Syracuse University
SURFACE at Syracuse University

Dissertations - ALL | SURFACE at Syracuse University

1-24-2024

On Energy-efficient Wireless Sensor Networks in the Presence of Byzantines

chen quan
Syracuse University

Follow this and additional works at: https://surface.syr.edu/etd

Recommended Citation

quan, chen, "On Energy-efficient Wireless Sensor Networks in the Presence of Byzantines" (2024). *Dissertations - ALL*. 1867.
https://surface.syr.edu/etd/1867

This Dissertation is brought to you for free and open access by the SURFACE at Syracuse University at SURFACE at Syracuse University. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE at Syracuse University. For more information, please contact surface@syr.edu.

ABSTRACT

Over the years, wireless communication systems have evolved into the most widely used framework for communication devices and networks. These wireless networks form the backbone of wireless sensor networks (WSNs) that have been employed in many applications such as military surveillance, autonomous driving systems and smart-homes. When designing WSNs, two crucial factors must be considered. The first factor is the security of WSNs, which is a concern due to the deployment of low-cost and potentially insecure sensors. The second critical factor is the energy efficiency of WSNs, as they often rely on battery-limited sensors. In this dissertation, we consider the design of various resilient energy-efficient WSNs for the inference task under Byzantine attacks, which are one of the most significant security threats faced by WSNs. When the system suffers from Byzantine attacks, some sensors in the network might be compromised and fully controlled by adversaries. Our goal is to design WSNs that are both energy-efficient and resilient.

The first part of this dissertation (Chapters 2 and 3) focuses on enhancing the resilience of WSNs that achieve energy-efficiency through quantization, particularly in scenarios where Byzantine nodes are prevalent, and the fusion center (FC) lacks knowledge of the attack strategy. Our in-depth exploration, analysis, and enhancements center around a promising energy-efficient mechanism known as the audit bit-based mechanism. For the traditional audit bit-based mechanism, we demonstrate how a simple attack strategy can compromise the entire system. To address this concern, we introduce an enhanced audit bit based mechanism, which effectively relaxes the stringent constraints on the attack strategies that this mechanism can withstand. Building upon the enhanced audit bit framework, we propose an advanced audit bit-based scheme that not only improves system robustness but also significantly reduces redundancy related to audit bits. Furthermore, drawing inspiration from both the audit bit-based mechanism and reputation-based mechanisms, we develop some advanced schemes designed to help systems effectively address challenges in scenarios where prior knowledge of attack strategies is unavailable, and Byzantine nodes are a prevailing threat.

In the next section of this dissertation (Chapters 4 and 5), we study the resilience of WSNs operating under constraints of limited power supply. Our research focuses on the security aspects of two promising energy-efficient frameworks: ordered transmission (Chapter 4) and compressed sensing (Chapter 5). In Chapter 4, we investigate the impact of Byzantine attacks on the performance of both the traditional order transmission based (OT-based) system and the communication-efficient OT-based (CEOT-based) system. We investigate the error probability and the number of saved transmissions for those OT-based systems under various Byzantine attack strategies. Furthermore, we derive upper and lower bounds on the number of transmissions saved for OT-based systems under various Byzantine attack strategies. A comparison of the resilience of CEOT-based and conventional OT-based systems is presented, offering guidance on implementing OT-based frameworks in potentially hostile environments.

In Chapter 5, we investigate the distributed detection problem of sparse stochastic signals with compressed measurements in the presence of Byzantine attacks. We propose two robust detectors based on the traditional Generalized Likelihood Ratio Test (GLRT) and traditional Quantized Locally Most Powerful Test (LMPT) detectors with adaptive thresholds, given that the sparsity degree and the attack strategy are unknown. The proposed detectors can achieve detection performance close to the benchmark likelihood ratio test (LRT) detector with perfect knowledge of the attack strategy and sparsity degree. Furthermore, we explore situations where the fraction of Byzantines in the networks is assumed to be known. In this context, two enhanced detectors building on the previous proposed robust detectors are proposed to further improve the detection performance of the system by filtering out potential malicious sensors.

In addition to our primary focus on traditional WSNs, our research extends to the domain of human-machine collaborative networks. These networks are particularly relevant in high-stake scenarios, such as remote sensing and emergency access systems, where automatic physical sensor-only decision-making may not be sufficient. A combination of human and machine inference networks leverages the cognitive strengths of humans and the sensing capabilities of sensors to enhance situational awareness of the systems. Chapter 6 introduces a belief-updating scheme de-

signed to enhance the resilience of these collaborative networks against potential attacks. The proposed belief-updating scheme, which builds on a human-machine hierarchical network, can also mimic the real-world decision-making process where the sensors' local decisions are collected by humans to make a final decision. Our research reveals that our proposed scheme can improve system performance, even in scenarios where a significant fraction of physical sensors in the system are compromised, and where knowledge about the exact fraction of malicious physical sensors is lacking. Additionally, we conduct an analysis of the impact of side information from individual human sensors, and compare different operations used to incorporate the side information.

ON ENERGY-EFFICIENT WIRELESS SENSOR NETWORKS IN THE PRESENCE OF BYZANTINES

By

Chen Quan
B.E., Nanjing University of Science and Technology, 2016
M.E., Syracuse University, 2018

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering

Syracuse University
December 2023

Copyright © Chen Quan, 2023
All Rights Reserved

ACKNOWLEDGMENTS

First of all, I would like to express my sincere gratitude to my advisor and life mentor, Professor Pramod K. Varshney. His guidance and inspiration have been invaluable throughout my doctoral studies. Without his patience, encouragement, and consistent support, I could not have reached this point and prepared this dissertation. I also cherish every moment spent in his wonderful research 'family', the 'Sensor Fusion Lab'. All the fellow members of this group have been incredibly kind and helpful, including Baocheng, Prashant, Pranay, Qunwei, Shan, Sai, Swatantra, Nandan, Anthony, and Hanne. The time we shared together will be eternally missed, and their assistance and companionship have brought immense joy and happiness to my life here.

I would like to extend my heartfelt appreciation to my defense committee members: Professor Venkata S.S. Gandikota, Professor Pinyuen Chen, and Professor M Cenk Gursoy. Their insightful suggestions and contributions to my research have been invaluable. I also want to express a special thanks to my family. Without their constant love and unwavering support, I would not be in the position I am today. Their encouragement has been a driving force behind my accomplishments, and I am profoundly grateful for their presence in my life. Additionally, I want to take this opportunity to convey my heartfelt thanks to Professor Varshney and Mrs. Varshney for their caring presence in my daily life and their warm invitations to holiday gatherings.

This journey would not have been the same without the support and camaraderie of these remarkable individuals, and for that, I am profoundly grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Over the years, wireless communication systems have evolved into the most widely used framework for communication devices and networks. These wireless networks form the backbone of wireless sensor networks (WSNs) that have been employed in many applications such as military surveillance, autonomous driving systems and smart-homes. One important factor to consider when designing WSNs is the security of their operation in carrying out their assigned tasks. Due to the distributed nature of wireless channels, wireless networks are vulnerable to various kinds of security threats, such as jamming, advanced persistent threats, spoofing, wiretap and Byzantine attacks [14, 26, 39, 54, 113, 119]. Attackers always aim to deteriorate the performance of wireless networks while carrying out their assigned functions, such as sensing performance. The security threat we are particularly interested in is Byzantine attacks, which are one of the most significant security threats faced by WSNs. Another important factor to consider when designing WSNs is the limited power supply. Reducing radio communications' energy consumption is key to sustainability and longevity of WSNs since radio communication is the major component of WSNs that consumes large amounts of energy. The amount of data transmitted over the network is the dominant factor that influences radio communications' energy consumption. Therefore, by quantizing the measurements, reducing the amount of data transmitted or optimizing communication processes, energy-efficiency can be achieved. For example, it can be done through data compression,

power-saving modes, and efficient routing algorithms. In the literature, some promising frameworks have been proposed for improving the energy efficiency of the WSNs such as censoring, ordered transmission and compressive sensing (e.g., [2, 5, 7, 9]). Energy efficiency is achieved by either reducing the number of transmissions in the networks (e.g., censoring and ordered transmission) or compressing the data sent by the sensors (e.g., quantization and compressive sensing). In this dissertation, the task of distributed detection using WSNs is considered and constructing energy-efficient WSNs for this task that are resilient to attacks is the main focus of our work.

1.1 Byzantine Attacks

Byzantine attack is a type of internal attack at the physical layer. This type of attack can be traced back to the issue of Byzantine generals, first introduced by [74], where traitors attempted to mislead other loyal generals by presenting false information. In WSNs, this term specifically refers to the malicious behaviors that occur within WSNs when certain sensors are compromised and transmit false data within the network. There are different types of Byzantine attacks on the distributed detection task in WSNs such as data modification attack, data omission attack and delayed attack. Malicious nodes can either selectively delay data (delayed attack), drop data (data omission attack) or alter data packets directly (data modification attack) to manipulate the network.

In the existing literature, there have been several studies of distributed WSNs under Byzantine attacks (e.g., [48]). The interactions between Byzantines and the WSNs can be viewed as games between attackers and the detection systems. Byzantines aim to undermine the integrity of data transmitted, thereby lowering the reliability of wireless sensor networks. Correspondingly, the FC can enhance the reliability of the network by identifying the Byzantines and making suitable use of information coming from Byzantines for mitigation purposes. Naturally, strategic Byzantine attackers strive to maximize their attack gains while attempting to avoid detection by the defense system.

The level of effort required for an effective mitigation varies depending on the data fusion

system architecture. In centralized fusion systems with a Fusion Center (FC), the system can better evaluate the behavior of all the sensors so that the attacks can be mitigated, especially when the majority of nodes are honest and the FC is trustworthy. However, in decentralized fusion, each sensor can only communicate with its neighbors to gather auxiliary information before making a decision. This decentralized approach makes the system more susceptible to attacks since false data can be stealthily incorporated into the decisions of neighboring nodes, and diffused throughout the network.

1.2 Typical Attack Models in WSNs

There are several factors that can be used to classify Byzantine attacks in WSNs. One such factor is the availability of additional information besides the sensing results at the Byzantine nodes. If no extra information is available, the attacks are referred to as *independent attacks*, meaning that the Byzantine nodes can only rely on their own sensing capabilities. On the other hand, if the attacks involve the acquisition of extra information by the Byzantine nodes, such as the current sensing results of other malicious nodes, fusion rules, and defense strategies, they are referred to as *dependent attacks*. The exchange of information in dependent attacks allows malicious nodes to increase their accuracy in sensing and the success rate of their attacks, making their collusion more effective. One approach to defend against these types of attacks is to use statistical methods to detect and identify malicious nodes that are demonstrating anomalous behavior, e.g., [117].

Another factor is the manner in which attacks are executed. If the attacks are launched with a certain probability, they are referred to as *probabilistic attacks*. Defense algorithms for these types of attacks usually identify attackers by analyzing the consistency of their attack behavior over time, such as reputation-based schemes (e.g., [117]) and cluster-based schemes (e.g., [10]). Conversely, if the attacks are launched based on specific conditions, such as when their posterior probability of being a malicious node exceeds a certain threshold, they are referred to as *non-probabilistic attacks*. These attacks are much harder to model compared to probabilistic attacks and can be very

difficult to defend against, as the Byzantine nodes are intentionally trying to appear normal while causing disruptions.

1.3 Existing Defense Schemes

A number of defense mechanisms have been proposed in the literature to mitigate the negative impact of Byzantine attacks on system performance. They either directly identify and isolate Byzantine attackers to reduce the impact of their attacks or design the system parameters to mitigate the effects of attacks on the system. Some works are based on statistical methods to build reputation so that the malicious nodes are identified (e.g., [70]). For example, the authors in [38] proposed an adaptive reputation-based clustering algorithm for spectrum sensing networks, which was effective in performing detection even in the presence of collaborative attacks. The authors in [70] addressed the problem of Byzantine attacks in distributed inference with M-ary quantized data and proposed a reputation-based defense mechanism, which enables the FC to detect various types of misbehaving nodes and improve detection accuracy. In addition, there are many other promising methods for dealing with Byzantine attacks in networks, such as game-theoretic techniques [63] and machine learning techniques [114]. Several consensus-based algorithms have been used in decentralized fusion to improve their robustness under attack. Efforts have been made to exclude nodes with significant deviations from consensus (e.g., [62]) and to design weights to mitigate the effect of data falsification attacks (e.g., [44, 69]). Additionally, trust-based mechanisms can also be utilized in decentralized fusion, where each node evaluates the trustworthiness of its neighboring nodes before exchanging data. There are also some works that have used the idea of quickest change detection to detect the presence of anomalous measurements due to the malicious sensors in the networks. A model of quickest change detection problems was proposed to detect the presence of Byzantines, who generate fake i.i.d. observations according to post-change and pre-change distributions before and after the change time (e.g., [22, 37]). In [22], the authors utilized a model of quickest change detection problems to detect the presence of Byzantines, who generate

fake i.i.d. observations according to post-change and pre-change distributions before and after the change time. The system can recognize any change due to any subset of affected sensors quickly and reliably. The results they obtained are useful for the robustness of existing multichannel procedures. In [37], the authors formulated and solved the multi-hypothesis Byzantine distributed quickest change detection problem where multiple post-change distributions are considered due to multiple types of attacks.

Aside from works that deal with performance analysis and robust design of networks with fixed sample sizes, there are also studies that deal with performance analysis and robust design of networks with unknown sample sizes, such as sequential hypothesis testing (e.g., [58, 111]). The authors in [111] designed a robust sequential hypothesis test for cooperative spectrum sensing in a mobile network. The authors in [58] investigated the effect of Byzantine attacks on sequential binary hypothesis testing problems in both centralized and fully distributed networks, and proposed asymptotically optimal algorithms to mitigate the effects of Byzantine attacks.

The previously discussed works have made strides in improving the resilience of systems against Byzantine attacks, however, they still have limitations in detecting distributed attacks when a significant number of nodes have been compromised. Some works, such as [33], [34] and [120], have successfully reduced the impact of Byzantine attacks on WSNs even when the majority of sensors are malicious. The authors in [120] proposed a robust framework for identifying Byzantine attackers in collaborative spectrum sensing, with the consideration of two cases: with and without prior knowledge of the attacker's behavior. The framework can still achieve good identification results when a majority of nodes were Byzantine, with the help of stale information[1]. The authors in [33] proposed an audit bit-based distributed detection scheme, where each sensor sends an additional bit to the fusion center (FC) along with its own decision to provide the FC with more information about the behavior of each sensor, was proposed under the Neyman-Pearson framework. The authors showed that the defense system could only be blinded by the attackers if all nodes in the network were Byzantine. The authors in [34] further extended this audit-based

[1]The stale information here is used to denote information that can reflect the real channel states but is outdated for spectrum sensing in the current slot, such as the transmit results.

mechanism to a Bayesian setting and a new decision rule was proposed, taking into account the design of the mitigation scheme over time. But those aforementioned works still have limitations, such as constraints on attack strategies and the need for knowledge of past true hypotheses.

In the literature, there is a large body of work on the performance analysis and robust design of networks that rely on data from all sensors to reach conclusive decisions in the presence of Byzantines. Nonetheless, there are still some areas of research that prioritize energy efficiency. To conserve energy, some approaches, such as censoring-based schemes, ordered transmission-based schemes, and sleep scheduling algorithms, require only a subset of sensors to actively transmit data. The reduced number of sensors needed for active data transmission appears to meet the growing demand for low energy consumption and long-lasting wireless sensor networks in various applications. However, there remains a need for further research efforts focused on the performance analysis and robust design of such energy-efficient wireless sensor networks. Compressed sensing is another representative scheme that achieves energy efficiency by compressing measurements from multiple sensors. However, the integration of data during this process sacrifices specific sensor information, making it challenging to pinpoint compromised sensors. These energy-efficient frameworks are still under investigation in terms of their robustness and their robust design in the context of error-prone environments and under attacks.

1.4 Major Contributions

Our research is primarily dedicated to enhancing the resilience of detection tasks against Byzantine attacks within energy-efficient wireless sensor networks. The increasing adoption of energy-efficient schemes, while beneficial in conserving the energy consumption of individual sensors, might introduce new vulnerabilities that demand robust security measures to be taken. Our work aims to achieve resilience in various types of energy-efficient networks with or without relying on prior knowledge of attack strategies and even when Byzantine nodes dominate the network landscape.

One key focus of our research is to strengthen the resilience of wireless sensor networks that achieve energy-efficiency through quantization, especially in scenarios where Byzantine nodes prevail, and the FC lacks knowledge of the attack strategy. The audit bit based framework proposed in [33, 34] is one promising and intriguing approach to address security threats that arise in scenarios where Byzantine nodes are prevalent. However, the audit bit-based framework discussed in [33, 34] requires prior knowledge of the attack strategy and imposes hard constraints on the behavior of malicious sensors. In our work, we take the traditional audit bit-based framework to the next level. We first conduct in-depth exploration, analysis, and enhancement of this mechanism, and then propose some advanced algorithms based on the idea of the traditional audit bit-based framework. Our contributions include the introduction of an enhanced audit bit-based mechanism, which relaxes the hard constraints on the attack strategies it can withstand. Building upon the enhanced audit bit framework, we propose an advanced audit bit-based scheme that not only enhances system robustness but also significantly reduces the redundancy associated with audit bits. Furthermore, we extend this work to tackle challenges in scenarios where prior knowledge of the attack strategies is unavailable. We introduce an adaptive algorithm that leverages reputation systems and employs advanced audit bit techniques to enhance the network's resilience and security. The proposed adaptive algorithms allow us to guarantee excellent performance even when Byzantine nodes are in the majority.

Another key focus of our research is to enhance the resilience of WSNs when limited power supply is available and energy efficiency is prioritized. Our work is concerned with the security aspects of two kinds of promising energy-efficient frameworks mentioned earlier: ordered transmission and compressed sensing. In ordered transmission-based (OT-based) schemes, energy efficiency is achieved by omitting transmission of less informative data. As only a fraction of sensor data is transmitted to the FC during each decision interval, these systems present greater challenges for ensuring security. The sacrifice of sensors' data introduces complexities in evaluating the reliability of sensors due to the limited availability of complete data records. However, the issue related to resilience of such systems has not been explored in the existing literature. In our

work, the effect of Byzantine attacks on the performance of the OT-based systems is investigated and a comparison of the robustness of two main OT-based systems is made, shedding light on how to employ OT-based frameworks in an attack-prone environment. Some possible countermeasures to mitigate the impact of Byzantines on OT-based systems are also discussed.

In the field of compressed sensing, the achievement of energy efficiency hinges on the compression of high-dimensional sparse data into a lower-dimensional format. However, the amalgamation of data during the compressed sensing process can introduce challenges in identifying compromised sensors, thereby posing a threat to system integrity. In addition, the unknown sparsity of the sparse signal increases the uncertainty of the model. Some promising works, such as [32,101,102], address the sparse signal detection problem within the context of compressed sensing, primarily in attack-free environments. In our research, we conduct a comprehensive evaluation of the impact of Byzantine attacks on the performance of two promising detectors in aforementioned works: the Generalized Likelihood Ratio Test (GLRT) detector introduced in [32], and the Quantized Locally Most Powerful Test (LMPT) detector presented in [101, 102]. Our results reveal the vulnerability of these detectors originally designed for attack-free environments to possible attacks. To address this issue, we propose robust detectors capable of withstanding Byzantine attacks, even in the presence of unknown sparse patterns and unknown attack strategies. The proposed detectors achieve detection performance close to the benchmark LRT detector with perfect knowledge of the attack strategy and sparsity degree.

Beyond our primary focus on traditional WSNs, our research also encompasses the domain of human-machine collaborative networks [87,104,107]. In some high stake scenarios such as remote sensing and emergency access systems, where human lives and assets are at risk, automatic physical sensor-only decision-making may not be sufficient. The emerging human-machine inference networks aim to combine humans' cognitive strength and sensors' sensing capabilities to improve system performance and enhance situational awareness. In our work on human-machine collaborative decision-making, our objective is to enhance the resilience of such collaborative networks to possible attacks. We introduce a belief-updating algorithm within a hierarchical framework which

mimics the real-world decision-making process. The hierarchical framework allows the use of local information collected by human agents and the final decision are made by those human agents. In the real world, human agents collect local decisions from physical sensors and use them as reference points to enhance the quality of human sensor decisions. The same idea is utilized in our proposed algorithm. This innovative strategy ensures the system's performance and significantly enhances the quality of decisions made by human sensors, even in scenarios where a majority of the physical sensors within the system are malicious.

1.5 Organization of the Dissertation

The rest of this dissertation is organized as follows. In Chapter 2, we analyze and improve the audit bit based mechanism, where the prior knowledge of the attacking strategies is assumed to be known. Chapter 3 proposes an adaptive reputation and audit bit based scheme, where the prior knowledge of attacking strategies is unknown and Byzantine nodes are in majority. In Chapter 4, we evaluate the performance of decision making in ordered transmission based systems under Byzantine attacks. In Chapter 5, we propose some resilient detectors for sparse signal detection. In Chapter 6, we propose a belief-updating algorithm based on hierarchical framework that is resilient to Byzantine attacks. In Chapter 7, we summarize the contributions made in this dissertation and present some future directions we intend to pursue.

1.6 Bibliographic Note

The research work appearing in this dissertation has either already been published at various venues or is currently under review. Following is a list of the published/submitted papers, based on the work during the course of my research work at Syracuse University.

Work Included in the Dissertation

Journal Papers:

- C. Quan, Y. S. Han, B. Geng and P. K. Varshney, "*Distributed Quantized Detection of Sparse Signals Under Byzantine Attacks*", accepted by IEEE Transactions on Signal Processing

- C. Quan, S. Bulusu, B. Geng, Y. S. Han, N. Sriranga and P. K. Varshney, "*On Ordered Transmission based Distributed Gaussian Shift-in-Mean Detection under Byzantine Attacks*", IEEE Transactions on Signal Processing, vol. 71, pp. 3343-3356, 2023

- C. Quan, N. Sriranga, H. Yang, Y. S. Han, B. Geng and P. K. Varshney, "*Efficient Ordered-Transmission Based Distributed Detection under Data Falsification Attacks*", IEEE Signal Processing Letters 30 (2023): 145-149.

Conference Papers:

- C. Quan, B. Geng, Y. S. Han and P. K. Varshney, "*Human-machine Hierarchical Networks for Decision Making in the Presence of Byzantine Attacks*", Proc. of the 57th Annual Conference on Information Sciences and Systems (CISS), 2023

- C. Quan, Y. S. Han, B. Geng and P. K. Varshney, "*Reputation and Audit Bit Based Distributed Detection in the Presence of Byzantines*", Proc. of the 56th Asilomar Conference on Signals, Systems, and Computers, August 2022.

- C. Quan, B. Geng and P. K. Varshney, "*Asymptotic performance of binary decision making in heterogeneous human-machine inference networks*", Proc. of the 54th Asilomar Conference on Signals, Systems, and Computers, Nov 2020.

Work not Included in the Dissertation

Journal Papers:

- A. Yadav, C. Quan, P. K. Varshney and H. V. Poor, "*On Performance Comparison of Multi-Antenna HD-NOMA, SCMA, and PD-NOMA Schemes*", IEEE Wireless Communications Letters, vol. 10, no. 4, pp. 715-719, April 2021.

- C. Quan, A. Yadav, B. Geng, P. K. Varshney, and H. V. Poor, "*A novel spectrally-efficient uplink hybrid-domain NOMA system*", IEEE Communications Letters, vol. 24, pp. 2609-2613, Nov, 2020.

- X. Wang , G. Li , C. Quan and P. K. Varshney. "*Distributed detection of sparse stochastic signals with quantized measurements: The generalized Gaussian case*", IEEE Transactions on Signal Processing 67.18 (2019): 4886-4898.

Conference Papers:

- B. Geng, C. Quan, T. Zhang, M. Fardad and P. K. Varshney, "*Loss Attitude Aware Energy Management for Signal Detection*", Proc. of the 56th Asilomar Conference on Signals, Systems, and Computers, August 2022.

- C. Quan, B. Geng and P. K. Varshney, "*On Strategic Jamming in Distributed Detection Networks*", Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4760-4764, May 2021.

- B. Geng, Q. Chen and P. K. Varshney, "*Cognitive Memory Constrained Human Decision Making based on Multi-source Information*", Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5325-5329, May 2021.

- C. Quan, B. Geng and P. K. Varshney, "*Asymptotic performance of binary decision making in heterogeneous human-machine inference networks*", Proc. of the 54th Asilomar Conference on Signals, Systems, and Computers, Nov 2020.

- X. Wang , G. Li , C. Quan and P. K. Varshney, "*Distributed Detection of Generalized Gaussian Sparse Signals with One-Bit Measurements*", Proc. of the 22th International Conference on Information Fusion (FUSION). IEEE, 2019.

CHAPTER 2
ENHANCED AUDIT BIT BASED DISTRIBUTED BAYESIAN DETECTION IN THE PRESENCE OF STRATEGIC ATTACKS

In this chapter, we study detection problem in the presence of Byzantines via an audit bit based approach in a binary hypothesis testing framework. In the traditional audit bit based scheme (TAS) [33, 34], all sensors are divided into groups of two and each sensor sends its local decisions to the FC via two paths, one is direct and another is through the sensor in the same group. However, TAS only considers the case in which each Byzantine node in the network utilizes the same attacking probability to falsify the decisions coming from their group member and its own decision. To consider a more realistic case, we relax the strong assumption of Byzantine nodes' attack behavior made in TAS, namely of equal probability, and we call this type of more general Byzantine nodes as strategic attackers. We evaluate the detection performance of the TAS under strategic attacks and show that it was possible for the strategic attackers to blind the FC as far as the information conveyed by the audit bits in TAS is concerned. To further improve the robustness and the detection performance of the system, we propose two new schemes in this chapter which are the enhanced audit bit based scheme (EAS) and the reduced audit bit based scheme (RAS).

2.1 Introduction

Distributed detection in wireless sensor networks (WSNs) has been studied over the last few decades [94, 95]. In distributed WSNs, instead of sending raw observations, the sensors send their quantized observations or their hard/soft decisions regarding the presence or absence of the phenomenon of interest (PoI) to the FC to make the final decision. This distributed framework is attractive for sensor networks that employ battery-limited sensors in bandwidth-limited environments. Because of the advantages of the distributed mechanism, it is widely used in many applications, such as IoT, cognitive radio networks, object detection networks, distributed spectrum sensing and military surveillance systems [16].

Security is an important issue for the distributed WSNs. The openness of the wireless networks and the distributed nature of such networks make the distributed system more vulnerable to various attacks. The security issues associated with distributed networks are increasingly being studied, e.g., jamming, wiretap, spoofing [14, 26, 39], advanced persistent threats [113] and Byzantine attacks [54, 119]. Here, we focus on Byzantine attacks. When the system suffers from Byzantine attacks, some sensors in the network might be compromised and fully controlled by strategic adversaries. We refer to these compromised sensors as Byzantine nodes. They may send falsified information to the FC. There are several types of Byzantine attacks, such as independent probabilistic attack [75], dependent probabilistic attack [42] and non-probabilistic attack [99]. In probabilistic attacks, the Byzantine nodes are in pursuit of long-term profits by launching attacks with a certain probability. In non-probabilistic attacks, the Byzantine nodes decide to launch attacks only when the observations satisfy some specific conditions. For example, a Byzantine node decides to launch attacks only when its observations are higher than threshold λ_1 or lower than threshold λ_0, where $\lambda_1 > \lambda_0$.

2.1.1 Related Work

There are several works that have studied Byzantine attack issues in distributed detection systems. In [46], optimal strategic data falsification attacks on distributed detection systems are studied. The smart attackers attempt to constrain their exposure to the defense mechanism and maximize the attacking efficacy. In [97], an adaptive algorithm at the FC is proposed to mitigate the impact of Byzantine attacks in the false discovery rate based distributed detection system when the Byzantine nodes know the true hypothesis. In [57, 82, 120], distributed detection problems are investigated in the context of collaborative spectrum sensing under Byzantine attacks. An abnormality-detection-based algorithm for the detection of attackers in collaborative spectrum sensing is proposed in [57]. In [82], the condition under which the Byzantine attackers totally blind the FC is investigated and an algorithm is proposed to detect Byzantine attacks by counting the mismatches between the local decisions and the global decision at the FC. Authors of [120] proposed a Byzantine attacker identification framework in collaborative spectrum sensing where two cases are considered: with and without the prior knowledge of attacker behavior. Good identification performances are achieved in both homogeneous and heterogeneous scenarios even when Byzantine nodes are in a majority. Similarly, in [47], the optimal attacking strategies are analyzed in general distributed network for the cases where the FC has the knowledge of the attackers' strategy and where the FC does not know the attackers' strategy. Audit bit based mechanisms are proposed to mitigate the effect of Byzantine attacks on the distributed WSNs [33,34]. In [33], the audit bit based distributed detection scheme is proposed in the Neyman-Pearson framework by utilizing Kullback-Leibler divergence (KLD) to characterize the detection performance of the system. Each sensor sends one additional audit bit to the FC which gives some information about the behavioral identity of each sensor and improves the detection and security performance of the system. Improved system robustness to Byzantine attacks is achieved at the expense of increased communication overhead. In [34], the audit bit based mechanism is utilized in the Bayesian setting. The detection performance of the system is evaluated in terms of the probability of error of the global decision at the FC, and the mitigation scheme over time is proposed by using the information coming from the audit bits.

