. We observe that the phases are almost the same shifted by $\frac{-2\pi}{3}$ and $\frac{2\pi}{3}$ and that the magnitudes are almost the same centered around 1. We also observe that the neutral-sequence and inverse-sequence components are very close to 0 in magnitude.

We simulated an attack on the direct-sequence measurements of the triplet of buses {23,35,36}, with the goal of minimizing the estimated power-flow on the line between buses 21 and 22, during a time interval of 100s (i.e. 5000 measurements) during which the system

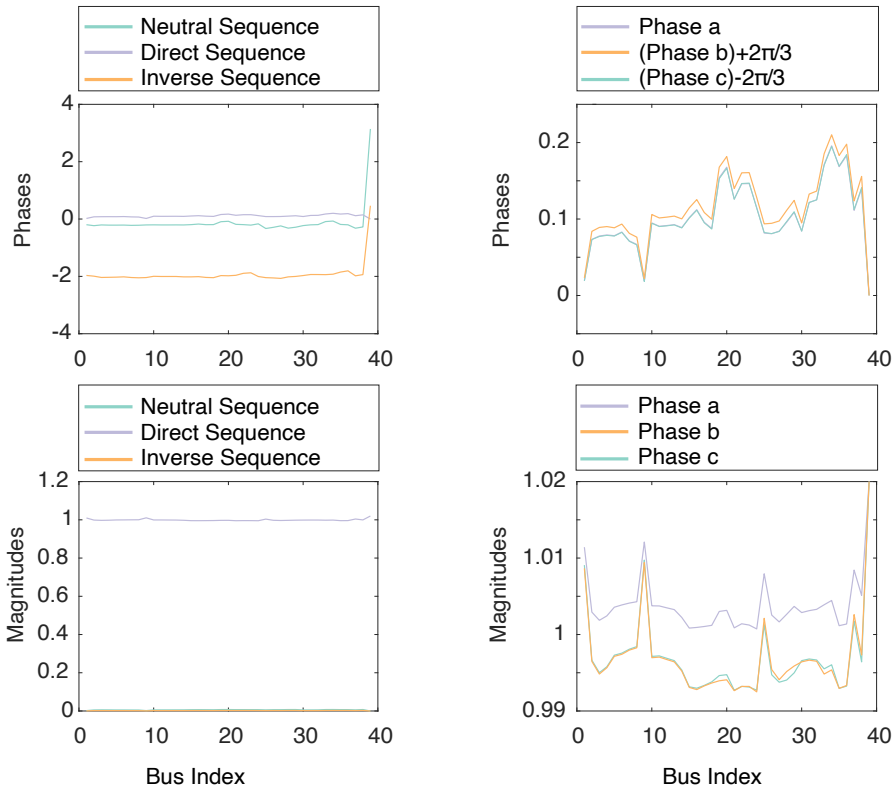


Figure 5.2 – Voltage phasors in a balanced system, phase representation and symmetric components, obtained after load-flow calculation at a single time instant

is practically balanced. We chose to attack this triplet because, as shown in Figure 5.3c, its direct-sequence vulnerability metric is equal to 0, which means that it is vulnerable to TSAs. We then applied the same attack on the three-phase measurements. Figure 5.3a shows the non-negligible effect of the attack on the magnitude of the estimated power-flow, with respect to the direct-sequence (left) and the three-phase measurements (right). The LNR values for the direct-sequence (left) and the three-phase (right) measurements are shown in the first row of Figure 5.3b. Notice that in both cases, the attacked and non-attacked LNR values are indistinguishable. In other words, the attack is undetectable whether the BDD algorithms take as input the direct-sequence measurements or the three-phase measurements. The first column of Figure 5.3c shows that the vulnerability metric computed for the triplet of buses with the direct-sequence measurements is equal to 0 during the entire simulation and is around $0.75e-3$ in the three-phase model. In both cases the metric is very low, showing that TSAs can be performed undetectably. Note that in Section 5.2.2, the vulnerability metric was introduced for pairs of sites but in the simulations, we chose to target three sites simultaneously. The metric computed is the same metric generalized to three sites: $1 - \mathrm{ERR}(\left[R^{S_{23}} z^{S_{23}} | R^{S_{35}} z^{S_{35}} | R^{S_{36}} z^{S_{36}}\right])$. The closer the latter is to 0, the more vulnerable the triplet of buses is.

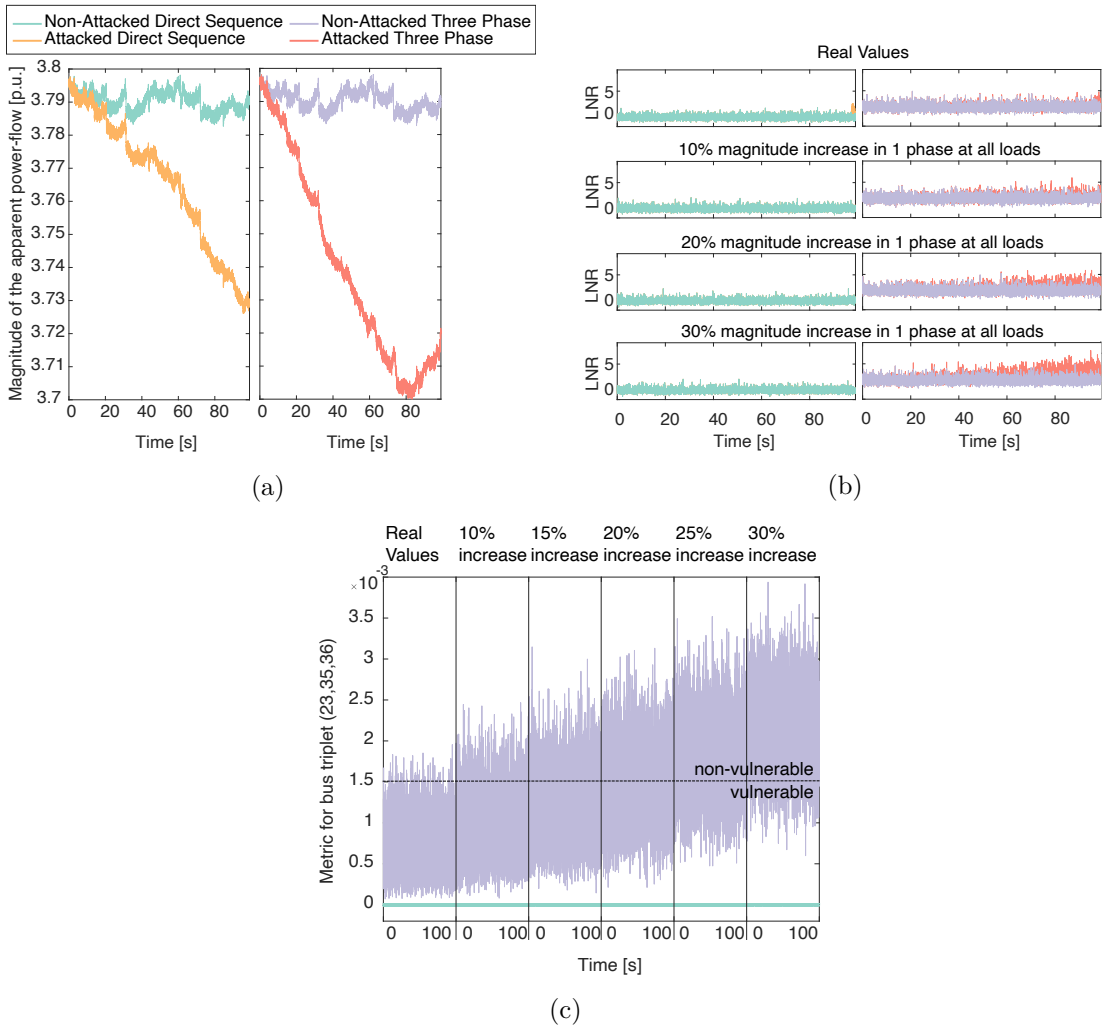


Figure 5.3 – Balanced measurements, attack on {23,35,36}: (a) Magnitude of the power-flow on the line between buses 21 and 22, with and without an attack on the direct-sequence and three-phase measurements; (b) LNR values with and without an attack with increasing unbalances: the attacked and non-attacked LNR values of the direct-sequence measurements are indistinguishable, whereas they are increasingly distinguishable for the three-phase measurements; (c) Vulnerability metric with increasing unbalances:the direct-sequence metric is constant at 0 (i.e. vulnerable) and the three-phase metric increases as the unbalances increase.

In order to investigate the level of unbalance required to create a discrepancy in the detectability of attacks applied to the direct-sequence and the three-phase measurements, we introduced unbalance in the power-flow measurements from the Lausanne grid. In general, unbalances can have various sources, including different amounts and types (inductive and capacitive) of loads, and line parameters. For clarity, we chose to use one source of unbalance: we gradually increased the magnitude of the loads on one phase by 10%, 15%, 20%, 25% and 30%. For each considered level of unbalance we computed an undetectable attack in the direct-sequence model, which we then applied to the three-phase model. Figure 5.3b shows that the LNR values of the attacked and non-attacked

direct-sequence measurements are always indistinguishable, which means that the attack on the direct sequence remains undetected as the unbalance increases. In contrast, Figure 5.3b shows that between the 10% and 30% magnitude increase, the LNR values of the attacked and non-attacked three-phase measurements are increasingly distinguishable. Hence, the attack on the three-phase measurements is becoming easier to detect. As expected, Figure 5.3c shows that the vulnerability metric remains unchanged for the direct-sequence measurements and gradually increases for the three-phase measurements. Depending on the LNR detectability threshold that we want to set, we can find the corresponding vulnerability metric threshold. For example, it is reasonable to say that an attack with the LNR values obtained with the 20% or 25% magnitude increase, can be easily detected. Therefore, a reasonable choice for the triplet-vulnerability threshold would be $1.5e-3$, as shown in Figure 5.3c.

5.3.3 Unbalanced Measurements

Next, we focus on actual unbalanced measured data from the Lausanne grid. Figure 5.4 shows the phase and magnitude of the voltage phasors at each bus at one time instant. We observe that both the angles and the magnitudes are quite different across phases and that the latter is not always close to 1. We also observe, as a sign of unbalance, that the neutral-sequence and inverse-sequence components are distinguishable from 0. We used this real dataset to compute TSAs targeting the same triplet of buses as previously {23,35,36} with the same objective of reducing the estimated power-flow on the line between buses 21 and 22.

Figure 5.5a shows the non-negligible effect of the attack on the magnitude of the estimated power flow on the line between buses 21 and 22. The attacked and non-attacked LNR values presented in Figure 5.5b are indistinguishable for the direct-sequence measurements (left), whereas they are clearly distinguishable for the three-phase measurements (right). As expected, Figure 5.5c shows that the vulnerability metric computed for the direct-sequence measurements is always very low, close to 0, and that the vulnerability metric computed for the three-phase measurements is always above the triplet-vulnerability threshold $1.5e-3$, which we empirically established in Section 5.3.2. Therefore, the use of the three-phase measurements as input to the BDD algorithms enables the detection of the attack, which is not the case if we use only the direct-sequence measurements.

5.3.4 Undetectable Attack on Three-Phase Measurements

In this section we consider an attacker that mounts its attack on the three-phase unbalanced measurements instead of on the direct-sequence measurements only. We show that even though employing a three-phase state estimator significantly reduces the vulnerability of power-system state estimation to TSAs, undetectable TSAs may

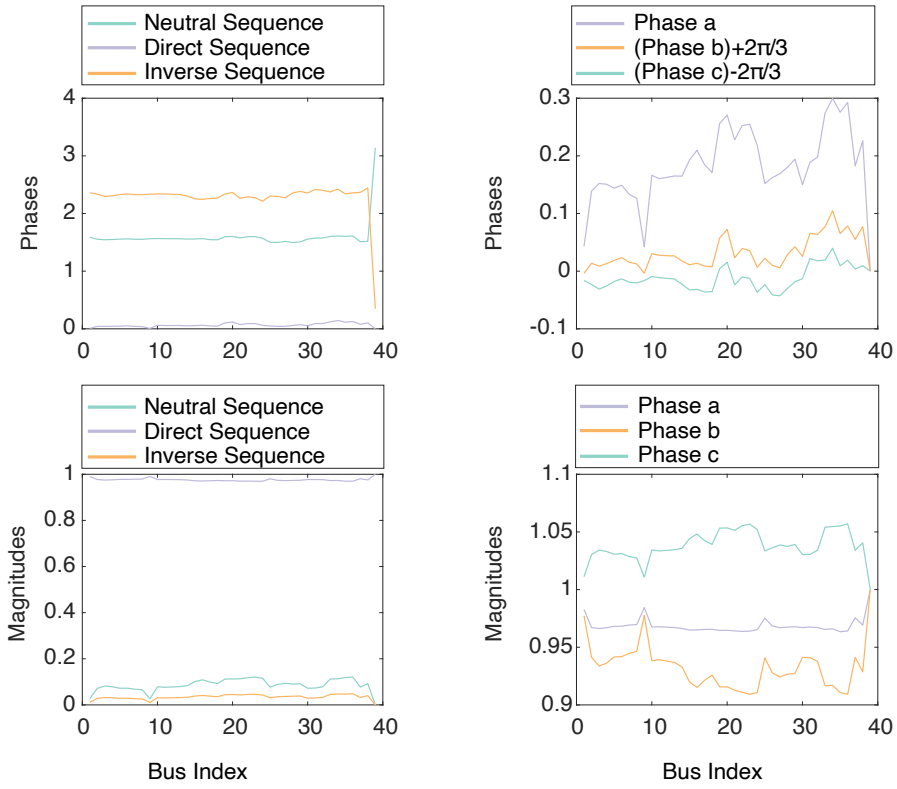


Figure 5.4 – Real unbalanced voltage phasors after the load flow at one time-instant according to the complete-sequence and three-phase models.

be possible despite using a three-phase state estimator. The measurements used in this section are the unbalanced measurements used in Section 5.3.3. For the PMU deployment shown in Figure 5.1 we found that the triplet of buses {26, 28, 38} has a low three-phase vulnerability metric value, as shown in Figure 5.6a. We observe that both the direct-sequence and the three-phase metrics are below the triplet-vulnerability threshold, and hence we expect that we can perform an undetectable attack against this triplet.

Figure 5.6b shows the results of the LNR test on the attacked and non-attacked three-phase measurements. We observe that the LNR values obtained with and without the attack are indistinguishable. Notably, Figure 5.6c shows that the TSA results in a very large (one order of magnitude) overestimation of the apparent power flow on the transmission line between buses 26 and 28. This scenario shows that even though a three-phase state estimator is harder to attack, it might still be vulnerable to undetectable TSAs. However, further analysis shows that $1 - \mathrm{ERR}(F^{[S_{26},S_{28},S_{38}]}) = 0$, which means that this triplet in the direct-sequence model is in fact structurally vulnerable. This means that this triplet is vulnerable in the direct-sequence whatever the values of the measurements. Hence, this attack is no longer feasible after securing the grid with Algorithm 4 from Chapter 4.

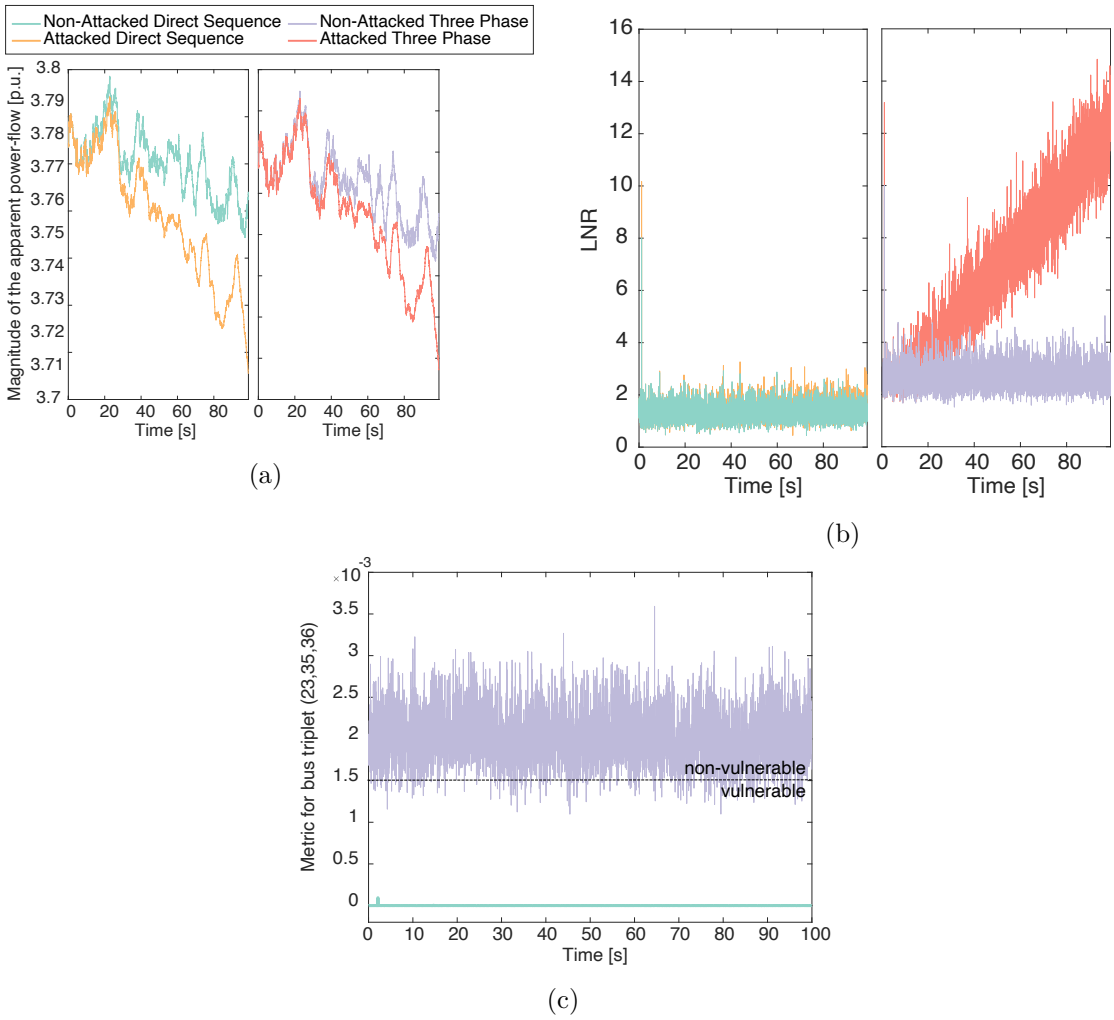


Figure 5.5 – Unbalanced system, attack on {23,35,36}: (a) Magnitude of the power-flow on the line between buses 21 and 22, with and without an attack on the direct-sequence and three-phase measurements; (b) LNR values with and without an attack: the attack is undetected using the direct-sequence model but is detected with three-phase estimation; (c) Vulnerability metric: the direct-sequence metric shows vulnerability and the three-phase metric is above the triplet-vulnerability threshold.

5.3.5 Vulnerability Analysis of Different PMU Deployments

The previous results demonstrate the potential of a three-phase state estimator in detecting TSAs against the IEEE 39-bus system for the PMU deployment shown in Figure 5.1. We now extend our analysis to other PMU deployments.

All PMUs measure voltages and branch currents on the IEEE 39-bus system:

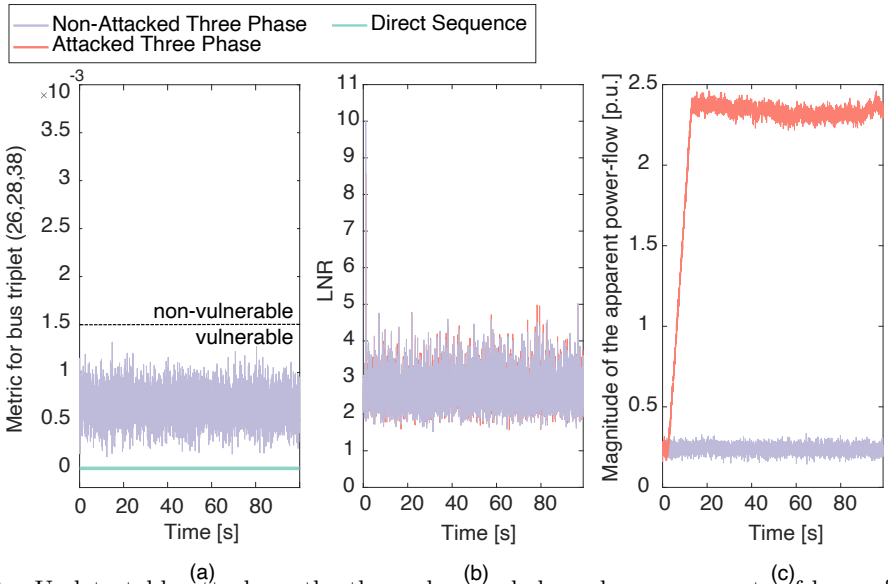


Figure 5.6 – Undetectable attack on the three-phase unbalanced measurements of buses {26,28,38}: (a)Vulnerability metric: both metrics are always below the triplet-vulnerability threshold; (b) Results of the LNR test with and without the attack are hardly distinguishable; (c) The attack results in a one order of magnitude misestimation of the magnitude of the power-flow on the transmission line between buses 26 and 28.

We also considered a scenario where every PMU can measure the bus voltage as well as the incident branch currents. Given this constraint, we found that the PMU allocation with the smallest number of PMUs such that the IEEE 39-bus system is observable contains 16 PMUs, placed at buses $\{3, 8, 12, 15, 20, 21, 23, 27, 29, 30, 31, 32, 33, 35, 37, 39\}$. With this allocation, we found that 24 pairs of buses are vulnerable to TSAs in the direct-sequence model, while there is no vulnerable pair of buses in the three-phase model. A PMU pair was considered vulnerable if the computed vulnerability metric was below 10^{-5}. The vulnerable pairs are: $(4, 23)$, $(4, 27)$, $(7, 23)$, $(7, 27)$, $(16, 23)$, $(16, 27)$, $(18, 23)$, $(18, 27)$, $(23, 24)$, $(23, 25)$, $(23, 26)$, $(23, 28)$, $(23, 29)$, $(23, 34)$, $(23, 36)$, $(23, 38)$, $(24, 27)$, $(25, 27)$, $(26, 27)$, $(27, 28)$, $(27, 29)$, $(27, 34)$, $(27, 36)$, $(27, 38)$.

Analysis on the 7-bus grid of the city of Lausanne: We tested our method on the real 7-bus grid of Lausanne, where all PMUs measure the bus voltage and the branch currents of all connected lines. A description of the grid can be found in [96]. This grid has an untransposed line, thus the sources of unbalance are both coming from the loads and from the line parameters. We had access to the following data: measurement values, admittance matrix and the noise covariance matrix. We computed our metric for all pairs of buses, and we show the results at a particular time instant in Figure 5.7. We observe that 6 pairs of buses have dangerously small metric values below 0.001 in the direct-sequence model. In contrast, we observe that all pairs of buses have high metric values above 0.32 in the three-phase model, which means that the 6 vulnerable pairs in the direct-sequence model are no longer vulnerable.

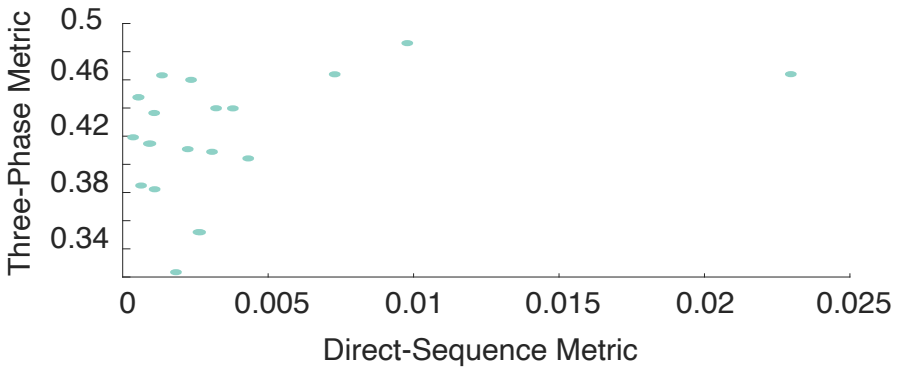


Figure 5.7 – Metric values for all pairs of buses of the 7-bus grid of the city of Lausanne: 6 pairs have a metric very close to 0 in the direct-sequence model and all metric values are far from 0 in the three-phase model.

5.4 Conclusion

In this chapter we analyzed the benefits of using a three-phase state estimator as a tool to detect TSAs. We showed that in a balanced three-phase system the vulnerability conditions of the three-phase and the direct-sequence state estimators are equivalent. In contrast, we proved that in an unbalanced system the vulnerability of direct-sequence state estimation does not imply the vulnerability of three-phase state estimation, which shows that three-phase state estimators are more resilient to TSAs than traditional direct-sequence state estimators. Our simulations performed with real load profiles on an IEEE test system confirmed these results and showed that as the unbalance grows, undetectable TSAs on the direct-sequence measurements may become detectable if three-phase state estimation is used. Although the use of a three-phase state estimator enables to detect more TSAs, our simulations also showed that it is not always sufficient to completely secure the grid. Nonetheless, even though not completely invulnerable, a three-phase state estimator is definitely a considerable step to enhance the security of future power grids.

PART II

SECURE LOCALIZATION

List of Variables of Part II

N	Number of network sensors
S_i	Sensor of index i
(x_i, y_i, z_i)	3D coordinates of sensor i
S	Unknown source
(x, y, z)	3D coordinates of the unknown source S
(x_g, y_g, z_g)	Geometrical estimate of the coordinates of S
$s(t)$	Continuous radio signal sent by the unknown source
$r_i(t)$	Signal received by sensor i
Δ_i	Signal travel time from source to sensor i
e_i	Gaussian noise contained in the received signal at sensor i
Δ_{ij}	True but unobservable TDOA between sensors i and j
$\widehat{\Delta_{ij}}$	Estimate of the TDOA between sensors i and j
e_ij	Error associated to $\widehat{\Delta_{ij}}$
c	Speed of light
γ_i	SNR at sensor i
γ_{ij}	Aggregated SNR of sensor pair (S_i, S_j)
σ_{ij}	Standard deviation of e_{ij}
$d_i \in \mathbb{R}$	Attack offset introduced at sensor i
μ_{ij}	Center of the distribution of $\widehat{\Delta_{ij}} - \Delta_{ij}$
w_{ij}	Weight associated to the trust level of the time-synchronization between sensors (S_i, S_j)
z_{ij}	Standardized statistic for sensor pair (S_i, S_j)
$v \in \mathbb{R}$	Exponent chosen for optimal weights computation
cfd	Confidence metric
D	Problem dimension (2D or 3D)
q	Proportion of attacked measurements

6

TDOA Source-Localization Technique Robust to Time-Synchronization Attacks

In this chapter, we focus on the time-difference-of-arrival (TDOA)-based localization of a passive source from a network of fixed sensors whose time references could be maliciously manipulated. TDOA measurements offer high precision hence are widely used. However, they are easily attacked because they are particularly sensitive to time-synchronization errors. Due to the high propagation speed of the signal, a small synchronization error can lead to a large range difference error between the two sensors and the source. For example, if $3\mu s$ are added to a TDOA measurement, then the corresponding range difference is increased by approximately 900m. Consequently, an attacked network could become unable to localize sources or an attacked vehicle could unintendedly enter a wrong territory.

This chapter first highlights the effect of time-synchronization attacks (TSAs) on the TDOA-based localization of an unknown source. We show that the delays between sensors can lead to a misestimation of the source location. We inject a few microseconds into one sensor and obtain an estimate that is approximately 1 km away from the true position of the source. We further explain how an attacker can compute positive or negative delays such that the localization process results in a specifically chosen misestimation. Recall that delays can be implemented by a delay-box insertion on the path between the master node and the slave node. A longer fiber on the path from the slave to the master, injects a negative delay in the time reference of the slave clock, and the reverse injects a positive delay. We also show that residual analysis does not enable the detection and identification of TSAs.

The main contribution of this chapter is to propose a TDOA-localization technique that is robust against TSAs. It secures against the misestimation of the location of unknown sources in the considered adversarial environment. Our technique works in two phases, the former analyses the error in TDOA measurements received from a known calibration source. As a result, it defines a weight for each pair of sensors; this reflects the confidence we have in their time synchronization. The latter phase of our solution

then uses the weighted least-squares (WLS) estimator with the newly created weights and the TDOA measurements received from the unknown source. Subsequently, our method either identifies the network as being too corrupt to localize, or gives a corrected estimate of the unknown position along with a confidence metric.

Our calibration technique requires the use of a trusted source of known coordinates. To our understanding, it is realistic to assume the existence of such a known source: a sensor of the localization network or a vehicle equipped with an emitter can be used for the calibration phase. In this first phase, we compare true TDOAs with observed TDOAs measured from signals emitted by the known source. We do not require the known source to be part of the synchronized network because our technique does not require the use of timestamps from the known source. In fact, our solution is immune to a TSA on the known source. However, nothing prevents an attacker from storing emitted calibration signals in order to replay them in the direction of the attacked sensors at times and locations of his choice. For example, he could replay them in a manner that compensates for the introduced attack delays. To counter such an attack, we propose an encrypted authenticated challenge-response scheme where the calibration source is triggered to emit a one-time response signal.

The rest of the chapter is structured as follows. In Section 6.1, we discuss related works. We describe our system model in Section 6.2, together with technical background on TDOA-localization in an unattacked environment. We define the attacker's capabilities and study the effect of TSAs in Section 6.3. We present our calibration-based robust localization technique in Section 6.4. We show how to counter replay attacks against our solution in Section 6.5. We present the numerical results of the evaluation of the performance of our solution and of the confidence metric in Section 6.6. We also compare the performance of our solution with usual tracking methods in Section 6.6. Finally, we conclude the chapter with Section 6.7.

6.1 Related Work

The effect of a TSA on the measurements is analoguous to the effect of a non-line-of-sight (NLOS) path, with the difference that the former adds positive or negative delays and the latter adds only positive delays. However, to the best of our knowledge, none of the existing NLOS mitigation techniques are applicable to our problem. A first category of solutions considers cooperative sources that are synchronized with the network of sensors [102, 103, 104, 105, 106]. This context is mostly significant for entities that wish to localize themselves using signals sent by antennas. Such techniques use as much additional information or measurement as possible. For example it was shown in [107] that knowledge on the layout of the surroundings can significantly improve the localization. A second category of solutions aims to track the trajectory of the source, using for example variants of Kalman filters, where prior information can be used to help localize the source

at a given instant [108, 109, 110, 111]. A third type of solution assumes that at most one measurement is impacted by a NLOS path [112]. In contrast, we consider *non-cooperative* sources that do not transmit any particular information willingly to the sensors. The TDOA measurements are obtained by overhearing radio signals emitted by the source and received by pairs of synchronized sensors. The sources that we consider can be fast, furtive objects that fly erratically, such as non-cooperative drones. Such objects are hard to track continuously as they can have a non-expected flight behaviour. Their localization is based on a few one-shot measurements with low or no redundancy and we assume that an arbitrary number of them is affected by a time-synchronization attack.We show through simulations in Section 6.6.5, that our robust solution outperforms typical tracking solutions used in LOS and NLOS environments.

In recent years, the subtleties of TDOA-based localization of passive sources sparked the interest of researchers. Due to the accuracy of this localization technique, its use is widespread. Its sensitivity to sensor location errors, oscillator-frequency-synchronization errors and time-synchronization errors between sensors has been the topic of various papers. The authors of [113, 114] studied the effect of phase and frequency-synchronization errors on the TDOA estimation, for different types of oscillators in the cases of single and multi-source localization. They propose a technique [115] to estimate both the TDOA measurement and the frequency error between sensors at low computational and memory complexities when the oscillator frequency error between two sensors is assumed to be non-zero and constant. Similarly, the authors of [116, 117] propose different techniques for estimating the TDOA between sensors, including oscillator phase and frequency errors. Their techniques are based on the Maximum Likelihood estimation of the TDOA, and one of them also estimates the frequency error of the oscillators. Then, the authors of [118, 119] focused on the localization of passive sources in systems of moving sensors that suffer from sensor position errors and from clock-synchronization bias between sensors. Their model, however, assumes that sensors are divided into groups within which sensors are time-synchronized, and that timing offsets are present only among different groups. In this chapter, we assume that sensors are fixed at known locations and that they are all spaced out, therefore we assume that there can be time offsets between all sensor pairs. Furthermore, unlike in the previously mentioned papers, we assume that the synchronization offsets are not due only to the use of inaccurate hardware but also to the presence of malicious activity. As explained in Section 6.3, we consider that an attacker is able to introduce time offsets in the clock of sensors in such a way that the resulting TDOA measurements seem plausible and intersect well, at a distant target location.

6.2 System Model

Consider a network of N time-synchronized sensors S_i, $1 \leq i \leq N$ with known coordinates (x_i, y_i, z_i). Suppose that a source S of unknown coordinates (x, y, z) produces a continuous signal $s(t)$. It is received by several network sensors in the following form: $r_i(t) = s(t - \Delta_i) + e_i$, where Δ_i is the time needed for the signal to travel from the source to sensor S_i and e_i is Gaussian noise. The receiving sensors then simply timestamp the received signals and transmit them to a centralised control center. In order to compute an estimate $\widehat{\Delta}_{ij}$ of the true but unobservable TDOA Δ_{ij} between each pair of receiving sensors (S_i, S_j), the control center then uses correlation techniques [120, 121, 122] on the signal samples

$$\widehat{\Delta}_{ij} = \Delta_{ij} + e_{ij} = \Delta_i - \Delta_j + e_{ij}, \tag{6.1}$$

where e_{ij} is the noise associated with the estimated delay, in other words, the difference between the true and the estimated delay. Note that e_{ij} is not equal to $e_i - e_j$. Each TDOA Δ_{ij} defines a hyperbola on which the source should lie. The variance of the estimated delay defines a zone of probable location of the source along the corresponding hyperbola. By aggregating several measurements, the source is then estimated to be in a probable zone defined by the intersecting hyperbolae. The estimation of the source location is not trivial as it is a quadratic non-convex problem. The relation between an estimated delay $\widehat{\Delta}_{ij}$ and the source coordinates is given by the following equation

$$\widehat{\Delta}_{ij} = \frac{\text{dist}(S_i, S) - \text{dist}(S_j, S)}{c} + e_{ij}, \tag{6.2}$$

where c is the propagation speed of the signal and
$\text{dist}(S_i, S) = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}$ is the distance between sensor S_i and the unknown source S. The unknowns in this equation are the source coordinates (x, y, z) and the noise e_{ij}. This noise is distributed according to a centered normal distribution $\mathcal{N}(\mu_{ij}, \sigma_{ij})$ with $\mu_{ij} = 0$ and where σ_{ij} is unknown.
Supposing that the noise of the received signal at various sensors is i.i.d with same signal-to-noise ratio (SNR), the covariance matrix is [123, 124]

$$K = \sigma^2 \begin{pmatrix} 1 & 1/2 & \cdots & 1/2 \\ 1/2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1/2 \\ 1/2 & \cdots & 1/2 & 1 \end{pmatrix}, \tag{6.3}$$

where σ^2 is the TDOA noise variance. For simulations in the litterature, the standard deviation is often set arbitrarily to a plausible value such as $1.83\mu s$ or $0.183ns$. Nevertheless, more precise formulas to compute the standard deviation as a function of the SNRs of sensors are given in [18, 125]

- An aggregated SNR (not in dB) is computed from the SNRs of the two sensors $\gamma_{ij} = \mathrm{SNR}(\gamma_i, \gamma_j)$

$$\frac{1}{\gamma_{ij}} = \frac{1}{2}\left(\frac{1}{\gamma_i} + \frac{1}{\gamma_j} + \frac{1}{\gamma_i \gamma_j}\right). \tag{6.4}$$

- For a low SNR value γ_{ij}, the standard deviation is given by

$$\sigma_{ij} = \sqrt{\frac{1}{8\pi^2}} \frac{1}{\gamma_{ij}} \frac{1}{\sqrt{T_{int} W}} \frac{1}{f_0} \frac{1}{\sqrt{1 + \frac{W^2}{12 f_0}}}, \tag{6.5}$$

  where T_{int} is the integration time of the signal for one measurement, $W = f_2 - f_1$ is the frequency bandwidth and f_0 is the center of frequency.

- For a high SNR value γ_{ij}, the standard deviation is given by

$$\sigma_{ij} = \sqrt{\frac{3}{4\pi^2 T_{int}}} \frac{1}{\sqrt{\gamma_{ij}}} \frac{1}{\sqrt{f_2^3 - f_1^3}}. \tag{6.6}$$

From noisy TDOA measurements, both geometrical and analytical techniques of localization can be found in the litterature [24, 124]. The non-linear least- squares (LS) estimator is a widespread localization technique that takes the noisy TDOA measurements as input and searches for a solution (x, y, z) minimizing the sum of squared errors

$$\arg\min_{x,y,z} \sum_{i>j} \left(\frac{\mathrm{dist}(S_i, S) - \mathrm{dist}(S_j, S)}{c} - \widehat{\Delta_{ij}}\right)^2.$$

This estimator can be modified to solve the weighted least-squares problem (WLS) using the covariance matrix of the TDOA measurements. The WLS estimator corresponds to the maximum likelihood estimator when the usual error can be modelled as drawn from a Gaussian distribution. Throughout the rest of the chapter, the Levenberg-Marquadt (LM) algorithm is used to solve this non-convex optimization problem, thus estimating the coordinates of unknown sources. More discussion on the LM algorithm is provided in Section 6.6.

In order to reduce the complexity of storage and of the estimation process, a widespread technique is to consider only linearly independent measurements by considering only the TDOA measurements with respect to a reference sensor. This reduces the number of equations from $\binom{N}{2}$ to $N - 1$, where N is the number of available sensors. However, this technique induces a loss of redundancy which can be fatal to the localization system in the case of an attack.

6.3 Impact of Time-Synchronization Attacks

In this section,we show how an attack on the time reference of one or more sensors can alter the localization process presented in Section 6.2. We begin by defining the capabilities of the attacker.

6.3.1 Attack Model

We consider two TSA models: one of them is referred to as the weak attack model, and the other is referred to as the strong attack model. In both cases, we suppose that an attacker is able to introduce an offset $d_i \in \mathbb{R}$, which can be positive or negative, to the time reference of sensor S_i. Recall that such an attack does not require physical access to the sensors as it can be achieved via signal spoofing or delay-box insertion, depending on the preferred synchronization technique. With this capability, the goal of the attacker is to introduce errors in TDOA measurements, thus provoking a misestimation of the location of an unknown source. In the weak attack model, the attacker is not able to choose the misestimation, his goal is to create errors in the localization. In contrast, in the strong attack model, we further suppose that the attacker knows the true source coordinates and the network topology, namely the sensor coordinates. In this case, the objective of the attacker is to ensure that the localization process results in a specific targeted misestimation.

6.3.2 Impact on Localization

Introducing delays d_i and d_j to sensors S_i and S_j respectively, adds components to Eq.(6.1)

$$\widehat{\Delta}_{ij} = \Delta_{ij} + d_i - d_j + e_{ij} = \Delta_{ij} + \mu_{ij} + e_{ij}, \tag{6.7}$$

where $\mu_{ij} = d_i - d_j$ is the introduced delay difference between the two sensors. Observe that $\widehat{\Delta}_{ij} - \Delta_{ij}$ is still distributed according to $\mathcal{N}(\mu_{ij}, \sigma_{ij})$, with $\mu_{ij} = 0$ if no delays are inserted and $\mu_{ij} = d_i - d_j$ otherwise. Therefore, introducing delays does not guarantee an impact on measurements and thus on the localization. In fact, the attack is meaningful only if there is a non-negligible delay difference between the time references of at least two sensors. If an attacker introduces the same delay to all sensors of the network, they will all be under attack yet remain synchronized with each other. Hence, the functionality of the system will not be altered and the presence of malicious activity will be undetected. When the attack is such that the difference $\mu_{ij} > \sigma_{ij}$ is non-negligible in comparison to the Gaussian noise, the measurement $\widehat{\Delta}_{ij}$ and its corresponding hyperbola are significantly modified. As a result, the localization process fails to give an accurate estimate. Next, we give two attack scenarios that illustrate the actions of an attacker as in the two proposed attack models of Section 6.3.1.