Our work is most related to the works in [33] and [34]. In [33] and [34], all the sensors in the network are divided into groups of two. Each sensor sends its local decisions to the FC via two paths, one is direct path and another is through the sensor in the same group (indirect path). The indirect decision bits that reach the FC via indirect path are referred to as audit bits which gives us extra information about the behavioral identity of each sensor. In [33] and [34], it is assumed that each Byzantine node falsifies its own local decisions and the decisions coming from its group member with the same probability.

2.1.2 Major Contributions

Different from the existing works in [33] and [34], we consider a more realistic case in which the strong assumption of Byzantine nodes' attack behavior made in [33] and [34], namely of equal probability, is relaxed. We call this type of Byzantine nodes as strategic attackers. We show that the traditional audit bit based scheme (TAS) is not robust enough in the presence of strategic attackers. Two new schemes, which are the enhanced audit bit based scheme (EAS) and the reduced audit bit based scheme (RAS), are proposed to improve the robustness and the detection performance of the system under strategic attacks. Then, we extend the above RAS for cluster based wide-area wireless sensor networks (CWSNs) [61, 65]. The cluster based framework has been proposed to deal with the significantly increased energy consumption of the sensors due to the long distance transmission in wide-area networks [67, 72]. This framework not only ensures higher data transmission efficiency, larger network scale, lower bandwidth consumption and prolonged network lifetime, but also efficiently reduces the amount of information transmission in the entire network and mitigates energy dissipation due to collisions. In CWSNs, sensors are divided into several clusters and each cluster is equipped with one cluster head (CH) which has ample energy and computation capacities for operation purposes. The CHs are responsible for collecting the data in the cluster and sending it to the FC. In this work, the sensors in each cluster are further divided into groups of two. Each sensor sends its own decisions via direct and indirect path to the corresponding CH just like the previously proposed audit-based system [33] and [34]. The data aggregation rule for

the CHs are designed according to RAS which prolongs the lifetime of the networks with the improved detection performance of the system.[1] We assume that CHs have ample energy to support the long distance transmission[2] and some protections against the attacks so that they can be trusted by the FC, e.g., tamper-resistant security module [76, 98]. The main contributions of this work are summarized as follows:

- We derive the detection performance of the system that employs TAS in the presence of strategic attackers. Instead of considering an identical attacking strategy in which each sensor utilizes the same attacking probability to falsify its own decisions and the decisions coming from their group member [33, 34], we consider attackers that can use different attacking strategies. The optimal attacking strategy of strategic attackers is investigated and we show that it is possible to degrade the performance of TAS to the system without audit bits.

- An EAS is proposed to deal with the security issues arising from the strategic attackers that may use different attacking strategies. We derive the optimal decision rule at the FC and evaluate its detection performance. Simulation results show that the proposed EAS outperforms TAS and the direct scheme under both strategic attacks and non-strategic attacks.

- The scheme EAS is further extended and a new scheme namely RAS is proposed based on our newly proposed EAS. We show that RAS is able to further improve the robustness and the detection performance of the system.

- A wide-area cluster-based WSN is considered. We extend the proposed RAS and design the data aggeration rule for the CHs. Simulation results show a significant reduction in the overall communication overhead between the FC and the CHs.

The key notations and symbols used in this chapter are listed in Table 2.1 for the convenience of readers.

[1] This framework is also suitable for sensor networks with mobile access points (SENMA) where the CHs traverse the network to collect information directly from the sensors [90].

[2] The CHs are assumed to be small base stations that can be charged or be unmanned aerial vehicles (UAVs) that are equipped with energy harvesting (EH) circuits which enable the CHs to harvest energy from renewable sources, e.g., vibration, solar and wind, to replenish their energy buffers [89].

Table 2.1: Glossary

N	number of sensors
G	number of sensor groups
For any sensor $i \in \{1, 2, \ldots, N\}$:	
v_i	the true local decision made by sensor i
u_i	the local decision sent to MMSD (or FC) by sensor i
z_i	the decision sent to MMSD (or FC) by sensor i which represents the decision made by its group member
w_i	the decision sent to the sensor in the same group by sensor i
P_d	the probability of detection for sensor i
P_f	the probability of false alarm for sensor i
For any sensor pair $i \in \{1, \ldots, N\}$ **and** $j \in \{1, \ldots, i-1, i+1, \ldots, N\}$**:**	
p_1	the probability of flipping v_i
p_2	the probability of flipping w_j
d_i	the status indicator which represents the MMS status of sensor i
$\underline{S}$	the set contains all the sensors whose status indicators are equal to 1
$\overline{S}$	the set contains all the sensors whose status indicators are equal to 0
$\underline{SS}$	the set contains all the sensors whose status indicators and group members' status indicators are both equal to 1
$\underline{S}\overline{S}$	the set contains all the sensors whose status indicators is equal to 1 and group members' status indicators is equal to 0
$\overline{S}\underline{S}$	the set contains all the sensors whose status indicators is equal to 0 and group members' status indicators is equal to 1
$\overline{SS}$	the set contains all the sensors whose status indicators and group members' status indicators are both equal to 0
$\underline{\mathcal{M}}$	the set contains all the sensors whose local decisions and group members' local decisions satisfy $u_i = u_j$
Key acronyms:	
TAS	traditional audit bit based scheme
EAS	enhanced audit bit based scheme
RAS	reduced audit bit based scheme
WSNs	wireless sensor networks
CWSNs	cluster based wide-area wireless sensor networks
MMS	match and mismatch
MMSD	match and mismatch detector

2.2 Traditional Audit Bit Based Scheme Under Strategic Attacks

In this section, we first give a brief introduction of the traditional Audit Bit based scheme (TAS). Then, we consider TAS in a more realistic case where the strong assumption of Byzantine nodes' attack behavior made in TAS, namely of equal probability, is relaxed. The performance of TAS is analyzed under the relaxed assumption.

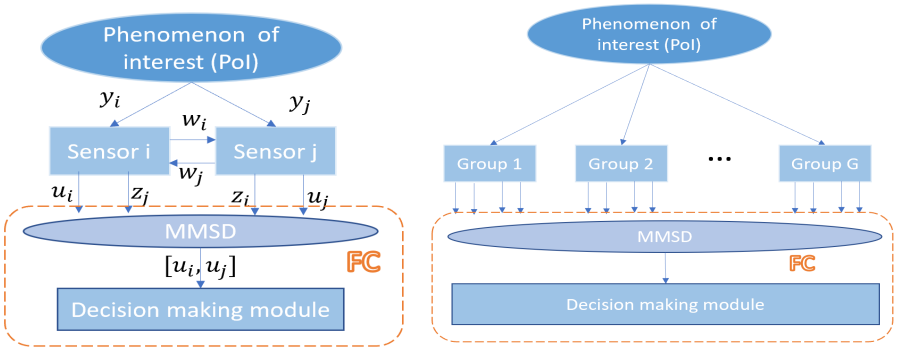
2.2.1 Traditional Audit Bit Based Scheme

We consider the binary hypothesis testing problem assuming that there are two possible hypotheses, H_0 (signal is absent) and H_1 (signal is present), regarding a PoI. Consider that we deploy a cluster of N sensors to determine which of the two hypotheses is true. Based on the local observations, each sensor $i \in \{1, \ldots, N\}$ makes a binary decision $v_i \in \{0, 1\}$ regarding the true hypothesis using the likelihood ratio (LR) test

$$\frac{P(y_i|\mathcal{H}_1)}{P(y_i|\mathcal{H}_0)} \underset{v_i=0}{\overset{v_i=1}{\gtrless}} \lambda, \tag{2.1}$$

where λ is the identical threshold used by all the sensors [91], and, $P(y_i|\mathcal{H}_m)$ denotes the conditional probability density function (PDF) of observation y_i under the hypothesis $\mathcal{H}_m$, for $m = 0, 1$. In the audit bit based framework [33] [34], the N sensors are partitioned into G groups where each group $g \in \{1, \ldots, G\}$ is composed of two sensors.[3] Let i and j represent the sensors in the same group, where $i \in \{1, 2, \ldots, N\}$ and $j \in \{1, 2, \ldots, i-1, i+1, \ldots, N\}$. Each sensor i sends its local binary decision to the FC via two paths, one is direct and the other is through sensor j in the same group. At the FC, we design a match and mismatch detector (MMSD) module that detects if the sensor's direct decision matches or mismatches the decision sent through sensor j (indirect decision).

[3]The sensors are divided into groups of two based on certain criteria, e.g., according to their distances from each other.



(a) The architecture of group k.

(b) The overall detection network for audit bit based scheme.

Fig. 2.1: (a) The architecture of a group $k \in \{1, 2, \ldots, G\}$ and (b) the overall system model.

The architecture of each group is shown as Fig. 2.1(a) and the overall detection network for TAS is shown as Fig. 2.1(b). As shown in Fig. 2.1(a), after making its own decision v_i, sensor i sends (i) u_i directly to the MMSD; (ii) w_i to the sensor j in the same group; (iii) z_j, corresponding to w_j coming from the sensor j in the same group, to the MMSD. Similarly, sensor j also sends two decisions u_j and z_i to the MMSD. If the sensor i is a Byzantine node, i.e., $i = B$, the decisions v_i, w_i and u_i are not necessarily the same and z_j are also not necessarily equal to u_j. Let $p(v_i \neq u_i | i = B)$, $p(v_i \neq w_i | i = B)$ and $p(w_j \neq z_j | i = B)$ denote the probabilities that the Byzantine node i flips its own decision, flips the decision sent to its group member and flips the decision coming from its group member, respectively. The probabilities $p_2 = p(w_j \neq z_j | i = B)$ and $p_1 = p(v_i \neq u_i | i = B) = p(v_i \neq w_i | i = B)$ are the attacking parameters the attackers want to optimize. If the sensor i is honest, i.e., $i = H$, we have $v_i = w_i = u_i$ and $z_j = u_j$. In other words, $p(v_i \neq u_i | i = H) = p(v_i \neq w_i | i = H) = 0$. We assume that a fraction α_0 of the N sensors are Byzantine nodes and the FC is not aware of the identity of Byzantine nodes in the network. Hence, each node has the probability of α_0 to be a Byzantine node. We also assume that each Byzantine node attacks the network independently with a certain probability and all the sensors are able to successfully receive the packets from their group members.

After collecting all the local decisions, the MMSD makes binary decisions regarding the match and mismatch (MMS) status of the two decisions corresponding to the same sensor received over

different paths, i.e., whether or not the decisions sent via different paths are the same, for all the sensors. Let d_i represent the MMS status of sensor i which is called the status indicator of sensor i. To give a concrete illustration, take one group of sensors (i, j) as an example. The MMSD sets $d_j = 1$ when $u_i = z_i$ and $d_j = 0$ when $u_i \neq z_i$. Similarly, the MMSD sets $d_i = 1$ if $u_j = z_j$ and $d_i = 0$ if $u_j \neq z_j$. The decisions d_i and d_j are the status indicators of sensor i and sensor j, respectively. According to the status indicator for each sensor, the FC places the sensors into two sets $\underline{S}$ and $\overline{S}$. Set $\underline{S}$ contains the sensors whose status indicators are equal to 1 and Set $\overline{S}$ contains the sensors whose status indicators are equal to 0. By employing the extra information coming from these status indicators, we are able to improve the detection performance of the system.

In the following two subsections, we discuss two different attack models and investigate the robustness of the traditional audit bit based mechanism under these two types of attacks. One attack model[4] is that the Byzantine nodes are assumed to flip their own decisions and all the decisions they received with the same probability p, i.e., $p_1 = p_2 = p$. The other model is that the Byzantine nodes use different probabilities to flip their own decisions and all the decisions they receive, i.e., $p_1 \neq p_2$. It is more general and practical to consider Byzantine nodes which relax the assumption of $p_1 = p_2 = p$ made in the traditional audit bit based mechanism. This allows the Byzantines to be strategic by optimally employing unequal probabilities p_1 and p_2.

2.2.2 Traditional Audit Bit based Scheme

In the traditional audit bit based mechanism, the Byzantine nodes are assumed to flip their own decisions and all the decisions they receive with the same probability p, i.e., $p_1 = p_2 = p$. Based on the status indicators $\{d_i\}_{i=1}^{N}$, we have the following two cases [33].

1. If $d_i = 1$, i is a Byzantine node with probability

[4]This attack model follows the work in [33] and [34].

$$\begin{aligned}\underline{\alpha} &= P(i = B|d_i = 1) \\ &= \frac{\alpha_0(1-p)[1-2\alpha_0 p(1-2p)]}{1-\alpha_0(3-2p)p+4\alpha_0^2(1-p)p^2}\end{aligned} \tag{2.2}$$

and the sensor i is placed in set $\underline{S}$.

2. If $d_i = 0$, i is a Byzantine node with probability

$$\begin{aligned}\overline{\alpha} &= P(i = B|d_i = 0) \\ &= \frac{1+2(1-p)(\alpha_0-2\alpha_0 p))}{1+2(1-p)(1-2\alpha_0 p))}\end{aligned} \tag{2.3}$$

and the sensor i is placed in set $\overline{S}$.

It has been proved in [34] (Lemma 1) that $\underline{\alpha} \leq \alpha_0 \leq \overline{\alpha}$. In other words, all the sensors are divided into two sets $\underline{S}$ and $\overline{S}$ in which the sensors have lower probability $\underline{\alpha}$ and higher probability $\overline{\alpha}$ of being Byzantine nodes, respectively, according to status indicators $\mathbf{d} = [d_1, d_2, \ldots, d_N]$. Let P_d, P_f be the probability of detection and the probability of false alarm for any sensor $i \in \{1, \ldots, N\}$, respectively, i.e., $P_d = P(v_i = 1|\mathcal{H}_1)$ and $P_f = P(v_i = 1|\mathcal{H}_0)$. Thus, the probability mass function (pmf) of local decision u_i is expressed as

$$P(u_i|\mathcal{H}_q) = \begin{cases} \underline{\pi}_{1q}^{u_i}(1-\underline{\pi}_{1q})^{1-u_i} & \text{for } i \in \underline{S} \\ \overline{\pi}_{1q}^{u_i}(1-\overline{\pi}_{1q})^{1-u_i} & \text{for } i \in \overline{S} \end{cases} \tag{2.4}$$

for q=0,1, where, for $i \in \underline{S}$,

$$\underline{\pi}_{11} = 1 - \underline{\pi}_{01} = P(u_i = 1|\mathcal{H}_1) = P_d(1-\underline{\alpha}p) + \underline{\alpha}p(1-P_d) \tag{2.5a}$$

$$\underline{\pi}_{10} = 1 - \underline{\pi}_{00} = P(u_i = 1|\mathcal{H}_0) = P_f(1-\underline{\alpha}p) + \underline{\alpha}p(1-P_f) \tag{2.5b}$$

and, for $i \in \overline{S}$,

$$\overline{\pi}_{11} = 1 - \overline{\pi}_{01} = P(u_i = 1|\mathcal{H}_1) = P_d(1 - \overline{\alpha}p) + \overline{\alpha}p(1 - P_d) \tag{2.6a}$$

$$\overline{\pi}_{10} = 1 - \overline{\pi}_{00} = P(u_i = 1|\mathcal{H}_0) = P_f(1 - \overline{\alpha}p) + \overline{\alpha}p(1 - P_f). \tag{2.6b}$$

Then the optimal decision rule when the attacking strategy p is assumed to be known is given as

$$\underline{WU} + \overline{WU} \gtrless \eta^{(A)}, \tag{2.7}$$

where $\underline{U} = \sum_{i\in\underline{S}} u_i$, $\overline{U} = \sum_{i\in\overline{S}} u_i$, $\underline{W} = \log(\frac{\underline{\pi}_{11}(1-\underline{\pi}_{10})}{\underline{\pi}_{10}(1-\underline{\pi}_{11})})$, $\overline{W} = \log(\frac{\overline{\pi}_{11}(1-\overline{\pi}_{10})}{\overline{\pi}_{10}(1-\overline{\pi}_{11})})$, $\eta^{(A)} = \log(\frac{\pi_0}{\pi_1}) + \underline{N}\log(\frac{1-\underline{\pi}_{10}}{1-\underline{\pi}_{11}}) + \overline{N}\log(\frac{1-\overline{\pi}_{10}}{1-\overline{\pi}_{11}})$, $\underline{N} = |\underline{S}|$, and $\overline{N} = |\overline{S}|$. Note that $\underline{U}$ and $\overline{U}$ are binomial distributed random variables with parameters $(N, \underline{\pi}_{10})$ and $(N, \overline{\pi}_{10})$, respectively, under $\mathcal{H}_0$, and with parameters $(N, \underline{\pi}_{11})$ and $(N, \overline{\pi}_{11})$, respectively, under $\mathcal{H}_1$. When N is large, $\underline{N}$ and $\overline{N}$ can be approximated by their expected value $NP(u_i = z_i)$ and $NP(u_i \neq z_i)$. $\eta^{(A)}$ is the threshold used by the FC for the traditional audit bit based system, where $\eta^{(A)} = \log(\frac{\pi_0}{\pi_1}) + NP(u_i = z_i)\log(\frac{1-\underline{\pi}_{10}}{1-\underline{\pi}_{11}}) + NP(u_i \neq z_i)\log(\frac{1-\overline{\pi}_{10}}{1-\overline{\pi}_{11}})$. Moreover, $\underline{U}$ and $\overline{U}$ can be approximated by the Gaussian distribution with parameters given as follows:

$$\begin{aligned}\mu_m^{(A)} =& E[U|\mathcal{H}_m] \\ =& N[P(u_i = z_i)\underline{\pi}_{1m}\underline{W} + P(u_i \neq z_i)\overline{\pi}_{1m}\overline{W}] \qquad (2.8a)\\ (\sigma_m^{(A)})^2 =& Var[U|\mathcal{H}_m] = N[P(u_i = z_i)\underline{\pi}_{1m}(1 - \underline{\pi}_{1m})\underline{W}^2 \\ &+ P(u_i \neq z_i)\overline{\pi}_{1m}(1 - \overline{\pi}_{1m})\overline{W}^2], \qquad (2.8b)\end{aligned}$$

for $m = 0, 1$. The detection performance, characterized by the probability of error $P_e^{(A)}$ for the system with TAS, is given as

$$P_e^{(A)} = \pi_0 Q\left(\gamma_f^{(A)}\right) + \pi_1 Q\left(\gamma_m^{(A)}\right), \tag{2.9}$$

where $\gamma_f^{(A)} = \frac{\eta^{(A)}-\mu_0^{(A)}}{\sigma_0^{(A)}}$ and $\gamma_m^{(A)} = \frac{\mu_1^{(A)}-\eta^{(A)}}{\sigma_1^{(A)}}$. Let $P_e^{(D)}$ denote the probability of error for the system with the direct scheme, which is expressed as (A.4). It has been shown in [34] (Theorem 3) that the probability of error of the traditional audit based system given any α_0 and p is always less than or equal to that of the system which relies only on direct decisions, i.e, $P_e^{(A)} \leq P_e^{(D)}$.

However, due to the strong assumption of $p_1 = p_2 = p$, TAS can accurately assess the behavioral identity of each sensor in the network so that it can improve the detection and security performances of the system. It is obvious that a higher p means a higher probability that the Byzantine nodes flip their own decisions and the decisions coming from their group members. Thus, the Byzantine nodes have a higher probability of being placed in the Set $\overline{S}$. In the next subsection, we relax the assumption of $p_1 = p_2 = p$ and investigate the detection performance of the traditional audit bit based system under the relaxed assumption.

2.2.3 The Strategic Attacker under Traditional Audit Bit based System

To make the model more general, we assume that the attackers are more strategic in that they can employ different values of p_1 and p_2 that are not necessarily equal. This allows the Byzantines to be strategic by optimally employing unequal probabilities p_1 and p_2. In this subsection, we analyze the detection performance of the traditional audit bit based system under such strategic attacks.

When the FC under strategic attacks makes use of the status indicators to place all the sensors into two sets, we have the following two cases.

- If $d_i = 1$, i is a Byzantine node with probability

$$\begin{aligned} \underline{\alpha}^I &= P(i = B|d_i = 1) \\ &= \frac{P(d_i = 1|i = B)P(i = B)}{P(d_i = 1)}, \end{aligned} \tag{2.10}$$

where

$$\begin{aligned} &P(d_i = 1|i = B) = P(u_j = z_j|i = B) \\ &= \alpha_0 p_1^2(1-p_2) + \alpha_0 p_1(1-p_1)p_2 + \alpha_0(1-p_1)^2(1-p_2) \\ &+ (1-\alpha_0)(1-p_2) + \alpha_0(1-p_1)p_1p_2 \\ &= -4\alpha_0 p_1^2 p_2 + 4\alpha_0 p_1 p_2 - 2\alpha_0 p_1 + 2\alpha_0 p_1^2 - p_2 + 1 \end{aligned} \tag{2.11}$$

and

$$\begin{aligned} &P(d_i = 1|i = H) = P(u_j = z_j|i = H) \\ &= \alpha_0 p_1^2 + \alpha_0(1-p_1)^2 + (1-\alpha_0) \\ &= 2\alpha_0 p_1^2 - 2\alpha_0 p_1 + 1. \end{aligned} \tag{2.12}$$

Thus, the unconditional probability of matching $p(u_j = z_j)$ is given as

$$\begin{aligned} P(d_i = 1) &= P(u_j = z_j|i = H)P(i = H) + P(u_j = z_j|i = B)P(i = B) \\ &= -4\alpha_0^2 p_1^2 p_2 + 4\alpha_0^2 p_1 p_2 + 2\alpha_0 p_1^2 - \alpha_0 p_2 - 2\alpha_0 p_1 + 1. \end{aligned} \tag{2.13}$$

In this case, the sensor i is placed in set $\underline{S}$ with

$$\underline{\alpha}^I = \frac{4\alpha_0^2 p_1^2 p_2 + 4\alpha_0^2 p_1 p_2 - 2\alpha_0^2 p_1 + 2\alpha_0^2 p_1^2 - \alpha_0 p_2 + \alpha_0}{4\alpha_0^2 p_1^2 p_2 + 4\alpha_0^2 p_1 p_2 + 2\alpha_0 p_1^2 - \alpha_0 p_2 - 2\alpha_0 p_1 + 1}. \tag{2.14}$$

- If $d_i = 0$, i is a Byzantine node with probability

$$\begin{aligned}\overline{\alpha}^I &= P(i=B|d_i=0)\\ &= P(i=B|u_j\neq z_j)\\ &= \frac{P(u_j\neq z_j|i=B)P(i=B)}{P(u_j\neq z_j)}\\ &= \frac{4\alpha_0p_1^2p_2-4\alpha_0p_1p_2+2\alpha_0p_1-2\alpha_0p_1^2+p_2}{4\alpha_0p_1^2p_2-4\alpha_0p_1p_2-2p_1^2+p_2+2p_1},\end{aligned} \tag{2.15}$$

where $p(u_j \neq z_j|i=B) = 1-p(u_j=z_j|i=B)$ and $p(u_j\neq z_j) = 1-p(u_j=z_j)$. In this case, the sensor i is placed in set $\overline{S}$.

We show two important properties of $\overline{\alpha}^I$ and $\underline{\alpha}^I$ in the next lemma.

Lemma 2.1. *We have the following two relationships in terms of $\underline{\alpha}^I$, $\overline{\alpha}^I$, and α_0.*

1. *Under strategic attacks, the probability of being a strategic node given the sensor in Set $\underline{S}$ is smaller than or equal to the one given the sensor in Set $\overline{S}$, i.e., $\underline{\alpha}^I \leq \alpha_0 \leq \overline{\alpha}^I$.*

2. *$\underline{\alpha}^I = \overline{\alpha}^I = \alpha_0$ when $p_2 = 0$.*

PROOF: According to (2.14) and (2.15), we show that $\frac{\partial\underline{\alpha}^I}{\mathrm{d}p_2}\leq 0$, and $\frac{\partial\underline{\alpha}^I}{\mathrm{d}p_1}\leq 0$. Due to the fact that $\alpha_0\in[0,1]$, $p_1\in[0,1]$, and $p_2\in[0,1]$, we have

$$\begin{aligned}\frac{\partial\underline{\alpha}^I}{\mathrm{d}p_2} &= \frac{(4\alpha_0^2p_1(1-p_1)-\alpha_0)(1-\alpha_0)(2\alpha_0p_1(p_1-1)+1)}{(4\alpha_0^2p_1^2p_2+4\alpha_0^2p_1p_2+2\alpha_0p_1^2-\alpha_0p_2-2\alpha_0p_1+1)^2}\\ &\overset{(a)}{\leq} \frac{\alpha_0(\alpha_0-1)(1-\alpha_0)(2\alpha_0p_1(p_1-1)+1)}{(4\alpha_0^2p_1^2p_2+4\alpha_0^2p_1p_2+2\alpha_0p_1^2-\alpha_0p_2-2\alpha_0p_1+1)^2}\\ &\overset{(b)}{\leq} 0\end{aligned} \tag{2.16a}$$

$$\frac{\partial\underline{\alpha}^I}{\mathrm{d}p_1} = -2\alpha_0^2(1-\alpha_0p_2)(1-2p_2)^2\leq 0. \tag{2.16b}$$

The equality in (a) is achieved when $p_1=\frac{1}{2}$. (b) is due to the fact that $2\alpha_0p_1(p_1-1)+1\geq 1-\frac{\alpha_0}{2}>0$ and the equality in (b) is achieved when $\alpha_0=1$. Thus, according to (A.11), we have

$$\gamma_f^I = \frac{\log(\frac{\pi_0}{\pi_1})/\sqrt{N} + \sqrt{N}(D_0(\overline{\alpha}^I, p_1, p_2)p(u_n = z_n) + D_0(\underline{\alpha}^I, p_1, p_2)p(u_n \neq z_n))}{\sqrt{p(u_i \neq z_i)g_0(\overline{\alpha}^I, p_1, p_2) + p(u_i = z_i)g_0(\underline{\alpha}^I, p_1, p_2)}} \tag{2.18a}$$

$$\gamma_m^I = \frac{\log(\frac{\pi_0}{\pi_1})/\sqrt{N} + \sqrt{N}(D_1(\overline{\alpha}^I, p_1, p_2)p(u_n = z_n) + D_1(\underline{\alpha}^I, p_1, p_2)p(u_n \neq z_n))}{\sqrt{p(u_i \neq z_i)g_1(\overline{\alpha}^I, p_1, p_2) + p(u_i = z_i)g_1(\underline{\alpha}^I, p_1, p_2)}}, \tag{2.18b}$$

the maximum value of $\underline{\alpha}^I$ when $p_1 = 0$ and $p_2 = 0$, i.e., $\underline{\alpha}^I(p_1, p_2) \leq \underline{\alpha}^I(p_1 = 0, p_2 = 0) = \alpha_0$. Since $p(u_i = z_i)\underline{\alpha}^I + p(u_i \neq z_i)\overline{\alpha}^I = \alpha_0$, we have

$$\begin{aligned} P(u_i = z_i)\alpha_0 + P(u_i \neq z_i)\overline{\alpha}^I &\geq \alpha_0 \\ P(u_i \neq z_i)\overline{\alpha}^I &\geq \alpha_0(1 - P(u_i = z_i)) \\ \overline{\alpha}^I &\geq \alpha_0 \end{aligned} \tag{2.17}$$

Based on the analysis above, we conclude that $\underline{\alpha}^I \leq \alpha_0 \leq \overline{\alpha}^I$. Note that the equality on both sides can be achieved when $p_2 = 0$. Hence, we get the results stated in Lemma 2.1. ■

Substituting $\underline{\alpha}$ and $\overline{\alpha}$ with $\underline{\alpha}^I$ and $\overline{\alpha}^I$, respectively, in (2.5) and (2.6), we can obtain $\underline{\pi}_{10}^I$, $\underline{\pi}_{11}^I$, $\overline{\pi}_{10}^I$, $\overline{\pi}_{11}^I$. After getting $\underline{\pi}_{10}^I$, $\underline{\pi}_{11}^I$ and $\overline{\pi}_{10}^I$, $\overline{\pi}_{11}^I$, we can calculate the pmfs of u_i according to (2.4). Hence, the probability of error for the system under strategic attack is given by $P_e^I = \pi_0 Q\left(\gamma_f^I\right) + \pi_1 Q\left(\gamma_m^I\right)$. γ_f^I and γ_m^I are shown in (2.18), where $D_0(\underline{\alpha}^I, p_1, p_2) = \underline{\pi}_{10}^{(I)} \log(\frac{\underline{\pi}_{10}^{(I)}}{\underline{\pi}_{11}^{(I)}}) + (1 - \underline{\pi}_{10}^{(I)}) \log(\frac{1-\underline{\pi}_{10}^{(I)}}{1-\underline{\pi}_{11}^{(I)}})$, $D_0(\overline{\alpha}^I, p_1, p_2) = \overline{\pi}_{10} \log(\frac{\overline{\pi}_{10}^{(I)}}{\overline{\pi}_{11}^{(I)}}) + (1 - \overline{\pi}_{10}^{(I)}) \log(\frac{1-\overline{\pi}_{10}^{(I)}}{1-\overline{\pi}_{11}^{(I)}})$, $D_1(\overline{\alpha}^I, p_1, p_2) = \overline{\pi}_{11}^{(I)} \log(\frac{\overline{\pi}_{11}^{(I)}}{\overline{\pi}_{10}^{(I)}}) + (1 - \overline{\pi}_{11}^{(I)}) \log(\frac{1-\overline{\pi}_{11}^{(I)}}{1-\overline{\pi}_{10}^{(I)}})$ and $D_1(\underline{\alpha}^I, p_1, p_2) = \underline{\pi}_{11}^{(I)} \log(\frac{\underline{\pi}_{11}^{(I)}}{\underline{\pi}_{10}^{(I)}}) + (1 - \underline{\pi}_{11}^{(I)}) \log(\frac{1-\underline{\pi}_{11}^{(I)}}{1-\underline{\pi}_{10}^{(I)}})$. We also have $g_0(\underline{\alpha}^I, p_1, p_2) = \underline{\pi}_{10}^{(I)}(1-\underline{\pi}_{10}^{(I)})\underline{W}^2$, $g_0(\overline{\alpha}^I, p_1, p_2) = \overline{\pi}_{10}^{(I)}(1-\overline{\pi}_{10}^{(I)})(\overline{W}^I)^2$ and $g_1(\underline{\alpha}^I, p_1, p_2) = \underline{\pi}_{11}^{(I)}(1 - \underline{\pi}_{11}^{(I)})\underline{W}^2$, $g_1(\overline{\alpha}^I, p_1, p_2) = \overline{\pi}_{11}^{(I)}(1 - \overline{\pi}_{11}^{(I)})(\overline{W}^I)^2$ where $\underline{W}^I = \log(\frac{\underline{\pi}_{11}^I(1-\underline{\pi}_{10}^I)}{\underline{\pi}_{10}^I(1-\underline{\pi}_{11}^I)})$ and $\overline{W}^I = \log(\frac{\overline{\pi}_{11}^I(1-\overline{\pi}_{10}^I)}{\overline{\pi}_{10}^I(1-\overline{\pi}_{11}^I)})$. The optimal attacking strategy is stated based on (2.18) in the following theorem.