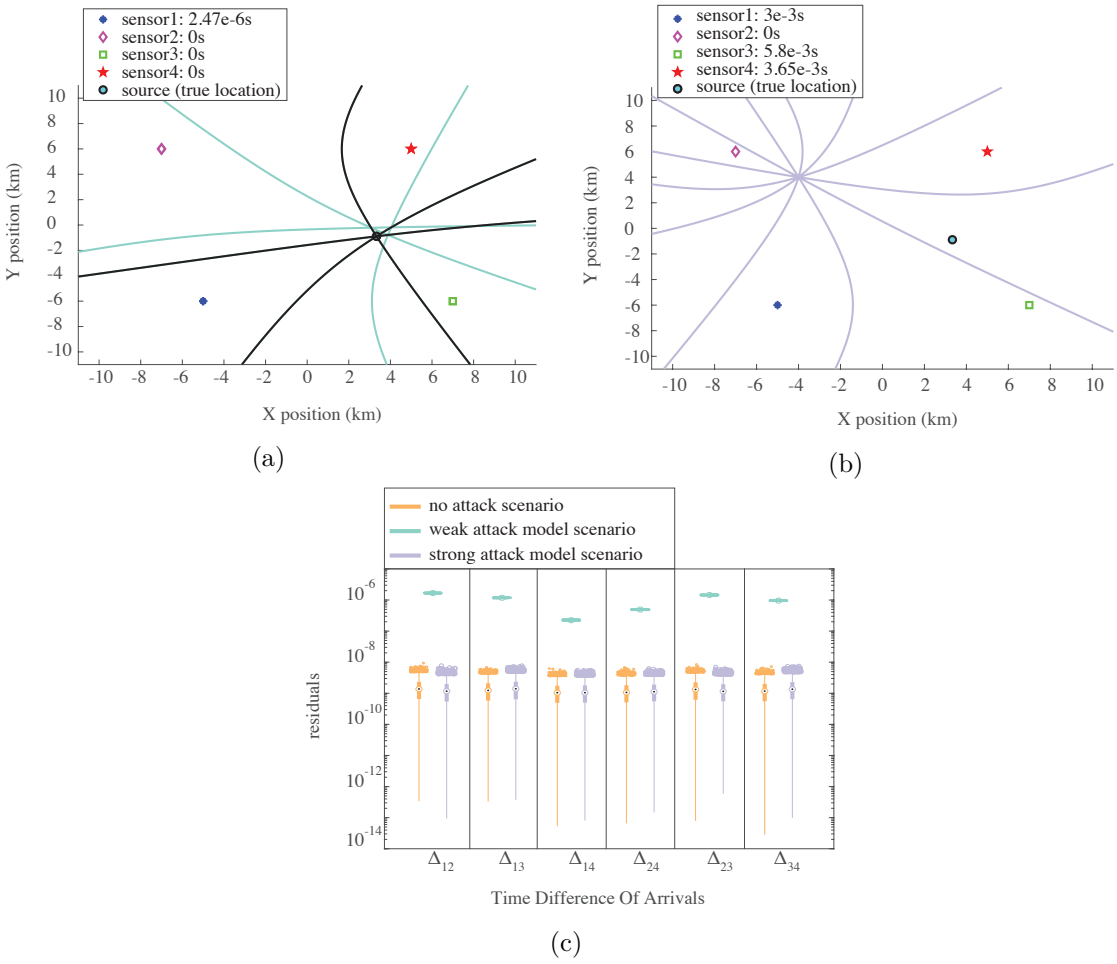


Figure 6.1 – (a): Weak-attack model scenario: sensor 1 is delayed by $2.47\mu s$. The green TDOA hyperbolae are shifted by this delay and the resulting source estimate is incorrect by approximately 1 km; (b): Strong-attack model scenario: delays are strategically computed which results in a specific misestimation almost 9 km away from the true source location; (c): TDOA residuals obtained for $10'000$ simulations of the previous two scenarios. The values are similar in the no-attack and strong attack-model scenarios: the strong attack is undetected by residual analysis. The values are increased by two to three orders of magnitude in the weak attack-model scenario, residual analysis detects malicious activity but does not identify it clearly.

In both scenarios, we consider a two-dimensional grid of side of 20 km with four sensors placed as in Figures 6.1a and 6.1b. The unknown source emits a signal that propagates to the four sensors. We suppose that the sensors send their received signal samples to a control center that processes them. The resulting TDOA measurements are given as input to the WLS estimator as described in Section 6.2. In this simulation, we set the noise standard deviation to $\sigma = 2.192ns$ for all TDOA measurements, this value is further discussed in Section 6.6. In the the weak attack-model scenario, the attacker delays the time reference of sensor S_1 by $2.47\mu s$, which modifies the three corresponding hyperbolae, drawn in green in Figure 6.1a. The resulting WLS estimate of the unknown source location is incorrect by approximately 1 km, thus illustrating that localization

from TDOA measurements is highly sensitive to time offsets. As the injected delay increases, the accuracy of the estimate decreases. Note that in this scenario, the full set of measurements was considered. Supposing that we considered only three linearly independent measurements with respect to sensor S_1, then the estimate would be even less accurate as the WLS estimator would have received only wrong measurements as input. In other words, only the green hyperbolae from Figure 6.1a would be taken into account and none of the black ones.

In the strong attack-model scenario, the attacker knows the coordinates of the sensors and the true coordinates of the source. With such knowledge, he is able to compute the delays to be injected such that the estimation process results in a specific targeted misestimation:

- From the source and sensor coordinates, he computes the true delays of propagation of the signal between the source and the sensors $\Delta_i = \frac{\text{dist}(S_i,S)}{c},\ \forall 1 \leq i \leq N$.
- Similarly, he computes the true delays of propagation of the signal between the targeted misestimation location and the sensors $\Delta_i^t,\ \forall 1 \leq i \leq N$.
- The attack is simply the difference between the two: the delay to inject to sensor S_i is $d_i = \Delta_i^t - \Delta_i$.

This attack is illustrated in Figure 6.1b, where the delays are computed specifically such that all hyperbolae are modified in a plausible manner, intersecting near the targeted misestimation location. The resulting estimate is, as chosen by the attacker, almost 9 km away from the true source location.

6.3.3 Residual Analysis

Once an estimate S_{est} is computed from measurement values, it is useful to compute and analyze the residuals in order either to assess the accuracy of the estimator or to attempt to detect and identify bad data in the measurements. For a pair of sensors (S_i, S_j), the corresponding residual is computed as follows

$$\frac{\text{dist}(S_i, S_{est}) - \text{dist}(S_j, S_{est})}{c} - \widehat{\Delta_{ij}}.$$

The residuals give insight on how well the estimate fits the measurements. Hence, if all hyperbolae intersect near the estimated location, then the residuals will have small values. However, if the points of intersection of the hyperbolae define a large probable zone of location, then the resulting estimate will be far from some or all hyperbolae, thus resulting in large residuals. The distribution of residual values are given in Figure 6.1c for

$10'000$ simulations of the no-attack scenario and of the scenarios depicted in Figures 6.1a and 6.1b that correspond to the two attack models. Observe that the strong attack-model scenario and the no-attack scenario have similar residuals of order of magnitude below a nanosecond. This is as expected because, in both cases, the estimate satisfies well all measurements. Against the strong attack model the residual analysis fails to detect the presence of malicious activity. In the weak attack-model scenario, Figure 6.1c shows that the residual values are by two to three orders of magnitude larger than their corresponding values in the no-attack scenario. However, these large residuals do not clearly identify sensor S_1 as being under attack. In this case, the residual analysis succeeds in detecting the presence of malicious activity in the system but fails to identify it clearly. In other weak attack-model scenarios with larger delays, the hyperbolae are even more dispersed, resulting in even larger residuals. Overall, residual analysis is misleading for this study as it either fails to detect the presence of an attack or fails to identify untrustworthy measurements. Another approach for building resilience against TSAs is proposed in the following section.

6.4 Calibration-Based Robust Localization

As mentioned above, the analysis of residuals during the localization of an unknown source is not sufficient to counter TSAs. In this section, we present a robust localization strategy that works in two phases. The first is a calibration phase that makes use of a known source to estimate the pairwise synchronized sensors of the network. The second phase of our strategy consists in the localization of an unknown source, given the results of the calibration process. We also show how to compute a confidence metric that gives insight on the accuracy of the estimated location.

6.4.1 Calibration Phase

In this first phase, we use authenticated received signals emitted by known sources of known coordinates. Our technique then compares the resulting TDOA measurements with the true delays that should be observed. These true values are easily computed from the known coordinates as shown in the previous section for the computation of specific attack delays. The aim of the calibration phase is to define weights w_{ij} for each pair of sensors (S_i, S_j), reflecting the confidence level of their time-synchronization. Recall that $\widehat{\Delta}_{ij} - \Delta_{ij}$ is distributed according to $\mathcal{N}(\mu_{ij}, \sigma_{ij})$ with $\mu_{ij} = 0$ if S_i and S_j are time-synchronized. Therefore, the weight w_{ij} must reflect the confidence with which we could declare that $\mu_{ij} = 0$ given the true delay Δ_{ij} and multiple samples of $\widehat{\Delta}_{ij}$, denoted $\widehat{\Delta}^1_{ij}$,...$\widehat{\Delta}^n_{ij}$. For example, the weights could have binary values in order to define hard

clusters within which sensors are time-synchronized:

$$w_{ij} = \begin{cases} 0 \text{ if } \mu_{ij} > \sigma_{ij} \\ 1 \text{ if } \mu_{ij} \leq \sigma_{ij} \end{cases}.$$

Nevertheless, given noisy observations, the true cluster can only be estimated with a certain level of confidence. In practice, hard clustering methods are not able to minimize both the probabilities of false positives and false negatives. Therefore, we use a soft clustering method, i.e., we allow non-binary values $w_{ij} \in [0, 1]$.

Hypothesis testing and the z-test in particular, are often used in order to determine whether a sample data-set is from a population with a specific mean. The z-test can be used only if the sample data is assumed to follow a normal distribution of known standard deviation as it is the case for the n samples of $\widehat{\Delta}_{ij} - \Delta_{ij}$. The test computes the standardized statistic

$$z_{ij} = \frac{\overline{(\widehat{\Delta}_{ij} - \Delta_{ij})}}{\frac{\sigma_{ij}}{\sqrt{n}}},$$

where $\overline{(\widehat{\Delta}_{ij} - \Delta_{ij})}$ corresponds to the sample mean of the n observed delay differences. If it is truly the case that $\mu_{ij} = 0$, then this standardized statistic z_{ij} must be distributed according to $\mathcal{N}(0, 1)$. Depending on the value of z_{ij} and a predefined threshold, the test either accepts or rejects the hypothesis that $\mu_{ij} = 0$. However, as mentioned above, the weights that we define are not constrained to have binary values. We define them as a function of the z-test p-values. Specifically, the weight w_{ij} is a function of the probability of observing a test statistic larger or equal to z_{ij} given that $\mu_{ij} = 0$. This probability is computed as $\text{erfc}(z_{ij}/\sqrt{2})$, where $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$ is the well-known complementary error-function. As the difference between the measured and the true delay decreases, the corresponding p-value increases. In order to amplify the weight differences between pairs of sensors with reasonably large and extremely small p-values, we define weights to be $w_{ij} = (pvalue_{ij})^{1/v}$, where $v \in \mathbb{R}$. The exponent v can be optimally chosen to maximize the weight difference for two specific p-values: $v = 15.0776$ maximizes the weight difference for p-values 10^{-4} and 10^{-10}. Nevertheless, our simulations show that all choices of $v \in [10, 30]$ give satisfactory localization results with negligible variance. The computed weights are used in the localization process of any unknown source until they are updated. This means that any TDOA measurement that results from the correlation of signals received by sensors S_i and S_j will be weighed by w_{ij}. Note that if all p-values are equal to zero, no sensor pair data will be trusted to be used in the localization process. If all p-values are very low but larger than zero, then all sensor data is taken into consideration. But our confidence in the synchronization of all sensor pairs is low, hence we expect low accuracy. Whereas if all p-values are high, the confidence in the synchronization is high throughout the network, and we expect to obtain accurate estimates.

In order to give insight about the level of accuracy with which the localization process is able to compute an estimate of the location of any unknown source, we propose to add the computation of a confidence metric to the calibration phase. As in two dimensions, the minimal number of measurements required to localize a source is two, the accuracy of a location estimate depends on how well the second most trustworthy sensor pair seems to be synchronized; this is captured by the second best p-value. Furthermore, the accuracy improves with redundancy, hence if the third best p-value is also high, we expect that the estimate will be even more accurate. Hence, our proposed confidence metric *cfd* is defined as the sum of the second and third best p-values to the power $1/v$, divided by two. Similarly in three dimensions, one level of redundancy is achieved by including the fourth best p-value and by dividing the sum by three instead of two. More discussion on this metric is provided in Section 6.6. The operations of the calibration phase are recapitulated in Algorithm 5. It shows how to compute the weights for each sensor pair and the confidence metric of the network at a given time.

Algorithm 5 Define-weights($\mathcal{N}, S_c, \sigma, c, \widehat{\Delta^1}, ..., \widehat{\Delta^n}, v, D$)

Input: $\mathcal{N}$ (network of sensors S_i, $1 \leq i \leq N$), S_c (known calibration source), σ (standard deviation of TDOA measurements), c (signal speed), $\widehat{\Delta^1}, ..., \widehat{\Delta^n}$ (n symmetric matrices of TDOA measurements from S_c), v (weight function exponent), D (dimension 2D or 3D)

for $S_i \in \mathcal{N}$ **do**
  $\Delta_i \leftarrow \frac{\text{dist}(S_i, S_c)}{c}$
end for
$weights \leftarrow \emptyset$
$pvals \leftarrow \emptyset$
for $(S_i, S_j) \in \mathcal{N}^2, i \neq j$ **do**
  $(e^1_{ij}, ..., e^n_{ij}) \leftarrow (\widehat{\Delta^1}_{ij} - \Delta_i + \Delta_j, ..., \widehat{\Delta^n}_{ij} - \Delta_i + \Delta_j)$
  $pvalue \leftarrow \text{z-test}(e^1_{ij}, ..., e^n_{ij}, \sigma_{ij})$
  $pvals \leftarrow pvals \cup pvalue$
  $weights \leftarrow weights \cup (pvalue)^{1/v}$
end for
if $D = 2$ **then**
  $cfd \leftarrow \frac{(\max_{2nd}(pvals))^{1/v} + (\max_{3rd}(pvals))^{1/v}}{2}$
else
  $cfd \leftarrow \frac{(\max_{2nd}(pvals))^{1/v} + (\max_{3rd}(pvals))^{1/v} + (\max_{4th}(pvals))^{1/v}}{3}$
end if
$weights \leftarrow \frac{weights}{\text{sum}(weights)}$

Output: $weights$, cfd

6.4.2 Robust-Localization Phase

The purpose of the second phase of our technique is to sporadically localize unknown sources by using data from sensors that are possibly suffering from a TSA. In Section 6.2, the WLS estimator was introduced with weights defined by the covariance matrix of the noise of the measurements. Our robust localization technique further weights the squared errors with the weights computed during the calibration phase. In other words,

our solution is to search for (x, y, z) minimizing

$$\sum_{i>j} \frac{w_{ij}}{\sigma_{ij}^2} \left(\frac{\text{dist}(S_i, S) - \text{dist}(S_j, S)}{c} - \widehat{\Delta_{ij}} \right)^2 .$$

Algorithm 6 describes how the robust localization phase works for a two-dimensional grid, when at least two sensor pairs have non-zero weights. It uses the function noise_std(γ_i, γ_j) to compute the noise standard deviation of TDOA measurement $\widehat{\Delta_{ij}}$ from the SNR values at sensors S_i and S_j according to equations 6.4, 6.5 and 6.6. In the case where all weights are set to zero, the algorithm states that the system is too corrupt to reliably estimate the location of the unknown source.

Algorithm 6 Robust-localization$(weights, \mathcal{N}, \widehat{\Delta}, \gamma, D)$

Input: $weights$ (computed by Algorithm 5), $\mathcal{N}$ (network of sensors), $\widehat{\Delta}$ (matrix of received TDOAs from unknown source), γ (vector of SNR values for each sensor), D (dimension 2D or 3D)

```
if |nonzero(weights)| ≥ D then
    σ ← ∅
    for (S_i, S_j) ∈ N^2, i ≠ j do
        σ ← σ ∪ noise_std(γ_i, γ_j)
    end for
    estimate ← WLS(weights/σ^2, M, N)
else
    estimate ← "corrupt_system"
end if
```

Output: estimate

Recall that in the weak attack-model scenario presented in Section 6.3, only sensor S_1 is attacked with a delay of $2.47\mu s$. When no defense strategy is in place, the LM algorithm on the full set of measurements gives a WLS estimate approximately 1 km away from the true source location. Using our robust localization technique with the LM algorithm, we obtain an estimate only 40cm away from the true source location. Furthermore, recall that in the strong attack-model scenario considered in Section 6.3, the attack delays were computed specifically such that the localization of a particular unknown source would result in a targeted location. The computed delays shown on Figure 6.1b are: $d_1 = 3ms$, $d_2 = 0s$, $d_3 = 5.8ms$ and $d_4 = 3.65ms$. The obtained estimate was, as chosen by the attacker, approximately 9 km away from the source. Whereas our technique flags all sensor pairs as not synchronized, and all weights are set to zero. As a result, our algorithm states that the system is too corrupt to give an estimate.

6.5 Countermeasures Against Replay Attacks

The calibration phase of our solution presented in Section 6.4.1, relies on signals received by the sensors and emitted by a calibration source of known coordinates. In this section,

we suppose that an attacker seeks to perform an attack on the calibration phase of our solution in order to make it attribute wrong weights. In case of such a successful attack, our robust localization would discard trustworthy measurements and/or trust attacked measurements. As a result, our technique would fail to detect malicious activity and would discard correct measurements.

In this section, we consider two additional attack models in which the attacker targets the calibration phase of our robust solution. His goal is to provoke a wrong weight attribution, thus maintaining the undetectability of his ongoing TSA or neutralising the localization system. In the weak calibration attack model, we suppose that the attacker is able to record signals from the calibration source in order to replay them at times and locations of his choice. For example, he could replay them in a manner that compensates for the introduced attack delays. In the strong calibration attack model, we further suppose that the attacker is able to jam signals emitted by the calibration source [126]. We assume that the attacker does not jam continuously but selectively. Note that a sensor being continuously jammed would be flagged as suspicious due to other identification methods, such as SNR analysis. The fact that an attacker with complete control on the flow of signals in the network would be all powerful, further justifies the assumption of a selective jamming.

In order to prevent such attacks, we propose an encrypted authenticated challenge-response scheme between the calibration source (CS) and the control center (CC). For the duration of this protocol, we assume that CS is stationary and emits signals continuously and that CC is responsible for triggering CS into embedding specific responses in the emitted signal. The first iteration of the scheme is depicted in Figure 6.2. We suppose that due to a key infrastructure, CC has a certificate binding its identity with its public key PK_{CC}. We further suppose that CS knows the public key of the certificate authority PK_{CA} and that CC and CS share two secrets p_1 and p_2 of large entropy. At the first iteration of the scheme, CC sends its certificate to CS who first verifies it using PK_{CA} and then extracts the public key PK_{CC}. The latter will be used to authenticate CC to CS, at all subsequent iterations of the scheme.

During the calibration phase, CC continuously sends the encryption of a one-time random challenge c concatenated with a flag $f \in \{0, 1\}$ that indicates whether or not CS should compute and embed a response r in the emitted signal. The use of this flag enables us to hide from the attacker the locations in the signal where there are embedded responses. These responses are used to compute the TDOA measurements that are analyzed during the calibration phase. Note that, in order to be successful, the attacker needs to succeed in delaying a number of these responses embedded in the signal. Using the flag, instead of sending useful responses continuously, decreases the probability that a useful response is delayed by an adversary. This is due to the fact that in order to be undetected, the adversary is required to jam only selectively. Therefore, if we assume that there is a maximal frequency with which he can jam signals while remaining undetected, if he does

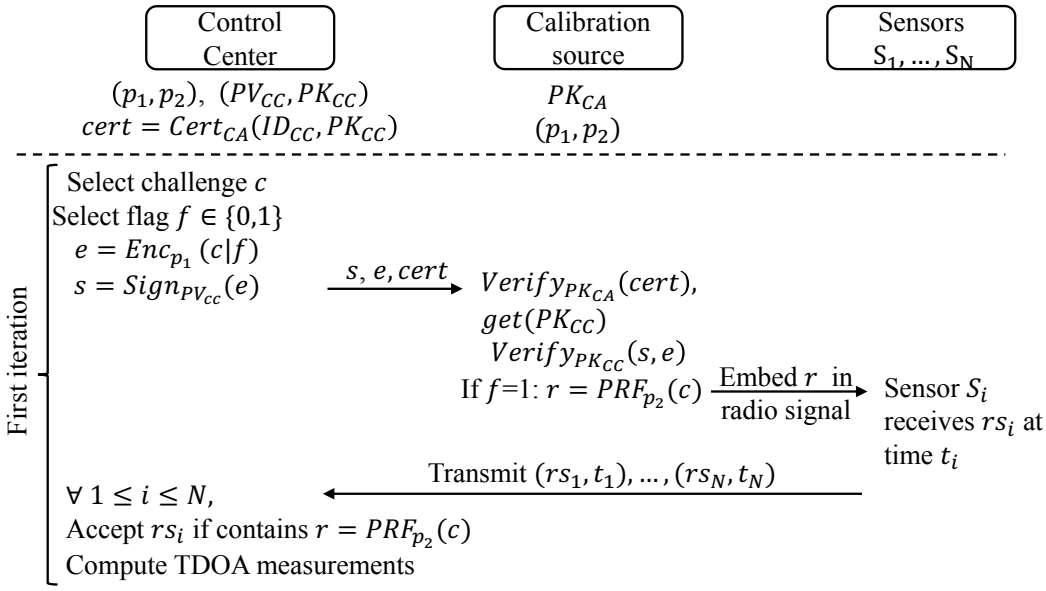


Figure 6.2 – First iteration of the encrypted authenticated challenge-response scheme that counters replay attacks against the calibration phase of our robust-localization technique; following iterations are identical but don't include the *cert* verification and the extraction of PK_{CC}.

not know which parts of the signal are useful and should be jammed, then the proportion of TDOA measurements that he successfully attacks is decreased. Note that, for this to be true, it is required that the flag be picked at random and that the signal emitted by CS without an embedded response be indistinguishable from when it contains a response.

The one-time challenge c concatenated with f, is encrypted into ciphertext e via a symmetric encryption scheme using secret key p_1. In order to authenticate itself to CS, CC also sends a signature of e computed using its private key PV_{CC}. Upon reception of (e, s), CS verifies the validity of the pair sent by CC. If it is valid, CS decrypts e using p_1 and extracts f. If the flag is equal to 1, then it computes a response r that corresponds to the result of a chosen pseudo-random function with c as input and secret p_2 as the key. Then, CS embeds r in the signal. In contrast, if the flag is equal to 0, CS does not embed anything in the signal. This signal is received by various known nearby sensors who simply timestamp everything before transmitting to CC. Finally, CC accepts only signals containing the valid response r that are transmitted by specific nearby sensors able to correctly decode signals from CS. It then correlates the accepted signals to obtain TDOA measurements. Note that to analyze the correlation of the signals, the signals are aligned according to the timestamp that was associated by the receiving sensors. Hence, introducing delays in the transmission of received signals between the sensors and CC does not affect the resulting TDOA measurements. This constitutes one iteration of the overall scheme. We denote by m, the number of iterations with f set to 1. Hence, at the end of the process, there are m resulting TDOA measurements per sensor pair. Recall that the calibration phase described in Algorithm 5 requires n observed measurements

for each sensor pair. Below, we discuss the selection of n among m measurements per sensor pair and analyze the security of our scheme firstly against a weak calibration attacker and secondly against a strong calibration attacker.

First, in order to enforce security against a weak calibration attacker, we propose to set $m = n$ and to discard signals containing a reoccurring valid specific response r. We show in Theorem 6.1 that this scheme is secure against a weak calibration attacker with overwhelming probability, i.e., the attacker is unable to inject delays in the TDOA measurements of the calibration phase. The resulting TDOA measurements are then given as input to Algorithm 5 as described in Section 6.4.

Theorem 6.1. *Assuming that challenges are unique, that p_1 and p_2 are secret, that signals with a reoccurring r are discarded, and that all relay attacks take more time than the direct signal to propagate to the sensors, then the encrypted authenticated challenge-response scheme with $m = n$ iterations is secure in the weak calibration attack model with overwhelming probability.*

Proof. Since p_1 and p_2 are secrets of large entropy and since c is unique and also a secret of large entropy, there is a negligible probability that an adversary manages to forge a valid response r that corresponds to the latest challenge c. Therefore, we assume that he can emit signals containing valid responses only by replaying those emitted by CS. As we assume that all relay attacks take more time than the direct signal takes to propagate to the sensors, all replayed signals will be received by CC with timestamps that are more recent than the timestamps of the direct signals. Hence, as they were already received by CC, replayed signals are all discarded. We conclude that the encrypted authenticated challenge-response scheme with $m = n$ iterations is secure in the weak calibration attack model, as long as the adversary is unable to forge an r, hence it is secure with overwhelming probability. □

Second, when we consider a strong calibration attacker able to jam signals, it is possible that the first occurring response r is a replayed signal. In this scenario, we assume that the attacker can selectively jam signals emitted by CS and replay them such that the direct signal is never received by the sensors. In order to enforce security in this strong attack model, we propose to set m much larger than n. In other words, we propose to select a small portion n of the received measurements m to use as input for Algorithm 5. As mentioned earlier, we suppose that the attacker does not jam continuously but selectively so that only qm measurements are successfully delayed by the adversary, where q is the proportion of attacked measurements. As the TDOA noise is distributed according to a Gaussian distribution, the $(1 - q)m$ unattacked measurements are distributed according to a Gaussian distribution centered in the observable TDOA value that depends on the coordinates of the sensors, the coordinates of CS and the possible time-synchronization offset between the sensors. In contrast, the qm attacked measurements are expected to

be distributed according to another Gaussian distribution centered around a value that depends on the delay difference with which the signal is replayed to the different sensors.

Our strategy is to analyze the distribution of the m observed measurements in order to extract the center of the tallest Gaussian distribution and to select the n measurements that are nearest to it. More specifically, we use a binning algorithm on all the received measurements; the algorithm returns b bins of uniform width covering the range of the m measurements. This reveals the shape of the sample distribution. We then extract the bin of highest density and iteratively extract its surrounding bins, until the sub-sampled dataset is of cardinality at least n. Then, we estimate the probability density function of the selected data and search for its peak value. As a result, we obtain an estimate of the center of the tallest underlying Gaussian distribution, in other words, of the TDOA value that should be observed. Finally, we select the n measurements that are nearest to the newly found estimate. Observe that this technique works in favour of the system only when there are fewer attacked than unattacked measurements during the calibration phase. Otherwise, the tallest Gaussian distribution would correspond to the underlying distribution of the attacked measurements. These selected measurements can then be given as input to Algorithm 5 as described in Section 6.4. We show the following claim through numerical analysis.

Claim 6.1. *As long as the proportion q of attacked measurements is lower or equal to* 0.45*, our strategy is secure against a strong calibration attacker.*

According to Theorem 6.1, if the attacker is unable to jam signals, our encrypted authenticated challenge-response scheme is secure against a replay attack and the resulting n measurements can then be given as input to Algorithm 5. Now, in order to show that Claim 6.1 is correct, we show that when an attacker is able to jam signals, successfully affecting qm measurements, our strategy to select n measurements is efficient, as long as $q \leq 0.45$. Specifically, we give numerical evidence that the weights that Algorithm 5 outputs from the n selected measurements with and without a calibration attack are similar. We analyze results in two scenarios.

In the first, the measured TDOA value comes from a perfectly synchronized sensor pair. Hence, we require that Algorithm 5 on the n selected measurements, outputs a weight close to 1. In this case, the goal of the attacker is to jam signals in order to replay them with introduced delays such that Algorithm 5 outputs a weight close to 0, thus making us discard trustworthy measurements. For every combination of calibration-attack size $p \in \{3, 6, 15000\}$ and calibration-attack proportion $q \in \{0, 0.1, 0.2, 0.3, 0.4, 0.45, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, we performed the following procedure ten thousand times:

- we create a data set DS_{160} of $m = 160$ i.i.d samples where a proportion q of the m samples are drawn from $\mathcal{N}(\mu + p\sigma, \sigma)$ and the remaining are drawn from $\mathcal{N}(\mu, \sigma)$,

with $\mu = 7e-7$ and $\sigma = 2.192e-9$,

- we use our selection technique with $b = 12$ bins, to extract the center of the highest Gaussian distribution and keep only the $n = 30$ nearest samples, resulting in DS_{30},
- for both DS_{160} and DS_{30}, we compute the p-value resulting from the z-test with $\mu = 7e-7$ and $\sigma = 2.192e-9$,
- we apply the weight function of Algorithm 5 to the two p-values: we exponentiate them to the power $1/v$ with $v = 15.0776$.

At the end of the procedure, for every combination of p and q, we obtain two data sets of ten thousand exponentiated p-values. Figure 6.3 shows the sample mean of the resulting weights with a confidence interval for every (p, q) pair, when using the entire $m = 160$ measurements and when using the selected $n = 30$ measurements. We observe that for a choice of $q \leq 0.45$, the weight obtained using the 30 selected measurements is always high as in the no-attack scenario, in other words, as when $q = 0$. Note that for $d = 3$ and $q = 0.45$, the weight is sometimes low although it is always large enough to ensure that the corresponding measurements are never incorrectly discarded. Such low weight values are due to the fact that the attack is small, comparable to large noise. Therefore the Gaussian distribution of the attacked measurements is merged with the Gaussian distribution of the unattacked measurements and the highest density peak is slightly shifted by the attack. When the attack is large enough for the intersection of the two Gaussian distributions to be empty, the attacked measurements are less likely to be mistaken for unattacked ones. Figure 6.3 also shows that the weight obtained with all measurements without any particular strategy, decreases drastically as the size of the calibration attack increases. From this analysis, we observe that our strategy is very efficient in protecting our trust in synchronized sensor pairs.

In the second scenario, the measured TDOA value comes from a sensor pair suffering from a TSA. Hence, we require that Algorithm 5 outputs a weight close to 0. In this case, the aim of the attacker is to attack calibration signals in order to compensate the synchronization delay between the attacked sensors. His goal is for Algorithm 5 to output a weight much higher than 0, thus making us trust attacked measurements. For every combination of calibration-attack size $p \in \{3, 6, 15000\}$ on the sensor pair, and calibration-attack proportion $q \in \{0, 0.1, 0.2, 0.3, 0.4, 0.45, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, we performed the following procedure ten thousand times:

- we create a data set DS_{160} of $m = 160$ i.i.d samples where a proportion q of the m samples are drawn from $\mathcal{N}(\mu, \sigma)$ and the remaining are drawn from $\mathcal{N}(\mu + p\sigma, \sigma)$, with $\mu = 7e-7$ and $\sigma = 2.192e-9$,
- we use our selection technique with $b = 12$ bins, to extract the center of the highest Gaussian distribution and keep only the $n = 30$ nearest samples, resulting in DS_{30},

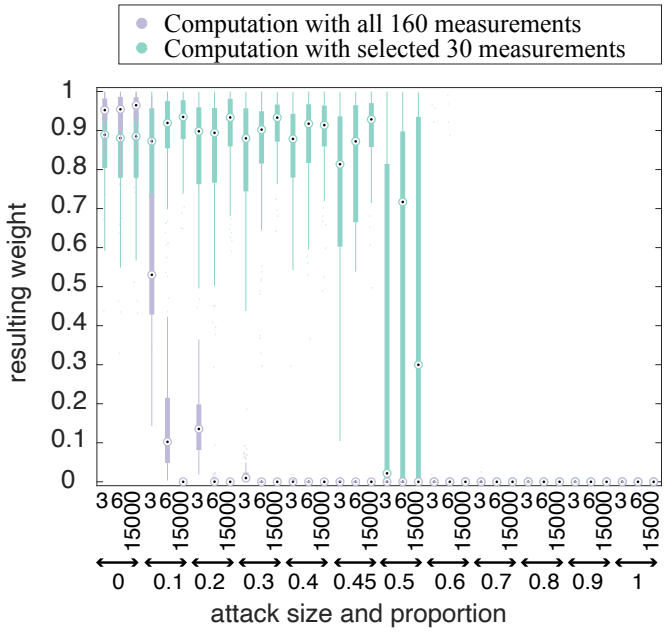


Figure 6.3 – Weight for an ideally synchronized sensor pair for different calibration-attack size and attack proportion from all 160 and the selected 30 measurements: when $q \leq 0.45$ the weight from the 30 selected measurements is always high as in the no-attack scenario; the weight from all measurements decreases drastically as the size of the calibration attack increases.

- for both DS_{160} and DS_{30}, we compute the p-value resulting from the z-test with $\mu = 7e-7$ and $\sigma = 2.192e-9$,

- we apply the weight function of Algorithm 5 to the two p-values: we exponentiate them to the power $1/v$ with $v = 15.0776$.

Figure 6.4 shows the sample mean of the resulting weights with a confidence interval for every (p, q) pair, when using the entire $m = 160$ measurements and when using the selected $n = 30$ measurements. We observe that for a choice of $q \leq 0.45$, both the weights obtained using the 30 and the 160 measurements are low. From this analysis, we observe that when $q \leq 0.45$, our strategy is still very efficient in protecting our distrust in non-synchronized sensor pairs.

In summary, the numerical evidence shows that as long as $q \leq 0.45$, our technique to select n out of the m received measurements is efficient and enables Algorithm 5 to define proper weights. We further recall that if q is selected too large, the attacker would be detected by other techniques such as SNR analysis.

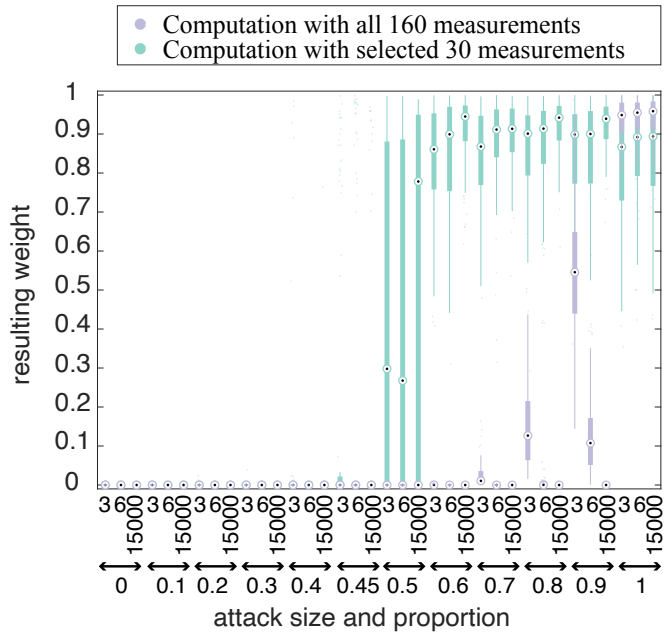


Figure 6.4 – Weight for a non-synchronized sensor pair for different TSA sizes and calibration-attack proportion from all 160 and from the selected 30 measurements: when $q \leq 0.45$, both the weights obtained using the 30 and 160 measurements are always low.

6.6 Performance Evaluation

In this section, we evaluate the performance of our solution presented in Section 6.4 by considering various attack scenarios. We show through Matlab simulations that our solution is robust to TSAs and that the confidence metric is reliable. We also show that in an adversarial environment, our solution outperforms usual localization and tracking methods. We start by defining the testing environment in two dimensions. We then consider a three-dimensional simulation.

6.6.1 Two-Dimensional Testing Environment

We consider a two-dimensional grid of side of 20 km on which we place four sensors, as illustrated in Figures 6.1a and 6.1b. We assume that at the time instant of the analysis, the unknown source is located at coordinates $[3333.3, -889.1111]$ and the known calibration source at $[0, -4000]$ all in meters, with respect to the center of the grid. In Section 6.6.4, we show a three-dimensional simulation on the same grid, to which we add altitude coordinates. In our simulations, we assume that the TDOA noise is i.i.d with a standard deviation of $2.192ns$ for all measurements. We computed this standard deviation from Eqs. (6.4), (6.5), (6.6) with the following parameters:

- the integration window of the signal $T_{int} = 60$ ms,

- the bandwidth of the signal $W = 1$ MHz,
- the center of frequency $f_0 = 30'000$ Hz,
- the SNR $\gamma_i = 3$ dB $\forall S_i$.

For the calibration phase, we created simulated TDOA measurements $\widehat{\Delta}_{ij}$ from the *known* calibration source for all sensor pairs (S_i, S_j) with the following procedure:

- we use the coordinates to compute the true time Δ_i taken by the signal to propagate from the calibration source to sensor S_i for both sensors,
- we compute the true TDOA $\Delta_{ij} = \Delta_i - \Delta_j$,
- depending on the attack scenario, we add delays $\Delta'_{ij} = \Delta_{ij} + d_i - d_j$,
- we add Gaussian noise $\widehat{\Delta}_{ij} = \Delta'_{ij} + e_{ij}$ with $e_{ij} \in \mathcal{N}(0, 2.192e-9)$,
- we repeat this $n = 15$ times in order to have 15 measurements for each pair of sensors.

Our calibration procedure defined by Algorithm 5 then uses the simulated measurements in the following way:

1. We compute the observed error $\widehat{e_{ij}} = \widehat{\Delta}_{ij} - \Delta_{ij}$ for each measurement, *using the known coordinates* of the calibration source.
2. For each pair of sensors, we compute the p-value resulting from the z-test with the 15 observed errors $\widehat{e_{ij}}$.
3. We compute the weights w by exponentiating all p-values to $1/v$ with $v = 15.0776$ and normalizing them.
4. We compute the confidence metric as the sum of the second and third largest weights before normalization, divided by two.

Recall that the exponent value was chosen to maximise the weight difference between p-values 10^{-4} and 10^{-10} and that we experimentally observed that all choices of $v \in [10, 30]$ give satisfactory results with low variance between them.

For each pair of sensors such that the corresponding weight is non-zero, we create a simulated TDOA measurement $\widehat{\Delta}_{ij}$ from the *unknown* source to localize:

- we use the coordinates to compute the true time Δ_i taken by the signal to propagate from the unknown source to sensor S_i for both sensors,

- we compute the true TDOA $\Delta_{ij} = \Delta_i - \Delta_j$,
- depending on the attack scenario, we add delays $\Delta'_{ij} = \Delta_{ij} + d_i - d_j$,
- we add Gaussian noise $\widehat{\Delta}_{ij} = \Delta'_{ij} + e_{ij}$ with $e_{ij} \in \mathcal{N}(0, 2.192e-9)$.

We implement the robust localization with the simulated measurements as in Algorithm 6:

1. We compute a geometrical estimate of the source location (x_g, y_g) as explained below.
2. We use the Matlab LM algorithm as a WLS estimator on all $\widehat{\Delta}_{ij}$ with weights w_{ij} computed at step (3). The initial step size is by default 0.01 and the initial solution is (x_g, y_g). We obtain the estimated robust source coordinates (x, y).

In all simulations, we analyze the confidence metric and the distance between our estimate and the true source coordinates.

Recall that the WLS estimator we use is the LM algorithm. It is a gradient descent algorithm that requires an initial solution and step size that are updated iteratively. When the gradient is small, the step size is chosen small so that we can move gradually closer to the minima without missing it; in this case the algorithm is similar to the Gauss-Newton method. In contrast, when the gradient is large, the step size is chosen large and the algorithm behaves similarly to the steepest descent method. The initial solution we use is a geometrical estimate computed as the coordinate-wise weighted median of intersection points of all hyperbolae. The weight of both coordinates of an intersection point corresponds to the smallest weight among the weights of the corresponding TDOAs.

6.6.2 Performance in Attack Scenarios

In order to test the performance of our technique, we apply it in five different scenarios of attack with increasing attack delays. Each attack scenario corresponds to an attack location, specifically to a subset of the sensors. In all of the scenarios, we perform attacks with 25 different delays ranging from 0 to 50 seconds. We simulated ten thousand times each attack scenario with each delay size. The $1'250'000$ results we obtained are presented in Figure 6.5, where each color corresponds to a specific attack scenario. More specifically, Figure 6.5a shows the confidence metric as a function of the distance between the true source position and our estimate in meters for each simulation. We refer to this distance as the estimate error. Figures 6.5b and 6.5c show the sample mean and confidence interval of the estimate error as a function of the delay size in seconds.

The first scenario is a control scenario in which no attack takes place, it is presented in beige on the figures. We observe that the estimate error is on average below 0.5m and that the confidence metric is always high, above 0.8.