Theorem 2.1. *In the traditional audit based system, if the strategic Byzantine attackers adopt the strategy given by $p_2 = 0$ when $\alpha_0 \in [0, 1]$, the system reduces to the one without audit bits and it can always be made blind by choosing p_1 such that $\alpha_0 p_1 = \frac{1}{2}$ if $\alpha_0 \geq 0.5$.*

PROOF: Please see Appendix A.1. ■

Note that the probability of error for the system under strategic attack is $P_e^I = \pi_0 Q\left(\gamma_f^I\right) + \pi_1 Q\left(\gamma_m^I\right)$. γ_f^I and γ_m^I are the arguments of function $Q(.)$ for the probability of false alarm and the argument of function $Q(.)$ for the probability of miss detection, respectively such that larger arguments mean better detection performance. Fig. 2.2 shows how γ_f^I and γ_m^I change with p_2. We can observe that both γ_f^I and γ_m^I achieve the minimum when $p_2 = 0$, which means that P_e^I achieves the maximum. We can also observe that arguments that attain this are equal to the ones in the system that does not use audit bits and thus P_e^I reduces to the probability of error of the system that does not use audit bits. Hence, Fig. 2.2 is in accordance with the result given in Theorem 2.1.

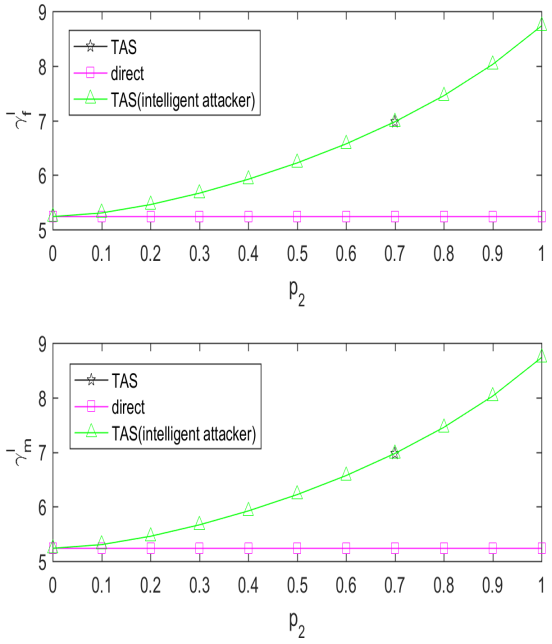


Fig. 2.2: γ_f^I and γ_m^I versus p_2 given $p_1 = 0.7$ and $\alpha_0 = 0.3$. Note that $p_1 = p_2 = 0.7$ in TAS.

Based on the analysis above, the assumption $p_1 = p_2$ given in [34] is not the optimal choice for the attackers in practice. The attackers can launch stronger attacks when they set $p_2 = 0$. Under this attacking strategy, there is no improvement in the detection performance of TAS compared with the direct scheme. Thus, we conclude that the strategic attackers can hide themselves by not flipping the decisions from their group members, i.e., $p_2 = 0$, according to Theorem 2.1 and Fig. 2.2. Moreover, when $p_2 = 0$, the detection error for TAS is the same as the one for the direct scheme. To enhance the robustness of the system, we propose a new scheme called enhanced audit

bit based scheme (EAS) in the next section.

2.3 Enhanced Audit Bit based Scheme

In this section, an enhanced audit bit based scheme (EAS) is proposed to improve the robustness of the system under strategic attacks. In TAS, the behavioral identity of each sensor is characterized by $\underline{\alpha}$ and $\overline{\alpha}$. The evaluations of the value of $\underline{\alpha}$ and $\overline{\alpha}$ only depends on its own status indicator as discussed in Section 2.2. However, in the newly proposed scheme, we utilize both the status indicators of the sensors in the same group to more accurately infer the behavioral identities of sensors in the network compared with TAS.

2.3.1 Audit Bits in the Same Group as Extra Information

The status indicators $\{d_i\}_{i=1}^N$ are again made by the MMSD. However, the sensors are no longer partitioned into two sets ($\underline{S}$ and $\overline{S}$). They are partitioned into four sets which are $\underline{SS}$ $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$ based on both status indicators of sensor i and sensor j in the same group. If $d_i = d_j = 1$, sensor i and sensor j are both placed in the set $\underline{SS}$. If $d_i = 0$ and $d_j = 1$, sensor i is placed in the set $\underline{S}\overline{S}$ and sensor j is placed in the set $\overline{S}\underline{S}$. If $d_i = d_j = 0$, sensor i and sensor j are both placed in the set $\overline{SS}$. We still assume a general attacking strategy which is $p_1 \neq p_2$. Then, we have the following four cases.

- If $i \in \underline{SS}$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_1 =&P(i=B|i,j\in\underline{SS})\\ =&P(i=B,j=H|i,j\in\underline{SS})+P(i=B,j=B|i,j\in\underline{SS})\\ =&\frac{P(i,j\in\underline{SS}|i=B,j=H)P(i=B,j=H)}{P(i,j\in\underline{SS})}\\ &+\frac{P(i,j\in\underline{SS}|i=B,j=B)P(i=B,j=B)}{P(i,j\in\underline{SS})}\\ =&\frac{\alpha_0^2 f_{BB}^{(1)}+\alpha_0(1-\alpha_0)f_{BH}^{(1)}}{P(i,j\in\underline{SS})},\end{aligned} \tag{2.19}$$

where

$$P(i,j\in\underline{SS})=\alpha_0^2 f_{BB}^{(1)}+\alpha_0(1-\alpha_0)(f_{HB}^{(1)}+f_{BH}^{(1)})+(1-\alpha_0)^2 f_{HH}^{(1)} \tag{2.20}$$

and $f_{BB}^{(1)}=[2p_1p_2(1-p_1)+(1-2p_1+2p_1^2)(1-p_2)]^2$, $f_{HB}^{(1)}=f_{BH}^{(1)}=(1-p_2)(1-2p_1+2p_1^2)$ and $f_{HH}^{(1)}=1$.

- If $i\in\underline{S}\overline{S}$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_2 =&P(i=B|i\in\underline{S}\overline{S},j\in\overline{S}\underline{S})\\ =&P(i=B,j=H|i\in\underline{S}\overline{S},j\in\overline{S}\underline{S})+P(i=B,j=B|i\in\underline{S}\overline{S},j\in\overline{S}\underline{S})\\ =&\frac{\alpha_0^2 f_{BB}^{(2)}+\alpha_0(1-\alpha_0)f_{BH}^{(2)}}{P(i\in\underline{S}\overline{S},j\in\overline{S}\underline{S})},\end{aligned} \tag{2.21}$$

where

$$P(i\in\underline{S}\overline{S},j\in\overline{S}\underline{S})=\alpha_0^2 f_{BB}^{(2)}+\alpha_0(1-\alpha_0)(f_{HB}^{(2)}+f_{BH}^{(2)})+(1-\alpha_0)^2 f_{HH}^{(2)} \tag{2.22}$$

and $f_{BB}^{(2)}=[2p_1p_2(1-p_1)+(1-2p_1+2p_1^2)(1-p_2)][1-2p_1p_2(1-p_1)-(1-2p_1+2p_1^2)(1-p_2)]$, $f_{HB}^{(2)}=p_2(1-2p_1+2p_1^2)$,$f_{BH}^{(2)}=2p_1(1-p_2)(1-p_1)$ and $f_{HH}^{(2)}=0$.

- If $i \in \overline{S}\underline{S}$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_3 =&P(i = B|i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S})\\ =&P(i = B, j = H|i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S}) + P(i = B, j = B|i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S})\\ =&\frac{\alpha_0^2 f_{BB}^{(3)} + \alpha_0(1-\alpha_0) f_{BH}^{(3)}}{P(i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S})},\end{aligned} \tag{2.23}$$

where

$$P(i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S}) =\alpha_0^2 f_{BB}^{(3)} + \alpha_0(1-\alpha_0)(f_{HB}^{(3)} + f_{BH}^{(3)}) + (1-\alpha_0)^2 f_{HH}^{(3)} \tag{2.24}$$

and $f_{BB}^{(3)} = [2p_1p_2(1-p_1)+(1-2p_1+2p_1^2)(1-p_2)][1-2p_1p_2(1-p_1)-(1-2p_1+2p_1^2)(1-p_2)]$, $f_{HB}^{(3)} = 2p_1(1-p_2)(1-p_1)$,$f_{BH}^{(3)} = p_2(1-2p_1+2p_1^2)$ and $f_{HH}^{(3)} = 0$.

- If $i \in \overline{SS}$ i is a Byzantine node with probability

$$\begin{aligned}\alpha_4 =&P(i = B|i, j \in \overline{SS})\\ =&P(i = B, j = H|i, j \in \overline{SS}) + P(i = B, j = B|i, j \in \overline{SS})\\ =&\frac{\alpha_0^2 f_{BB}^{(4)} + \alpha_0(1-\alpha_0) f_{BH}^{(4)}}{P(i, j \in \overline{SS})},\end{aligned} \tag{2.25}$$

where

$$P(i, j \in \overline{SS}) =\alpha_0^2 f_{BB}^{(4)} + \alpha_0(1-\alpha_0)(f_{HB}^{(4)} + f_{BH}^{(4)}) + (1-\alpha_0)^2 f_{HH}^{(4)} \tag{2.26}$$

and $f_{BB}^{(4)} = [2p_1(1-p_2)(1-p_1) + p_2p_1^2]^2$, $f_{HB}^{(4)} = f_{BH} = 2p_1p_2(1-p_1)$ and $f_{HH}^{(4)} = 0$.

The next lemma shows that our proposed EAS performs a more accurate evaluation of the behavioral identity of each sensor compared with TAS.

Lemma 2.2. *The probability of sensor i being a Byzantine node when $i \in \underline{S}$ in TAS is equal to the weighted average of the probabilities of sensor i being a Byzantine node when $i, j \in \underline{SS}$ and $i \in \underline{S}\overline{S}, j \in \overline{S}\underline{S}$, respectively. That is*

$$\begin{aligned} &P(i = B | i \in \underline{S}) \\ &= \alpha_1 P(d_j = 1 | d_i = 1) + \alpha_2 P(d_j = 0 | d_i = 1) \end{aligned} \tag{2.27}$$

A similar result can be obtained for sensor $i \in \overline{S}$.

PROOF: The right hand side (RHS) of (2.27) is the same as $P(i = B|d_i = 1, d_j = 1)P(d_j = 1|d_i = 1) + P(i = B|d_i = 1, d_j = 0)P(d_j = 0|d_i = 1)$. According to the Bayes' rule, we have

$$\begin{aligned} &\sum_{x=0,1} P(i = B | d_i = 1, d_j = x) P(d_j = x | d_i = 1) \\ &= \sum_{x=0,1} P(i = B | d_i = 1, d_j = x) \frac{P(d_i = 1, d_j = x)}{P(d_i = 1)} \\ &= \sum_{x=0,1} \frac{\alpha_0^2 f_{BB}^{(x)} + \alpha_0(1-\alpha_0) f_{BH}^{(x)}}{P(d_i = 1, d_j = x)} \frac{P(d_i = 1, d_j = x)}{P(d_i = 1)} \\ &= \sum_{x=0,1} \frac{\alpha_0^2 f_{BB}^{(x)} + \alpha_0(1-\alpha_0) f_{BH}^{(x)}}{P(d_i = 1)} \\ &= P(i = B | i \in \underline{S}) \end{aligned} \tag{2.28}$$

We can also show that $i \in \overline{S}$ is the weighted average of the probabilities of sensor i being a Byzantine node when $i, j \in \overline{SS}$ and $i \in \overline{S}\underline{S}, j \in \underline{S}\overline{S}$ by following a similar procedure and, therefore, the details of its proof are omitted here. ■

Fig. 2.3 corroborates the results in Lemma 2.2. Note that each sensor placed in $\underline{S}$ (or $\overline{S}$) is a Byzantine node with probability of $\underline{\alpha}$ (or $\overline{\alpha}$) for TAS. We can observe that the value of $\underline{\alpha}$ (or $\overline{\alpha}$) is in the middle of the values of α_1 and α_2 (or α_3 and α_4) for the proposed scheme. It shows that taking both the status indicators from the same group into consideration can give us more information about the behavioral identities of the sensors in the network. Hence, our proposed EAS outperforms TAS that only utilizes the averaged probabilities ($\underline{\alpha}$ or $\overline{\alpha}$) to assess the behavioral

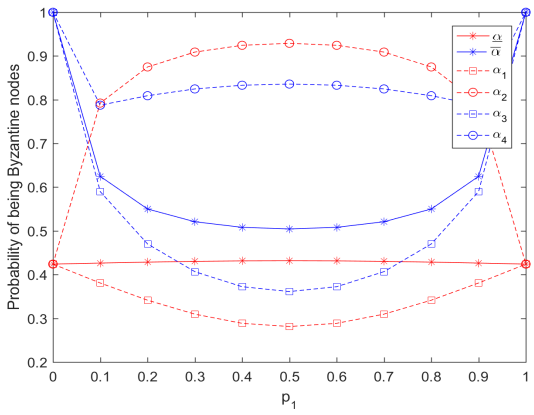


Fig. 2.3: The probability of being Byzantine nodes for sensors in sets $\underline{S}$, $\overline{S}$, $\underline{SS}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$ when $p_2 = 0.1$.

identity for each sensor.

2.3.2 Optimal Decision Rule

From the analysis above, the pmf of local decision u_i for our proposed EAS is expressed as

$$P(u_i|\mathcal{H}_q) = \begin{cases} \pi_{1q,1}^{u_i}(1-\pi_{1q,1})^{1-u_i} & \text{for } i \in \underline{SS} \\ \pi_{1q,2}^{u_i}(1-\pi_{1q,2})^{1-u_i} & \text{for } i \in \underline{S}\overline{S} \\ \pi_{1q,3}^{u_i}(1-\pi_{1q,3})^{1-u_i} & \text{for } i \in \overline{S}\underline{S} \\ \pi_{1q,4}^{u_i}(1-\pi_{1q,4})^{1-u_i} & \text{for } i \in \overline{SS} \end{cases} \tag{2.29}$$

for $q = 0, 1$, where

$$\pi_{11,e} = 1 - \pi_{10,e} = P_d(1-\alpha_e p_1) + \alpha_e p_1(1-P_d) \tag{2.30a}$$

$$\pi_{10,e} = 1 - \pi_{00,e} = P_f(1-\alpha_e p_1) + \alpha_e p_1(1-P_f) \tag{2.30b}$$

for $e = 1, 2, 3, 4$. $\pi_{11,e}$ and $\pi_{10,e}$ are the probabilities of sending the local decision $u_i = 1$ given hypothesis $\mathcal{H}_1$ and given hypothesis $\mathcal{H}_0$, respectively, for $e = 1, 2, 3, 4$ which are corresponding to the sensors being in $\underline{SS}$ $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$. The new optimal decision rule is provided in Theorem 2.2.

Theorem 2.2. *The new decision rule for the proposed EAS, given the Byzantine flipping probabilities p_1, p_2 and α_0 fraction of Byzantine nodes, is expressed as*

$$\sum_{e=1}^{4} W_e U_e \gtrless \eta^{(En)}, \tag{2.31}$$

where $U_1 = \sum_{i\in\underline{S}\underline{S}} u_i$, $U_2 = \sum_{i\in\underline{S}\overline{S}} u_i$, $U_3 = \sum_{i\in\overline{S}\underline{S}} u_i$, $U_4 = \sum_{i\in\overline{S}\overline{S}} u_i$ and $W_e = \log(\frac{\pi_{11,e}(1-\pi_{10,e})}{\pi_{10,e}(1-\pi_{11,e})})$ for $e = 1, 2, 3, 4$. $\eta^{(En)}$ is the threshold used by the FC for EAS, where $\eta^{(En)} = \log(\frac{\pi_0}{\pi_1}) + \sum_{e=1}^{4} N_e \log(\frac{1-\pi_{10,e}}{1-\pi_{11,e}})$. N_1, N_2, N_3 and N_4 are the cardinalities of sets $\underline{S}\underline{S}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{S}\overline{S}$, respectively, where $N_1 = |\underline{S}\underline{S}|$,$N_2 = |\underline{S}\overline{S}|$, $N_3 = |\overline{S}\underline{S}|$ and $N_4 = |\overline{S}\overline{S}|$.

PROOF: We know that the local decisions are independent given the hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$ and the information about the sets where all the sensors are placed in. Hence, the optimal decision rule, which is given in (2.32), can be further simplified. Substituting (2.29) in (2.32), and taking the logarithm on both sides, we obtain the fusion rule in the theorem.

$$\prod_{i\in\underline{S}\underline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \prod_{i\in\underline{S}\overline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \prod_{i\in\overline{S}\underline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \prod_{i\in\overline{S}\overline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \gtrless \frac{\pi_0}{\pi_1} \tag{2.32}$$

■

Note that U_e is binomial distributed random variables with parameters $(N, \pi_{11,e})$ under $\mathcal{H}_1$, and with parameters $(N, \pi_{10,e})$ under $\mathcal{H}_0$ for $e = 1, 2, 3, 4$. When N is large, N_1, N_2, N_3 and N_4 can be approximated by their expected value $NP(i \in \underline{S}\underline{S})$, $NP(i \in \underline{S}\overline{S})$, $NP(i \in \overline{S}\underline{S})$ and $NP(i \in \overline{S}\overline{S})$, respectively. For any sensor $i \in \{1, 2, \ldots, N\}$, the probability of being placed in $\underline{S}\underline{S}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{S}\overline{S}$ are $P(i \in \underline{S}\underline{S}) = P(d_i = d_j = 1)$, $P(i \in \underline{S}\overline{S}) = P(i \in \overline{S}\underline{S}) = P(d_i = 1, d_j = 0) = P(d_i = 0, d_j = 1)$ and $P(i \in \overline{S}\overline{S}) = P(d_i = d_j = 0)$, respectively. The threshold used by the FC becomes $\eta^{(En)} = \log(\frac{\pi_0}{\pi_1}) + NP(i \in \underline{S}\underline{S}) \log(\frac{1-\pi_{10,1}}{1-\pi_{11,1}}) + NP(i \in \underline{S}\overline{S}) \log(\frac{1-\pi_{10,2}}{1-\pi_{11,2}}) + NP(i \in \overline{S}\underline{S}) \log(\frac{1-\pi_{10,3}}{1-\pi_{11,3}}) + NP(i \in \overline{S}\overline{S}) \log(\frac{1-\pi_{10,4}}{1-\pi_{11,4}})$. Thus, the PDF of the global static $U = \sum_{e=1}^{4} W_e U_e$

can be approximated by the Gaussian distribution with parameters given as follows.

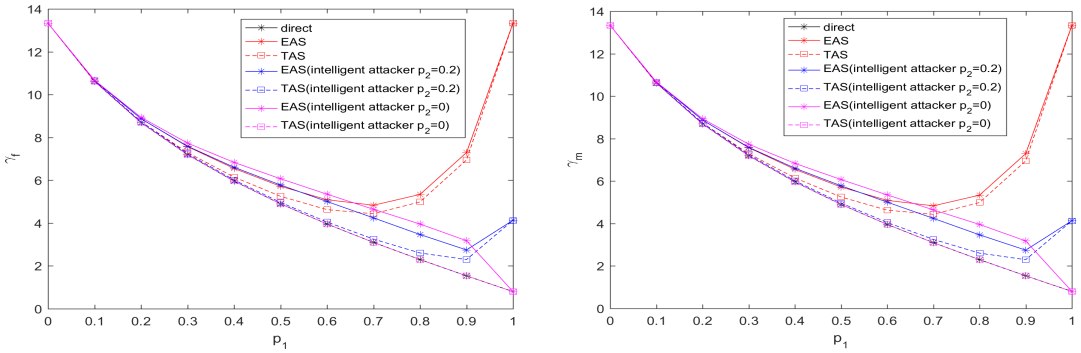
$$\begin{aligned}
\mu_0^{(En)} =& E[U|\mathcal{H}_0] \\
=& N(P(i \in \underline{SS})\pi_{10,1}W_1 + P(i \in \underline{S}\overline{S})\pi_{10,2}W_2 \\
&+ P(i \in \overline{S}\underline{S})\pi_{10,3}W_3 + P(i,j \in \overline{SS})\pi_{10,4}W_4) \qquad (2.33a) \\
\mu_1^{(En)} =& E[U|\mathcal{H}_1] \\
=& N(P(i \in \underline{SS})\pi_{11,1}W_1 + P(i \in \underline{S}\overline{S})\pi_{11,2}W_2 \\
&+ P(i \in \overline{S}\underline{S})\pi_{11,3}W_3 + P(i \in \overline{SS})\pi_{11,4}W_4) \qquad (2.33b) \\
(\sigma_0^{(En)})^2 =& Var[U|\mathcal{H}_0] \\
=& N(P(i \in \underline{SS})\pi_{10,1}(1-\pi_{10,1})W_1^2 + P(i \in \underline{S}\overline{S})\pi_{10,2}(1-\pi_{10,2})W_2^2 \\
&+ P(i \in \overline{S}\underline{S})\pi_{10,3}(1-\pi_{10,3})W_3^2 + P(i \in \overline{SS})\pi_{10,4}(1-\pi_{10,4})W_4^2) \qquad (2.33c) \\
(\sigma_1^{(En)})^2 =& Var[U|\mathcal{H}_1] \\
=& N(P(i \in \underline{SS})\pi_{11,1}(1-\pi_{11,1})W_1^2 + P(i \in \underline{S}\overline{S})\pi_{11,2}(1-\pi_{11,2})W_2^2 \\
&+ P(i \in \overline{S}\underline{S})\pi_{11,3}(1-\pi_{11,3})W_3^2 + P(i \in \overline{SS})\pi_{11,4}(1-\pi_{11,4})W_4^2) \qquad (2.33d)
\end{aligned}$$

The detection performance, characterized by the probability of error of the system, is given as

$$P_e^{(En)} = \pi_0 Q\left(\gamma_f^{(En)}\right) + \pi_1 Q\left(\gamma_m^{(En)}\right), \qquad (2.34)$$

where $\gamma_f^{(En)} = \frac{\eta^{(En)} - \mu_0^{(En)}}{\sigma_0^{(En)}}$ and $\gamma_m^{(En)} = \frac{\mu_1^{(En)} - \eta^{(En)}}{\sigma_1^{(En)}}$. Fig. 2.4 shows that the detection performance of the proposed scheme in terms of $\gamma_f^{(En)}$ and $\gamma_m^{(En)}$ is better than the detection performance of the traditional one, TAS, under both strategic attacks and non-strategic attacks. We can observe that the detection performance of TAS is the same as the direct scheme when the system is under strategic attacks ($p_2 = 0$). This is in accordance with the results shown in Theorem 2.1. However, the proposed EAS prevents it from happening. As shown in Fig. 2.4, the worst case from the perspective of the FC is that the strategic attackers take the attacking strategy of $p_1 = 1$ and $p_2 = 0$, i.e. , the Byzantine nodes always send falsified data to the MMSD and their group members and do

not forge data from their group members. In this case, the proposed EAS has the same detection performance as the direct scheme. In the next section, another new scheme is proposed which achieves better detection performance and higher robustness compared with EAS.



(a) γ_f as a function of flipping probability p_1 given $p_2 = 0$ and $p_2 = 0.2$.

(b) γ_m as a function of flipping probability p_1 given $p_2 = 0$ and $p_2 = 0.2$.

Fig. 2.4: The probability of error is characterized by the argument of function $Q(.)$ for the probability of false alarm shown in (a) and the argument of function $Q(.)$ for the probability of miss detection shown in (b). Smaller values of the argument result in higher probabilities of error.

2.4 Reduced Audit Bit based Scheme

In this section, we propose a new framework and a new fusion rule for the audit bit based system. In this framework, we focus on the practical scenario in which the Byzantine nodes are in a minority due to the limited attacking resources, i.e., $\alpha_0 \leq 1/2$. We will first start with a network with one cluster, then we will move on to a wide-area network with multiple clusters.

2.4.1 A Single-cluster Network

As before, the sensors are partitioned into sets $\underline{SS}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$ by the MMSD based on both status indicators of sensor i and sensor j in the same group. Moreover, the local decisions (u_i, u_j) sent from the same group are also compared to give us additional information about the behavioral identity of sensors in the networks. Each sensor again transmits its decision to the MMSD via two paths, namely the direct path and indirect path to the FC. After collecting all the local decisions,

the MMSD places the sensors into sets $\underline{SS}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$. These steps are the same as the ones in EAS. However, the MMSD also considers the MMS of the decisions u_i and u_j from the same group: if the sensor decisions for sensors i and j are the same, i. e., $u_i = u_j$, they are placed in the Set $\underline{\mathcal{M}}$ and the others are placed in the Set $\overline{\mathcal{M}}$. The MMSD only transmits the local decisions of the sensors with the sensor index i given by $\{i : (\underline{SS} \bigcap \underline{\mathcal{M}}) \bigcup \underline{S}\overline{S} \bigcup \overline{S}\underline{S}\}$ to the decision making module to make the final decision. In other words, the local decisions from the sensors in Set $\underline{SS} \bigcap \overline{\mathcal{M}}$ or Set $\overline{SS}$ are not used to make the final decision which correspond to the two conditions stated as below.

- Condition 1: The sensor i and its group member j are both in the set $\overline{SS}$.
- Condition 2: The sensor i and its group member j are both in the set $\underline{SS}$ and $u_i \neq u_j$.

In the next lemma, we show the reasons why not using the decisions of sensors that satisfy one of the above two conditions improves the detection performance of the system.

Lemma 2.3. *1. When the sensor pair (i, j) satisfies Condition 1, i. e., sensors i and$j$ belong to $\overline{SS}$, removing this sensor pair results in the removal of two Byzantine nodes when $p_2 = 0$.*

2. When we remove the sensor pairs that satisfy Condition 2, the ability of removing the Byzantine nodes for the proposed RAS increases with the increase of p_1 given specific p_2 and α_0.

PROOF:

1. Let E be the event that at least one node in sensor pair (i, j) is a Byzantine node. When $i, j \in \overline{SS}$, it is obvious that $P(E|i, j \in \overline{SS}) = 1$. Thus, we can obtain $P(i, j \notin \overline{SS}|\overline{E}) = 1$ due to the fact that the contrapositive of the conditional statement is also true. So we can conclude that there is at least one Byzantine node in the sensor pair. Moreover, it is easy to conclude that all the sensors are Byzantine nodes in the Set $\overline{SS}$ when the attackers take the strategy of $p_2 = 0$ according to (A.9). Thus, removing the decisions of sensors in this set can remove at least one Byzantine node in each pair, and it can even remove two Byzantine nodes in each pair when the attackers employ the strategy of $p_2 = 0$.