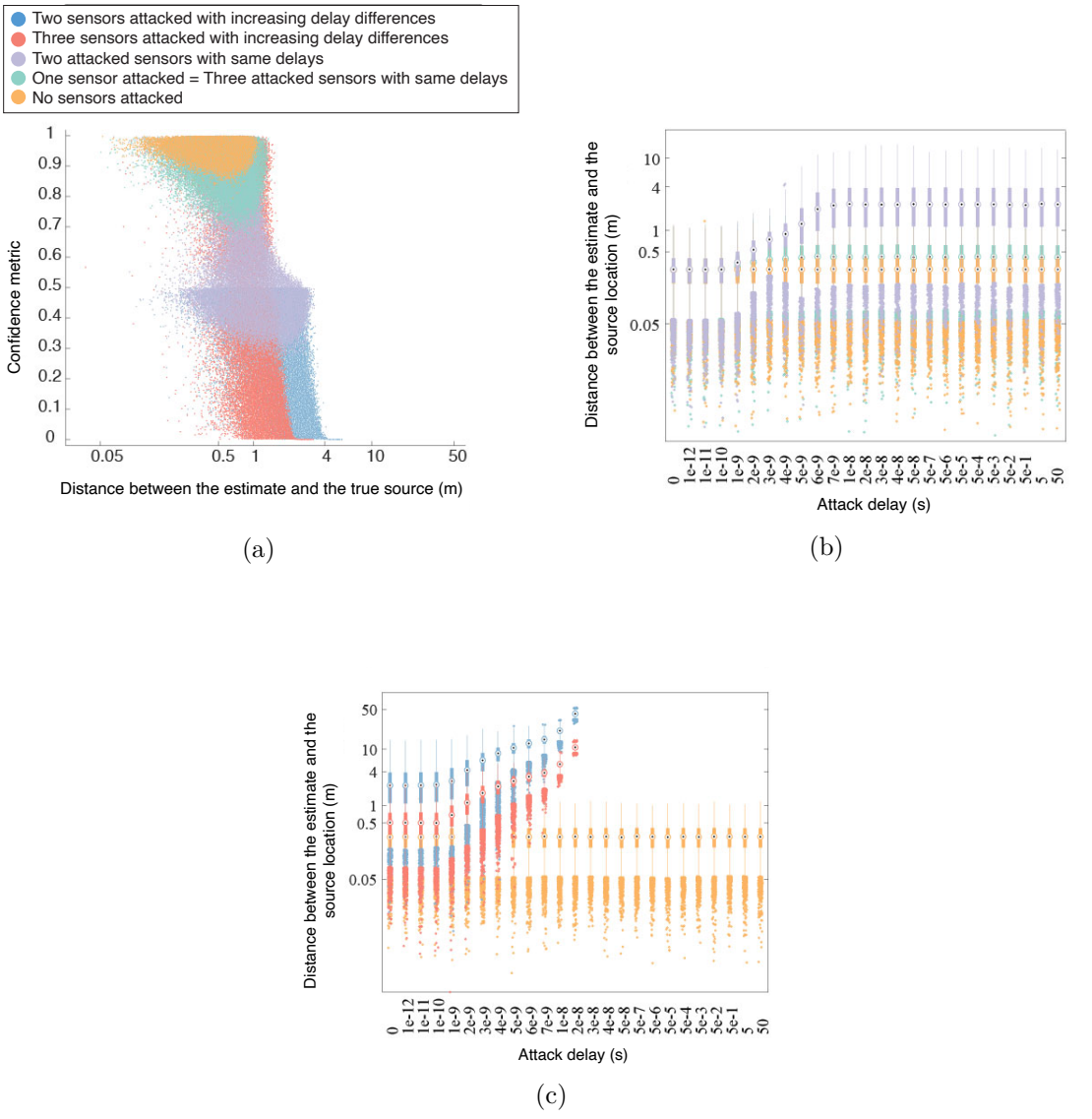


Figure 6.5 – Results from five different TSA scenarios each with 25 different delays, each simulated 10′000 times: (a) confidence metric as a function of the distance between the true source position and the estimate provided by our robust solution, we observe that the metric is related to the accuracy and shows if there is redundancy in the measurements. There are some false-negative cases but no false positives. (b) the mean and confidence interval of the estimate error for each attack delay for three different scenarios. (c) the mean and confidence interval of the estimate error for each attack delay for two other scenarios: when the system is too corrupt, it stops localizing.

Then, we consider a scenario where only sensor S_1 is under attack, it is presented in pastel green on the figures. We observe that whatever the delay injected, the estimate is always quite accurate with an error on average slightly above the one from the no-attack scenario. The estimate error is still in the proximity of 0.5m and remains below 4m. Figure 6.5a shows that the confidence metric is also quite high as it remains above 0.7. The slight reduction of estimate error comes from the fact that discarding signals from

S_1 reduces the redundancy but still provides enough signals to locate accurately with one level of redundancy.

The third scenario, presented in purple, consists of attacking two sensors, S_1 and S_2, with the same delay. We observe that when the injected delay is below the standard deviation of the noise, the estimate error is as in the no-attack scenario. Then, as the delay increases, the distance between the estimate and the true source position also increases until it stabilises. This is explained by the fact that as the delay increases, the impacted TDOAs are less trusted but are still taken into account with a small weight, until they are completely discarded. At some point, only two TDOA values are trusted and used for localization. This is exactly enough, as the simulation is in 2D, but removes all redundancy. Therefore, the estimate grows less accurate. Nevertheless, the estimate error stays well below 4m on average. This decrease in accuracy is accompanied by a decrease of the confidence metric value that is concentrated around 0.5 and remains between 0.35 and 0.7. This illustrates that the source can be localized fairly correctly but with less accuracy as there is no redundancy.

The fourth scenario, presented in light blue, is performed by attacking S_1 and S_2 with the same delay of 500s, such that only two TDOAs are used from the start of the simulation. Then, we increase slightly the delay difference between the sensors of synchronized pairs. Specifically, we attack in the following way:

- delay for S_1: 500s,
- delay for S_2: $(500 + d)$s,
- delay for S_3: 0s,
- delay for S_4: $(0 + d)$s,

where d is the delay difference that takes values from the same 25 different delays considered above. We observe on Figure 6.5c that the distance between the true source position and our estimate starts as in the previous scenario, which is as expected because only two TDOAs are trusted. Then, as the delay differences increase, the two TDOAs are increasingly affected, thus the localization relies solely on wrong TDOAs and the resulting estimate grows less accurate. We observe that when the delay differences are around $30n$s, the two TDOAs are declared as untrustworthy and the overall system as too corrupt to localize. In the worst case scenario, the estimate error is approximately 50m but the corresponding confidence metric is around 2×10^{-21}, which is extremely low. For this simulation, Figure 6.5 shows that the estimates are less accurate and that the confidence metric is low, below 0.35. Nevertheless, we can identify a grey zone, where the estimate error is below 4m and the confidence metric is also low. In this case, a user of our solution might want to discard the estimate when in fact it is not very far from the

true position of the source. Although this is unfortunate, it constitutes a false negative that is not as fatal as trusting a very inaccurate estimate.

We identified more cases of false negatives in the fifth scenario. In this setting, we start by attacking sensor S_4 with a delay of 500s. In this way, the TDOAs with respect to S_4 are discarded from the start. Then, similarly to the fourth scenario, we increase slightly the delay differences between the three synchronized sensors:

- delay for S_1: 0s,
- delay for S_2: $(0 + d)$s,
- delay for S_3: $(0 + 2d)$s,
- delay for S_4: 500s.

The results of this scenario are depicted in coral red in Figure 6.5. The accuracy of the estimate decreases in a fashion similar to the previous scenario but with lower error values. This is due to the fact that three TDOA values are used instead of two. Even though they are under attack, they include one level of redundancy. The largest distance between the source and the estimate remains below 10m. After that, the system is declared as too corrupt. Although the accuracy is better in this scenario than in the previous one, the confidence metric ranges also between 0 and 0.35. Similarly, there are cases of accurate estimation but low confidence, which constitutes false negatives.

In summary, our method gives an estimate which is always quite accurate, considering the fact that blindly trusting all TDOAs would lead to errors of many kilometers. When the system is too corrupt, it declares that the confidence level is too low to localize. We showed that the confidence metric gives useful insight on the accuracy of the estimate, although it can lead to some false-negative cases. We have not found any corner cases of false positives, in other words, our solution never trusts a highly inaccurate solution. If we were to recommend a course of action depending on the confidence metric, it would be the following:

- confidence metric $\in [0.75, 1]$: trust the estimate to be as accurate as it can be because it includes at least one level of redundancy,
- confidence metric $\in [0.3, 0.75]$: probably computed with no redundancy, trust the estimate to be fairly correct but slightly less accurate,
- confidence metric $\in [0, 0.3]$: the true source is in a probable zone around the estimate, this result is not very accurate and trusting it depends on the application,
- algorithm 6 output is *"corrupt_system"*: the attacks are too important to define even a probable zone of location.

6.6.3 Comparison of Robust and Naive Estimates

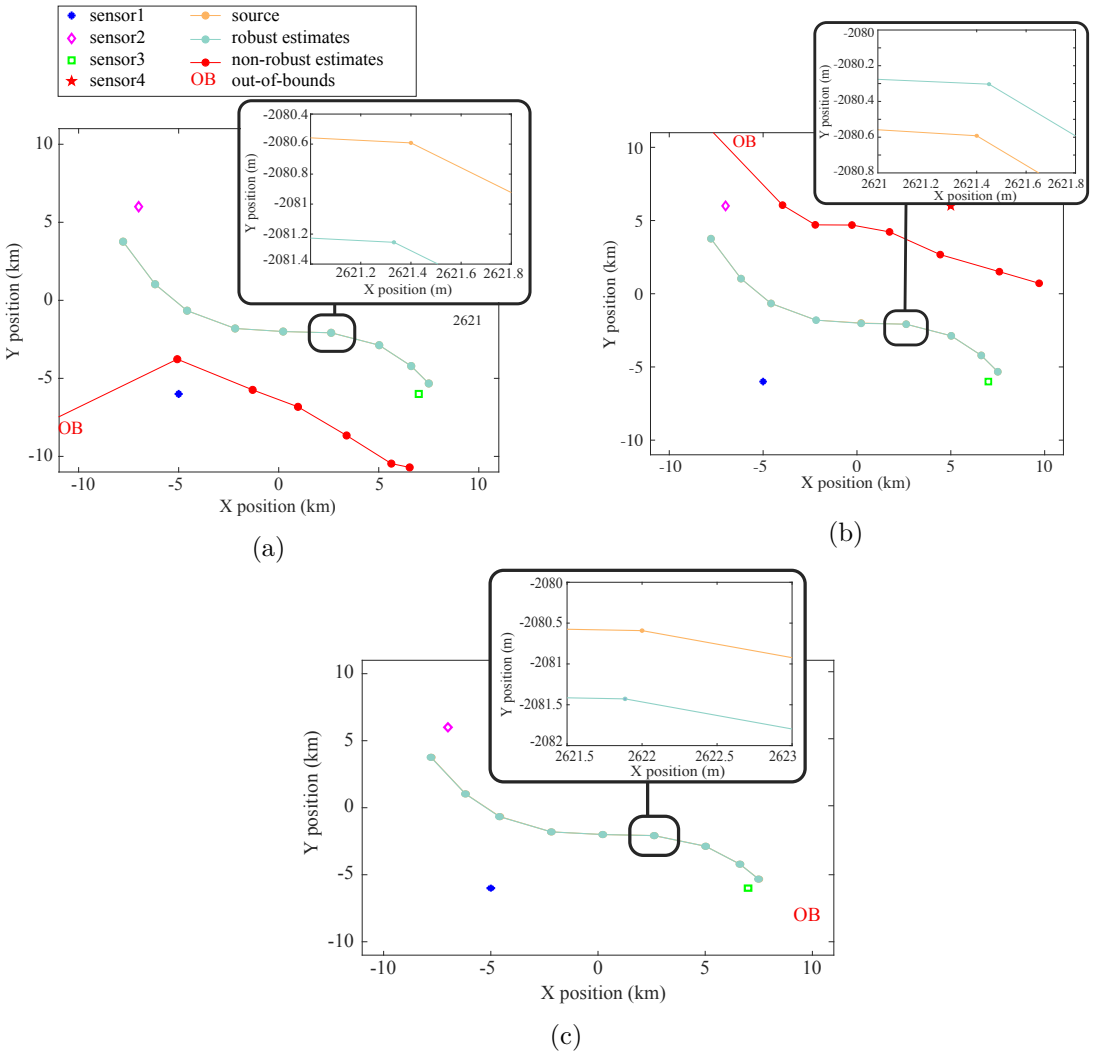


Figure 6.6 – Estimation results at nine different time instants when (a) S_4 is attacked by 30μs, the confidence metric is above 0.86: the first three naive estimates are out-of-bounds and then wrong by 2 km, whereas our solution provides estimates indistinguishable from the source; at the sixth time instant, our robust estimate is less than 60cm away from the source; (b) S_1 and S_3 are attacked by 30μs, the confidence metric is above 0.38: the naive estimate is out-of-bounds at first and then wrong by more than 5 km, whereas our solution provides estimates indistinguishable from the source; at the sixth time instant, our robust estimate is less than 50cm away from the source; (c) S_1, S_2 and S_4 are respectively attacked with a delay of 5s, $(5 + 3e - 9)$s and $3e - 9$s; the confidence metric is below 0.25: the naive estimates are out-of-bounds at all time instants, whereas, our solution provides estimates indistinguishable from the source; at the sixth time instant, our robust estimate is less than 1m away from the source.

Next, we simulate the localization of an unknown source at nine different time instants and compare the true estimates, the estimates obtained with our robust solution, and the estimates obtained naively by trusting all measurements. The naive estimates are found by the WLS estimator with all weights set to 1. This section shows that our solution substantially improves the accuracy of the estimates when compared with the accuracy

of the naive estimates obtained by ignoring the presence of TSAs. We do so in three different attack scenarios, one for each confidence metric interval identified above.

In the first scenario we consider, only sensor S_4 is attacked with a fixed delay of 30μs. The results are given in Figure 6.6a. For ease of understanding, we connected the estimates with straight lines of the same color. Shown as out-of-bounds, the naive estimates are more than $100'000$ km away from the true source position at the first three time instants. Such obviously incorrect estimates would be flagged as bad data as it is not plausible to consider an estimate that far. Then, for the six remaining time instants, the naive estimates are off by more than two kilometers. In contrast, when we zoom closer to the source at the sixth time instant, we can observe that the distance between the true source and our robust estimate is always under 60 cm. Note that the confidence metric is always above 0.86 in this scenario.

In the second scenario, sensors S_1 and S_3 are both attacked with a fixed delay of 30μs. The results given in Figure 6.6b, show that the estimate obtained naively is out-of-bounds at the first time instant. In fact, it is incorrect by more than 20 km at first and by approximately 5 km afterwards. In contrast, our solution provides estimates that are much more accurate, as they are all under 2m away from the true position. In this scenario, the confidence metric is always above 0.38.

In the last scenario, we attack sensors S_1 and S_2 with a delay of 5 seconds and we add a delay of $3n$s to S_2 and S_4. Namely, sensor pairs (S_1, S_2) and (S_3, S_4) are believed to be time synchronized when in fact, they have a delay difference slightly above the usual noise standard deviation $2.192n$s. In this scenario, the confidence metric is always well below 0.25. Figure 6.6c shows that the naive estimates are always out-of-bounds as they are wrong by more than $1'000'000$ km. Figure 6.6c also shows that our estimate is always very close to the source. More precisely, our solution provides estimates that are always under 5m away from the true source position, except for a corner case at the third time instant where our estimate is off by 15m.

6.6.4 Three-Dimensional Simulation

We performed a three-dimensional simulation where sensor S_1 is attacked by 30μs as before. We added altitude coordinates 600m, 1250m, 900m and 700m to S_1, S_2, S_3 and S_4, respectively. In order to achieve the same level of redundancy as before, we placed a fifth sensor on the grid at an altitude of 400 meters. From each TDOA, we computed the corresponding hyperboloids and used the coordinate-wise weighted median of intersection points as initial solution (x_g, y_g, z_g) to the WLS estimator. The results are illustrated in Figure 6.7. We observe that the altitude of the naive estimates fluctuate far from the true altitude of the source. Whereas, our robust solution provides accurate estimates indistinguishable from the source. Similarly, Figure 6.7 shows that the estimates obtained

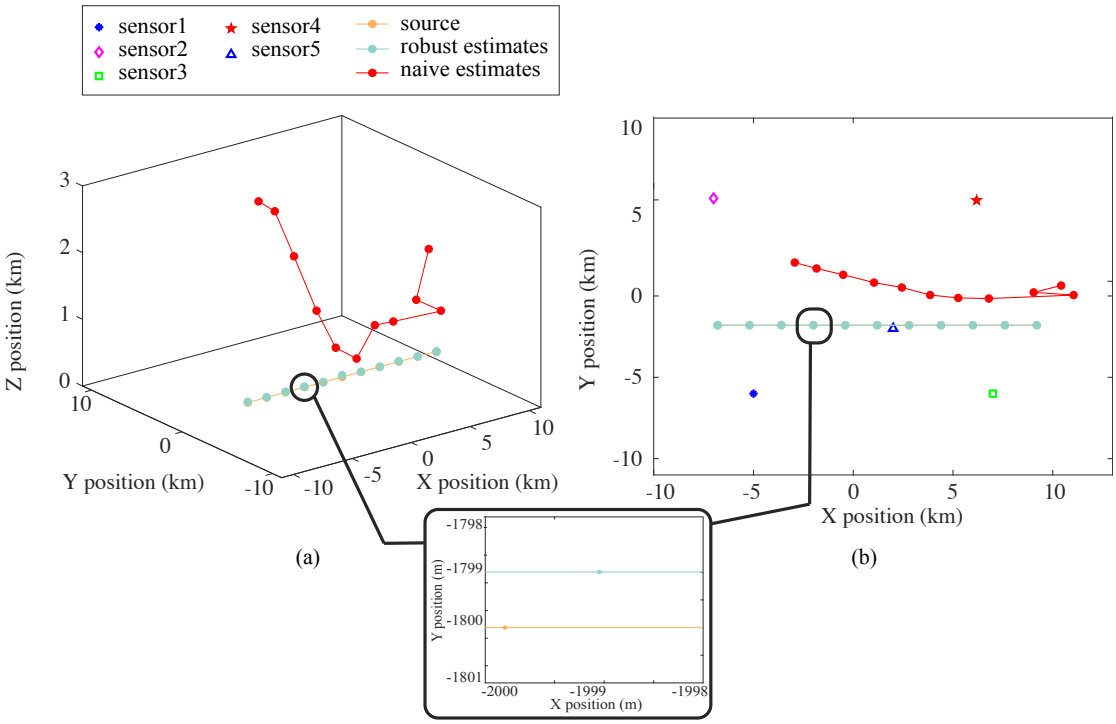


Figure 6.7 – Estimation results in $3D$ at eleven different time instants when S_1 is attacked by $3e-5$s; the confidence metric is high at 0.94: (a) The altitude of the naive estimates is very inaccurate at all time instants, our robust solution is indistinguishable from the source. (b) in the xy-plane, the naive estimates are more than 2km wrong and our robust estimates are accurate. Zoom-in at time instant 4: our robust estimate is less than 2m away from the source on the xy-plane.

with our solution on the xy-plane, matches with the source, whereas the naive estimates are always more than 2km wrong. Finally, a close-up look in 2D at time instant number 4 shows that our estimate is less than 2m away from the source. However, not shown on the figures, at time instant number 4, our estimate's altitude is off by approximately nine meters, whereas the naive estimate's altitude is off by more than a kilometer. Note that for this simulation, with weight function exponent $v = 15.0776$, we obtain a high confidence metric equal to 0.94. In this simulation, the true source is at a constant altitude of 350m, which is below the height of approximately the twenty highest towers on Earth. The minimal altitude found with the naive estimates is at 752m, which is above all towers in the region of interest. A naive estimator would fail to detect a potential collision danger in this case. In contrast, the altitude of our robust estimates is always between 339m and 377m, which allows us to detect a dangerous flight behaviour.

6.6.5 Comparison with Traditional NLOS Tracking Solutions

Lastly, we show that our robust solution outperforms classical solutions used for mobile source tracking in LOS and NLOS environments. The other localization techniques that we consider are [127]:

- an Unscented Kalman Filter (UKF) [128]: this filter offers better accuracy than the classical Extended Kalman Filter (EKF) because it does not require a first order Taylor approximation to linearize the non-linear measurement-to-state function.
- a Particle Filter (PF) [129]: this solution allowed us to track the general trajectory of the source but did not offer very accurate estimates. In our environment, we observed that this solution works when using a large number of particles, we chose 5000 particles. In a smaller environment, such as an indoor environment, it was shown that the PF offers much better accuracy. The accuracy is also improved when the initial estimate is closer to the truth, this is typically the case when tracking a source that we control, i.e. when the control vector is known.
- the Residual Weighting Algorithm (Rwgh): this solution was originally proposed by Chen in 1999 [130] in order to estimate the position of a source in an NLOS environment from time of arrival (TOA) measurements. Since then, various modifications [131] have been proposed either to reduce its complexity or to include different types of measurements, such as time difference of arrival (TDOA) measurements. This NLOS solution works as follows:
  - compute the weighted least squares (WLS) source-position estimates from all combinations of at least two measurements (in 2D): we call them the temporary estimates,
  - for all temporary estimate, compute the mean of their corresponding temporary residuals,
  - compute the final estimate as a weighted mean of all the temporary estimates weighted by the mean of their temporary residuals.

As mentioned in Section 6.1, we are interested in furtive, fast objects flying erratically along unpredictable trajectories. However, in order to provide a thorough comparison, we compare the results obtained on both predictable and unpredictable trajectories. To this end, we first compare the four solutions in unattacked scenarios on three different trajectories: one predictable straight-line trajectory illustrated on Figure 6.8a, one predictable curved trajectory illustrated on Figure 6.8b and one completely unpredictable trajectory illustrated on Figure 6.8c. In all cases, we consider a moving source flying at 300km/h and that there are 5s between each observation time, hence, there are approximately 417m between each observation point. The initial point we give to all estimators is $[-3500, -4500]$, we chose it because the coordinates are round numbers and it is close enough to the true first point that the PF can find it with a reasonable number of particles. These simulations are computed on the same grid as in the 2D simulations presented above in Sections 6.6.2 and 6.6.3. On Figures 6.8a, 6.8b and 6.8c, the true trajectory of the source, the estimates offered by our robust solution and the estimates offered by Rwgh are indistinguishable. The median of the error offered by our robust solution and by Rwgh are approximately 0.2m and 1m, respectively for the three

trajectories. Hence, they appear to be reliable and unaffected by the predictability of the trajectory. This is not the case for the UKF and PF: the median of the error on predictable trajectories are approximately 0.3m and 90m, respectively, whereas they are 9m and 240m for the unpredictable trajectory.

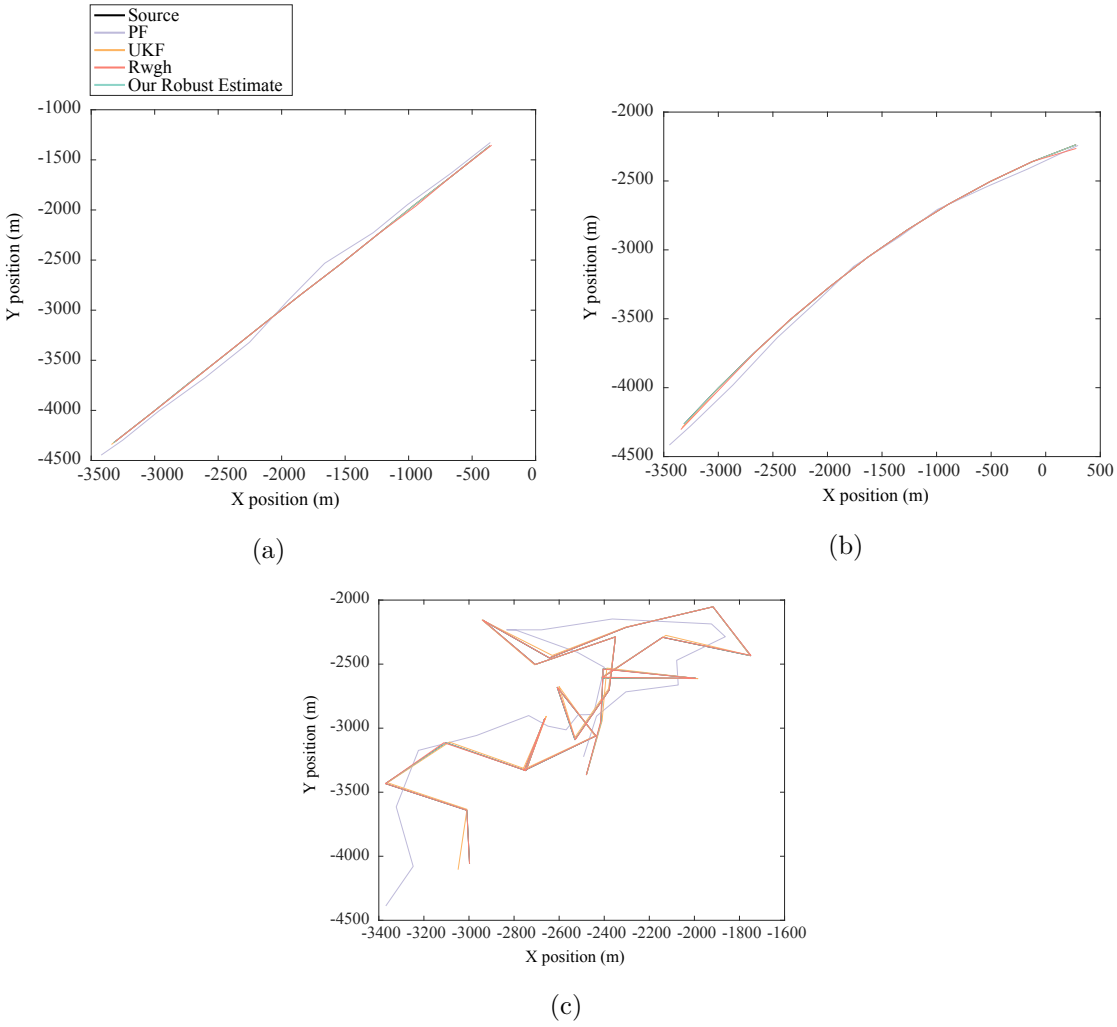


Figure 6.8 – Considered trajectories in unattacked scenarios and comparison of the tracking results according to the PF, UKF, Rwgh and our robust solution: (a) predictable straight-line trajectory; (b) predictable curved trajectory; (c) unpredictable trajectory. Our robust solution and the Rwgh algorithm offer the best accuracy on all trajectories. The UKF and the PF are much less accurate on unpredictable trajectories.

On the predictable curve and on the unpredictable trajectories, we compare the mean and confidence interval of errors of the estimates obtained with our robust solution, the UKF, the PF and Rwgh as we increase the attack delay in the same attack scenarios as in Section 6.6.2:

- Predictable trajectory and increasing delays on 1 sensor: Figure 6.9a shows the

error of the four considered estimators as a function of the increasing attack delay when only Sensor 1 is attacked. We notice that the error of our robust solution is only slightly increased as the attack delay increases because we loose one level of redundancy in the measurements. The UKF and the Rwgh solutions are also accurate until the attack reaches 50ns, then the estimates become drastically erroneous. The solution of the Rwgh regains a reasonable accuracy as the attack increases even more because the unreliable sets of measurements will have weights close to 0. In comparison, the UKF only looses accuracy and becomes unreliable as the attack delay gets too large. We observe that the error of the PF remains constant around 100m, this lack of accuracy is due to the large area not easily covered by a reasonable amount of particles. Then, when the delay reaches 50μs, the error increases to approximately 1km. Observe that in NLOS scenarios, the delays are often larger than micro-seconds and impact a small amount of measurements. According to Figure 6.9, in such scenarios, the Rwgh solution has a reasonable accuracy because it is expected that the weights given to the impacted sets of measurements are already close to 0.

- Predictable trajectory and increasing delays on 2 sensors: When attacking two sensors, we still have enough correct measurements to localize but without any redundancy. As expected, the results shown on Figure 6.9b have a similar behaviour as when only one sensor is attacked. The only difference is that the estimate errors are slightly increased as there is no redundancy left in the non-attacked measurements.

- Predictable trajectory, only two pairs of synchronized sensors and gradual increase of the delay between the sensors of each pair: Figure 6.9c shows that the error of our robust solutions increases to a maximum of 50m before it declares the system is too corrupt to be trusted for localization. The Rwgh algorithm has a similar behaviour but does not recognize the system as too corrupt and therefore continues to propose very erroneous estimates. The UKF is out-of-bounds from the start and the error of the PF remains stable around 1km as the delay increases.

- Predictable trajectory, only three synchronized sensors and gradual increase of the delays between them: Figure 6.9d shows similar behaviours as in the previous scenario.

- All above four scenarios on the unpredictable trajectory: Figure 6.10 shows the same general behaviours as for the predictable trajectories: our robust solution is the best in all scenarios and declares that the system is too corrupt to localize if it is the case. The Rwgh looses accuracy as the delay increases until the weights of attacked measurements drop close to 0 and produces very erroneous estimates when the system is too corrupt to localize. The estimates of the UKF and PF are always less accurate than when the source follows a predictable trajectory. The UKF drastically looses accuracy as the attack delay increases above three times

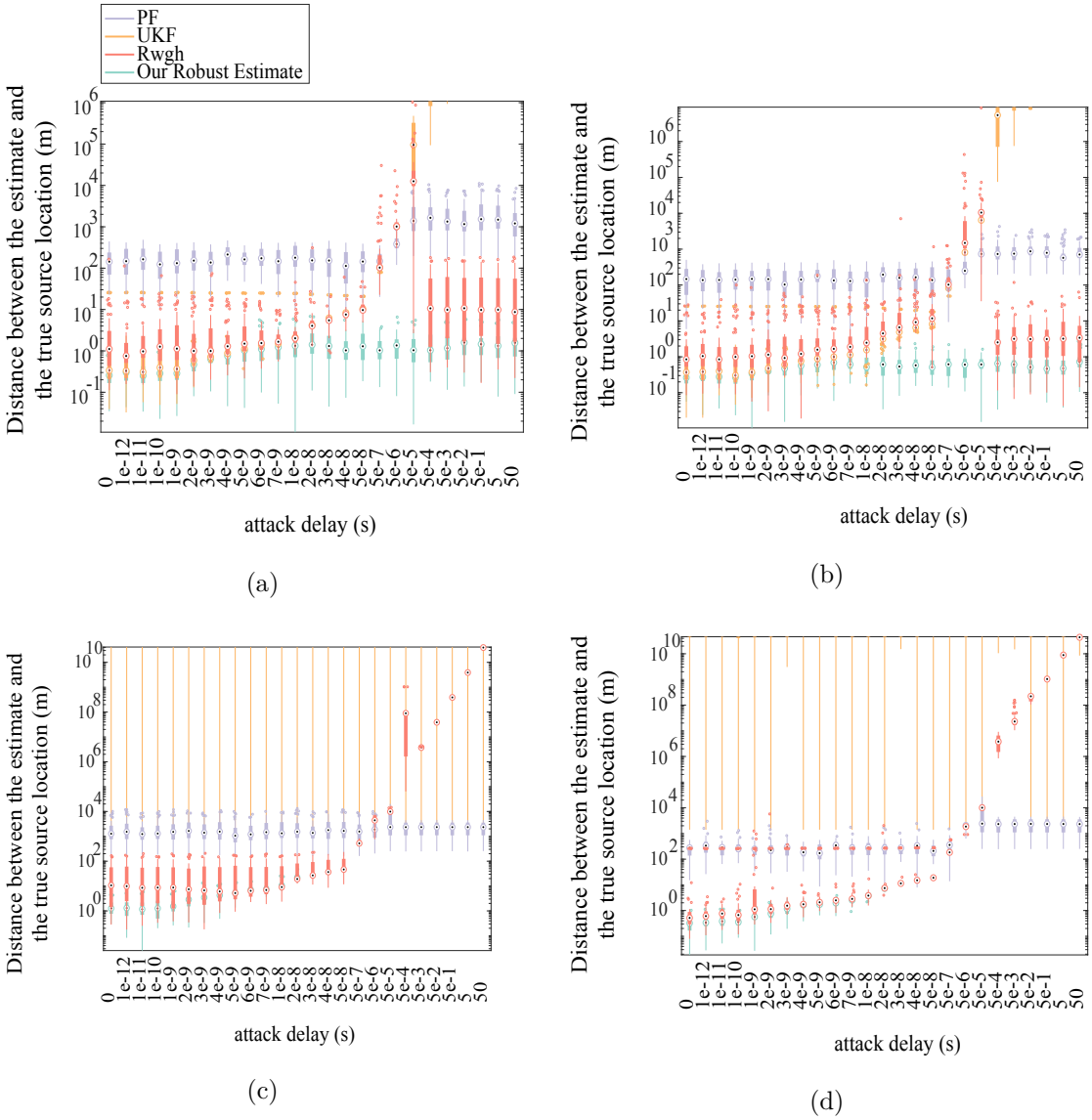


Figure 6.9 – Mean and confidence interval of estimates obtained with the four considered estimators on the predictable curve trajectory as a function of the attack delay: (a) when the attack delay is applied only to sensor 1; (b): when the attack delay is applied to sensors 1 and 2; (c) when the delay increases between the sensors of two synchronized pairs; (d): when the attack delay increases between three synchronized sensors. Our robust solution remains accurate and stops localizing when not enough measurements can be trusted, the Rwgh solution looses accuracy while the weights of attacked measurements are non-zero and gives unreliable estimates when the system is too corrupt to localize, the UKF looses accuracy similarly to Rwgh but does not regain accuracy later, the PF is not accurate from the start and keeps a stable accuracy at around 1km as the delay increases.

the standard deviation of the measurement noise. The error of the PF estimates is stable around 300m and above 1km when the attack delay is below and above three times the standard deviation of the measurement noise, respectively.

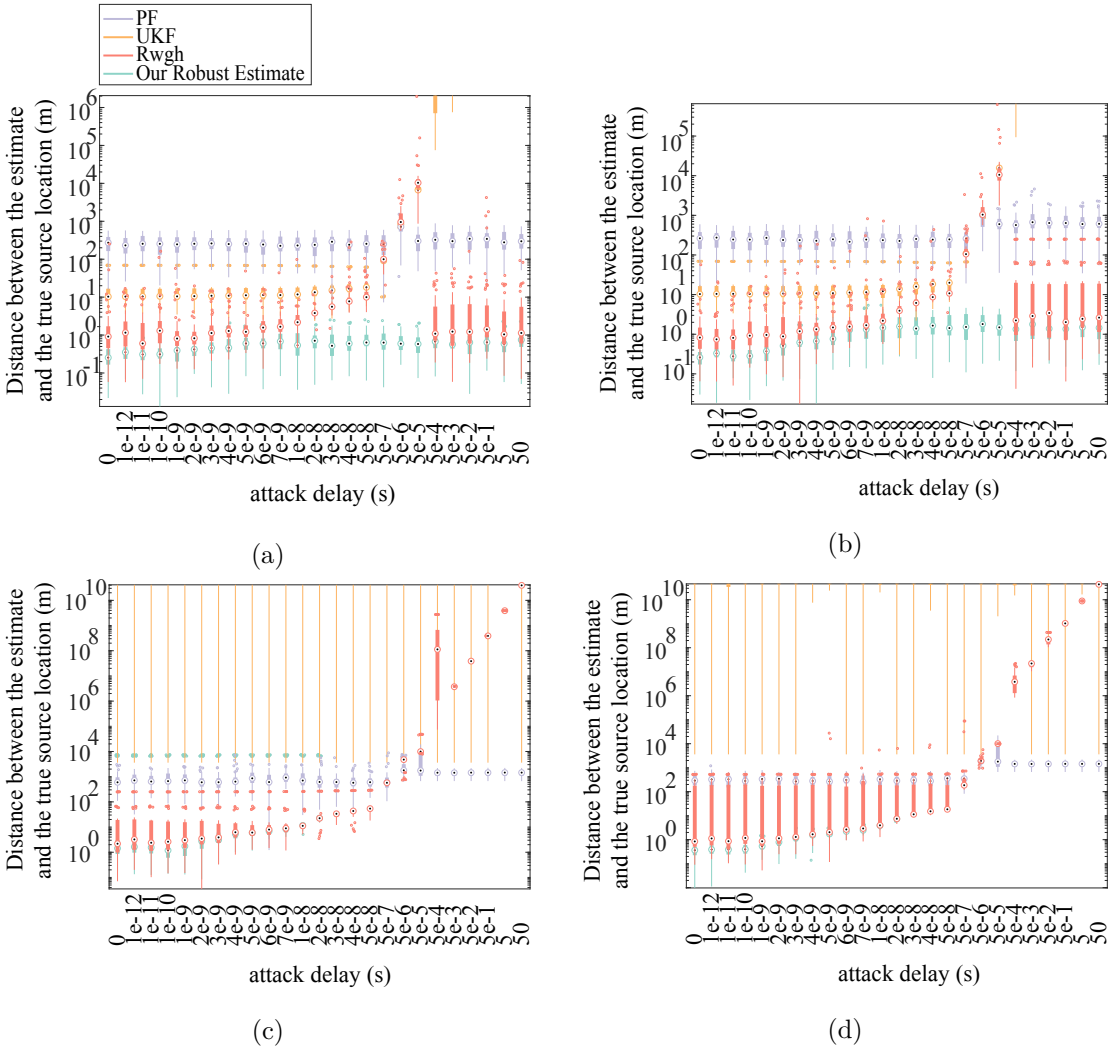
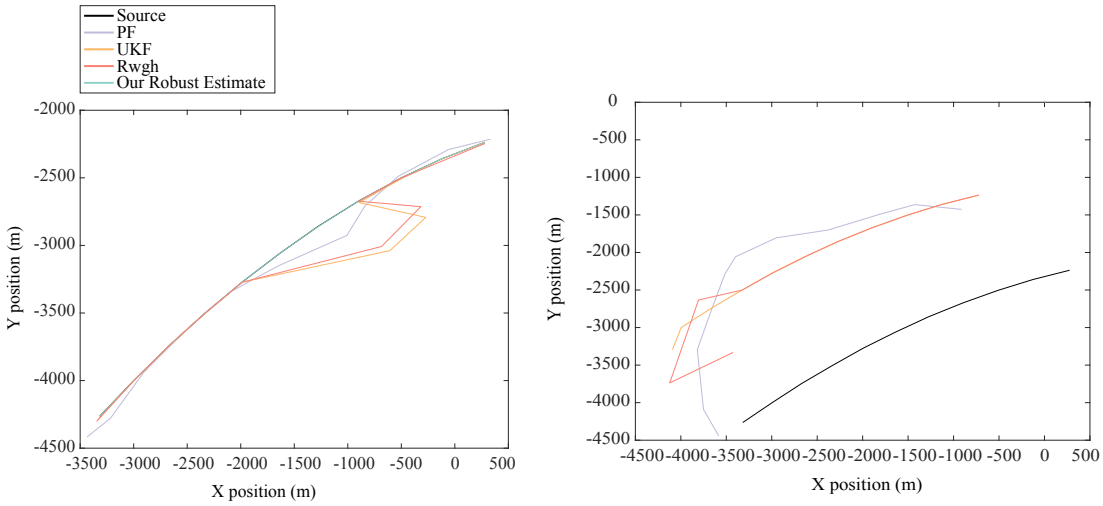


Figure 6.10 – Mean and confidence interval of estimates obtained with the four considered estimators on the unpredictable trajectory as a function of the attack delay: (a) when the attack delay is applied only to sensor 1; (b): when the attack delay is applied to sensors 1 and 2; (c) when the delay increases between the sensors of two synchronized pairs; (d): when the attack delay increases between three synchronized sensors. Our robust solution remains accurate and stops localizing when not enough measurements can be trusted, the Rwgh solution looses accuracy while the weights of attacked measurements are non-zero and gives unreliable estimates when the system is too corrupt to localize, the UKF is less accurate than our robust solution and looses accuracy as the delay increases, the PF is not accurate from the start and keeps a stable accuracy above 1km as the delay increases.

We also analyse the effect of a sporadic attack occurring in the middle of the trajectory. We give enough time for all estimators to initialize as best they can, we then perform an attack on two sensors at two iterations in the middle of the simulation and remove the attack for the remaining of the simulation. During the two iterations where the attack is occurring, we chose to inject 5μs to the two sensors. We chose this specific delay because Figure 6.9b shows that it has a large impact on the estimates of the UKF and the Rwgh. Figure 6.11a shows that our solution is not impacted, whereas the PF, the UKF and the

Rwgh are impacted and are able to get back to the right trajectory as soon as the attack is over, as expected. Notice that although the general accuracy of the PF is not perfect, the attacks create smaller variations in the error of this estimator than in the error of the UKF and of the Rwgh.



(a) (b)

Figure 6.11 – Estimated trajectories when (a) a sporadic attack of 5μs occurs on sensors 1 and 2 at only two iterations in the middle of the simulation: all estimators have had enough time to initialize, our solution is indistinguishable with the source, the UKF, PF and Rwgh estimates are impacted by the attack and are able to recover an accurate trajectory after the attack, as expected; (b) an attack according to our strong attack model occurs: the attacker injects delays such that the misestimation is at $[x_{source} - 1000, y_{source} + 1000]$, our robust solution declares that the system is too corrupt to localize and doesn't give any estimates (the attacker fails) and all the other estimators give solutions around the targeted misestimation (the attacker succeeds).