2. To evaluate the impact of removing the unequal local decisions of sensor pairs on the performance of removing Byzantine nodes, we utilize the ratio $F = \frac{P(E,u_i=u_j|i,j\in\underline{SS})}{P(E|i,j\in\underline{SS})}$ to characterize that performance. The numerator of ratio F is the probability of the joint event that there exists at least one Byzantine node and the event $u_i = u_j$ given $i, j \in \underline{SS}$. The denominator is the probability of at least one Byzantine node given $i, j \in \underline{SS}$. The ratio $F = P(u_i = u_j|i, j \in \underline{SS}, E)$ gives the probability of $u_i = u_j$ given event E and $i, j \in \underline{SS}$. We have

$$
\begin{aligned}
&P(E, u_i = u_j|i, j \in \underline{SS}) \\
&= P(E|u_i = u_j, i, j \in \underline{SS})P(u_i = u_j|i, j \in \underline{SS}) && \text{(2.35a)} \\
&= (1 - P(i = H, j = H|i, j \in \underline{SS}, u_i = u_j))P(u_i = u_j, |i, j \in \underline{SS}) && \text{(2.35b)} \\
&= P(u_i = u_j|i, j \in \underline{SS}) - P(u_i = u_j|i, j \in \underline{SS}, \\
&\quad i = H, j = H)P(i = H, j = H|i, j \in \underline{SS}) && \text{(2.35c)} \\
&= P(u_i = u_j|i, j \in \underline{SS}) - \frac{(1-\alpha_0)^2}{P(i, j \in \underline{SS})}P(u_i = u_j|i, j \in \underline{SS}, i = H, j = H) && \text{(2.35d)}
\end{aligned}
$$

and

$$
\begin{aligned}
P(E|i, j \in \underline{SS}) &= 1 - P(i = H, j = H|i, j \in \underline{SS}) && \text{(2.36a)} \\
&= 1 - \frac{(1-\alpha_0)^2}{P(i, j \in \underline{SS})}, && \text{(2.36b)}
\end{aligned}
$$

where $P(u_i = u_j|i, j \in \underline{SS}) = P(u_i = u_j|i, j \in \underline{SS}, \mathcal{H}_0)P(\mathcal{H}_0) + P(u_i = u_j|i, j \in \underline{SS}, \mathcal{H}_1)P(\mathcal{H}_1) = \pi_1[\pi_{11,1}^2 + (1 - \pi_{11,1})^2] + \pi_0[\pi_{10,1}^2 + (1 - \pi_{10,1})^2]$ and $P(u_i = u_j|i, j \in \underline{SS}, i = H, j = H) = [P_d^2 + (1 - P_d)^2]\pi_1 + [P_f^2 + (1 - P_f)^2]\pi_0$.

■

The relationship among p_1, p_2, α_0 and F is shown in Fig. 2.5. Note that a small F means a lower probability of existence of Byzantine nodes in the sensor pair given $i, j \in \underline{\mathcal{M}} \bigcap \underline{SS}$. We can observe from Fig. 2.5 that the value of F has a significant decrease when p_1 is large. It can also

be observed that the value of F decreases with the increase of α_0 given $p_1 \geq 0.5$ and p_2, and it slightly changes with different α_0 and p_2 given $p_1 < 0.5$. From the analysis in Section 2.2, it is evident that p_1 affects the final decision making by mainly affecting the local decisions (u_i) used to make the final decision, while both α_0 and p_2 only affect the final decision making by affecting the evaluated probability of one sensor being a Byzantine node (α_1, α_2, α_3 or α_4). Intuitively, changing p_1 has greater effect on the final decision making. Hence, when $p_1 < 0.5$, p_1 is not large enough to enable us to observe a distinct difference in F for different p_2 and α_0. In general, Fig. 2.5 shows that the ability of removing the Byzantine nodes increases with the increase of p_1 for a given p_2 by removing the sensor pairs which satisfy Condition 2.

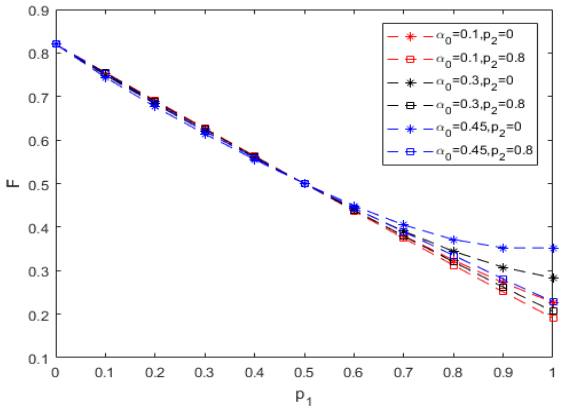


Fig. 2.5: F versus p_1 given $p_2 = 0.1$ for different α_0 and $N = 100$.

According to Theorem 2.1, the attackers' optimal attacking strategy in TAS is to choose $p_2 = 0$. In the scenario where p_2 is very small (close to 0), however, Fig. 2.2 has shown that the detection performance of TAS significantly degrades for a large value of p_1. The proposed scheme in this section achieves better detection performance compared with TAS when the attackers adopt the strategy of $p_2 = 0$ with $\forall p_1 \in [0, 1]$. It is because when p_2 is small, the Byzantine nodes have high probabilities of being placed in the set $\underline{SS}$ in our proposed scheme. If the attacker chooses p_1 to be large, there is a high probability that the group containing a Byzantine node satisfies Condition 2. Hence, the decision of the Byzantine node is likely to be blocked by the MMSD and not transmitted to the FC. As a result, our scheme prevents the attacker from designing p_1 to be very large and p_2 to be very small. On the other hand, when p_1 is not so large, each Byzantine node has a relatively

higher probability, i.e., $1-p_1$, to act honestly. Through such a trade off, the detection accuracy of the proposed scheme outperforms TAS under strategic attacks.

Based on the analysis above, we can show that the proposed scheme can effectively remove the decisions coming from Byzantine nodes. Hence, in the proposed RAS, we have the following relations for sensor i.

$$P(u_i = 1 | i \in \underline{S}\overline{S}, \mathcal{H}_q) = \pi_{1q,2} \tag{2.37a}$$

$$P(u_i = 1 | i \in \overline{S}\underline{S}, \mathcal{H}_q) = \pi_{1q,3} \tag{2.37b}$$

where $q = 0, 1$. Although u_i and u_j are dependent given $i, j \in \underline{SS} \bigcap \underline{\mathcal{M}}$, they are independent given $i, j \in \underline{SS}$. Hence, we have

$$\begin{aligned} & P(u_i = 1, u_j = 1 | i, j \in \underline{SS} \bigcap \underline{\mathcal{M}}, \mathcal{H}_q) \\ = & \frac{P(u_i = 1 | i, j \in \underline{SS}, \mathcal{H}_q) P(u_j = 1 | i, j \in \underline{SS}, \mathcal{H}_q)}{P(u_i = u_j | i, j \in \underline{SS}, \mathcal{H}_q)} \\ = & \frac{\underline{\pi}_{1q}^2}{\underline{\pi}_{1q}^2 + (1 - \underline{\pi}_{1q})^2} = \pi_{1q,5} \end{aligned} \tag{2.38}$$

for $q = 0, 1$. To simplify the analysis, we consider the group votes instead of the individual votes for the sensors in set $\underline{SS} \bigcap \underline{\mathcal{M}}$. Let z_g denote the group vote for group $g \in \underline{T}$, where $\underline{T}$ is the set of group whose sensors are in set $\underline{SS} \bigcap \underline{\mathcal{M}}$. Due to the fact that the sensors in the same group in set $\underline{SS} \bigcap \underline{\mathcal{M}}$ has the same decisions, we have $z_g = \{0, 2\}$. Hence, we obtain the following pdfs

$$f(u_i | \mathcal{H}_q) = \begin{cases} \pi_{1q,2}^{u_i} (1 - \pi_{1q,2})^{1-u_i} & \text{for } i \in \underline{S}\overline{S} \\ \pi_{1q,3}^{u_i} (1 - \pi_{1q,3})^{1-u_i} & \text{for } i \in \overline{S}\underline{S} \end{cases} \tag{2.39}$$

for sensor $i \in \underline{S}\overline{S} \bigcup \underline{S}\overline{S}$, and

$$f(z_g | \mathcal{H}_q) = \pi_{1q,5}^{z_g/2} (1 - \pi_{1q,5})^{1 - z_g/2} \tag{2.40}$$

for group $g \in \underline{T}$, where $q = 0, 1$. Thus, the proposed new decision rule is shown in Theorem 2.2.

Theorem 2.3. *The new optimal decision rule, given the Byzantine flipping probabilities p_1, p_2 and α_0 fraction of Byzantine nodes, is expressed as*

$$W_5 \sum_{g \in \underline{T}} \frac{z_g}{2} + W_2 \sum_{i \in \underline{S}\overline{S}} u_i + W_3 \sum_{i \in \overline{S}\underline{S}} u_i \gtrless \eta^{(RA)}, \tag{2.41}$$

where $W_2 = \log(\frac{\pi_{11,2}(1-\pi_{10,2})}{\pi_{10,2}(1-\pi_{11,2})})$, $W_3 = \log(\frac{\pi_{11,3}(1-\pi_{10,3})}{\pi_{10,3}(1-\pi_{11,3})})$, $\eta^{(RA)} = \log(\frac{\pi_0}{\pi_1}) + N_{re}^{LL}\log(\frac{1-\pi_{10,5}}{1-\pi_{11,5}}) + N_{re}^{L}\log(\frac{1-\pi_{10,2}}{1-\pi_{11,2}}) + N_{re}^{U}\log(\frac{1-\pi_{10,3}}{1-\pi_{11,3}})$. *$N_{re}^{L}$, N_{re}^{U} and N_{re}^{LL} are the cardinalities of sets $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\underline{T}$, respectively, where $N_{re}^{L} = |\underline{S}\overline{S}|$,$N_{re}^{U} = |\overline{S}\underline{S}|$, and $N_{re}^{LL} = |\underline{T}|$. W_5 denotes the rearranged weight for group decisions in set $\underline{T}$ which is given as*

$$W_5 = \frac{\pi_{11,5}(1-\pi_{10,5})}{\pi_{10,5}(1-\pi_{11,5})}. \tag{2.42}$$

PROOF: We know that all groups of sensors whose decisions are sent to the FC are elements of one of the three sets $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\underline{S}\underline{S}\bigcap\underline{\mathcal{M}}$. Thus, the optimal decision rule is given as (2.43) due to the fact that the sensors in sets $\underline{S}\overline{S}$ or $\overline{S}\underline{S}$ independently send their local decisions to the FC given the hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$. Even though the decisions coming from the sensors in the same group in set $\underline{S}\underline{S}\bigcap\underline{\mathcal{M}}$ are dependent, the group votes are independent of each other. Hence, the optimal decision rule can be reformulated as (2.44). Substituting (2.37), (2.38), (4.23), (2.40) in (2.44), and taking the logarithm on both sides, we can get the fusion rule stated in the theorem.

$$\prod_{i,j \in \underline{S}\underline{S}\bigcap\underline{\mathcal{M}}} \frac{P(u_i,u_j|\mathcal{H}_1)}{P(u_i,u_j|\mathcal{H}_0)} \prod_{i \in \underline{S}\overline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \prod_{i \in \overline{S}\underline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \gtrless \frac{\pi_0}{\pi_1} \tag{2.43}$$

$$\prod_{g \in \underline{T}} \frac{P(z_g|\mathcal{H}_1)}{P(z_g|\mathcal{H}_0)} \prod_{i \in \underline{S}\overline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \prod_{i \in \overline{S}\underline{S}} \frac{P(u_i|\mathcal{H}_1)}{P(u_i|\mathcal{H}_0)} \gtrless \frac{\pi_0}{\pi_1} \tag{2.44}$$

■

Let U denote the left-hand side of the optimal decision rule in (2.41) which is given as

$$U = W_5U_5 + W_2U_2 + W_3U_3, \tag{2.45}$$

where $U_5 = \sum_{g\in\underline{T}} z_g/2$, $U_2 = \sum_{i\in\underline{S}\overline{S}} u_i$ and $U_3 = \sum_{i\in\overline{S}\underline{S}} u_i$. U_2 and U_3 are all Binomial distributed variables and U_5 is equivalent to a Binomial distributed variable. When N is large, the expected number of sensors in $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and the expected number of groups in $\underline{T}$ are $NP(i \in \underline{S}\overline{S})$, $NP(i \in \overline{S}\underline{S})$ and $GP(u_i = u_j|i, j \in \underline{SS})P(i, j \in \underline{SS})$, respectively. $P(i \in \underline{S}\overline{S})$ and $P(i \in \overline{S}\underline{S})$ are defined in (2.33), and $P(i, j \in \underline{SS})$ is defined in (2.20). $P(u_i = u_j|i, j \in \underline{SS})$ is given as

$$\begin{aligned} P(u_i = u_j|i, j \in \underline{SS}) = &\sum_{q=0,1} P(\mathcal{H}_q) \sum_{t=0,1} P(u_i = t|i \in \underline{SS}, \mathcal{H}_q) \\ &P(u_j = t|j \in \underline{SS}, \mathcal{H}_q) && \text{(2.46a)} \\ =&(\underline{\pi}_{11}^2 + (1 - \underline{\pi}_{11})^2)\pi_1 \\ &+ (\underline{\pi}_{10}^2 + (1 - \underline{\pi}_{10})^2)\pi_0 && \text{(2.46b)} \end{aligned}$$

Hence, U, which is the sum of Binomial distributed variables, can be approximated as the Gaussian distribution with parameters as follows:

$$\begin{aligned} \mu_0^{(RA)} =&E[U|\mathcal{H}_0] \\ =&GP(u_i = u_j|i, j \in \underline{SS})P(i, j \in \underline{SS})\pi_{10,5}W_5 \\ &+ N(P(i \in \overline{S}\underline{S})\pi_{10,3}W_3 + P(i \in \underline{S}\overline{S})\pi_{10,2}W_2) && \text{(2.47a)} \\ \mu_1^{(RA)} =&E[U|\mathcal{H}_1] \\ =& GP(u_i = u_j|i, j \in \underline{SS})P(i, j \in \underline{SS})\pi_{11,5}W_5 \\ &+ N(P(i \in \overline{S}\underline{S})\pi_{11,3}W_3 + P(i \in \underline{S}\overline{S})\pi_{11,2}W_2) && \text{(2.47b)} \\ (\sigma_0^2)^{(RA)} =&Var[U|\mathcal{H}_0] \\ =&GP(u_i = u_j|i, j \in \underline{SS})P(i, j \in \underline{SS})\pi_{10,5}(1 - \pi_{10,5})W_5^2 \end{aligned}$$

$$+ N(P(i \in \overline{S}\underline{S})\pi_{10,3}(1-\pi_{10,3})W_3^2 + P(i \in \underline{S}\overline{S})\pi_{10,2}(1-\pi_{10,2})W_2^2) \quad (2.47c)$$

$$\begin{aligned}(\sigma_1^2)^{(RA)} =& Var[U|\mathcal{H}_0] \\ =& GP(u_i = u_j|i,j \in \underline{SS})P(i,j \in \underline{SS})\pi_{11,5}(1-\pi_{11,5})W_5^2 \\ &+ N(P(i \in \overline{S}\underline{S})\pi_{11,3}(1-\pi_{11,3})W_3^2 + P(i \in \underline{S}\overline{S})\pi_{11,2}(1-\pi_{11,2})W_2^2) \end{aligned} \quad (2.47d)$$

The threshold η for large N is given as

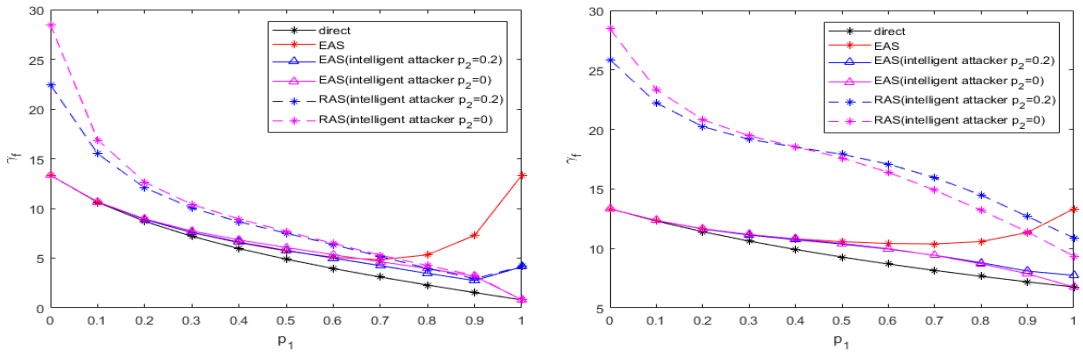
$$\begin{aligned}\eta^{(RA)} =& \log(\frac{\pi_0}{\pi_1}) + E(N_{re}^{LL})\log(\frac{1-\pi_{10,5}}{1-\pi_{11,5}}) \\ &+ E(N_{re}^{L})\log(\frac{1-\pi_{10,2}}{1-\pi_{11,2}}) + E(N_{re}^{U})\log(\frac{1-\pi_{10,3}}{1-\pi_{11,3}}),\end{aligned} \quad (2.48)$$

where $E(N_{re}^{L}) = NP(i \in \underline{S}\overline{S})$, $E(N_{re}^{U}) = NP(i \in \overline{S}\underline{S})$ and $E(N_{re}^{LL}) = GP(u_i = u_j|i,j \in \underline{SS})$. Thus, the probability of error $P_e^{(RA)}$ for the system is expressed as

$$P_e^{(RA)} = \pi_0 Q\left(\gamma_f^{(RA)}\right) + \pi_1 Q\left(\gamma_m^{(RA)}\right), \quad (2.49)$$

where $\gamma_f^{(RA)} = \frac{\eta^{(RA)} - \mu_0^{(RA)}}{\sigma_0^{(RA)}}$ and $\gamma_m^{(RA)} = \frac{\mu_1^{(RA)} - \eta^{(RA)}}{\sigma_1^{(RA)}}$ is the argument of function $Q(.)$ for the probability of false alarm and the argument of function $Q(.)$ for the probability of miss detection for the new proposed fusion rule. Fig. 2.6 shows how argument $\gamma_f^{(RA)}$ changes with p_1 given specific p_2 and α_0 when $N = 100$, $P_d = 0.9$ and $P_f = 0.1$. We can observe that the argument $\gamma_f^{(RA)}$ of RAS is larger than that of EAS under strategic attacks. Since the argument $\gamma_m^{(RA)}$ has similar properties, we only include the simulation results of $\gamma_f^{(RA)}$ here. Note that the larger arguments mean better detection performance. Fig. 2.7 shows how the probabilities of detection and false alarm of the system change with p_1 given specific $p_2 = 0$ and $\alpha_0 = 0.3$ when $N = 10$, $P_d = 0.9$ and $P_f = 0.1$. From Fig. 2.6 and Fig. 2.7, we can observe that our proposed RAS has a significant improvement on the detection performance of the system when α_0 is small. Even though the detection performance of the proposed scheme gets close to EAS when α_0 approaches 0.5 and p_1 is large, the proposed RAS still outperforms EAS and the direct scheme. This improvement

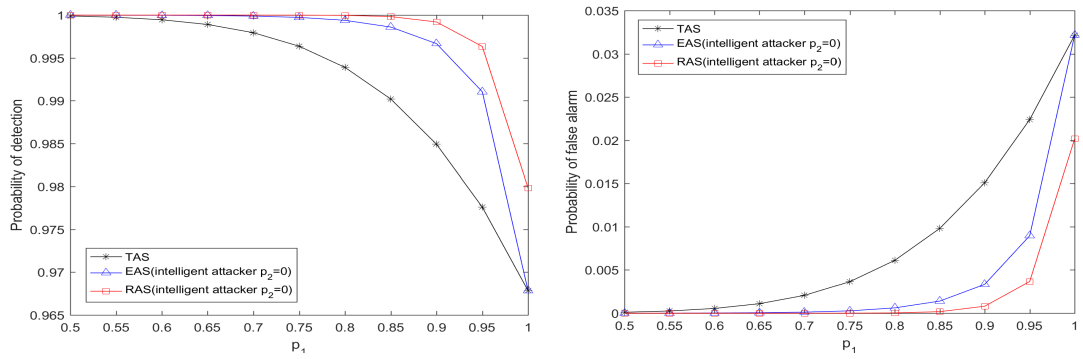
becomes more prominent when p_1 is relatively small. Moreover, in both EAS and RAS, a large p_1 can always make it harder for Byzantine nodes to evade the detection system. In this case, the FC has the history of all the local decisions it received in the past to identify Byzantine nodes. And some reputation-based schemes can help the FC to identify the Byzantine nodes [82] [91].



(a) γ_f as a function of flipping probability p_1 given $p_2 = 0$ and $p_2 = 0.2$ when $\alpha_0 = 0.45$.

(b) γ_f as a function of flipping probability p_1 given $p_2 = 0$ and $p_2 = 0.2$ when $\alpha_0 = 0.15$.

Fig. 2.6: The argument for the probability of false alarm function for different values of α_0.



(a) The probability of detection versus p_1 given $p_2 = 0$ for $\alpha_0 = 0.3$ and $N = 10$.

(b) The probability of false alarm versus p_1 given $p_2 = 0$ for $\alpha_0 = 0.3$ and $N = 10$.

Fig. 2.7: The probability of false alarm and the probability of detection for the system.

2.4.2 The Network with Multiple Clusters

In this subsection, we extend our work from the single cluster case to the case of multiple clusters in the wide-area network. We show that the proposed RAS can not only improve the detection

performance of the system, but also reduce the communication overhead[5] between the clusters and the FC. In a cluster based network as shown in Fig. 2.8, the N sensors in the network are grouped into T clusters and the sensors in each cluster are further divided into groups of two. Each cluster is equipped with one MMSD which serves as a data integration processor for this cluster. Note that the MMSD is no longer a part of the FC.

Based on the local observations, each sensor makes a binary decision regarding the absence or presence of the PoI. Then, the sensors send both their own decisions and their group member's decision to the corresponding MMSDs. By comparing the MMS of the direct and indirect decisions, the MMSDs are able to obtain the status indicators for all the sensors in the corresponding clusters. Based on these status indicators, each MMSD partitions the sensors in the cluster into sets $\underline{SS}$, $\underline{S}\overline{S}$, $\overline{S}\underline{S}$ and $\overline{SS}$. In addition, the sensors are placed into $\underline{\mathcal{M}}$ if the local decisions of the sensors in the same group are the same. The flow chart to illustrate the decision making and communication process of a cluster $t \in \{1, \ldots, T\}$ is shown in Fig. 2.9.

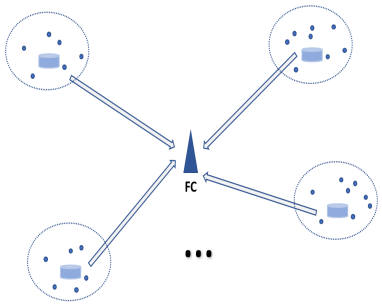


Fig. 2.8: System model of a distributed CWSN. The blue cylinders represent MMSDs in each cluster and the small blue circles represent low-cost sensors.

Let $N_t^{(RA)}$ and $N_t^{(A)}$ denote the number of local decisions sent by the MMSDs to the FC for the proposed RAS and the number of local decisions sent by the sensors to the FC, respectively. Note that the MMSDs only transmit the direct decisions, and they do not transmit the ones that satisfy Condition 1 or Condition 2. Thus, the number of direct decisions $N_t^{(RA)}$ sent by the MMSDs to the FC is smaller than that of TAS $N_t^{(A)}$, where $N_t^{(RA)} = |\underline{SS} \bigcap \underline{\mathcal{M}}| + |\underline{S}\overline{S}| + |\overline{S}\underline{S}|$ and $N_t^{(A)} = 2N$.

[5]In this section, we measure the overall communication overhead of the system by the number of bits in all communication messages sent.

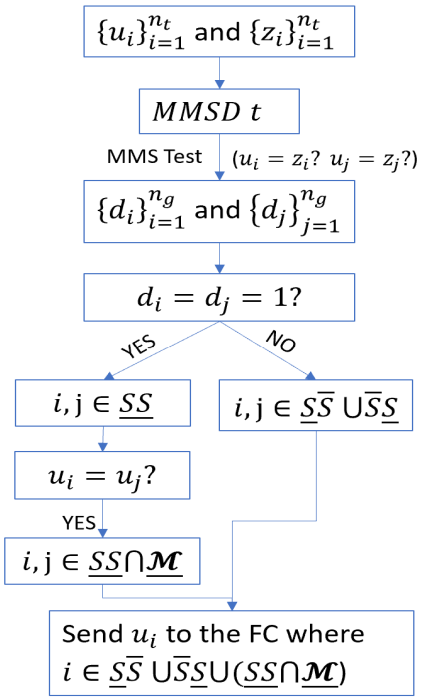


Fig. 2.9: The flow chart of the decision making and communication processes of cluster t. n_t is the number of sensors in cluster t. Sensor j is the group member of sensor i.

Let r represent different sets as follows. If $r = 00$, it refers to the set $\underline{SS} \bigcap \underline{\mathcal{M}}$; If $r = 01$ it refers to the set $\underline{S}\overline{\underline{S}}$; If $r = 10$, it refers to the set $\overline{\underline{S}}\underline{S}$. Each MMSD sends three data packets which contain r and the direct decisions from the sensors in the sets $\underline{SS} \bigcap \underline{\mathcal{M}}$, $\underline{S}\overline{\underline{S}}$ and $\overline{\underline{S}}\underline{S}$, respectively. For example, if sensor 1 to sensor 4 are in $\underline{SS} \bigcap \underline{\mathcal{M}}$, sensor 5 to sensor 8 are in $\overline{\underline{S}}\underline{S}$ and sensor 9 to sensor 12 are in $\underline{S}\overline{\underline{S}}$. The three data packets contain $[r = 00, u_1, \ldots, u_4]$, $[r = 10, u_5, \ldots, u_8]$ and $[r = 01, u_9, \ldots, u_{12}]$. Upon receiving these data packets, the FC is able to determine which sets those sensors belong to so that it can make the final decision based on those transmitted direct decisions.

When N is large, we are able to calculate the expected number of bits transmitted to the FC from all the MMSDs, which is $E(N_t^{(RA)}) = E(N_{re}^{LL}) + E(N_{re}^{L}) + E(N_{re}^{U})$, according to (2.48). Fig. 2.10 shows the expected number of bits transmitted to the FC when $N = 100$ and $N_t^{(A)} = 2N = 200$. We can observe that the expected number of bits transmitted to the FC for the proposed RAS significantly decreases compared with the one for TAS. It is due to fact that the MMSDs only send the direct decisions of sensors which do not satisfy Condition 1 or Condition2. We can also

observe that the expected number of bits decreases with an increased α_0 given a specific p_2. It is due to the fact that the number of sensors temporarily removed by the MMSDs increases when the fraction of Byzantine nodes α_0 increases with a given attacking probability p_2. Hence, the proposed new fusion rule is able to reduce the energy cost of the sensors to half of the traditional case which prolongs the lifetime of the network, especially for the wide area network.

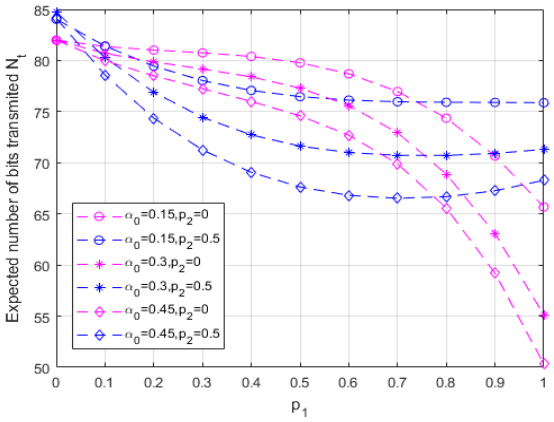


Fig. 2.10: The expected number of bits transmitted to the FC N_t versus p_1 given different value of α_0 and p_2.

2.5 Summary

In this chapter, an audit based mechanism was utilized to mitigate the effect of Byzantine attacks in the networks. Instead of employing the identical attacking strategy of TAS where each sensor utilizes the same attacking probability to falsify the decisions coming from their group member and its own decision, we considered strategic attackers that can use different attacking strategies. We showed the that it was possible for the strategic attackers to blind the FC as far as the information conveyed by the audit bits in TAS is concerned. To overcome this problem, we proposed an enhanced audit bit based scheme, namely EAS. Our results showed that the proposed scheme outperforms TAS. Furthermore, we proposed a reduced audit bit based scheme (RAS) based on our new proposed EAS. We showed that RAS is able to further improve the robustness and the detection performance of the system. We extended our work for the wide-area CWSNs. In wide-

area cluster-based WSNs, we showed that the proposed RAS is able to significantly reduce the communication overhead between the clusters and the FC.

CHAPTER 3
REPUTATION AND AUDIT BIT BASED DISTRIBUTED DETECTION IN THE PRESENCE OF BYZANTINES

In this chapter, we deal with the Byzantine attack problem when the FC has no prior knowledge of the attacking strategy of Byzantine nodes. Under this assumption, two reputation based algorithms called Reputation and audit based clustering (RAC) algorithm and Reputation and audit based clustering with auxiliary anchor node (RACA) algorithm are proposed to defend against Byzantine attacks in distributed detection networks. These two algorithms enable the FC to accurately identify Byzantine nodes and significantly improve the robustness of the system. The proposed RACA algorithm could still work well even when the number of Byzantine nodes exceeds half of the total number of sensors in the network.

3.1 Introduction

In distributed wireless sensor networks (WSNs), local sensors send their decisions regarding the presence or absence of the phenomenon of interest (PoI) to the fusion center (FC) and the FC

makes a final decision regarding the presence or absence of the PoI. Due to its energy-efficiency, distributed framework is widely adopted in many bandwidth-limited scenarios, e.g., IoT, cognitive radio networks and military surveillance systems. However, the open nature of WSNs makes the distributed system vulnerable to various attacks such as Byzantine attacks, wiretap, jamming and spoofing [26, 39, 54]. In this paper, we focus on Byzantine attacks where the sensors in a network may be compromised and controlled by adversaries and send falsified decisions to the FC.

3.1.1 Related Work

The Byzantine attack problem in distributed detection systems has been studied in the literature [33, 34, 46, 47, 82, 85, 97, 117]. In [85], an adaptive reputation based clustering algorithm is proposed for spectrum sensing networks to achieve good detection performance. In [117], a reputation-based scheme is proposed for cooperative spectrum sensing networks to improve the robustness of the networks. In [82], the authors investigated the condition under which the Byzantine attackers totally blind the FC and an algorithm is proposed to detect Byzantine attacks by counting the mismatches between the decisions and the global decision at the FC in collaborative spectrum sensing networks. In [46], the optimal attacking strategy is investigated for distributed detection systems where the smart attackers maximize the attacking efficacy by finding a trade-off between their exposure to the defense mechanism and the error probability of the system. In [97], an adaptive algorithm is proposed to defend against Byzantine attacks in the false discovery rate based distributed detection system when the Byzantine nodes are omniscient and know the true hypothesis. In [47], the optimal attacking strategies are analyzed for the cases where the FC has knowledge of the attackers' strategy and where the FC does not know the attackers' strategy. In [33], the audit bit based distributed detection scheme is proposed in the Neyman-Pearson framework. Each sensor sends one additional audit bit to the FC which gives more information about the behavioral identity of each sensor. The proposed scheme significantly improves the robustness and the detection performance of the system. In [34], the audit bit based mechanism is considered in the Bayesian setting and the mitigation scheme over time is also proposed. In Chapter 2, a more

general attacking strategy than investigated in [33] and [34] was proposed. A number of enhanced audit bit based schemes are proposed in this chapter which further improve the robustness and the detection performance of the system.

3.1.2 Major Contributions

In this work, we further consider the presence of strategic attackers that employ the general attacking strategy utilized in Chapter 2. Different from previous works, we consider that the FC does not have prior knowledge of the attacking strategy of Byzantine nodes, namely the flipping probabilities. We propose two reputation based algorithms to mitigate the effect of Byzantine attacks. In both proposed algorithms, we utilize the reputation indexes of sensors to represent the trustworthiness of sensors in the network. The reputation indexes of sensors are updated at each time step according to their behaviors. Sensors with low reputation indexes are usually identified as Byzantine nodes and are excluded from the decision-making process. In particular, the audit bit based mechanism and the Partitioning Around Medoid (PAM) algorithm are developed to update the reputation indexes of sensors in the network and to identify potential Byzantine nodes. The ability to identify Byzantine nodes can be further enhanced by the use of anchor nodes even when the number of Byzantine nodes exceeds half of the total number of sensors in the networks. The robustness of proposed algorithms is tested both in dynamic (attacking parameters change dynamically over time) and static (attacking parameters remain the same) scenarios. Simulation results show that our proposed algorithms are capable of defending against attackers in both scenarios.

3.2 System model

Consider a binary hypothesis testing problem with the two hypotheses denoted by H_0 and H_1. A WSN is comprised of N sensors and one FC, where the FC makes a final decision on which hypothesis is true based on the sensor's local decisions. The sensors are divided into groups of two and there are a total of $G = N/2$ groups in the network. The sensors make binary decisions

on whether H_0 or H_1 is true by utilizing the likelihood ratio (LR) test. For ease of notation, let us assume that each sensor has the same probabilities of detection and false alarm, i.e., $P_d = P(v_i = 1|\mathcal{H}_1)$ and $P_f = P(v_i = 1|\mathcal{H}_0)$ for $i \in \{1, \ldots, N\}$, where v_i is the decision made by sensor i. In addition to its local decision v_i, each sensor also sends one more decision, which comes from its group member, to the FC and we call this additional decision the audit bit as described in [33, 34, 78].

For simplicity, let i and j represent the sensors in the same group. As shown in Fig. 2.1(a), after making its own decision v_i, sensor i sends (i) u_i directly to the FC; (ii) w_i to the sensor j in the same group; (iii) z_j, corresponding to w_j coming from the sensor j in the same group, to the FC. Similarly, sensor j also sends two decisions u_j and z_i to the FC. If sensor i is a Byzantine node, i.e., $i = B$, the decisions v_i, w_i and u_i are not necessarily the same and z_j are also not necessarily equal to u_j. If sensor i is honest, i.e., $i = H$, it sends genuine or uncorrupted information to the FC. Hence, given a Byzantine node i, the attacking parameters p_1 and p_2 are given by

$$p_1 = p(v_i \neq u_i|i = B) = p(v_i \neq w_i|i = B) \tag{3.1}$$

$$p_2 = p(w_j \neq z_j|i = B). \tag{3.2}$$

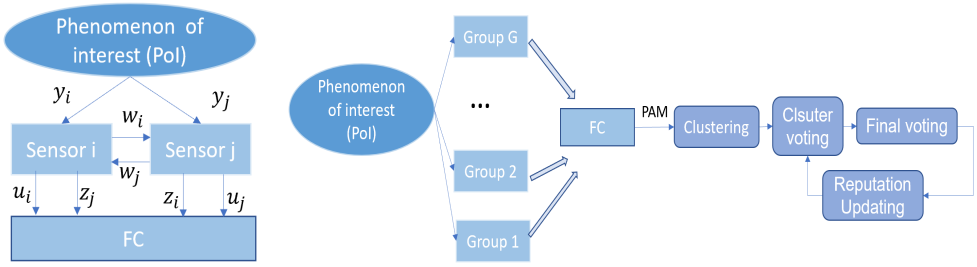
Given an honest node i, we have

$$p_1 = p(v_i \neq u_i|i = B) = p(v_i \neq w_i|i = B) = 0 \tag{3.3}$$

$$p_2 = p(w_j \neq z_j|i = B) = 0. \tag{3.4}$$

In other words, p_1 and p_2 represent the probabilities that a node flips its own decision and flips the decision coming from its group member, respectively. If $u_i = z_i$, we have a 'match' for sensor j, otherwise, we have a 'mismatch' for sensor j. Similarly, for sensor i, we have a 'match' when $u_j = z_j$ and a 'mismatch' when $u_j \neq z_j$. We assume that a fraction α_0 of the N sensors are Byzantine nodes and they attack independently. The FC is not aware of the identity or the attacking strategy of Byzantine nodes in the network. Hence, each node has the probability α_0

to be a Byzantine node and the FC does not know the values of p_1 and p_2. We consider the more general and practical attacking strategy as stated in [78] which allows the Byzantines to be strategic by optimally employing unequal probabilities p_1 and p_2.



(a) The architecture of group k.

(b) The block diagram of the proposed algorithm.

Fig. 3.1: The architecture of any group k is shown in (a). The block diagram of the proposed algorithm is shown in (b).

3.3 Proposed Reputation and Audit Bit Based Clustering Algorithms

In this section, we present the proposed robust defense algorithms for the system under attack when the FC does not possess the knowledge of the attacking strategy, namely p_1 and p_2, used by Byzantine nodes. We also evaluate the performance of our proposed algorithms in this section.

3.3.1 Reputation and Audit Bit based Clustering Algorithm

Upon receiving measurements $\{u_i\}_{i=1}^N$ and $\{z_i\}_{i=1}^N$, the FC is able to determine the match and mismatch (MMS) results ($u_i = z_i$ or $u_i \neq z_i$) for all the sensors in the network. Based on the received measurements and the MMS results, we propose a robust reputation based algorithm to defend against Byzantine attacks. The proposed reputation and audit bit based clustering (RAC) algorithm consists of four successive phases and the flow chart is shown in Fig.3.1(b).

- Macro clustering phase: At time step t, T most recent decisions of each node are utilized. The FC keeps a $N(2T+1)$ dimensional vector[1] to store the information corresponding to each sensor in the network, which consists of the records of local decisions, the records of MMS results and the updated reputation index. We make use of the MMS results to cluster or partition the sensors into two different sets, which are $\underline{\mathcal{T}}$ and $\overline{\mathcal{T}}$. If the MMS results for both sensors in the same group are always 'match', the sensors in this group are placed in set $\underline{\mathcal{T}}$, otherwise, they are placed in set $\overline{\mathcal{T}}$.

- Micro clustering phase: After partitioning all the sensors into two sets, in each set, we employ the Partitioning Around Medoid (PAM) algorithm[2] [50] to partition the sensors in the same set into several clusters or subsets based on the decisions $\{\mathbf{u}_i\}_{i=1}^{N}$. We assume that the sensors are grouped into K clusters in each set via PAM.[3] Hence, we have a total $2K$ clusters in the network.

- Voting phase: The Voting phase contains two successive steps, i.e., Intra-cluster voting and Inter-cluster voting.

  – Intra-cluster voting: After each sensor makes the decision at time step t, we perform cluster voting by weighting the decisions of the sensors in that cluster with their impact factors. The impact factor of sensor i is inversely proportional to the Hamming distance between the decision vector of sensor i and the decision vector of the medoid of that cluster for the most recent T time steps. Let $I_i(t)$ denote the impact factor of sensor i at time step t and it is given as

$$I_i(t) = \begin{cases} \frac{1}{d_t(i,m_k)} & \text{if sensor } i \text{ is in cluster } k \\ 0 & \text{if sensor } i \text{ is not in cluster } k \end{cases} \tag{3.5}$$

[1] The information of each sensor consists of T local decisions, T MMS results, and one reputation index.

[2] PAM is one possible algorithm to implement K-medoid clustering. K-medoid clustering is a prominent clustering technique which attempts to minimize the distance between points assigned to a cluster and a point designated as the center of that cluster, namely the medoid of that cluster.

[3] The number of clusters in different sets are assumed to be the same for simplicity.

where $d_t(i, m_k)$ denotes the Hamming distance between the decision record of sensor i and the decision record of the medoid of that cluster for the most recent T time steps, and, m_k represents the index of the medoid of cluster $k \in \{1, \ldots, 2K\}$. Note that the first K clusters consist of the sensors in set $\underline{\mathcal{T}}$ and the rest of the clusters consist of the sensors in set $\overline{\mathcal{T}}$. According to (3.6), the cluster vote for cluster k at time t is either 0 or 1 to represent the absence or presence of the PoI, respectively, at the cluster level.

$$\mathcal{V}_k(t) = \left\lfloor \frac{\sum_{i=1}^{N} I_i(t) u_i(t)}{\sum_{i=1}^{N} I_i(t)} \right\rceil, \tag{3.6}$$

where $\lfloor x \rceil$ means rounding x to the nearest integer.

– Inter-cluster voting: After receiving all the cluster decisions, the FC makes a decision regarding the behavioral status, i. e., Byzantine or not, of each cluster based on the cluster reputation. Assume the initial reputation for all the sensors in the network is r_{init} and the cluster reputation is defined as the averaged reputation of the sensors in that cluster. If the cluster reputation is below a threshold λ_{valid}, the cluster is temporarily considered to be Byzantine and the cluster decision from that cluster is not taken into consideration when the FC makes the final decision regarding the hypothesis that is true. The decision rule used by the FC is expressed as

$$\gamma_1 \sum_{k=1}^{K} \beta_k \mathcal{V}_k(t) + \gamma_2 \sum_{k=K+1}^{2K} \beta_k \mathcal{V}_k(t) \underset{v_{fc}(t)=0}{\overset{v_{fc}(t)=1}{\gtrless}} \lambda_{fc}, \tag{3.7}$$

where λ_{fc} is the threshold used by the FC, $v_{fc}(t)$ is the final decision at time step t, and, γ_1 and γ_2 are the weights of the cluster decisions for the first K clusters (clusters in set $\underline{\mathcal{T}}$) and the weights of the cluster decisions for the rest of the clusters, respectively. Based on Lemma 3.1, we give appropriate values to γ_1 and γ_2 to emphasize different importance of the cluster decisions from different sets. β_k is the weight of decisions

from cluster k in the corresponding set and it is given by

$$\beta_k = \begin{cases} \frac{Y_k}{\sum_{k=1}^{K} Y_k} & \text{if cluster } k \in \{1, 2, \ldots, K\} \\ \frac{Y_k}{\sum_{k=K+1}^{2K} Y_k} & \text{if cluster } k \in \{K+1, K+2, \ldots, 2K\} \end{cases} \tag{3.8}$$

where $Y_k = \frac{n_k F_k}{\sum_{k=1}^{2K} n_k F_k}$, n_k is the number of sensors in cluster k and F_k is the cluster behavioral identity indicator for cluster k. $F_k = 1$ represents the cluster k is not considered Byzantine and $F_k = 0$ represents the cluster k is considered Byzantine.

We set $\gamma_1 > \gamma_2$ to emphasize that the importance of cluster decisions coming from set $\underline{\mathcal{T}}$ is greater than the ones coming from set $\overline{\mathcal{T}}$ according to Lemma 3.1.

Lemma 3.1. *The sensors in set $\underline{\mathcal{T}}$ have a higher probability being Byzantine nodes than the sensors in set $\overline{\mathcal{T}}$ when $\alpha_0 \leq 0.8$.*

PROOF: Please see Appendix A.2. ■

- Reputation updating phase: At the end of each time step, the reputations of all the sensors are updated. The final decision of the FC is propagated back to cluster level and further to the individual level to update the reputation of each sensor. If the final decision is the same as a cluster decision, that cluster gets a positive feedback; otherwise, it gets a negative feedback. Similarly, if the cluster decision is the same as a sensor decision in that cluster, that sensor gets a positive feedback; otherwise, it gets a negative feedback. The reputation updating rule of sensor i is given as

$$r_i = r_i + M(v_{fc}(t), \mathcal{V}_k(t)) g_k \frac{H_i(t)}{\sum_{i=1}^{N} H_i(t)}, \tag{3.9}$$

where

$$g_k = \frac{\sum_{i=1}^{N} M(u_i(t), \mathcal{V}_k(t)) I_i(t)}{\sum_{i=1}^{N} I_i(t)} \tag{3.10}$$

represents the step size to penalize or reward a sensor in cluster k and $M(a, b)$ is an indicator function that returns 1 if a equals b and returns -1 otherwise. Let $H_i(t)$ denote the reputation impact factor of sensor i at time step t and it is given by

$$H_i(t) = \begin{cases} \frac{1}{D_t(i,m_k)} & \text{if sensor } i \text{ is in cluster } k \\ 0 & \text{if sensor } i \text{ is not in cluster } k \end{cases} \tag{3.11}$$

where $D_t(i, m_k)$ is the Hamming distance between the MMS result record of sensor i and the MMS result record of the medoid of that cluster for the most recent T time steps. The reputation updating rule of cluster k is given as

$$R_k = \frac{\sum_{i \in \mathcal{E}_k} r_i}{n_k}, \tag{3.12}$$

for $k = 1, 2, \ldots, K$, where $\mathcal{E}_k$ is the set of indices of the sensors in cluster k. If R_k is smaller than a threshold τ, we temporarily remove all the sensors in cluster k and go back to voting phase.[4]

3.3.2 Proposed Algorithm with Auxiliary Anchor Node

In the above algorithm, simulation results show an improved detection performance of the system. However, as we will see later, the simulation results in Fig. 3.4 show that the system employing RAC algorithm breaks down when the Byzantine nodes adopt the strategy that p_1 approaches 1, p_2 approaches 0 and $\alpha_0 \geq 0.5$. Hence, we further propose an algorithm with auxiliary anchor nodes to overcome that problem. We use the same procedure in the above algorithm except the Reputation updating phase. Assume there are J ($J \ll N$) anchor nodes in the network which can be trusted by the FC, and P_d and P_f are the same as the other sensors in the network. Let $A(t)$ represent the final decision according to the local decisions from anchor nodes at time step t and

[4]Because it is possible that several honest nodes are grouped into a Byzantine cluster or the cluster is wrongly identified as Byzantine, we just temporarily remove all the sensors in that cluster.

the final decision is decided by majority vote if more than 1 anchor node is used. The reputation updating rule of sensor i given in (3.9) is reformulated as

$$r_i = r_i + M(v_{fc}(t), \mathcal{V}_k(t)) g_k f \frac{H_i(t)}{\sum_{i=1}^{N} H_i(t)}, \tag{3.13}$$

in the proposed algorithm with auxiliary anchor node, where f is given by

$$f = \frac{\sum_{q=t-T+1}^{t} Q(A(t), A(q))}{T} M(A(t), v_{fc}(t)), \tag{3.14}$$

where $Q(a, b)$ is an indicator function that returns 1 if a equals b and returns 0 otherwise. f can be regarded as a reward (or punishment) step size of reputation when the decision of the anchor node is the same as the final decision (or different from the final decision). Note that although the anchor nodes are reference nodes, they still have a chance to make the wrong decisions and we assume the hypothesis does not change here. In this algorithm, we are able to accurately identify most of the Byzantine nodes in the system and obtain excellent detection performance with the help of anchor nodes even when the number of Byzantine nodes is greater than half of the total number of sensors in the network.

3.4 Performance Analysis

In this section, we evaluate the robustness of the system and determine the optimal attack strategy for Byzantines when employing our proposed algorithms to make the FC completely blind. Since Macro clustering is performed, we need to consider two cases: (i) The sensors are in set $\underline{\mathcal{T}}$; (ii) The sensors are in set $\overline{\mathcal{T}}$. The probabilities of detection and false alarm are different for the sensors in different sets. Let $\underline{\pi}_{11}$, $\underline{\pi}_{10}$, ans $\overline{\pi}_{11}$, $\overline{\pi}_{10}$ denote the probabilities of detection and false alarm for the sensors in set $\underline{\mathcal{T}}$ and the sensors in set $\overline{\mathcal{T}}$, respectively. We have

$$\underline{\pi}_{11} = 1 - \underline{\pi}_{01} = P(u_i = 1|\mathcal{H}_1) = P_d(1 - \underline{\alpha} p_1) + \underline{\alpha} p_1 (1 - P_d) \tag{3.15a}$$

$$\underline{\pi}_{10} = 1 - \underline{\pi}_{00} = P(u_i = 1|\mathcal{H}_0) = P_f(1 - \underline{\alpha}p_1) + \underline{\alpha}p_1(1 - P_f) \tag{3.15b}$$

for any sensor i in $\underline{\mathcal{T}}$, and

$$\overline{\pi}_{11} = 1 - \overline{\pi}_{01} = P(u_i = 1|\mathcal{H}_1) = P_d(1 - \overline{\alpha}p_1) + \overline{\alpha}p_1(1 - P_d) \tag{3.16a}$$

$$\overline{\pi}_{10} = 1 - \overline{\pi}_{00} = P(u_i = 1|\mathcal{H}_0) = P_f(1 - \overline{\alpha}p_1) + \overline{\alpha}p_1(1 - P_f) \tag{3.16b}$$

for any sensor i in $\overline{\mathcal{T}}$. $\underline{\alpha}$ is the probability that one sensor in set $\underline{\mathcal{T}}$ is a Byzantine node and $\overline{\alpha}$ is the probability that one sensor in set $\overline{\mathcal{T}}$ is a Byzantine node. $\underline{\alpha}$ is the probability that one sensor in set $\underline{\mathcal{T}}$ is a Byzantine node and $\overline{\alpha}$ is the probability that one sensor in set $\overline{\mathcal{T}}$ is a Byzantine node and they are given by

$$\underline{\alpha} = \frac{\alpha_0^2 f_1 + \alpha_0(1-\alpha_0)f_2}{\alpha_0^2 f_1 + 2\alpha_0(1-\alpha_0)f_2 + (1-\alpha_0)^2}, \tag{3.17a}$$

$$\overline{\alpha} = \frac{\alpha_0 - (\alpha_0^2 f_1 + \alpha_0(1-\alpha_0)f_2)}{1 - (\alpha_0^2 f_1 + 2\alpha_0(1-\alpha_0)f_2 + (1-\alpha_0)^2)}, \tag{3.17b}$$

where $f_1 = [2p_1p_2(1-p_1) + (1-2p_1+2p_1^2)(1-p_2)]^2$ and $f_2 = (1-p_2)(1-2p_1+2p_1^2)$. To totally blind the FC, the adversaries need to ensure that the following equalities simultaneously hold.

$$D\left(\underline{\alpha}, p_1, p_2\right) = 0, D\left(\overline{\alpha}, p_1, p_2\right) = 0 \tag{3.18}$$

where $D(\cdot)$ represents the Kullback–Leibler divergence (KLD), and $D\left(\underline{\alpha}, p_1, p_2\right) = \underline{\pi}_{11} \log \frac{\underline{\pi}_{11}}{\underline{\pi}_{10}} + \underline{\pi}_{01} \log \frac{\underline{\pi}_{01}}{\underline{\pi}_{00}}$ and $D\left(\overline{\alpha}, p_1, p_2\right) = \overline{\pi}_{11} \log \frac{\overline{\pi}_{11}}{\overline{\pi}_{10}} + \overline{\pi}_{01} \log \frac{\overline{\pi}_{01}}{\overline{\pi}_{00}}$. The equations in (3.18) always hold only when $\underline{\pi}_{11} = \underline{\pi}_{10}$ and $\overline{\pi}_{11} = \overline{\pi}_{10}$, respectively, which yields $\underline{\alpha}p_1 = \frac{1}{2}$ and $\overline{\alpha}p_1 = \frac{1}{2}$. Moreover, since we assign different weights to different cluster decisions, i.e., γ_1 and γ_2, according to the fusion rule shown in (3.7), the optimal attacking strategy is $p_1 = 1$, $p_2 = 0$, and $\alpha_0 = 0.5$.

Let P_{HH}^{diff} and P_{BH}^{diff} denote the probabilities that two honest nodes differ in their sensing reports and the probability that one honest node and one Byzantine node differ in their sensing reports, respectively, at any time step t. Obviously, P_{HH}^{diff} and P_{BH}^{diff} are given as $P_{HH}^{diff} = 2\pi_0 P_f(1 - P_f) +$

$2\pi_1 P_d(1-P_d)$ and $P_{BH}^{diff} = \pi_0[\kappa_{10}(1-P_f)+\kappa_{00}P_f]+\pi_1[\kappa_{11}(1-P_d)+\kappa_{01}P_d]$, respectively, where $\kappa_{10} = (1-P_f)p_1 + P_f(1-p_1)$ and $\kappa_{11} = (1-P_d)p_1 + P_d(1-p_1)$ are the probabilities of detection and false alarm for the Byzantine nodes in the network. If the adversaries want to totally deceive the FC so that the FC misplaces Byzantine nodes and honest nodes in the same cluster in Micro clustering phase, we should have

$$D\left(P_{HH}^{diff}|P_{BH}^{diff}\right) = P_{HH}^{diff}\log_2\left(P_{HH}^{diff}/P_{BH}^{diff}\right) = 0 \tag{3.19}$$

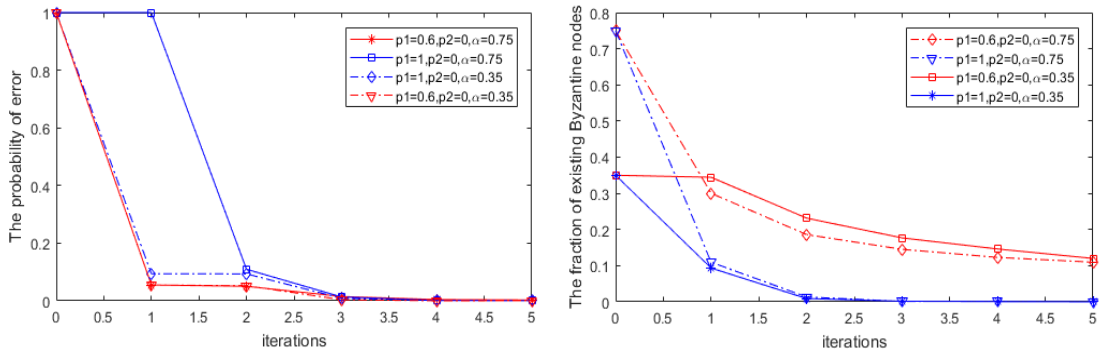
where $D\left(P_{HH}^{diff}|P_{BH}^{diff}\right)$ is the KLD. The solutions of equation (3.19) are $p_1 = 0$ or $P_d = P_f = \frac{1}{2}$ which means that the adversaries can totally deceive the FC in Micro clustering phase only when $p_1 = 0$ or $P_d = P_f = \frac{1}{2}$.

So in conclusion, the proposed mechanism pushes the Byzantine nodes to choose a large p_1 and a small p_2 to blind the FC in Macro clustering phase. It is due to the fact that a small p_2 increases the probability that the Byzantine nodes are placed into set $\underline{\mathcal{T}}$, whose sensor decisions have more impact on the final decision. However, a large p_1 also increases the exposure to our defense mechanism in Micro clustering phase which guarantees a good detection performance. Benefiting from our proposed scheme, we are also able to achieve a good detection performance even when $\alpha_0 \geq 0.5$. It should also be noted that in prior work (for e.g., [33, 34, 45, 82, 97]), the FC can be made blind with only 50% of Byzantine nodes in the network.

3.5 Simulation Results and Discussion

Some numerical results are presented in this section. We assume that identical sensors are utilized in the networks. Hence, we have $P_d = 0.9$, $P_f = 0.1$ for sensor $i \in \{1, \ldots, N\}$ and anchor node $j \in \{1, \ldots, J\}$. We set $N = 500$, $r_{init} = 0.5$, $\lambda_{valid} = 0.5$, $\tau = 0.5$, $\gamma_1 = 1.5$, $\gamma_2 = 0.5$ and $\lambda_{fc} = 1$.

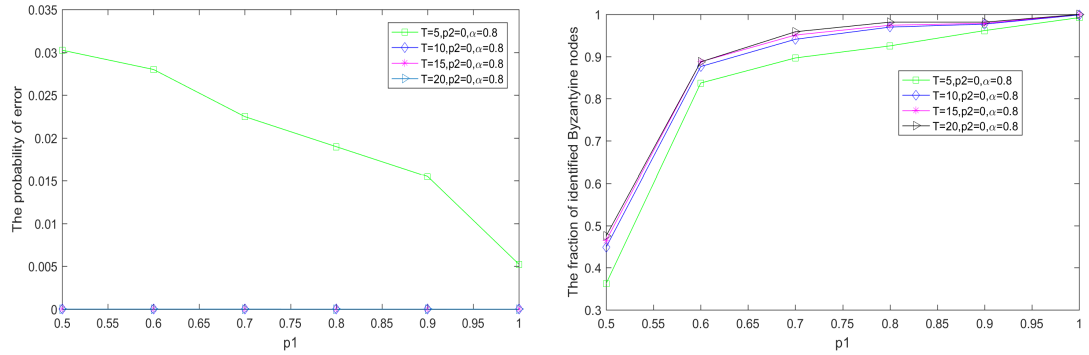
Fig. 3.2 shows that our proposed algorithms are able to quickly identify Byzantine nodes so that we can obtain an excellent detection performance. Fig. 3.3 shows that the starting dimension



(a) The probability of error versus the number of iterations of the proposed algorithm.

(b) The fraction of identified Byzantine nodes versus the number of iterations of the proposed algorithm.

Fig. 3.2: The probability of error and the fraction of identified Byzantine nodes for the proposed algorithm with one anchor node given $\alpha_0 = 0.35$ and $\alpha_0 = 0.75$.



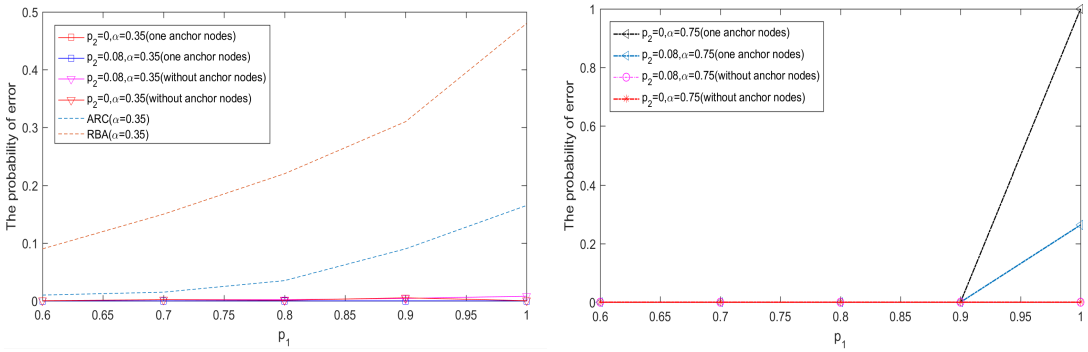
(a) The probability of error versus p_1 given different starting starting dimension T

(b) The fraction of identified Byzantine nodes versus p_1 given different starting starting dimension T

Fig. 3.3: The probability of error and the fraction of identified Byzantine nodes given different value of T for the proposed algorithm with one anchor node.