Lastly, we simulate an attack according to our strong attack model. At every time instant, we attack three sensors such that the measurements are all plausible and intersect close to a targeted misestimation located at $[x_{source} - 1000, y_{source} + 1000]$. The results from Figure 6.11b show that our robust solution directly declares that the system is too corrupt to localize, therefore no curve corresponding to our solution is visible. We also observe that the UKF and the Rwgh estimates are close to the targeted misestimation, in other words the attacker succeeds. This is expected because in this scenario, the residuals are very small since the measurements seem plausible. The PF takes more time to fluctuate towards the targeted misestimation because the initial point is far from the targeted wrong trajectory.

From the above comparisons, we conclude that on a predictable trajectory, our robust solution, the Rwgh and the UKF offer a similar accuracy if the attack delay is lower than approximately 3 times the standard deviation of the measurement noise. With an unpredictable trajectory, the UKF has a lower accuracy and does not offer as accurate results as our solution or the Rwgh. We observe that if there are enough non-attacked

measurements to localize with true measurements, our robust solution is always very accurate. In comparison the Rwgh drastically looses accuracy as the attack delay grows and then regains a very good accuracy once the weights of sets with attacked measurements get close to 0. The PF does not offer a good accuracy but remains fairly stable as attack delays grow. If the system is very corrupt, the Rwgh continues to give very erroneous estimates while our robust solution declares that the system is too corrupt to localize. Against the strong attack model, only our robust solution is able to detect the attack.

6.7 Conclusion

We have highlighted that TSAs on the time reference of the sensors of the network are a threat to TDOA localization. By injecting a few micro-seconds into the clock of a sensor, the network estimates the source to be located kilometers away from the true source position. We have also shown that a strong attacker with knowledge of the sensor coordinates and of the source coordinates, is able to choose the delays to inject such that the resulting misestimation is at a specific targeted location. To counter such TSAs, we have proposed a robust technique that attributes weights to all sensor pairs by relying on signals from a known calibration source of a known position. These weights are computed to reflect the confidence we have in the time synchronization of the corresponding sensor pairs. Subsequently, our localization technique uses these weights either to identify the network as too corrupt to localize, or to give an accurate estimate of the unknown source location. Our technique also provides a confidence metric that gives insight on the accuracy of the estimate. The calibration phase of our proposed solution is, however, vulnerable to replay attacks. In such attacks, the calibration signal is replayed at times and locations of the attacker's choice, thus possibly affecting the attributed weights. In order to counter these replay attacks, we have provided an encrypted authenticated challenge-response scheme that ensures that the measurements used for calibration are trustworthy. Numerical evidence in 2D and 3D show that our technique is efficient and that the confidence metric is trustworthy although it might lead to false negatives. Numerical results also show that our solution outperforms usual localization and tracking methods used in LOS and NLOS environments.

7

Conclusions

In this thesis, we studied time-synchronization attacks (TSAs) and their mitigation in the setting of networks that rely on the timely analysis of observations. More specifically, we considered physical TSAs that are undetectable and not thwarted by the usual cyber-security tools. Such attacks lead to an unsynchronized network, which in turn may lead to its malfunction. Our proposed mitigation techniques make use of the non-plausibility of observations taken by an attacked system. Whatever the studied network, this work highlights that the more redundancy there is in the measurements, the harder it is to effectively attack the network, as is expected intuitively. In this thesis, we leverage the in-depth knowledge of the system's operation and its normal dynamics to strategically maintain the good-functioning of the system even if under attack. We specifically dived into the maths both of the state estimation of smart grids and of the localization of a passive source from time-difference of arrival (TDOA) measurements. In both settings, we showed how to mount undetectable attacks. In the former setting, we further showed how to position measurement points in order to achieve a high level of security against TSAs and how to monitor the grid for potential unlikely vulnerabilities. In the latter setting, we proposed a solution either to localize accurately in spite of the attack or to detect the presence of an attack that is too strong to continue using the network in its current state.

The study of TSAs and their mitigation in the first setting constitutes the first part of this thesis. More specifically, we investigated the practical feasibility of TSAs under the realistic assumptions that the time reference of a PMU is managed by a clock controller, and that several measurement points may share the same time reference. We can regroup our results into two categories depending on whether our aim is to mount smart undetectable attacks, or to secure the grid against TSAs.

In favour of attacking the grid. We showed in Chapter 3, that vulnerable sets of PMUs of arbitrary size can be found by grouping PMUs in equivalence classes with respect to the IoS. However, because the IoS depends on the measurement values, equivalence classes would have to be updated at every time instant. In contrast, the

infimum IoS^* of IoS over all measurement values does not depend on the measurement values and thus can be used to find sets of PMUs that are vulnerable to TSAs irrespective of their measurement values. In Chapter 4, we extend our theory to account for the fact that several sites may share the same time reference, in which case an attack angle will impact all corresponding measurement points in the same manner. We found a sufficient structural vulnerability condition that does not depend on the measurement values and that is able to find more vulnerable sets than IoS^*. Specifically, the best known strategy is to search for groups of sites in which the submatrices of F for each pair of sites, have ERR values that are higher than a threshold that can be predefined experimentally as in Section 5.3.2. Once vulnerable sets of sites are identified, we proposed attack strategies that take into account the practical undetectability conditions posed by the clock servo. Specifically, for clock servos to implement attack angles without modifying them, is is required that the injected attack angles are small. Such small incremental delays can be achieved when the solution set of undetectable attacks forms a continuum instead of a finite set, which happens if at least three vulnerable sites are targeted simultaneously. Finally, Chapter 5 shows that if the system has unbalances, then all undetectable TSAs on the three-phase system are undetectable on the direct-sequence system but that the converse is not always true. As a result, unless the attacker is sure that only the direct-sequence system is used for the state estimation and for BDD, vulnerable sets of sites and undetectable TSAs should be found and computed using the three-phase model in order to remain stealthy.

In favour of securing the grid. We showed that the vulnerability of sets of arbitrary number of sites reduces to the vulnerability of single sites and pairs of sites. Hence, it is sufficient to ensure the security of every site and pair of sites in order to secure the grid. Our work shows that all undetectable TSAs presented in the literature, were possible because the targeted sites were structurally vulnerable to TSAs. We established a security requirement to prevent this vulnerability. We also provided an offline recursive greedy algorithm that enforces our security requirement by strategically increasing the number of measurement points in the vulnerable areas of the observable grid. As a result, after using our algorithm, all of the existing TSAs became detectable through residual analysis. A limitation of our security requirement is that even if satisfied, it is still possible that measurement values are such that attacks are feasible, although we reason that it is unlikely. We identified sufficient and necessary vulnerability conditions for single sites and for pairs of sites, each measuring an arbitrary number of phasors. We recommend the monitoring of two metrics associated to these conditions in order to continuously check the non-vulnerability of the grid. In order to secure the grid against structural vulnerabilities, we are required to slightly modify the PMU allocation, which may include increasing the number of measurements. Such modifications to the grid may be costly. In Chapter 5, we show that by leveraging the available three-phase measurements in unbalanced systems, we are able to increase the security of the grid without modifying

its structure. However, as our numerical results show, although the security is increased by using the three-phase model instead of the direct-sequence model, it is still possible that certain sets of sites are vulnerable in three-phase. Nevertheless, all of the found sets that are vulnerable in the three-phase model, are in fact structurally vulnerable to TSAs. Therefore, by combining our results presented in Chapters 4 and 5 we can achieve the security of the grid at a reduced cost. Specifically, we recommend to implement the security recommendations established in Chapter 4, the state estimation and the BDD on the three-phase system. This requires less modifications to the grid than is required for securing the grid in the direct-sequence model.

Other solutions for the identification of TSAs. As mentioned in the introduction, the identification of an attack requires in-depth knowledge of the system's operation and its normal dynamics in order to differentiate observations with and without an attack. In this thesis we focused on identifying attacks through the traditional BDD mechanisms. However, there are alternatives that could be investigated. For example, our numerical results show that our undetectable clock-servo aware attack strategies become momentarily detectable in the presence of an inrush. The effects of sudden changes to the grid could be investigated further to differentiate normal grid dynamics from the ones of an attacked grid, possibly using machine learning methods. Other physical quantities that are impacted by TSAs could be monitored and correlated in order to detect TSAs. For example, an alternative approach for the detection of TSAs against PMUs, proposed in [132], relies on monitoring the correlation between adjustments made to the PMU clock frequency and changes in the phase angles measured by the PMU.

In the second part of this thesis, we proposed a solution to localize a passive source accurately from TDOA measurements in spite of a TSA. Our method works under the realistic assumption that a known non-synchronized calibration source can be used. If the system is too corrupt to maintain its localizing purpose, then our solution detects it. Our numerical results also show that our proposed solution outperforms the usual tracking and localization methods used in line-of-sight and non-line-of-sight (NLOS) environments, for both predictable and unpredictable source trajectories. Hence, it could be used not only in our adversarial setting of TSAs but also as a reliable localization mechanism in NLOS environments. As for the smart grids setting, other TSA identification techniques could be investigated further. For instance, other signal properties such as the frequency-difference of arrival (FDOA) or the angle-of-arrival (AOA), could be analysed and correlated. Due to the relatively low speed of a source with respect to the speed of the radio signals, the impact of a TSA on the AOA of the signal is much less important than the impact on a TDOA measurement. However, the AOA of a signal is often obtained after some post-processing which involves the analysis of TDOA measurements from different receptors of a receiving antenna. Assuming that a TSA impacts all receptors of an antenna in the same manner, then the AOA won't be too impacted by a TSA. In a non-adversarial setting, TDOA measurements enable a more accurate localization than

AOA measurements do. Their combination could lead to an efficient attack detection and misestimation correction solution.

Bibliography

[1] "IEEE standard for synchrophasor measurements for power systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, Dec 2011.

[2] P. Romano and M. Paolone, "Enhanced interpolated-dft for synchrophasor estimation in fpgas: Theory, implementation, and validation of a pmu prototype," *IEEE Trans. on Instrumentation and Measurement*, vol. 63, no. 12, pp. 2824–2836, 2014.

[3] Y. hua Tang ; Gerard N. Stenbakken ; Allen Goldstein, "Calibration of phasor measurement unit at nist," *IEEE Trans. on Instrumentation and Measurement*, vol. 62, no. 6, pp. 1417–1422, 2013.

[4] G. Barchi, D. Fontanelli, D. Macii, and D. Petri, "On the accuracy of phasor angle measurements in power networks," *IEEE Trans. on Instrumentation and Measurement*, vol. 64, no. 5, pp. 1129–1139, 2015.

[5] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proceedings of the 16th ACM Conf. on Computer and Communications Security*, ser. CCS '09. New York, NY, USA: ACM, 2009, pp. 21–32. [Online]. Available: http://doi.acm.org/10.1145/1653662.1653666

[6] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. on Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec 2011.

[7] G. Dan and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *2010 First IEEE Int. Conf. on Smart Grid Communications*, Oct 2010, pp. 214–219.

[8] A. Anwar, A. N. Mahmood, and Z. Tari, "Identification of vulnerable node clusters against false data injection attack in an ami based smart grid," *Inf. Syst.*, vol. 53, no. C, pp. 201–212, Oct. 2015. [Online]. Available: http://dx.doi.org/10.1016/j.is.2014.12.001

[9] A. Bretas, N. Bretas, S. Braunstein, A. Rossoni, and R. Trevizan, "Multiple gross errors detection, identification and correction in three-phase distribution systems wls state estimation: A per-phase measurement error approach," *Electric Power Systems Research*, vol. 151, pp. 174 – 185, 2017.

[10] O. Vukovic, K. C. Sou, G. Dan, and H. Sandberg, "Network-aware mitigation of data integrity attacks on power system state estimation," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1108–1118, July 2012.

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

[12] T. Zou, A. S. Bretas, C. Ruben, S. C. Dhulipala, and N. Bretas, "Smart grids cyber-physical security: Parameter correction model against unbalanced false data injection attacks," *Electric Power Systems Research*, vol. 187, p. 106490, 2020.

[13] A. S. Bretas, N. G. Bretas, and B. E. Carvalho, "Further contributions to smart grids cyber-physical security as a malicious data attack: Proof and properties of the parameter error spreading out to the measurements and a relaxed correction model," *International Journal of Electrical Power Energy Systems*, vol. 104, pp. 43 – 51, 2019.

[14] S. Barreto, M. Pignati, G. Dan, J.-Y. Le Boudec, and M. Paolone, "Undetectable pmu timing-attack on linear state-estimation by using rank-1 approximation," *IEEE Trans. on Smart Grid*, vol. 9, no. 4, pp. 3530–3542, 2018.

[15] E. Shereen, M. Delcourt, S. Barreto, G. Dán, J.-Y. Le Boudec, and M. Paolone, "Feasibility of time synchronization attacks against pmu-based state-estimation," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2019.

[16] M. Delcourt and J.-Y. L. Boudec, "Security measures for grids against rank-1 undetectable time-synchronization attacks," *arXiv, eess.SY*, vol. 2002.12607, 2020.

[17] B. Brinkmann, K. Bicevskis, R. Scott, and M. Negnevitsky, "Evaluation of single- and three-phase state estimation in distribution networks," in *2017 Australasian Universities Power Engineering Conference (AUPEC)*, 2017, pp. 1–5.

[18] S.-K. Lin, "Electronic warfare target location methods, second edition. edited by richard a. poisel, artech house, 2012; 422 pages. price: £99.00, isbn 978-1-60807-523-2," *Sensors*, vol. 13, no. 1, p. 1151–1157, Jan 2013.

[19] X. Liu, J. Yin, S. Zhang, B. Ding, S. Guo, and K. Wang, "Range-based localization for sparse 3-d sensor networks," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 753–764, Feb 2019.

[20] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 864–870.

[21] B. Triggs, "Model-Based Sonar Localization for Mobile Robots," *Robotics and Autonomous Systems*, vol. 12, no. 3-4, pp. 173–186, Apr. 1994, originally appeared in International Conference on Intelligent Robot Systems, Zakopane, Poland, 199.

[22] H. Zou, Z. Chen, H. Jiang, L. Xie, and C. Spanos, "Accurate indoor localization and tracking using mobile phone inertial sensors, wifi and ibeacon," in *2017 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, March 2017, pp. 1–4.

[23] A. Mikhalev and R. Ormondroyd, "Passive emitter geolocation using agent-based data fusion of aoa, tdoa and fdoa measurements," in *2007 10th International Conference on Information Fusion*, July 2007, pp. 1–6.

[24] F. Gustafsson and F. Gunnarsson, "Positioning using time-difference of arrival measurements," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, vol. 6, April 2003, pp. VI–553.

[25] M. Delcourt and J. L. Boudec, "TDOA source-localization technique robust to time-synchronization attacks," *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2020.

[26] E. A. Curtis, "Synchronized to an optical atomic clock," *Science*, vol. 368, no. 6493, pp. 825–826, 2020.

[27] P. Rochat, F. Droz, Q. Wang, and S. Froidevaux, "Atomic clocks and timing systems in global navigation satellite systems," 2012.

[28] A. Carta, N. Locci, C. Muscas, and S. Sulis, "A flexible gps-based system for synchronized phasor measurement in electric distribution networks," *IEEE Trans. on Instrumentation and Measurement*, pp. 1547–1552, 2006.

[29] M. Lixia, A. Benigni, A. Flammini, C. Muscas, F. Ponci, and A. Monti, "A software-only ptp synchronization for power system state estimation with pmus," *IEEE Trans. on Instrumentation and Measurement*, 2012.

[30] D. P. Shepard, T. E. Humphreys, and A. A. Fansler, "Evaluation of the vulnerability of phasor measurement units to gps spoofing attacks," *International Journal of Critical Infrastructure Protection*, vol. 5, no. 3, pp. 146 – 153, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1874548212000480

[31] X. Jiang, J. Zhang, B. J. Harding, J. J. Makela, and A. D. Domınguez-Garcıa, "Spoofing gps receiver clock offset of phasor measurement units," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3253–3262, 2013.

[32] C. Konstantinou, M. Sazos, A. S. Musleh, A. Keliris, A. Al-Durra, and M. Maniatakos, "Gps spoofing effect on phase angle monitoring and control in a real-time digital simulator-based hardware-in-the-loop environment," *IET Cyber-Physical Systems: Theory Applications*, vol. 2, no. 4, pp. 180–187, 2017.

[33] P. Risbud, N. Gatsis, and A. Taha, "Vulnerability analysis of smart grids to gps spoofing," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3535–3548, 2019.

[34] K. M.G., "An asymmetric security mechanism for navigation signals," *Lecture Notes in Computer Science*, vol. 3200, 2004.

[35] K. Jansen, N. O. Tippenhauer, and C. Pöpper, "Multi-receiver gps spoofing detection: Error models and realization," in *Proceedings of the 32nd Annual Conference on Computer Security Applications*, ser. ACSAC '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 237–250. [Online]. Available: https://doi.org/10.1145/2991079.2991092

[36] P. Montgomery, T. Humphreys, and B. Ledvina, "Receiver-autonomous spoofing detection: Experimental results of a multi-antenna receiver defense against a portable civil gps spoofer," 2009, pp. 124–130.

[37] D.-Y. Yu, A. Ranganathan, T. Locher, S. Capkun, and D. Basin, "Short paper: Detection of gps spoofing attacks in power grids," in *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless amp; Mobile Networks*, ser. WiSec '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 99–104.

[38] J. S. Warner and R. Johnston, "Gps spoofing countermeasures," 2003.

[39] A. Ranganathan, H. Ólafsdóttir, and S. Capkun, "Spree: A spoofing resistant gps receiver," ser. MobiCom '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 348–360.

[40] B. Ledvina, W. Bencze, B. T. Galusha, and I. Miller, "An in-line anti-spoofing device for legacy civil gps receivers," 2010, pp. 698–712.

[41] A. Jafarnia-Jahromi, T. Lin, A. Broumandan, J. Nielsen, and G. Lachapelle, "Detection and mitigation of spoofing attacks on a vector-based tracking gps receiver," 2012.

[42] S. Narain, A. Ranganathan, and G. Noubir, "Security of gps/ins based on-road location tracking systems," in *2019 IEEE Symposium on Security and Privacy (SP)*, 2019, pp. 587–601.

[43] Y. Yang and J. Xu, "Gnss receiver autonomous integrity monitoring (raim) algorithm based on robust estimation," *Geodesy and Geodynamics*, vol. 7, no. 2, pp. 117 – 123, 2016.

[44] B. W. PARKINSON and P. AXELRAD, "Autonomous gps integrity monitoring using the pseudorange residual," *NAVIGATION*, vol. 35, no. 2, pp. 255–274, 1988.

[45] K. Jansen, M. Schäfer, D. Moser, V. Lenders, C. Pöpper, and J. Schmitt, "Crowd-gps-sec: Leveraging crowdsourcing to detect and localize gps spoofing attacks," in *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2018, pp. 1018–1031.

[46] A. Derviškadić, R. Razzaghi, Q. Walger, and M. Paolone, "The white rabbit time synchronization protocol for synchrophasor networks," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 726–738, Jan 2020.

[47] N. M. Freris, S. R. Graham, and P. R. Kumar, "Fundamental limits on synchronizing clocks over networks," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1352–1364, 2011.

[48] O. Gurewitz, I. Cidon, and M. Sidi, "One-way delay estimation using network-wide measurements," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2710–2724, 2006.

[49] A. Ceccarelli, E. Lisova, M. Gutiérrez, W. Steiner, E. Uhlemann, J. Åkerberg, R. Dobrin, and M. Björkman, "Protecting clock synchronization: Adversary detection through network monitoring," *Journal of Electrical and Computer Engineering*, vol. 2016, p. 6297476, 2016. [Online]. Available: https://doi.org/10.1155/2016/6297476

[50] S. Barreto, A. Suresh, and J. Y. L. Boudec, "Cyber-attack on packet-based time synchronization protocols: The undetectable delay box," in *2016 IEEE Int. Instrumentation and Measurement Technology Conf. Proceedings*, May 2016, pp. 1–6.

[51] 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.

[52] J. Tournier and O. Goerlitz, "Strategies to secure the ieee 1588 protocol in digital substation automation," in *2009 Fourth International Conference on Critical Infrastructures*, 2009, pp. 1–8.