T that is needed to get a relatively good detection performance is greater than 10. Hence, to start with the algorithm, we maintain at least the latest 10 decision records to guarantee a relatively good detection performance.

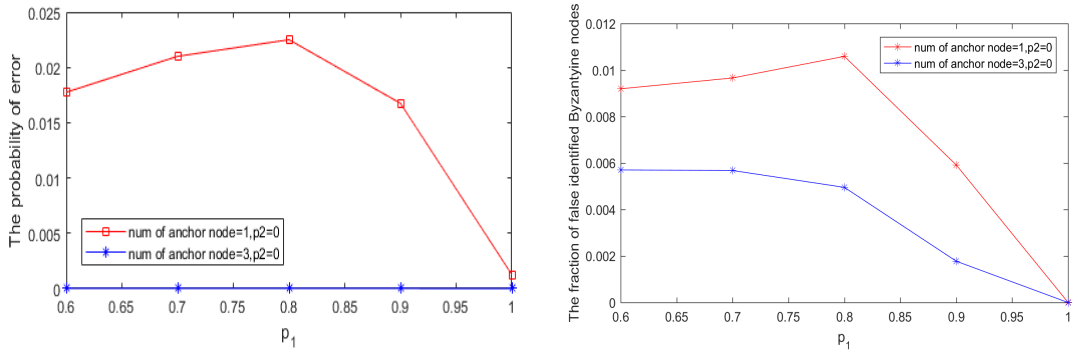
In Fig. 3.4(a), we can observe that the proposed algorithm with one anchor node and the one without anchor nodes both have outstanding detection performance when the fraction of Byzantine nodes $\alpha_0 = 0.35$. Fig. 3.4(b) shows that the system can still obtain a good detection performance with the help of anchor nodes even when the fraction of Byzantine nodes is greater than 0.5. We can also observe that the detection performance of the algorithm without anchor nodes degrades



(a) The probability of error versus p_1 when $\alpha_0 = 0.35$.

(b) The probability of error versus p_1 when $\alpha_0 = 0.75$.

Fig. 3.4: The probability of error versus p_1 given different value of p_2 and α_0 for the proposed algorithm with one anchor node and without anchor nodes.



(a) The probability of error versus p_1.

(b) The fraction of identified Byzantine nodes versus p_1.

Fig. 3.5: The probability of error and the fraction of identified Byzantine nodes for the proposed algorithm with different number of anchor nodes.

significantly only when p_1 approaches 1 and p_2 approaches 0 if the fraction of Byzantine nodes is greater than 0.5. We can also observe that our proposed algorithms outperform Adaptive reputation clustering algorithm (ARC) proposed in [97] and Reputation based algorithm (RBA) proposed in [82]. Note that those algorithms (ARC and RBA) break down when over half of the sensors are Byzantine nodes. Furthermore, we examined the impact of the number of anchor nodes used on the performance of the system in Fig. 3.5. It shows that an increasing number of anchor nodes significantly enhances the detection performance of the system since a larger number of anchor nodes could provide the FC better reference decisions to identify the Byzantine nodes. Fig. 3.6 shows the detection performance of our proposed algorithms under Byzantine attacks with dynam-

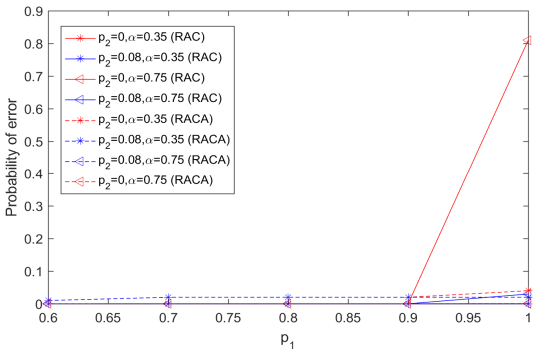


Fig. 3.6: The probability of error versus dynamically changing p_1 given different value of p_2 and α_0 for our proposed algorithms.

ically changing attacking parameter p_1. In each time step, we assume that the real value of p_1 is uniformly generated from $[p_1 - 0.05, p_1 + 0.05]$ in order to represent the dynamically changing attacking parameter p_1. The average error probability we obtain for a specific dynamically changing $p_1 \in [p_1 - 0.05, p_1 + 0.05]$ is the error probability that corresponds to p_1 in Fig. 3.6. We can observe that our proposed algorithms are able to defend against attackers whose attacking parameters are dynamically changing. It is due to the fact that the performance improvement of our proposed algorithms is directly affected by the deviation of Byzantine nodes' decision records from those of honest nodes, and the deviation is the reflection of p_1 and p_2 in the system. Hence, the dynamically changing attacking parameters in each iteration do not result in a significant impact on the ability of our proposed algorithms to defend against attacks.

3.6 Summary

In this chapter, we proposed the RAC algorithm and the RACA algorithm to defend against Byzantine attacks in sensor networks when the FC is not aware of the attacking strategy. We utilized the history of local decisions and MMS results to update the reputation index of sensors and help the system accurately identify Byzantine nodes. Our simulation results showed that we are able to achieve superior detection performance and the enhanced ability of identifying Byzantine nodes by employing anchor nodes even when the Byzantines exceed half of the total number of sensors

in the network. Furthermore, we showed that our algorithms are capable of defending against attackers whose attacking parameters change dynamically over time.

CHAPTER 4
ORDERED TRANSMISSION-BASED DETECTION IN DISTRIBUTED NETWORKS IN THE PRESENCE OF BYZANTINES

In this chapter, we consider the Byzantine attack problem in ordered transmission based (OT-based) schemes for the binary hypothesis testing problem. Ordered transmission (OT) is a promising technique which reduces the number of transmissions needed in a distributed detection network without any loss in the probability of error performance. Here, we discuss two main types of systems: the conventional OT-based system and the communication-efficient OT-based (CEOT-based) system. We investigate the performance of the aforementioned two OT-based systems in the presence of additive Byzantine attacks in Gaussian shift in mean problems, focusing on the detection performance and the number of transmissions saved. Moreover, we conduct a performance comparison between the conventional OT-based system and the CEOT-based system to reveal the robustness of these two systems.

4.1 Introduction

Energy-efficiency is an important aspect to consider while designing a wireless sensor network (WSN) with prolonged lifetime [23]. Several notable schemes have been proposed to improve energy efficiency by reducing the number of transmissions in the networks [3, 4, 7, 80]. In this chapter, we consider two such schemes called the conventional ordered transmission based (OT-based) scheme [7] and CEOT-based scheme [88]. In the conventional OT-based scheme, all the sensors in the network transmit their data in decreasing order of their respective absolute values of the log-likelihood ratios (LLRs). In the CEOT-based scheme, informative sensors transmit binary decisions to the FC, improving communication efficiency in the distributed setup, rather than sending raw LLR values. In both schemes, the starting time of transmission at each sensor is proportional to the inverse of the absolute value of its LLR. Hence, the more informative sensors (sensors with larger magnitudes of the value of LLR) transmit earlier than the less informative ones (sensors with smaller magnitudes of the value of LLR). When the FC has received enough observations to make the final decision of desired quality, the FC broadcasts a stop signal to stop the sensors from further transmitting. The sensors that have not yet transmitted their observations reset their timers for the next decision interval after they receive the stop signal. For simplicity, in the rest of this chapter, the OT-based scheme refers to the conventional OT-based scheme.

4.1.1 Related Work

The OT-based scheme for distributed networks was first proposed in [7], where only informative sensors in the network transmitted their LLRs to the FC instead of sending raw data. The concept was extended to an ordering approach for a class of noncoherent signal detection problems where the LLRs at each sensor could only take nonnegative values in [81]. The authors in [8] demonstrated that a single observation was sufficient to make a final decision for an OT-based system with a large number of sensors. In [35], sequential detection along with OT was considered for cooperative spectrum sensing to obtain fast and reliable decisions regarding primary user activities

over the spectrum. The sequential test was run at the FC with a constraint on the maximum number of sensors that reported their LLRs. This constraint was incorporated using the OT-based scheme. Furthermore, the authors in [11] considered the quickest change detection problem to detect the change in the distribution of independent observations by proposing a new approach where the transmissions from the sensors were ordered and stopped when sufficient information was accumulated at the FC. The authors showed that the proposed approach achieved the optimal performance equivalent to the centralized cumulative sum (CUSUM) algorithm with less sensor transmissions. In [12], the OT-based scheme is employed in the quickest change detection problem with dependent observations to reduce communications without increasing detection delays. The dependence among sensor observations is characterized using a decomposable graphical model (DGM). The authors showed that the proposed algorithm is able to achieve identical performance to the non-OT based scheme in terms of the worst-case average detection delay. In order to eliminate some sensor-to-FC uplink communications, the authors in [13] considered an ordered gradient approach where the timer at each sensor was set inversely proportional to the magnitude of the gradient of the loss function. This resultant gradient-based approach achieved the same order of convergence rate as the gradient descent approach for nonconvex smooth loss functions. Also, the work in [6] considered an OT-based algorithm for the discretized estimation problem with a latency constraint. The authors showed that the proposed algorithm can greatly reduce latency without loss of estimation accuracy. In [31], the OT-based scheme was incorporated along with energy harvesting in the WSNs in order to improve energy efficiency of the sensors. A correlation-aware OT-based scheme was proposed in [30] where spatial correlation between the sensors was considered. The OT-based framework was applied to determine a shift in the mean and covariance, and the decision rule was proposed accordingly. The CEOT-based scheme was proposed in [88], where informative sensors transmit binary decisions to the FC, improving communication efficiency in the distributed setup, rather than sending raw LLR values. The above works show that the OT-based schemes are capable of efficiently reducing the number of transmissions needed for decision-making while maintaining the same inference performance.

However, due to the large number of low-cost sensors and the vulnerability of WSNs to failures and adversarial attacks, the robustness of the OT-based and CEOT-based systems under attack is an important aspect to consider. Here, we consider these systems operating under Byzantine attacks [25, 60, 64, 66, 66, 96, 103, 110].

4.1.2 Major Contributions

Unlike the previous OT-based systems that only considered honest sensors in the networks [6–8, 11–13, 19, 30, 31, 35, 81, 88], we aim to evaluate the performance of the OT-based systems via both the detection performance and the number of transmissions saved in the presence of Byzantine sensors. We investigate the effect of two types of attacks on the OT-based systems: decision-falsifying Byzantine (DF-Byzantine) attack, where only the local decisions are altered by Byzantine sensors; and additive Byzantine attacks, where the Byzantine sensors can alter not only their data but also the order in which data is transmitted by altering their LLRs. Specifically, we investigate the effect of DF-Byzantine attack on the detection performance and the number of transmissions saved for the CEOT-based system; and the effect of different types of additive Byzantine attacks on the detection performance and the number of transmissions saved for both CEOT-based and OT-based system. The following are our major contributions:

- We investigate the performance of the OT-based distributed detection system in the presence of two kinds of additive Byzantine attacks. The first type involves shifting the mean of the actual observations, which is referred to as the mean-shift attack model. The second type involves shifting both the mean and the variance of the actual observations, known as the mean-variance-shift attack model. We also determine the optimal attack strategy for such Byzantine sensors. The attack strategy is determined by utilizing the deflection coefficient (DC) as a surrogate for the probability of error. Moreover, we evaluate the performance of CEOT-based distributed detection systems in the presence of DF-Byzantine attack and different additive Byzantine attacks.

- We derive the probabilities of error for the OT-based and CEOT-based systems under additive Byzantine attacks and the probability of error for the CEOT-based system under DF-Byzantine attacks. The number of transmissions saved in both OT-based systems are evaluated in the presence of different types of Byzantine sensors. We also derive upper bounds (UB) and lower bounds (LB) on the number of transmissions saved in various types of OT-based networks under various types of attacks.

- A comparison between the OT-based system and the CEOT-based system is made. We observe that the CEOT-based system is more robust to Byzantine attacks since the impact of Byzantine attacks on the CEOT-based system is reduced by the quantization of data.

4.2 System Model

In this section, we consider a binary hypothesis testing problem where hypothesis $\mathcal{H}_1$ indicates the presence of the signal and $\mathcal{H}_0$ indicates the absence of the signal. We consider a distributed network consisting of N sensors and one FC. Furthermore, the OT-based scheme is considered to reduce the number of transmissions in the network. Let y_i be the received observation at sensor $i \in \{1, 2, \ldots, N\}$. We assume that all the observations are independent and identically distributed (i.i.d) conditioned on the hypotheses. For sensor i, the observation y_i is modeled as

$$y_i = \begin{cases} n_i & \text{under } \mathcal{H}_0 \\ s + n_i & \text{under } \mathcal{H}_1, \end{cases} \tag{4.1}$$

where s is non-negative and it is the signal strength at each sensor, and n_i is the Gaussian noise with zero mean and variance σ^2. We assume that s and n_i are independent. Note that y_i is Gaussian with mean s and variance σ^2 under hypothesis $\mathcal{H}_1$, and is Gaussian with mean 0 and variance σ^2 under hypothesis $\mathcal{H}_0$.

Next, we review two OT-based schemes where all the sensors are assumed to be honest. One is

the OT-based scheme proposed in [7] where the local sensors send their LLRs to the FC. The other is the CEOT-based scheme proposed in [88] where the local sensors transmit binary decisions to the FC.

4.2.1 Network with OT-based Scheme

Let L_i denote the LLR for sensor i given by

$$L_i = \log\left(\frac{f_{Y_i}(y_i|\mathcal{H}_1)}{f_{Y_i}(y_i|\mathcal{H}_0)}\right), \tag{4.2}$$

where $f_{Y_i}(y_i|\mathcal{H}_h)$ is the probability density function (PDF) of y_i given hypothesis $\mathcal{H}_h$, for $h = 0, 1$. The prior probabilities of hypotheses $\mathcal{H}_h$ are $P(\mathcal{H}_h) = \pi_h$, for $h \in \{0, 1\}$. Recall that the LLR-based optimal Bayesian hypothesis test at the FC for an unordered system is given by $\sum_{i=1}^{N} L_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda = \log\left(\frac{\pi_0}{\pi_1}\right)$, where λ is the threshold used by the FC. In the OT-based system, the sensor transmissions are ordered based on the magnitude of their respective LLRs. We denote the magnitude of the ordered transmissions as $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$, where $|L_{[i]}|$ indicates the i^{th} largest LLR. Hence, the sensor with LLR $L_{[1]}$ transmits first, the sensor with LLR $L_{[2]}$ transmits second, and so on. The optimal decision rule of the FC [7] becomes

$$\begin{cases} \sum_{i=1}^{k} L_{[i]} > \lambda + n_{UT}|L_{[k]}| & \text{decide } \mathcal{H}_1 \\ \sum_{i=1}^{k} L_{[i]} < \lambda - n_{UT}|L_{[k]}| & \text{decide } \mathcal{H}_0, \end{cases} \tag{4.3}$$

for an OT-based system, where n_{UT} is the number of sensors that have not yet transmitted at time k. The FC waits for the next transmission if it can not make the decision with desired accuracy. In this work, we assume that both the sensors and the FC are aware of the relationship between the transmission time t of the sensors and the corresponding magnitude of their LLRs, i.e, $t \propto 1/|L_i|, \forall i \in 1, 2 \ldots, N$. Note that the following assumption was also made in [7] for their analysis.

Assumption 4.1. $Pr(L_i > 0|\mathcal{H}_1) \to 1$ *and* $Pr(L_i < 0|\mathcal{H}_0) \to 1$ *when* s *is sufficiently large.*

The assumption states that the true hypothesis can be decided easily based on L_i for any sensor

i if the distance (dependent on s) between the distributions of the observations of sensor i occurring under the two hypotheses becomes large.

4.2.2 Network with CEOT-based Scheme

Let $u_i \in \{0, 1\}$ denote the binary local decision regarding the true hypothesis for sensor i given by $L_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \log\left(\frac{\pi_0}{\pi_1}\right)$ [91]. The sensor transmissions are still assumed to be ordered based on the magnitude of their LLRs here. Recall that the magnitudes of the LLRs are ordered as $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$. Then, the sensors transmit their local decisions to the FC in the order determined by their magnitude-ordered LLRs, i.e., in the order of $u_{[1]}, u_{[2]}, \ldots, u_{[N]}$, where $u_{[k]}$ is the local decision of the sensor with k^{th} largest LLR.[1] The optimal decision rule [88] becomes

$$\begin{cases} \sum_{i=1}^{k} u_{[i]} \geq T & \text{decide } \mathcal{H}_1 \\ \sum_{i=1}^{k} u_{[i]} < T - (N-k) & \text{decide } \mathcal{H}_0, \end{cases} \tag{4.4}$$

for an CEOT-based system, which follows the T out of N counting rule. The following assumption is made in [88] for the CEOT-based scheme similar to the OT-based scheme made in [7].

Assumption 4.2. $Pr(u_i = 1|\mathcal{H}_1) \to 1$ *and* $Pr(u_i = 0|\mathcal{H}_0) \to 1$ *when* s *is sufficiently large.*

Remark 4.1. *Note that large* s *is key to proving the result that the average number of transmissions saved by utilizing both the OT-based scheme and the CEOT-based scheme is lower bounded by* $N/2$ *(see [7, Theorem 2] and [88]). However, when* s *is small or when there are Byzantine sensors in the system, Assumptions 1 and 2 are no longer valid.*

In the following sections, we analyze the performance of the OT-based system and the CEOT-based system when confronted with Byzantine sensors employing different attack strategies, i.e, additive Byzantine attacks and DF-Byzantine attack.

[1]Note that the magnitude-ordered LLRs do not imply that local decisions are also magnitude-ordered, i.e., $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$ does not imply $u_{[1]} > u_{[2]} > \ldots > u_{[N]}$.

4.3 Performance of OT-based System with additive Byzantine Attacks

In this section, we first introduce the additive Byzantine attack models we considered in the OT-based system. Then we analyze and evaluate the performance of the OT-based system in the presence of additive Byzantine attacks. Note that only passive systems are considered, i.e. the system is unaware of the presence of attackers. The performance evaluation under different additive Byzantine attacks is investigated, including the evaluation of detection performance of the system and the evaluation of the average number of transmissions saved in the system. Moreover, the optimal attack strategy of Byzantine nodes is found.

4.3.1 Additive Byzantine Attack Models

In our setup, a sensor i can be honest (H) or Byzantine (B). The Byzantine sensors are assumed to have perfect knowledge of the underlying true hypothesis. Admittedly, it is hard to realize in practice but it is useful to consider this case as it provides the impact of Byzantines in the worst case. We also assume that each sensor can be a Byzantine with probability α. Two types of additive Byzantine attacks are considered here, which are mean-shift attack and mean-variance-shift attack.

Mean-shift Attack Model

In the mean-shift attack model, we assume that the Byzantine sensors falsify data by controlling the attack strength. The falsified observation $\tilde{y}_i$ for Byzantine node i is given by

$$\tilde{y}_i = \begin{cases} s + n_i - D & \text{under } \mathcal{H}_1 \\ n_i + D & \text{under } \mathcal{H}_0, \end{cases} \tag{4.5}$$

where D is the attack strength and it is a non-negative constant value. The above attack strategy adopted by Byzantine nodes is equivalent to launching attacks by generating falsified observations

from another distribution, and it is commonly used in the literature [43, 45, 66, 71, 121]. Here, the distribution used by Byzantine nodes to generate falsified observations is obtained by shifting the mean of the actual distribution with a constant value D. For an honest node i, the observation is y_i as given in (4.1).

Mean-variance-shift Attack Model

The mean-variance-shift attack strategy is a more general attack strategy where the signal is perturbed by random noise, instead of a constant noise, when compared to the mean-shift attack strategy. But this kind of attack strategy can be easily extended from the mean-shift attack strategy. For the sake of simplicity in performance analysis, we consider the scenario where the actual data is perturbed by Gaussian noise.

Recall that the mean-shift attack strategy assumes that Byzantines falsify their observations with constant values D and $-D$ as shown in (4.5). Consequently, the LLR for Byzantine sensor i can be expressed as

$$L_i = \begin{cases} \frac{2(y_i - D)s - s^2}{2\sigma^2} = L_{i,true} + f_1(D) & \text{under } \mathcal{H}_1 \\ \frac{2(y_i + D)s - s^2}{2\sigma^2} = L_{i,true} + f_0(D) & \text{under } \mathcal{H}_0, \end{cases} \tag{4.6}$$

where $f_1(D) = -\frac{Ds}{\sigma^2}$, $f_0(D) = \frac{Ds}{\sigma^2}$ and $L_{i,true} = \frac{2y_i s - s^2}{2\sigma^2}$ is the actual value of sensor i's LLR. We can easily observe that Byzantines falsify their actual observations y with D and $-D$, which can be equivalently viewed as falsifying their actual LLRs with constant values $f_1(D)$ and $f_0(D)$. In this case, the falsified LLRs are generated from another Gaussian distribution with a different mean and the same variance. If we assume a more general attack strategy, where the actual observations are perturbed by Gaussian noise, both the mean and variance of the Byzantines' LLRs will be altered.

Assuming that the actual LLR of compromised sensor $i \in \{1, 2, \ldots, N\}$ is perturbed by a random noise component that follows a Gaussian distribution, the perturbed LLR is given by:

$$L_i = \begin{cases} L_{i,true} + n_{1,i,w} & \text{under } \mathcal{H}_1 \\ L_{i,true} + n_{0,i,w} & \text{under } \mathcal{H}_0, \end{cases} \tag{4.7}$$

where $n_{1,i,w}$ represents the perturbation noise under hypothesis $\mathcal{H}_1$ that follows a Gaussian distribution with mean $f_1(D)$ and variance σ_w^2, and $n_{0,i,w}$ represents the perturbation noise under hypothesis $\mathcal{H}_0$ that follows a Gaussian distribution with mean $f_0(D)$ and variance σ_w^2.

Remark 4.2. *Node that Assumptions 1 and 2 made in [7] and [88], respectively, are no longer valid.*

Remark 4.3. *Note that both the sensors and the FC are aware of the relationship between the transmission time t of the sensors and the corresponding magnitude of their LLRs, i.e, $t \propto 1/|L_i|, \forall i \in 1, 2 \ldots, N$. If an attacker deviates from the ordered-transmission protocol, they introduce an additional dimension of adversarial behavior. This non-compliance makes their malicious actions more conspicuous and susceptible to identification by the FC. In other words, it increases the possibility of being easily detected by the system as being malicious. Here, we assume that the Byzantines follow the ordered-transmission protocol. By making this assumption, we are assuming a more challenging situation where attackers attempt to hide their malicious actions within the prescribed protocol.*

4.3.2 Detection Performance

Mean-shift Attack Model

We begin our analysis of the detection performance of the OT-based scheme in the presence of Byzantine sensors by first presenting the following Lemma which states that the OT-based system can achieve the same detection performance as the one without ordering.[2]

[2]Note that both the OT-based and unordered systems have the same probability of error in the presence of Byzantine sensors. However, the number of transmissions saved is significantly impacted by the presence of Byzantine sensors for the OT-based system as discussed later.

Lemma 4.1. *Under the optimum Bayesian decision rule, the detection performance remains the same whether or not the system uses the OT-based scheme in the presence of Byzantine sensors.*

PROOF: The proof is relegated to Appendix A.3. ■

Thus, based on the above Lemma, we can obtain the detection performance of the OT-based system by evaluating the detection performance of the system without ordering. For the system without ordering, we have $L_i = \frac{2y_i s - s^2}{2\sigma^2}$ when sensor i is honest $(i = H)$. The PDF of L_i conditioned on hypothesis $\mathcal{H}_h$ follows Gaussian distribution with mean μ_h and variance σ_h^2 for $h = 0, 1$, where $\mu_1 = \frac{s^2}{2\sigma^2}$, $\mu_0 = \frac{-s^2}{2\sigma^2}$, $\sigma_1^2 = \sigma_0^2 = \frac{s^2}{\sigma^2} \triangleq \beta$. When sensor i is Byzantine $(i = B)$, according to (4.6), the PDF of L_i conditioned on hypothesis $\mathcal{H}_h$ follows Gaussian distribution with mean η_h and variance ν_h^2 for $h = 0, 1$, where $\eta_0 = \frac{s^2 - 2Ds}{2\sigma^2}$, $\eta_1 = \frac{2Ds - s^2}{2\sigma^2}$, $\nu_0^2 = \nu_1^2 = \frac{s^2}{\sigma^2} \triangleq \beta$. Therefore, the PDF of L_i given hypothesis $\mathcal{H}_h$ is expressed as

$$\begin{aligned} f_L(l_i|\mathcal{H}_h) &= \alpha f_L(l_i|\mathcal{H}_h, i = B) + (1-\alpha) f_L(l_i|\mathcal{H}_h, i = H) \\ &= \alpha \mathcal{N}(\eta_h, \nu_h^2) + (1-\alpha)\mathcal{N}(\mu_h, \sigma_h^2), \end{aligned} \tag{4.8}$$

for $h = 0, 1$. Here, α denotes the probability of a node being Byzantine. Let $\mathcal{K} = \{A_1, \ldots A_t, \ldots, A_{2^N}\}$ denote the power set that contains all possible subsets of set $\{1, \ldots, N\}$ and A_t be the t^{th} subset of the combination of honest sensors. Also, $|A_t|$ is the cardinality of set A_t. Let $Z = \sum_{i=1}^{N} L_{[i]}$ denote the global test statistic and $f(Z|\mathcal{H}_h)$ denote the Gaussian mixture with PDF given by $f(Z|\mathcal{H}_h) = \sum_{t=1}^{2^N} (1-\alpha)^{m_t} \alpha^{N-m_t} \mathcal{N}((\mu_h)_{A_t}, N\beta)$ for $h = 0, 1$, where $(\mu_h)_{A_t} = \mu_h |A_t| + \eta_h (N - |A_t|)$ and m_t denotes the cardinality of set A_t, i.e., $m_t = |A_t|$.

Therefore, the detection performance can be evaluated in terms of the probability of detection P_d^{FC} and the probability of false alarm P_f^{FC} of the FC given as $P_d^{FC} = \sum_{t=1}^{2^N} (1-\alpha)^{N-m_t} \alpha^{m_t} Q\left(\frac{\lambda - (\mu_1)_{A_t}}{\sqrt{N\beta}}\right)$ and $P_f^{FC} = \sum_{t=1}^{2^N} (1-\alpha)^{N-m_t} \alpha^{m_t} Q\left(\frac{\lambda - (\mu_0)_{A_t}}{\sqrt{N\beta}}\right)$ by following steps that are similar to those outlined in [45], where $Q(.)$ is the tail distribution function of the standard normal distribution.

From the analysis above, we can calculate the system's error probability, given by $P_e = \pi_1(1 - P_d^{FC}) + \pi_0 P_f^{FC}$. Nevertheless, as the size of $\mathcal{K}$ grows exponentially with increasing N, it becomes

intractable to evaluate the worst-case system performance using P_e. Therefore, we employ the DC [92] as a surrogate to determine the most effective attack strategy adopted by Byzantines so that the worst-case system performance is evaluated. By minimizing DC, P_e is maximized.

The DC is defined as $D(\tilde{Z}) = \frac{(\mathbb{E}(\tilde{Z}|H_1)-\mathbb{E}(\tilde{Z}|H_0))^2}{Var(\tilde{Z}|H_0)}$. For the system without ordering, let $\tilde{Z} = \sum_{i=1}^{N} L_i$ denote the global statistic. Therefore, we have $\mathbb{E}(\tilde{Z}|H_1) = -\mathbb{E}(\tilde{Z}|H_0) = N\frac{s^2-2Ds\alpha}{2\sigma^2}$. From Lemma 4.1 and the above discussion, to maximize the probability of error of the system with ordering, we can minimize the DC of the system without ordering. For a specific value of α, the value of D which minimizes DC is the optimal attack strength D^*. Since the DC is always non-negative, the optimal strategy for the Byzantine sensors is to make $D(\tilde{Z}) = 0$. From the definition of DC, when $\mathbb{E}(\tilde{Z}|H_1) = \mathbb{E}(\tilde{Z}|H_0)$, we have $D(\tilde{Z}) = 0$. Hence, for a given α, the optimal attack strength D^* is given by

$$D^* = \frac{s}{2\alpha}, \tag{4.9}$$

which is the minimum attack strength to blind the FC, i.e., to make the probability of error equal to 1/2.

Mean-variance-shift Attack Model

According to (4.7), it is easy to obtain that the falsified L_i of Byzantine sensor i follows a Gaussian distribution with its mean given by

$$E[L_i|\mathcal{H}_h] = \mu_h + f_h(D), \tag{4.10}$$

where $\mu_1 = \frac{s^2}{2\sigma^2}$ and $\mu_0 = \frac{-s^2}{2\sigma^2}$, and the variance given by

$$Var[L_i|\mathcal{H}_h] = \frac{s^2}{\sigma^2} + \sigma_w^2 \triangleq \nu_{h,w}^2 \tag{4.11}$$

for $h = 0, 1$. To evaluate the performance of the system under such general attacks, we only need to replace ν_0^2 and ν_1^2 with $\nu_{0,w}^2$ and $\nu_{1,w}^2$, respectively, in the discussions above regarding the detection

performance of the mean-shift attack model.

4.3.3 Average Number of Transmissions Saved under Additive Byzantine Attack

Mean-shift Attack Model

When the system is under mean-shift attack, we derive an expression for the average number of transmissions $\bar{N}_t$ in the following theorem. Let k^* denote the minimum number of transmissions needed to make a final decision with desired accuracy. Let $F_{|L|}(l_i|\mathcal{H}_h)$ be the cumulative distribution function (CDF) of $|L_i|$ for $h = 0, 1$ provided as

$$\begin{aligned} F_{|L|}(l_i|\mathcal{H}_h) =& \alpha \left(Q\left(\frac{-l_i - \eta_h}{\nu_h}\right) - Q\left(\frac{l_i - \eta_h}{\nu_h}\right)\right) \\ &+ (1-\alpha)\left(Q\left(\frac{-l_i - \mu_h}{\sigma_h}\right) - Q\left(\frac{l_i - \mu_h}{\sigma_h}\right)\right). \end{aligned} \tag{4.12}$$

Theorem 4.1. *The average number of transmissions $\bar{N}_t$ is given as*

$$\bar{N}_t = \sum_{k=1}^{N} \pi_1 Pr(k^* \geq k|\mathcal{H}_1) + \pi_0 Pr(k^* \geq k|\mathcal{H}_0) \tag{4.13}$$

where

$$Pr(k^* \geq k|\mathcal{H}_h) = E_{\mathbf{L}_{k-1}}\left[F_{|L|}(L_{k-1}|\mathcal{H}_h)^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}} \frac{N!}{(N-k+1)!}\right], \tag{4.14}$$

for $h = 0, 1$. The indicator function $\mathbf{1}_{\{\mathcal{J}\}}$ is 1 when $\mathbf{L}_{k-1} = \{L_1, L_2 \ldots, L_{k-1}\}$ is in the region $\mathcal{J}$, and 0 otherwise. Here, $\mathcal{J}$ is a hyperplane with $k-1$ dimensions formed by the intersection of three hyperplanes, $\mathcal{J} = \mathcal{L} \bigcap \mathcal{U} \bigcap \mathcal{D}$, which are given below

$$\mathcal{L} = \left\{ \mathbf{L}_{k-1} : \sum_{i=1}^{k-1} L_i \leq \lambda + (N-k+1)|L_{k-1}| \right\} \tag{4.15}$$

$$\bar{N_s}^U = \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr\left(|L_{[k]}| \leq \frac{g_U - \lambda}{N-k} | \mathcal{H}_h \right) + Pr\left(|L_{[k]}| \leq \frac{\lambda - g_L}{N-k} | \mathcal{H}_h \right) \right.$$
$$\left. - Pr\left(|L_{[k]}| \leq \min\left(\frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k} \right) | \mathcal{H}_h \right) \right] \tag{4.18}$$

$$\bar{N_s}^L = \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr\left(|L_{[k]}| < \frac{g_L - \lambda}{(N-k)} | \mathcal{H}_h \right) + Pr\left(|L_{[k]}| < \frac{\lambda - g_U}{(N-k)} | \mathcal{H}_h \right) \right] \tag{4.19}$$

$$\mathcal{U} = \left\{ \mathbf{L}_{k-1} : \sum_{i=1}^{k-1} L_i \geq \lambda - (N-k+1)|L_{k-1}| \right\} \tag{4.16}$$

$$\mathcal{D} = \left\{ \mathbf{L}_{k-1} : |L_1| > |L_2| > \cdots > |L_{k-1}| \right\}. \tag{4.17}$$

PROOF: Please see Appendix A.4. ■

Note that the set $\mathcal{L}$ is the set of $\mathbf{L}_{k-1}$ such that the FC can not decide hypothesis $\mathcal{H}_1$. Also, the set $\mathcal{U}$ is the set of $\mathbf{L}_{k-1}$ such that the FC can not decide hypothesis $\mathcal{H}_0$. Furthermore, the set $\mathcal{D}$ is the set of $\mathbf{L}_{k-1}$ such that $L_1, L_2, \ldots, L_{k-1}$ are ordered in magnitude. For a given k, we evaluate (4.14) numerically using the Monte Carlo approach as the following. We generate M i.i.d. realizations of $L_1, L_2, \ldots, L_{k-1}$, where the PDF of L_i is given in (4.8) for $\forall i \in \{1, 2, \ldots, k-1\}$. From our experiments, we observe that when N increases, the number of samples M needed to get an accurate evaluation of (4.14) significantly increases.

Next, we derive the upper bound (UB) and the lower bound (LB) for the number of the transmissions saved by the OT-based scheme under Byzantine attack in the following Theorem. Let $\bar{N_s}^U$ and $\bar{N_s}^L$ denote the UB and the LB of the average number of transmissions saved.

Theorem 4.2. *When N is sufficiently large, the average number of transmissions saved $\bar{N}_s$ can be bounded as $\bar{N_s}^L \leq \bar{N}_s \leq \bar{N_s}^U$, where $\bar{N_s}^U$ and $\bar{N_s}^L$ are given in* (4.18) *and* (4.19)*, respectively. Furthermore, we have $Pr\left(|L_{[k]}| < W | \mathcal{H}_h\right) = \int_{-W}^{W} f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h) \mathrm{d}l_{[k]}$ for $W \in \{\frac{g_U-\lambda}{N-k}, \frac{\lambda-g_L}{N-k}, \min(\frac{g_U-\lambda}{N-k}, \frac{\lambda-g_L}{N-k}), \frac{g_L}{N}$ and $f_{|L_{[k]}|}(|l_{[k]}||\mathcal{H}_h)$ is shown in* (A.44)*. We have $g_L = -[\sum (c_i - \bar{c})^2 N \zeta_h^2]^{\frac{1}{2}} + k\delta_h$ and $g_U = [\sum (c_i - \bar{c})^2 N \zeta_h^2]^{\frac{1}{2}} + k\delta_h$, where δ_h and ζ_h^2 are shown in* (A.41)*. Here, $\bar{c} = \frac{\sum_{i=1}^{N} c_i}{N}$ where $c_i = 1$ if $i \leq k$ and $c_i = 0$ if $i > k$.*

PROOF: Please see Appendix A.5. ■

Mean-variance-shift Attack Model

Similarly, to evaluate the number of transmissions saved in the system under mean-variance-shift attacks, we only need to replace ν_0^2 and ν_1^2 with $\nu_{0,w}^2$ and $\nu_{1,w}^2$, respectively, in the discussions above regarding the saving performance of the mean-shift attack model.

4.4 CEOT-based System with Byzantine Attacks

In this section, we discuss another OT-based framework, called CEOT-based scheme. It was first proposed by [88] where the sensors send binary decisions, rather than LLRs, to the FC. The performance of the CEOT-based system under two types of Byzantine sensors was evaluated: decision-falsifying (DF-Byzantine) sensors, which perform pure decision flipping, and additive Byzantine sensors, which not only flip decisions but also change the transmission order.

To demonstrate the basic concepts of CEOT-based schemes, we again consider a binary hypothesis testing problem. Based on the local observations $\{y_i\}_{i=1}^N$, each sensor $i \in \{1, \ldots, N\}$ makes a binary decision $u_i \in \{0, 1\}$ regarding the true hypothesis using the LLR test $L_i \underset{u_i=0}{\overset{u_i=1}{\gtrless}} \log\left(\frac{\pi_0}{\pi_1}\right)$. Notably, sensor transmissions remain ordered according to the magnitude of their LLRs. Specifically, if the magnitudes of the LLRs are sorted as $|L_{[1]}| > |L_{[2]}| > \ldots > |L_{[N]}|$, the sensors transmit their local decisions to the FC in the order determined by their magnitude-ordered LLRs, i.e., in the order of $u_{[1]}, u_{[2]}, \ldots, u_{[N]}$, where $u_{[k]}$ refers to the k^{th} transmitted local decision.

Thus, the optimal decision rule is given by [88]

$$\begin{cases} \sum_{i=1}^{k} u_{[i]} \geq T & \text{decide } \mathcal{H}_1 \\ \sum_{i=1}^{k} u_{[i]} < T - (N-k) & \text{decide } \mathcal{H}_0, \end{cases} \tag{4.20}$$

which follows the T out of N counting rule.

Attack Model

Two possible types of security threats are considered in this framework: DF-Byzantine attacks and additive Byzantine attacks. In the former, the Byzantine sensors perform pure decision flipping, while in the latter, the sensors not only flip decisions but also change the order of the transmitted decisions. Each sensor is assumed to have a probability α of being a Byzantine sensor, and the Byzantines are assumed to have perfect knowledge of the true hypothesis.

- **DF-Byzantine attack** For a DF-Byzantine sensor $i \in \{1, 2 \ldots, N\}$, we have

$$\begin{cases} v_i = 1 - u_i & \text{with probability } \alpha p \\ v_i = u_i & \text{with probability } 1 - \alpha p \end{cases}, \tag{4.21}$$

  where v_i is the actual local decision made by sensor i, u_i is the local decision sent to the FC by sensor i and p is the probability that the sensor i flips its local decisions. If sensor i is honest, v_i is the same as u_i.

- **Additive Byzantine attacks** Attackers falsify data in the same way as stated in Section 4.3.1.

4.4.1 Performance of CEOT-based System with DF-Byzantines

The performance of the CEOT-based system with DF-Byzantine sensors is analyzed in terms of the detection performance and the number of transmissions saved in the network.

4.4.2 Detection Performance

The following lemma always holds for the CEOT-based scheme in the presence of DF-Byzantines:

Lemma 4.2. *When the FC follows the Bayesian decision rule, the detection performance of systems with and without the use of the CEOT-based scheme is the same in the presence of data falsification attacks.*

PROOF: Please see Appendix A.7. ■

The lemma shows that the CEOT-based system can achieve the same detection performance in the presence of DF-Byzantine attacks as an unordered system. Thus, we evaluate the detection performance of the CEOT-based system in the presence of DF-Byzantine attacks by evaluating the detection performance of the corresponding distributed system without ordering. For the system without ordering, the probabilities of $v_i = 1$ and $v_i = 0$ given $\mathcal{H}_h$ are expressed as

$$\pi_{1,h} = P(v_i = 1|\mathcal{H}_h) = Q\left(\frac{\lambda - \mu_h}{\nu_h}\right) \tag{4.22}$$

and $\pi_{0,h} = P(v_i = 0|\mathcal{H}_h) = 1 - \pi_{1,h}$, respectively, for $h = 0, 1$, where $Q(.)$ is the tail distribution function of the standard normal distribution, $\mu_0 = 0$, $\mu_1 = s$ and $\nu_0 = \nu_1 = \sigma$. Hence, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed as

$$\begin{aligned}\widetilde{\pi}_{1,h} =& P(u_i = 1|\mathcal{H}_h) \\ =& P(u_i = 1|\mathcal{H}_h, i = B)P(i = B) + P(u_i = 1|\mathcal{H}_h, i = H)P(i = H) \\ =& \alpha p \pi_{0,h} + (1 - \alpha p)\pi_{1,h}\end{aligned} \tag{4.23}$$

and $\widetilde{\pi}_{0,h} = P(u_i = 0|\mathcal{H}_h) = 1 - \widetilde{\pi}_{1,h}$, respectively, for $h = 0, 1$. From Assumption 4.2, we have $\pi_{1,0} = \pi_{0,1} \approx 0$ and $\pi_{0,0} = \pi_{1,1} \approx 1$. Thus, we have $\widetilde{\pi}_{1,0} = \widetilde{\pi}_{0,1} \approx \alpha p$ and $\widetilde{\pi}_{1,1} = \widetilde{\pi}_{0,0} \approx 1 - \alpha p$.

The fusion rule for the distributed system when all of the sensor decisions are used is given by

$$\sum_{i=1}^{N} u_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T, \tag{4.24}$$

which follows [72]. Using the decision rule in (4.24), the detection performance can be evaluated in terms of the probability of detection $P_{d,CEOT}^{fc}$ and the probability of false alarm $P_{f,CEOT}^{fc}$ of the FC given below as

$$P_{d,CEOT}^{FC} = \sum_{i=T+1}^{N} \binom{N}{i} \pi_{1,1}^{i} \pi_{0,1}^{N-i} \tag{4.25}$$

and

$$P_{f,CEOT}^{FC} = \sum_{i=T+1}^{N} \binom{N}{i} \pi_{1,0}^{i} \pi_{0,0}^{N-i}, \tag{4.26}$$

respectively. Next, we aim at finding the value of optimal T used by the FC in (4.24) which minimizes the probability of error of both the unordered system and the CEOT-based system. Let $Z = \sum_{i=1}^{N} u_i$ denote the number of local decisions that decided 1. Note that $Z \geq 0$. The optimal decision rule at the FC, which is $\frac{\prod_{i=1}^{N} P(u_i|\mathcal{H}_1)}{\prod_{i=1}^{N} P(u_i|\mathcal{H}_0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\pi_0}{\pi_1}$, can be rewritten as

$$\left(\frac{\widetilde{\pi}_{1,1}}{\widetilde{\pi}_{1,0}}\right)^{Z} \left(\frac{1-\widetilde{\pi}_{1,1}}{1-\widetilde{\pi}_{1,0}}\right)^{N-Z} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\pi_0}{\pi_1}. \tag{4.27}$$

We make the reasonable assumption that the probability of a sensor being malicious is less than 0.5, i.e., $\alpha < 0.5$, and $0 \leq p \leq 1$ which implies that $\alpha p < 0.5$. This implies that $\widetilde{\pi}_{1,1} > \widetilde{\pi}_{1,0}$ (and $\widetilde{\pi}_{0,0} > \widetilde{\pi}_{0,1}$).

Taking the logarithm of both sides of (4.27), the optimal decision rule can be rewritten as

$$Z \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \left[\log\left(\frac{\pi_0}{\pi_1}\right) + N \log\left(\frac{1-\widetilde{\pi}_{1,0}}{1-\widetilde{\pi}_{1,1}}\right)\right] / \log\left(\frac{\widetilde{\pi}_{1,1}(1-\widetilde{\pi}_{1,0})}{\widetilde{\pi}_{1,0}(1-\widetilde{\pi}_{1,1})}\right). \tag{4.28}$$

Therefore, the optimal threshold T^* at the FC is equal to the right hand side of (4.28), i.e., $T^* = \left[\log\left(\frac{\pi_0}{\pi_1}\right) + N \log\left(\frac{1-\widetilde{\pi}_{1,0}}{1-\widetilde{\pi}_{1,1}}\right)\right] / \log\left(\frac{\widetilde{\pi}_{1,1}(1-\widetilde{\pi}_{1,0})}{\widetilde{\pi}_{1,0}(1-\widetilde{\pi}_{1,1})}\right)$.

Average Number of Transmissions Saved under DF-Byzantine Attacks

Next, we consider the effect of DF-Byzantine attack on the average number of transmissions required by the CEOT scheme. In order to simplify the computation, we find the upper bound (UB) of the average number of transmissions required by finding the lower bound (LB) of the average number of transmissions saved. When the system is under attack, we derive the LB for the average number of transmissions saved in the CEOT-based system. We first consider the case when the FC decides $\mathcal{H}_1$. It has been derived in [88] that the average number of transmissions saved to make a final decision is lower bounded by $N/2$ in the absence of the data falsification attacks. Here, we

investigate the effect that the attacks have on the lower bound of the expected number of transmissions saved and finding a lower bound for the system under data falsification attacks. It is also shown in this chapter that the lower bound we obtain is tight.

We define k_L^* as the minimum number of transmissions required to decide $\mathcal{H}_1$ in the presence of data falsification attacks and it is given in (A.65). Without any loss of generality, let $\lceil T \rceil$ denote the rounding up of T to the closest integer that is greater than or equal to T. The average number of transmissions saved when the FC decides $\mathcal{H}_1$ is given as

$$\begin{aligned}
\bar{N}_{s,1}(\beta) =& E(N-k_L^*) = \sum_{k=1}^{N}(N-k)Pr(k_L^*=k) && \text{(4.29a)}\\
=& \sum_{k=1}^{\lceil T\rceil+\beta}(N-k)Pr(k_L^*=k)\\
&+\sum_{k=\lceil T\rceil+\beta+1}^{N}(N-k)Pr(k_L^*=k) && \text{(4.29b)}\\
\geq& \sum_{k=1}^{\lceil T\rceil+\beta}(N-k)Pr(k_L^*=k) && \text{(4.29c)}\\
\geq& (N-\lceil T\rceil-\beta)Pr(k_L^*\leq \lceil T\rceil+\beta) && \text{(4.29d)}\\
=& (N-\lceil T\rceil-\beta)[Pr(\sum_{k=1}^{\lceil T\rceil+\beta}u_{[k]}\geq T|\Gamma_1\geq T)Pr(\Gamma_1\geq T)\\
&+Pr(\sum_{k=1}^{\lceil T\rceil+\beta}u_{[k]}\geq T|\Gamma_1\leq T)Pr(\Gamma_1\leq T)] && \text{(4.29e)}\\
\geq& (N-\lceil T\rceil-\beta)\sum_{h=0,1}Pr(\sum_{k=1}^{\lceil T\rceil+\beta}u_{[k]}\geq T|\Gamma_1\geq T,\mathcal{H}_h)\\
&\times Pr(\Gamma_1\geq T|\mathcal{H}_h)Pr(\mathcal{H}_h) && \text{(4.29f)}\\
=& (N-\lceil T\rceil-\beta)Pr(\sum_{k=1}^{\lceil T\rceil+\beta}u_{[k]}\geq T|\Gamma_1\geq T,\mathcal{H}_1)\\
&\times Pr(\Gamma_1\geq T|\mathcal{H}_1)Pr(\mathcal{H}_1) && \text{(4.29g)}\\
=& (N-\lceil T\rceil-\beta)Pr(\sum_{k=1}^{\lceil T\rceil+\beta}u_{[k]}\geq T|\Gamma_1\geq T,\mathcal{H}_1)Pr(\mathcal{H}_1) && \text{(4.29h)}\\
\triangleq& f_1(\beta), && \text{(4.29i)}
\end{aligned}$$

where $\Gamma_1 = \sum_{i=1}^{\lceil T \rceil + \beta} v_{[i]}$ and $Pr(\sum_{k=1}^{\lceil T \rceil + \beta} u_{[k]} \geq T | \Gamma_1 \geq T, \mathcal{H}_1)$ can be expressed as

$$Pr(\sum_{k=1}^{\lceil T \rceil + \beta} u_{[i]} \geq T | \Gamma_1 \geq T, \mathcal{H}_1) = \sum_{i=0}^{\beta} \binom{\lceil T \rceil + \beta}{i} \widetilde{\pi}_{0,1}^{i} \widetilde{\pi}_{1,1}^{\lceil T \rceil + \beta - i}. \tag{4.30}$$

Substituting (4.30) in (4.29h), we are able to obtain the LB of the average number of transmissions saved in the network when the FC decides $\mathcal{H}_1$. In going from (4.29b) to (4.29c), the second summation term, which is positive, is dropped. As the difference between the actual average number of transmissions saved and its LB is dependent on the number of terms in the dropped second summation term in (4.29b), an appropriate number of terms should be chosen in order to reduce that difference and tighten the LB. Thus, we introduce a variable β in (4.29c) and try to find an appropriate β later to prevent the dropped second part of (4.29b) from being too large so that the LB is tight when the FC decides $\mathcal{H}_1$.

Next, we consider the case when the FC decides $\mathcal{H}_0$. Define k_U^* as the minimum number of transmissions required to decide $\mathcal{H}_0$ and it is given in (A.64). Let $\lfloor T \rfloor$ denote the rounding down of T to the next lowest integer. Similarly, the average number of transmissions saved when the FC decides $\mathcal{H}_0$ is given as

$$\begin{aligned}
\bar{N}_{s,2}(\beta) =& E(N - k_U^*) = \sum_{k=1}^{N} (N-k) Pr(k_U^* = k) && \text{(4.31a)}\\
=& \sum_{k=1}^{N-\lceil T \rceil + \beta} (N-k) Pr(k_U^* = k) \\
&+ \sum_{k=N-\lceil T \rceil + \beta + 1}^{N} (N-k) Pr(k_U^* = k) \\
\geq& \sum_{k=1}^{N-\lceil T \rceil + \beta} (N-k) Pr(k_U^* = k) && \text{(4.31b)}\\
\geq& (\lceil T \rceil - \beta) Pr(k_U^* \leq N - \lceil T \rceil + \beta) && \text{(4.31c)}\\
=& (\lceil T \rceil - \beta)[Pr(\sum_{k=1}^{N-\lceil T \rceil + \beta} u_{[k]} < \kappa | \Gamma_2 > T) Pr(\Gamma_2 > T) \\
&+ Pr(\sum_{k=1}^{N-\lceil T \rceil + \beta} u_{[k]} < \kappa | \Gamma_2 < T) Pr(\Gamma_2 < T)] && \text{(4.31d)}
\end{aligned}$$

$$\geq(\lceil T\rceil-\beta)\sum_{h=0,1}Pr(\sum_{k=1}^{N-\lceil T\rceil+\beta}u_{[k]}<\kappa|\Gamma_2<T)$$

$$\times Pr(\Gamma_2<T) \tag{4.31e}$$

$$=(\lceil T\rceil-\beta)Pr(\sum_{k=1}^{N-\lceil T\rceil+\beta}u_{[k]}<\kappa|\Gamma_2<T,\mathcal{H}_0)$$

$$\times Pr(\Gamma_2<T|\mathcal{H}_0)Pr(\mathcal{H}_0) \tag{4.31f}$$

$$=(\lceil T\rceil-\beta)Pr(\sum_{k=1}^{N-\lceil T\rceil+\beta}u_{[k]}<\kappa|\Gamma_2<T,\mathcal{H}_0)Pr(\mathcal{H}_0) \tag{4.31g}$$

$$\triangleq f_2(\beta), \tag{4.31h}$$

where $\Gamma_2=\sum_{i=1}^{N-\lceil T\rceil+\beta}v_{[i]}$, $\kappa=T-(\lceil T\rceil-\beta)$ and $Pr(\sum_{k=1}^{N-\lceil T\rceil+\beta}u_{[k]}<\kappa|\Gamma_2<T,\mathcal{H}_0)$ can be expressed as

$$Pr(\sum_{k=1}^{N-\lceil T\rceil+\beta}u_{[k]}<\kappa|\Gamma_2<T,\mathcal{H}_0)$$

$$=\sum_{i=0}^{\lfloor T\rfloor-\lceil T\rceil+\beta}\binom{N-\lceil T\rceil+\beta}{i}\widetilde{\pi}_{1,0}^{i}\widetilde{\pi}_{0,0}^{N-\lceil T\rceil+\beta-i}. \tag{4.32}$$

Substituting (4.32) in (4.31g), we are able to obtain the LB of the average number of transmissions saved in the network when the FC decides $\mathcal{H}_0$. In a manner similar to the one employed earlier, variable β is introduced to ensure that the LB of the average number of transmissions saved is tight when the FC decides $\mathcal{H}_0$. Since only one of the two hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$ can occur at any given time, the events $k_L^*=k$ and $k_U^*=k$ given hypothesis $\mathcal{H}_1$ or $\mathcal{H}_0$ are disjoint. Hence the total average number of transmissions saved is $N_{s,CEOT}(\beta)=\sum_{k=1}^{N}(N-k)\sum_{h=0}^{1}[Pr(k_U^*=k|\mathcal{H}_h)+Pr(k_L^*=k|\mathcal{H}_h)]Pr(\mathcal{H}_h)=\sum_{k=1}^{N}(N-k)[Pr(k_U^*=k)+Pr(k_L^*=k)]$ and the LB of the average number of transmissions saved is $N_{s,CEOT}^{L}(\beta)=f_1(\beta)+f_2(\beta)$. When $\beta=0$, the LB derived here reduces to the LB obtained in [88]. However, $\beta=0$ might not be an appropriate value that allows us to get a tight LB in the presence of attacks. Thus, we aim at finding an optimal β so that $N_{s,CEOT}^{L}(\beta)$ is maximized and the LB becomes tighter. Upon solving the optimization problem given in (4.33), we are able to find the optimal β^*. We denote the set of integers by Z and

cast the optimization problem as:

$$\max_{\beta} \quad f_1(\beta) + f_2(\beta) \tag{4.33a}$$

$$\text{s.t.} \quad 0 \leq \beta \leq min(N - \lceil T \rceil, \lceil T \rceil) \tag{4.33b}$$

$$\beta \in \mathbb{Z}, \tag{4.33c}$$

The constraint in (4.33b) is due to the fact that the value of β must satisfy both (4.34) and (4.35), which are derived from (4.29h) and (4.31g), respectively:

$$\lceil T \rceil + \beta \leq N \tag{4.34}$$

$$N - \lceil T \rceil + \beta \leq N \tag{4.35}$$

This is due to the fact that the upper index of the summations in (4.29h) and (4.31g) should be less or equal to N. As the optimization problem in (4.33) is an integer programming (IP) problem, it is a non-convex optimization problem. However, we have the following theorem which helps us obtain the optimal solution to the optimization problem in (4.33).

Theorem 4.3. *$N^L_{s,CEOT}(\beta)$ as a function of β satisfies either*

1. *$N^L_{s,CEOT}(\beta)$ is a non-increasing function, $\forall \beta \in [0, \min(N - \lceil T \rceil, \lceil T \rceil)]$.*

   or

2. *There exists a $\beta_l \in \mathbb{Z}$ such that $N^L_{s,CEOT}(\beta)$ is an increasing function $\forall \beta \in [0, \beta_l - 1]$ and a non-increasing function $\forall \beta \in [\beta_l, \min(N - \lceil T \rceil, \lceil T \rceil)]$.*

PROOF: Let $g_1(\beta) = \sum_{i=0}^{\beta} \binom{\lceil T \rceil + \beta}{i} \widetilde{\pi}^i_{0,1} \widetilde{\pi}^{\lceil T \rceil + \beta - i}_{1,1}$ and $g_2(\beta) = \sum_{i=0}^{\lfloor T \rfloor - \lceil T \rceil + \beta} \binom{N - \lceil T \rceil + \beta}{i} \widetilde{\pi}^i_{1,0} \widetilde{\pi}^{N - \lceil T \rceil + \beta - i}_{0,0}$. Hence, we have

$$g_1(\beta + 1) = \sum_{i=0}^{\beta+1} \binom{\lceil T \rceil + \beta + 1}{i} \widetilde{\pi}^i_{0,1} \widetilde{\pi}^{\lceil T \rceil + \beta + 1 - i}_{1,1} \tag{4.36}$$

$$g_2(\beta+1) = \sum_{i=0}^{T_d+\beta+1} \binom{N-\lceil T\rceil+\beta+1}{i} \widetilde{\pi}_{1,0}^{i} \widetilde{\pi}_{0,0}^{N-\lceil T\rceil+\beta+1-i}, \tag{4.37}$$

where $T_d = \lfloor T\rfloor - \lceil T\rceil = -1$, $\widetilde{\pi}_{0,1} = \widetilde{\pi}_{1,0} = \alpha p$ and $\widetilde{\pi}_{0,0} = \widetilde{\pi}_{1,1} = 1-\alpha p$ based on Assumption 4.2. Hence, we have $g_1(\beta) = \sum_{i=0}^{\beta} \binom{\lceil T\rceil+\beta}{i} (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta-i}$ due to Assumption 4.2. $g_1(\beta+1)$ can be given by

$$\begin{aligned}
g_1(\beta+1) =& \sum_{i=0}^{\beta+1} \binom{\lceil T\rceil+\beta+1}{i} (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta+1-i} && \text{(4.38a)}\\
=& (1-\alpha p)^{\lceil T\rceil+\beta+1} + \sum_{i=1}^{\beta+1} \left[\binom{\lceil T\rceil+\beta}{i} + \binom{\lceil T\rceil+\beta}{i-1} \right] \\
& \times (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta+1-i} && \text{(4.38b)}\\
=& \sum_{i=0}^{\beta+1} \binom{\lceil T\rceil+\beta}{i} (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta+1-i} \\
& + \sum_{i=1}^{\beta+1} \binom{\lceil T\rceil+\beta}{i-1} (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta+1-i} && \text{(4.38c)}\\
=& \binom{\lceil T\rceil+\beta}{\beta+1} (\alpha p)^{\beta+1} (1-\alpha p)^{\lceil T\rceil} + (1-\alpha p) g_1(\beta) \\
& + \sum_{i=0}^{\beta} \binom{\lceil T\rceil+\beta}{i} (\alpha p)^i (1-\alpha p)^{\lceil T\rceil+\beta-i} (\alpha p) && \text{(4.38d)}\\
=& g_1(\beta) + \binom{\lceil T\rceil+\beta}{\beta+1} (\alpha p)^{\beta+1} (1-\alpha p)^{\lceil T\rceil} && \text{(4.38e)}
\end{aligned}$$

based on Pascal's rule. Following a similar sequence of steps, we can obtain

$$g_2(\beta+1) = g_2(\beta) \binom{N-\lceil T\rceil+\beta}{\beta+T_d+1} (\alpha p)^{\beta+T_d+1} (1-\alpha p)^{N-\lceil T\rceil-T_d}. \tag{4.38f}$$

Hence, $g_1(\beta+1)$ and $g_2(\beta+1)$ can be expressed in terms of $g_1(\beta)$ and $g_2(\beta)$ that are respectively given as

$$g_1(\beta+1) = g_1(\beta) + A(\beta) \tag{4.39}$$

and

$$g_2(\beta+1) = g_2(\beta) + B(\beta), \tag{4.40}$$

where $A(\beta) = \binom{\lceil T \rceil + \beta}{\beta+1}(\alpha p)^{\beta+1}(1-\alpha p)^{\lceil T \rceil}$ and $B(\beta) = \binom{N-\lceil T \rceil + \beta}{\beta+T_d+1}(\alpha p)^{\beta+T_d+1}(1-\alpha p)^{N-\lceil T \rceil - T_d}$. It is evident that if

$$[f_1(\beta+1) + f_2(\beta+1)] - [f_1(\beta) + f_2(\beta)] < 0, \tag{4.41}$$

then $N^L_{s,CEOT}(\beta) > N^L_{s,CEOT}(\beta+1)$. By rewriting (4.41) using (4.39) and (4.40), we obtain the inequality

$$D(\beta) > D_2(\beta), \tag{4.42}$$

where $D(\beta) = \pi_1 g_1(\beta) + \pi_0 g_2(\beta)$, $D_2(\beta) = \pi_1 h_1(\beta) + \pi_0 h_2(\beta)$, $\pi_1 = Pr(\mathcal{H}_1)$, $\pi_0 = Pr(\mathcal{H}_0)$, $h_1(\beta) = (N - \lceil T \rceil - \beta - 1)A(\beta)$ and $h_2(\beta) = (\lceil T \rceil - \beta - 1)B(\beta)$.