[53] A. Treytl and B. Hirschler, "Security flaws and workarounds for ieee 1588 (transparent) clocks," in *2009 International Symposium on Precision Clock Synchronization for Measurement, Control and Communication*, 2009, pp. 1–6.

[54] A. Treytl, G. Gaderer, B. Hirschler, and R. Cohen, "Traps and pitfalls in secure clock synchronization," in *2007 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication*, 2007, pp. 18–24.

[55] 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 *2019 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control, and Communication (ISPCS)*, 2019, pp. 1–6.

[56] S. Frankel and S. Krishnan, "Ip security (ipsec) and internet key exchange (ike) document roadmap," Informational, RFC 6071, February 2011. [Online]. Available: https://www.rfc-editor.org/rfc/rfc6071.txt

[57] "Ieee standard for local and metropolitan area networks–port-based network access control," *IEEE Std 802.1X-2020 (Revision of IEEE Std 802.1X-2010 Incorporating IEEE Std 802.1Xbx-2014 and IEEE Std 802.1Xck-2018)*, pp. 1–289, 2020.

[58] E. Rescorla, "The transport layer security (tls) protocol version 1.3," Proposed Standard, RFC 8446, August 2018. [Online]. Available: https://www.rfc-editor.org/rfc/rfc8446.txt

[59] M. Ullmann and M. Vögeler, "Delay attacks — implication on ntp and ptp time synchronization," in *2009 International Symposium on Precision Clock Synchronization for Measurement, Control and Communication*, 2009, pp. 1–6.

[60] T. Mizrahi, "A game theoretic analysis of delay attacks against time synchronization protocols," in *2012 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control and Communication Proceedings*, 2012, pp. 1–6.

[61] T. Mizrahi, "Security requirements of time protocols in packet switched networks," in *RFC 7384*, 2014.

[62] "Ieee standard for a precision clock synchronization protocol for networked measurement and control systems," *IEEE Std 1588-2008 (Revision of IEEE Std 1588-2002)*, pp. 1–300, 2008.

[63] N. Moreira, J. Lázaro, J. Jimenez, M. Idirin, and A. Astarloa, "Security mechanisms to protect ieee 1588 synchronization: State of the art and trends," in *2015 IEEE International Symposium on Precision Clock Synchronization for Measurement, Control, and Communication (ISPCS)*, 2015, pp. 115–120.

[64] J. Tsang and K. Beznosov, "A security analysis of the precise time protocol (short paper)," in *Information and Communications Security*, P. Ning, S. Qing, and N. Li, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 50–59.

[65] Qingyu Yang, Dou An, and Wei Yu, "On time desynchronization attack against ieee 1588 protocol in power grid systems," in *2013 IEEE Energytech*, 2013, pp. 1–5.

[66] R. Annessi, J. Fabini, and T. Zseby, "Securetime: Secure multicast time synchronization," 2017.

[67] L. Narula and T. Humphreys, "Requirements for secure clock synchronization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, pp. 749–762, 2018.

[68] *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

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

[70] L. Zanni, "Power-system state estimation based on pmus static and dynamic approaches - from theory to real implementation," p. 189, 2017. [Online]. Available: http://infoscience.epfl.ch/record/228451

[71] K. C. Sou, H. Sandberg, and K. H. Johansson, "Computing critical k-tuples in power networks," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1511–1520, 2012.

[72] M. Göl and A. Abur, "Observability and criticality analyses for power systems measured by phasor measurements," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3319–3326, 2013.

[73] A. Wehenkel, A. Mukhopadhyay, J. Le Boudec, and M. Paolone, "Parameter estimation of three-phase untransposed short transmission lines from synchrophasor measurements," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2020.

[74] M. Göl and A. Abur, "Lav based robust state estimation for systems measured by pmus," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1808–1814, 2014.

[75] M. R. Irving, R. C. Owen, and M. J. H. Sterling, "Power-system state estimation using linear programming," *Proceedings of the Institution of Electrical Engineers*, vol. 125, no. 9, pp. 879–885, 1978.

[76] Y. Lin and A. Abur, "A highly efficient bad data identification approach for very large scale power systems," *IEEE Trans. on Power Systems*, vol. 33, no. 6, pp. 5979–5989, 2018.

[77] M. Göl and A. Abur, "A modified chi-squares test for improved bad data detection," *IEEE Eindhoven PowerTech*, pp. 1–5, 2015.

[78] Y. Al-Eryani and U. Baroudi, "An investigation on detecting bad data injection attack in smart grid," *Int. Conf. on Computer and Information Sciences (ICCIS)*, 2019.

[79] A. Abur and A. Exposito, *Power system state estimation: theory and implementation*. CRC, 2004, vol. 24.

[80] Z. Zhang, S. Gong, A. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Trans. on Smart Grid*, vol. 4, no. 1, pp. 87–98, March 2013.

[81] S. Wang, J. Zhao, Z. Huang, and R. Diao, "Assessing gaussian assumption of pmu measurement error using field data," *IEEE Trans. on Power Delivery*, vol. 33, no. 6, pp. 3233–3236, 2018.

[82] K. A. Clements and P. W. Davis, "Multiple bad data detectability and identifiability: A geometric approach," *IEEE Transactions on Power Delivery*, vol. 1, no. 3, pp. 355–360, 1986.

[83] D. Fontanelli, D. Macii, S. Rinaldi, P. Ferrari, and A. Flammini, "A servo-clock model for chains of transparent clocks affected by synchronization period jitter," *IEEE Trans. on Instrumentation and Measurement*, vol. 63, no. 5, pp. 1085–1095, 2014.

[84] S. Barreto, E. Shereen, M. Pignati, G. Dan, J.-Y. Le Boudec, and M. Paolone, "A continuum of undetectable timing-attacks on pmu-based linear state-estimation," in *IEEE Smart Grid Comm*, 2017, pp. 473–479.

[85] D. Kushner, "The real story of stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, March 2013.

[86] L. E. Dickson, *Modern algebraic theories*, 1930.

[87] K. Correll and N. Barendt, "Design considerations for software only implementations of the ieee 1588 precision time protocol," in *In Conf. on IEEE 1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*, 2006.

[88] "IEEE std1588-2008, IEEE standard for a precision clock synchronization protocol for networked measurement and control systems," *IEEE Instrumentation and Measurement Society*, July 24, 2008.

[89] D. J. Brueni and L. S. Heath, "The pmu placement problem," *SIAM Journal on Discrete Mathematics*, vol. 19, no. 3, pp. 744–761, 2005.

[90] W. Yuill, A. Edwards, S. Chowdhury, and S. P. Chowdhury, "Optimal pmu placement: A comprehensive literature review," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–8.

[91] S. M. M. Niyaragh, A. J. Irani, and H. Shayeghi, "Modeling of zero injection buses based to optimal placement of pmus for full observability of power systems," *Journal of Electrical Engineering & Technology*, vol. 15, no. 6, pp. 2509–2518, 2020. [Online]. Available: https://doi.org/10.1007/s42835-020-00536-0

[92] B. Xu and A. Abur, "Observability analysis and measurement placement for systems with pmus," in *IEEE PES Power Systems Conference and Exposition, 2004.*, 2004, pp. 943–946 vol.2.

[93] N. M. Manousakis, G. N. Korres, and P. S. Georgilakis, "Taxonomy of pmu placement methodologies," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1070–1077, 2012.

[94] M. Picallo, A. Anta, and B. D. Schutter, "Efficient convex optimization for optimal pmu placement in large distribution grids," in *2019 IEEE Milan PowerTech*, 2019, pp. 1–6.

[95] A. Derviškadić, P. Romano, M. Pignati, and M. Paolone, "Architecture and experimental validation of a low-latency phasor data concentrator," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2885–2893, 2018.

[96] L. Zanni, A. Derviškadić, M. Pignati, C. Xu, P. Romano, R. Cherkaoui, A. Abur, and M. Paolone, "Pmu-based linear state estimation of lausanne subtransmission network: Experimental validation," *Electric Power Systems Research*, vol. 189, p. 106649, 2020.

[97] C. Muscas, S. Sulis, A. Angioni, F. Ponci, and A. Monti, "Impact of different uncertainty sources on a three-phase state estimator for distribution networks," *IEEE Trans. on Instrumentation and Meas.*, vol. 63, no. 9, pp. 2200–2209, 2014.

[98] A. P. Meliopoulos, G. J. Cokkinides, and G. K. Stefopoulos, "Numerical experiments for three-phase state estimation performance and evaluation," in *Proc. of IEEE Russia Power Tech*, 2005, pp. 1–7.

[99] Yue Yang and S. Roy, "Pmu placement for optimal three-phase state estimation performance," in *Proc. of IEEE SmartGridComm*, 2013, pp. 342–347.

[100] M. Göl and A. Abur, "A robust PMU based three-phase state estimator using modal decoupling," *IEEE Trans. on Power Syst.*, vol. 29, pp. 2292–2299, 2014.

[101] B. Zargar, A. Angioni, F. Ponci, and A. Monti, "Multiarea parallel data-driven three-phase distribution system state estimation using synchrophasor measurements," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6186–6202, 2020.

[102] S. Mazuelas, F. A. Lago, J. Blas, A. Bahillo, P. Fernandez, R. M. Lorenzo, and E. J. Abril, "Prior nlos measurement correction for positioning in cellular wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 5, pp. 2585–2591, Jun 2009.

[103] J. Prieto, A. Bahillo, S. Mazuelas, R. M. Lorenzo, J. Blas, and P. Fernandez, "Nlos mitigation based on range estimation error characterization in an rtt-based ieee 802.11, indoor location system," in *2009 IEEE International Symposium on Intelligent Signal Processing*, Aug 2009, pp. 61–66.

[104] E. Arias-de-Reyna and P. M. Djurić, "Indoor localization with range-based measurements and little prior information," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1979–1987, May 2013.

[105] T. Hillebrandt, H. Will, and M. Kyas, *Quantitative and Spatial Evaluation of Distance-Based Localization Algorithms.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 173–194.

[106] I. Guvenc, C. Chong, and F. Watanabe, "Analysis of a linear least-squares localization technique in los and nlos environments," in *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*, April 2007, pp. 1886–1890.

[107] Y. Zhao, X. Fan, C. Xu, and X. Li, "Er-crlb: An extended recursive cramér rao lower bound fundamental analysis method for indoor localization systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 2, pp. 1605–1618, Feb 2017.

[108] B. Chen, C. Yang, F. Liao, and J. Liao, "Mobile location estimator in a rough wireless environment using extended kalman-based imm and data fusion," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1157–1169, March 2009.

[109] J. Prieto, S. Mazuelas, A. Bahillo, P. Fernandez, R. M. Lorenzo, and E. J. Abril, "Adaptive data fusion for wireless localization in harsh environments," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1585–1596, April 2012.

[110] J. Zhang, F. Dong, G. Feng, and C. Shen, "Analysis of the nlos channel environment of tdoa multiple algorithms," in *2015 IEEE SENSORS*, Nov 2015, pp. 1–4.

[111] M. Boccadoro, G. De Angelis, and P. Valigi, "Tdoa positioning in nlos scenarios by particle filtering," *Wireless Networks*, vol. 18, 07 2012.

[112] L. Cong and Weihua Zhuang, "Non-line-of-sight error mitigation in tdoa mobile location," in *GLOBECOM'01. IEEE Global Telecommunications Conference (Cat. No.01CH37270)*, vol. 1, Nov 2001, pp. 680–684 vol.1.

[113] J. FaIk, P. Handel, and M. Jansson, "Effects of frequency and phase errors in electronic warfare tdoa direction-finding systems," in *IEEE Military Communications Conference, 2003. MILCOM 2003.*, vol. 1, Oct 2003, pp. 118–123 Vol.1.

[114] J. Falk, P. Handel, and M. Jansson, "Multisource time delay estimation subject to receiver frequency errors," in *The Thrity-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, vol. 1, Nov 2003, pp. 1156–1160 Vol.1.

[115] J. FaIk, P. Handel, and M. Jansson, "Estimation of receiver frequency error in a tdoa-based direction-finding system," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, vol. 2, Nov 2004, pp. 2079–2083 Vol.2.

[116] S. Zhong, W. Xia, and Z. He, "Approximate maximum likelihood time differences estimation in the presence of frequency and phase consistence errors," in *IEEE International Symposium on Signal Processing and Information Technology*, Dec 2013, pp. 000 305–000 308.

[117] S. Zhong, W. Xia, and Z. He, "Joint estimation of time delay and clock error in the incoherent reception systems," *Circuits, Systems, and Signal Processing*, vol. 35, no. 9, pp. 3284–3309, Sep 2016.

[118] X. Chen, D. Wang, J. Yin, and Y. Wu, "Performance analysis and dimension-reduction taylor series algorithms for locating multiple disjoint sources based on tdoa under synchronization clock bias," *IEEE Access*, vol. 6, pp. 48 489–48 509, 2018.

[119] D. Wang, J. Yin, T. Tang, X. Chen, and Z. Wu, "Quadratic constrained weighted least-squares method for tdoa source localization in the presence of clock synchronization bias: Analysis and solution," *Digital Signal Processing*, vol. 82, pp. 237 – 257, 2018.

[120] A. Piersol, "Time delay estimation using phase data," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 471–477, June 1981.

[121] J. Falk, P. Handel, and M. Jansson, "Direction finding for electronic warfare systems using the phase of the cross spectral density," *Radiovetenskap och Kommunikation Stockholm*, 2002.

[122] E. Weinstein and D. Kletter, "Delay and doppler estimation by time-space partition of the array data," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 6, pp. 1523–1535, December 1983.

[123] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, Aug 1994.

[124] S. R. Drake and K. Dogancay, "Geolocation by time difference of arrival using hyperbolic asymptotes," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, May 2004, pp. ii–361.

[125] A. Quazi, "An overview on the time delay estimate in active and passive systems for target localization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 527–533, June 1981.

[126] S. Amuru and R. M. Buehrer, "Optimal jamming against digital modulation," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 2212–2224, Oct 2015.

[127] I. Guvenc and C. Chong, "A survey on toa based wireless localization and nlos mitigation techniques," *IEEE Communications Surveys Tutorials*, vol. 11, no. 3, pp. 107–124, 2009.

[128] M. Bosch and M. Nájar, "Unscented kalman filter for location in non-line-of-sight," in *2006 14th European Signal Processing Conference*, 2006, pp. 1–5.

[129] M. Boccadoro, G. De Angelis, and P. Valigi, "Tdoa positioning in nlos scenarios by particle filtering," *Wireless Networks*, vol. 18, no. 5, pp. 579–589, 2012. [Online]. Available: https://doi.org/10.1007/s11276-012-0420-9

[130] Pi-Chun Chen, "A non-line-of-sight error mitigation algorithm in location estimation," in *WCNC. 1999 IEEE Wireless Communications and Networking Conference (Cat. No.99TH8466)*, vol. 1, 1999, pp. 316–320 vol.1.

[131] N. El Gemayel, S. Meier, and F. K. Jondral, "On the applicability of the residual weighting algorithm for tdoa," in *2012 IV International Congress on Ultra Modern Telecommunications and Control Systems*, 2012, pp. 143–147.

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

Marguerite DELCOURT

PhD Student at EPFL in Information Security

Avenue de France, 38
1004 Lausanne, Vaud
Tel. +41 (0)768041228
marguerite.delcourt@epfl.ch

12.12.1992
French
C permit

EDUCATION

From 15.04.2021	**Security and Network Engineer**, R&D at Siemens Mobility Ltd., Wallisellen, Switzerland
02.2017 – 04.2021	**PhD** in Information Security at École Polytechnique Fédérale de Lausanne (EPFL) 4th year PhD at LCA2 with Pf. Jean-Yves Le Boudec Time-Synchronization Attacks on Critical Infrastructures (Smart-Grids, sensor networks for source Localization)
09.2014 – 02.2017	**Master Degree** at École Polytechnique Fédérale de Lausanne (EPFL) Communication System Master Specialization in Information Security
09.2011 – 07.2014	**Bachelor Degree** at Université Bordeaux 1 Mathematics with Computer Science options with Honors (1st in the section)
Until 2010	Lycée Français de Singapour Baccalauréat Général Scientifique with Honors (mention Bien)

WORK EXPERIENCE & PROJECTS

2017-present	PhD research projects • Feasibility of Time-Synchronization Attacks Against PMU-Based State Estimation • TDOA Source-Localization Technique Robust to Time-Synchronization Attacks • Security Measures for Grids against Rank-1 Undetectable Time-Synchronization Attacks • Time-Synchronization Attack Detection in Unbalanced Three-Phase Systems Collaborations with: DESL lab (EPFL), KTH and Armasuisse.
08.2016 – 02.2017	Security engineer intern at Amazon, (Seattle) • Project on promising cryptographic families for post-quantum crypto, studying security Parameters and existing attacks; mostly for LWE, Ring-LWE and supersingular ellitic curve isogenies • Three granted patents for post-quantum secure authenticated key exchange scheme
09.2014 – 08.2016	Research Assistant at LACAL, (Laboratory for Cryptologic Algorithms), EPFL • Complexity minimization of Coppersmith's modifications to the GNFS to get optimal parameters; • How to use the cloud to determine key strengths • Breaking RSA keys on large database
09.2015 – 12.2015	Semester project at LCA1, EPFL • Secure meta-analysis of genomic data: decentralized protocol to perform not only linear secure computations
02.2015 – 06.2015	Semester project at LASEC, (Security & Cryptography Laboratory), EPFL • Lattice based cryptography: the LWE problem (state of the art to prove its hardness) and the Renyi divergence to modify a proof and gain better constraints.
01.2014 – 06.2014	Bachelor semester project • Pseudo-random generators using hardcore bits • Understanding the AKS algorithm.

TEACHING

2019	Teaching Assistant Award, (EPFL, IC Section)
	Master-level courses: TCPIP and Smart grid technologies
2018, 2019, 2020	Teaching Assistant
	Master-level courses: TCPIP and Smart grid technologies
2019	Student semester project supervision
	Master and Bachelor level
2018, 2019	Study Advisor
	For International EPFL Master Students
11.2018	Presentation for the high school students day
	Sécurité de l'information: concepts et enjeux.
2017	Teaching Assistant for EPFL 1st year bachelors
	C programming
05.2015	Conference speaker at STIL 2015
	Introduction to cryptography and security issues in new technologies
09.2014 – 12.2014	Teaching Assistant for EPFL 1st year bachelors
	Introductory course to Computer Sciences

SKILLS

Matlab	Advanced Cryptology & Cryptanalysis	Localization Methods
C	TCP/IP networking	Post-Quantum Cryptography
Python	Smart Grids	Network Security
SageMaths	Estimation theory	Information Theory
Latex	Bad Data Detection	Privacy Protection
Adobe Illustrator	Power-Systems State Estimation	Advanced Algorithms
Microsoft Office	Sensor Synchronization	Machine Learning techniques

LANGUAGES

French	Native speaker
English	Fluent – C2 level (TOEFL in 2009)
Serbian	B2 level
German	A2 level

HOBBIES

Reading Novels	Rock Climbing
Boxing	Skiing
Running	Travelling
Hiking	…

PUBLICATIONS & PATENTS

E. Shereen, M. Delcourt, S. Barreto, G. Dán, J. Le Boudec and M. Paolone, "Feasibility of Time-Synchronization Attacks Against PMU-Based State Estimation," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 6, pp. 3412-3427, June 2020, doi: 10.1109/TIM.2019.2939942.

M. Delcourt and J. L. Boudec, "TDOA Source-Localization Technique Robust to Time-Synchronization Attacks," in IEEE Transactions on Information Forensics and Security, doi: 10.1109/TIFS.2020.3001741.

M. Delcourt and J. L. Boudec, " Security Measures for Grids against Rank-1 Undetectable Time-Synchronization Attacks", conditionally accepted for publication in IEEE Transactions on Control of Network Systems.

M. Delcourt, E. Shereen, G. Dán, J. Le Boudec and M. Paolone, " Time-Synchronization Attack Detection in Unbalanced Three-Phase Systems", currently under review in IEEE Transactions on Smart Grid.

M. Delcourt, T Kleinjung, AK Lenstra, S Nath, D Page, NP Smart, «Using the Cloud to Determine Key Strengths-Triennial Update. », IACR Cryptol. ePrint Arch. 2018, 1221.

M. Delcourt, M. Campagna, « Implicit certificates using ring learning with errors », US Patent 10,798,086

M. Delcourt, M. Campagna, « Generation of shared secrets using pairwise implicit certificates », US Patent 10,511,591

M. Delcourt, M. Campagna, « Communication protocol using implicit certificates», US Patent 10,516,543