We proceed to show that if $D(\beta) > D_2(\beta)$ is true, then $D(\beta+1) > D_2(\beta+1)$ is also true. Using the expressions in (4.39), (4.40) and (4.42), we rewrite $D(\beta+1) > D_2(\beta+1)$ as

$$D(\beta) + \pi_1 A(\beta) + \pi_0 B(\beta) > D_2(\beta) - \pi_1 A(\beta) - \pi_0 B(\beta). \tag{4.43}$$

By reformulating (4.43), we have

$$2[\pi_1 A(\beta) + \pi_0 B(\beta)] > D_2(\beta) - D(\beta). \tag{4.44}$$

Due to the assumption that $D(\beta) > D_2(\beta)$, $A(\beta) \geq 0$ and $B(\beta) \geq 0$, the left hand side of (4.44) is greater than or equal to 0 and the right hand side of (4.44) is smaller than 0. Therefore, (4.44) is always true if $D(\beta) > D_2(\beta)$ is true. In other words, if $N^L_{s,CEOT}(\beta) > N^L_{s,CEOT}(\beta+1)$, we always have $N^L_{s,CEOT}(\beta+1) > N^L_{s,CEOT}(\beta+2)$.

Let $\beta = \beta_s$ be the smallest β for which $N^L_{s,CEOT}(\beta) > N^L_{s,CEOT}(\beta+1)$. If $\beta_s = 0$, $N^L_{s,CEOT}(\beta)$ is a decreasing function of β. If $\beta_s > 0$, $N^L_{s,CEOT}(\beta)$ is a non-decreasing function of β when $\beta \in [0, \beta_s - 1]$ and a decreasing function of β when $\beta \in [\beta_s, \min(N - \lceil T \rceil, \lceil T \rceil)]$. Let $\beta = \beta_l$ be the largest β for which $N^L_{s,CEOT}(\beta) < N^L_{s,CEOT}(\beta + 1)$. The above statement is equivalent to that made in Theorem 4.3 about the monotonicity of $N_{s,CEOT}(\beta)$. This completes our proof. ■

According to Theorem 4.3, the optimal solution β^* to the optimization problem in (4.33) is the smallest β for which the inequality $D(\beta) \geq D_2(\beta)$ holds, and the LB of the number of transmissions saved is then given as

$$N^L_{s,CEOT} = f_1(\beta^*) + f_2(\beta^*). \tag{4.45}$$

Therefore, the tight UB of the average number of transmissions required is $N^U_{t,CEOT} = N - N^L_{s,CEOT}$.

4.4.3 Performance of CEOT-based System with Additive Byzantines

Detection Performance

Similar to the case of DF-Byzantine sensors, the detection performance of the CEOT-based system is not affected by ordering in the presence of additive Byzantine sensors as stated in Lemma 4.2. We can evaluate the detection performance of the CEOT-based system in the presence of additive Byzantine attacks by analyzing the detection performance of the corresponding distributed unordered system. If the system is under mean-shift attack, for the system without ordering, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ for an honest sensor i are expressed as $\pi^H_{1,h} = P(u_i = 1|\mathcal{H}_h, i = H) = Q\left(\frac{\lambda - \mu_h}{\sigma_h}\right)$ and $\pi^H_{0,h} = P(u_i = 0|\mathcal{H}_h, i = H) = 1 - \pi^H_{1,h}$, respectively, for $h = 0, 1$. If the sensor is Byzantine $i = B$, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed as $\pi^B_{1,h} = P(u_i = 1|\mathcal{H}_h, i = B) = Q\left(\frac{\lambda - \eta_h}{\nu_h}\right)$ and $\pi^B_{0,h} = P(u_i = 0|\mathcal{H}_h, i = B) = 1 - \pi^B_{1,h}$, respectively, for $h = 0, 1$. Therefore, the probabilities of $u_i = 1$ and $u_i = 0$ given $\mathcal{H}_h$ are expressed

as

$$\pi_{1,h} = P(u_i = 1|\mathcal{H}_h) = \alpha\pi_{1,h}^B + (1-\alpha)\pi_{1,h}^H,$$

and $\pi_{0,h} = P(u_i = 0|\mathcal{H}_h) = \alpha\pi_{0,h}^B + (1-\alpha)\pi_{0,h}^H$, respectively, for $h = 0, 1$.

The fusion rule of the distributed system without ordering is given as $\sum_{i=1}^{N} u_i \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} T$, by noting that we can consider the unordered scheme and taking $k = N$ in (4.20). Based on the fusion rule of unordered system, the detection performance can be evaluated in terms of the probability of detection $P_{d,CEOT}^{FC}$ and the probability of false alarm $P_{f,CEOT}^{FC}$ given as $P_{d,CEOT}^{FC} = \sum_{i=T}^{N} \binom{N}{i} \pi_{1,1}^i \pi_{0,1}^{N-i}$ and $P_{f,CEOT}^{FC} = \sum_{i=T}^{N} \binom{N}{i} \pi_{1,0}^i \pi_{0,0}^{N-i}$. If the system is under mean-variance-shift attack, we only need to replace ν_h^2 with $\nu_{h,w}^2$ for $h = 0, 1.$ in above discussions.

Average Number of Transmissions Saved under additive Byzantine Attacks

Let $\bar{N}_{s,CEOT}$ denote the average number of transmissions saved in the CEOT-based scheme given as

$$\bar{N}_{s,CEOT} = E(N - k^*) = \sum_{k=1}^{N} (N-k) Pr(k^* = k) = \sum_{k=1}^{N-1} Pr(k^* \leq k), \tag{4.46}$$

where k^* denotes the minimum number of transmissions needed to make a final decision with desired accuracy. However, the computation of $Pr(k^* \leq k)$ is intractable. Hence, we derive the UB and LB of $\bar{N}_{s,CEOT}$ by considering the best case and the worst case scenarios for the number of transmissions saved in the network in the presence of Byzantines, respectively. The information of global statistic of the distributed system without ordering, which is given as $\Gamma = \sum_{i=1}^{N} u_i$, is utilized to derive both LB and UB. It is easy to conclude that $\Gamma < T$ means that there exists a k^* such that $\sum_{i=1}^{k^*} u_{[i]} < T - (N - k^*)$ and $\Gamma \geq T$ means that there exists a k^* such that $\sum_{i=1}^{k^*} u_{[i]} \geq T$ according to Lemma 4.2. In order to find the LB and UB of $\bar{N}_{s,CEOT}$, we consider the worst and best cases as follows.

When we consider the worst case, we try to find the maximum k^* needed to make a final decision for a given set of local decisions $\{u_i\}_{i=1}^{N}$. Therefore, the worst case given $\Gamma < T$ would

be that the magnitude of local decisions are ordered in descending order expressed as

$$|z_{[1]}| \geq |z_{[2]}| \cdots \geq |z_{[N]}|, \tag{4.47}$$

where $z_{[k]} \in \{0, 1\}$, for $\forall k \in \{1, 2, \ldots, N\}$ is the k^{th} largest local decision.[3] This is due to the fact that $\Gamma < T$ implies that the unordered system (i.e., the system where the FC receives all local decisions) has more '0' decisions than '1' decisions[4], and the detection performance of the unordered system is the same as the ordered system, as stated in Lemma 4.2. The worst case scenario would occur if the magnitudes of local decisions are ordered in descending order. Similarly, the worst case given $\Gamma \geq T$ would be that the magnitude of local decisions are ordered in ascending order expressed as

$$|z_{(1)}| \leq |z_{(2)}| \cdots \leq |z_{(N)}|, \tag{4.48}$$

where $z_{(k)} \in \{0, 1\}$ for $\forall k \in \{1, 2, \ldots, N\}$ is the k^{th} smallest local decision.

Similar to the above discussion, for the best case, we try to find the minimum k^* needed to make a final decision for a given set of local decisions $\{u_i\}_{i=1}^N$. The best case given $\Gamma < T$ would be that the magnitude of local decisions are ordered in ascending order as shown in (4.48), The best case given $\Gamma \geq T$ would be that the magnitude of local decisions are ordered in descending order as shown in (4.47). Based on the above analysis, we have the following theorem.

Theorem 4.4. *The average number of transmissions saved $\bar{N}_{s,CEOT}$ can be bounded as $\bar{N}^L_{s,CEOT} \leq \bar{N}_{s,CEOT} \leq \bar{N}^U_{s,CEOT}$. Here, the upper bound $\bar{N}^U_{s,CEOT}$ and the lower bound $\bar{N}^L_{s,CEOT}$ are given in* (4.49) *and* (4.50)*, respectively, where $P(\Gamma \geq T|\mathcal{H}_h) = \sum_{i=T}^{N} \binom{N}{i} \pi_{1,h}^i \pi_{0,h}^{N-i}$, $P(\Gamma < T|\mathcal{H}_h) =$*

[3]Note that $z_{[k]}$ is not the same as $u_{[k]}$. The values $u_{[1]}, u_{[2]}, \ldots, u_{[N]}$ are ordered based on the magnitude of their LLRs, while $z_{[1]}, z_{[2]}, \ldots, z_{[N]}$ are ordered based on the magnitude of local decisions $\{u_i\}_{i=1}^N$.

[4]More specifically, the number of '1's should be smaller than T.

$$\bar{N}^{U}_{s,CEOT}=\sum_{h=0}^{1}\sum_{k=1}^{N-1}\pi_h\left[P\left(k_0^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)P(\Gamma\geq T|\mathcal{H}_h)+P\left(k_1^*\leq k|\Gamma<T,\mathcal{H}_h\right)P(\Gamma<T|\mathcal{H}_h)\right], \tag{4.49}$$

$$\bar{N}^{L}_{s,CEOT}=\sum_{h=0}^{1}\sum_{k=1}^{N-1}\pi_h\left[P\left(k_1^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)P(\Gamma\geq T|\mathcal{H}_h)+P\left(k_0^*\leq k|\Gamma<T,\mathcal{H}_h\right)P(\Gamma<T|\mathcal{H}_h)\right], \tag{4.50}$$

$1-P(\Gamma\geq T|\mathcal{H}_h)$, *and*

$$P\left(k_0^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)=\sum_{i=0}^{N-T}\binom{N}{i}\pi_{0,h}^{i}\pi_{1,h}^{N-i}, \tag{4.51}$$

$$P\left(k_1^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)=\sum_{i=0}^{min(N-T,k-T)}\binom{N}{i}\pi_{0,h}^{i}\pi_{1,h}^{N-i}, \tag{4.52}$$

when $k\geq T$, *and*

$$P\left(k_1^*\leq k|\Gamma<T,\mathcal{H}_h\right)=\sum_{i=0}^{T-1}\binom{N}{i}\pi_{1,h}^{i}\pi_{0,h}^{N-i}, \tag{4.53}$$

$$P\left(k_0^*\leq k|\Gamma<T,\mathcal{H}_h\right)=\sum_{i=0}^{min(T-1,k-(N-T+1))}\binom{N}{i}\pi_{1,h}^{i}\pi_{0,h}^{N-i}, \tag{4.54}$$

when $k>N-T$. *Otherwise,* $P\left(k_1^*\leq k|\Gamma<T,\mathcal{H}_h\right)$, $P\left(k_1^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)$, $P\left(k_0^*\leq k|\Gamma\geq T,\mathcal{H}_h\right)$ *and* $P\left(k_0^*\leq k|\Gamma<T,\mathcal{H}_h\right)$ *are equal to 0. Here,* k_0^* *and* k_1^* *denote the minimum number of transmissions needed to make a final decision for descending and ascending ordered local decisions, respectively.*

PROOF: Please see Appendix A.6. ■

4.5 Simulation Results and Comparison of OT-based and CEOT-based Systems under Byzantine Attacks

OT-based and CEOT-based Systems under Additive Byzantine Attacks

In this section, we present the numerical results to corroborate our theoretical results. We set the channel noise variance $\sigma^2 = 1$ and the prior probabilities $\pi_1 = \pi_0 = 0.5$. The detection performance in Fig. 4.1 and the actual average number of transmissions saved in Figs. 4.3, 4.5, 4.7, 4.8 are obtained via Monte Carlo method with 10^4 trials and the average number of transmissions saved in Figs. 4.2 and 4.4 are obtained via Monte Carlo method with 10^7 trials. In order to obtain an accurate evaluation of the average number of transmissions saved in the network as shown in Figs. 4.2 and 4.4, we need to significantly increase the number of trials as the number of sensors increases. Note that the other parameters like the perturbation noise variance σ_w^2, signal strength s, attack strength D, total number of sensors N, and the probability of each sensor being Byzantine α required for the simulations are included in the respective captions of the figures.

Detection performance comparison of OT-based and CEOT-based systems: Fig. 4.1 shows the effect of additive Byzantine attacks on the detection performance of the OT-based system and the CEOT-based system. In Fig. 4.1, we compare the probability of error of the CEOT-based system to the OT-based system and we observe that the CEOT-based system is more robust to additive Byzantine attacks with the same attack parameters. This is due to the fact that the global statistic of an OT-based system is a summation of LLRs, and some of these could be falsified to very large values when D/s is large. In this case, a large deviation is generated from the actual summation of LLRs. However, the global statistic of the CEOT-based system is the summation of quantized LLRs. Although some Byzantine nodes may falsify data, it is unlikely to lead to a significant deviation in the sum of quantized LLR values, even if D/s is large. Hence, D has less negative impact on the probability of error of the CEOT-based system than the OT-based system.

Effect of additive Byzantine attacks on $\bar{N}_s/N$ in the OT-based system: Figs. 4.2 and 4.3 show the effect of additive Byzantine attacks on the average percentage of savings for the OT-based sys-

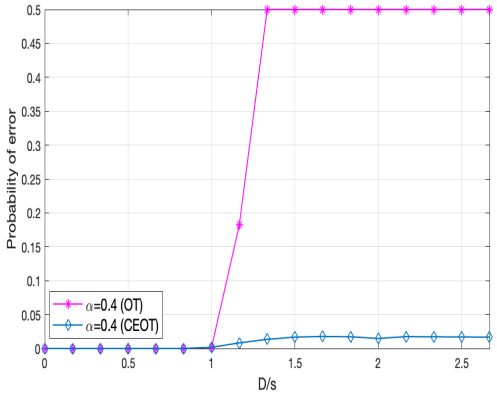


Fig. 4.1: P_e as a function of D/s in the CEOT-based system and the OT-based system when $s = 3$ and $N = 300$.

tem. Fig. 4.2 presents the average percentage of saving $\bar{N}_s/N$ in an OT-based system as a function of D/s for different values of α. Initially, $\bar{N}_s/N$ decreases when D/s increases. However, when D/s increases further, the FC starts to make wrong decisions and the number of transmissions needed to make the final decision starts to decrease and the savings start to increase. We compare the results obtained via simulation using the Monte Carlo method with our analysis using (4.13), and observe a good match. In Fig. 4.3, we observe that the attack strength D^* obtained in (4.9)

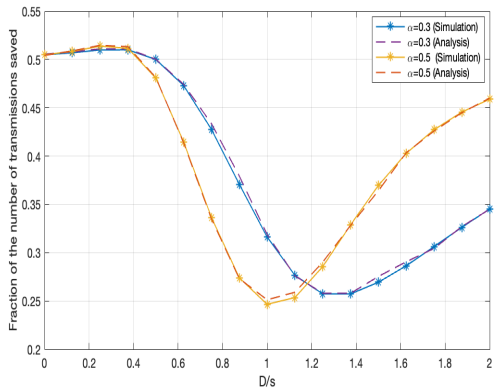


Fig. 4.2: Comparison of $\bar{N}_s/N$ as a function of D/s for different values of α when $s = 4$ and $N = 10$ in the OT-based system.

for the OT-based system is near the point where the average percentage of savings is minimum. Compared with the OT-based system when no Byzantines are present, i.e., D=0, the system in the presence of attacks needs more transmissions to make a final decision. Therefore, the attack

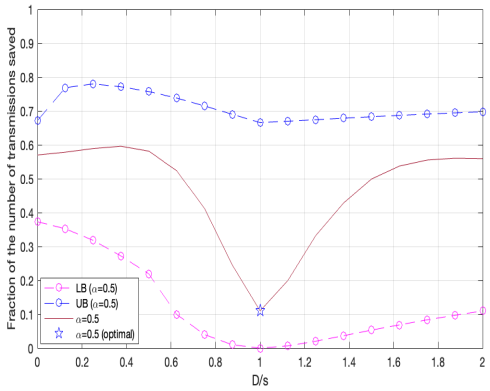
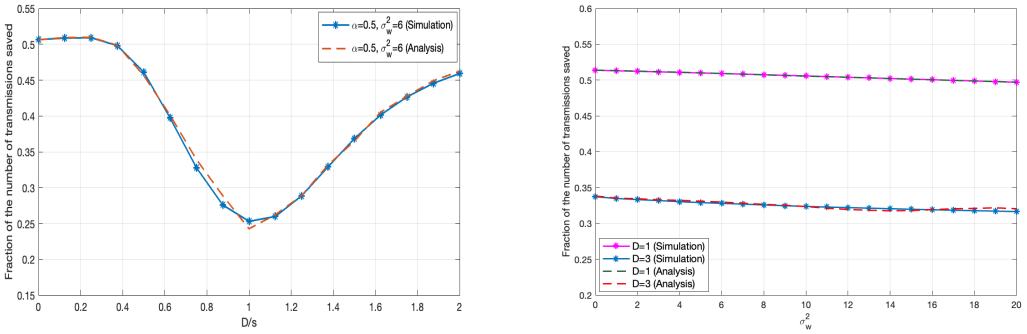


Fig. 4.3: Benchmarking UB and LB for $\bar{N}_s/N$ as a function of D/s for $\alpha = 0.5$ when $s = 4$ and $N = 300$ in the OT-based system.

strength D^* from (4.9) not only blinds the FC but also leads to a smaller average percentage of savings. Fig. 4.3 also shows the UB and LB for the average percentage of savings as a function of D/s in an OT-based system. We observe that both the LB obtained in (4.19) and the UB obtained in (4.18) show a similar trend as that of the average percentage of saving, i.e., the UB and LB track the change in actual average number of transmissions that have been saved. Compared to the UB, the LB performs better in tracking the changes, which enables us to infer what the optimal attack strategy for the attacker is, i.e., what is the value of D that the attacker will employ to cause the greatest damage to the system. As for the UB, it provides more information regarding the maximum number of average transmissions saved in the network as well as alerts about the existence of outliers. For example, if the average number of transmissions saved is larger than the maximum value of UB, we can determine that there are potential outliers and they deviate far from the actual data, i.e., the attackers use an extremely large value of D.

Figs. 4.4 and 4.5 illustrate the impact of mean-variance-shift attacks on the average percentage of savings for the OT-based system. Fig. 4.4 shows that the error probability obtained via simulation using the Monte Carlo method and our error probability analysis have a good match. Fig. 4.5 shows the UB and LB we obtained that show a similar trend as that of the average percentage of saving. As we can observe, Figs. 4.4 (a) and 4.5 (a) demonstrate very similar trends as Figs. 4.2 and 4.3 when the mean of perturbation noise changes. In Figs. 4.4 (b) and 4.5 (b), we can observe

that the values of the variance of the perturbation noise do not significantly affect the average number of transmissions saved. This phenomenon may arise from the fact that the change in variance value only affects the extent to which the noise samples deviate from the mean. Consequently, samples of perturbation noise might fall below or exceed the mean perturbation noise value. Given that the FC's global statistic is the summation of received LLRs, the overall perturbation to this statistic corresponds to the accumulation of perturbation noise from malicious nodes. The average perturbation for each malicious node will tend to the mean of the perturbation noise as the perturbation noise samples below the mean value will balance out the samples above the mean value. Therefore, the change in variance does not have as significant an effect as the change in mean on the number of transmissions saved in the network.

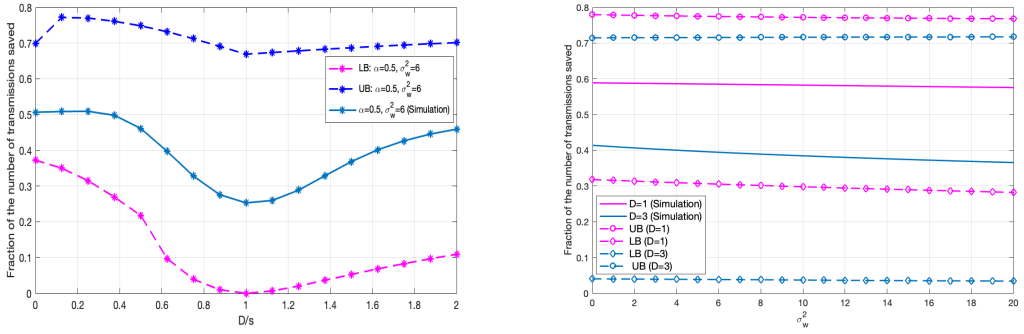


(a) $\bar{N}_s/N$ as a function of D/s for $\sigma_w^2 = 6$. (b) $\bar{N}_s/N$ as a function of σ_w^2.

Fig. 4.4: $\bar{N}_s/N$ when $\alpha = 0.5$, $s = 4$, and $N = 10$ in the OT-based system under mean-variance-shift attacks.

When we consider the case where the sensor observations follow an exponential distribution $f(y) = \frac{1}{\lambda}e^{-\frac{y}{\lambda}}$(a non-Gaussian distribution) with $\lambda = 2$ under hypothesis $\mathcal{H}_0$ and $\lambda = 8$ under hypothesis $\mathcal{H}_1$, we can observe a trend in Fig. 4.6 similar to that shown in Fig. 4.2 regarding the fraction of the number of transmissions saved as a function of D/s.

Effect of additive Byzantine attacks on $\bar{N}_s/N$ in the CEOT-based system: Figs. 4.7, and 4.8 show the effect of additive Byzantine attacks on the average percentage of savings for the CEOT-based system. Fig. 4.7 shows the UBs obtained in (4.49) and LBs obtained in (4.50) for the average percentage of saving $\bar{N}_{s,CEOT}/N$ as a function of D/s for different values of α. We observe that



(a) $\bar{N}_s/N$ as a function of D/s. (b) $\bar{N}_s/N$ as a function of σ_w^2.

Fig. 4.5: Benchmarking UBs and LBs when $\alpha = 0.5$, $s = 4$ and $N = 300$ in the OT-based system under the mean-variance-shift attacks.

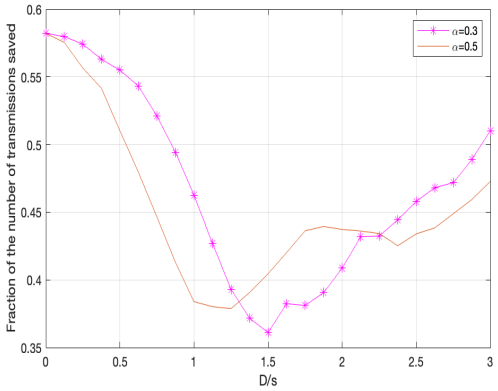


Fig. 4.6: $\bar{N}_s/N$ as a function of D/s in the OT-based system for different values α when $\lambda = 2$ under hypothesis $\mathcal{H}_0$, $\lambda = 8$ under hypothesis $\mathcal{H}_1$ for exponentially distributed observations and $N = 50$.

Byzantine sensors have more negative impact on the final decision making process with an increasing D/s. However, the additive Byzantine attacks have limited negative impact on the number of transmissions saved in the CEOT-based system compared to the OT-based system. When D/s is large enough, the first several local decisions received by the FC are most likely from Byzantine sensors which is the worst case scenario in terms of the performance for the system. With further increase of D/s, the impact of Byzantines on the number of transmissions saved in the network does not further increase since the LLRs are quantized, which limits the negative impact of Byzantine sensors on the system. In Fig. 4.8(a), the average percentage of saving, the UB and the LB are shown for a system with a relatively weak signal $s = 3$. Furthermore, Fig. 4.8(b) shows the plots

for a system with a relatively strong signal $s = 6$. By comparing Fig. 4.8(a) and Fig. 4.8(b), we observe that the LB gets tighter when we increase the signal strength s. This is reasonable due to the facts that i) Assumption 4.2 always works for honest sensors when s is sufficiently large; ii) the first several local decisions received by the FC are most likely from Byzantine sensors when D/s is sufficiently large. The above two reasons make the error probability of the CEOT-based system with a sufficiently large D/s approach the LB of the error probability we obtained in Theorem 4.4.

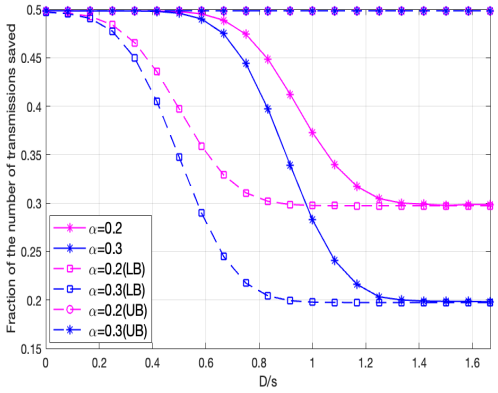


Fig. 4.7: Benchmarking UBs and LBs for $\bar{N}_{s,CEOT}/N$ as a function of D/s with different values of α when $s = 6$ and $N = 300$ in the CEOT-based system.

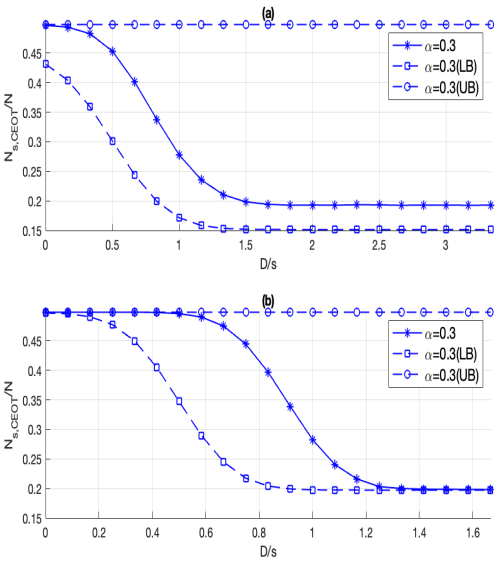


Fig. 4.8: Benchmarking UBs and LBs for $\bar{N}_{s,CEOT}/N$ as a function of D/s (a) when $s = 3$ and $N = 300$; (b) when $s = 6$ and $N = 300$ in the CEOT-based system.

CEOT-based Systems under DF-Byzantine Attack

We assume that $N = 100$ and $s = 20$. Fig. 4.9 shows the probability of error as a function of p in the CEOT-based system. The system with the threshold closest to the optimal threshold T^* (T^* is roughly $N/2$), as compared to other systems, has the lowest error probability, which is in accordance with the conclusion that we obtained about the optimal threshold T^*.[5] We can also observe that for the same parameter values, both CEOT-based and unordered systems have the same probabilities of error. This is in accordance with Lemma 4.2.

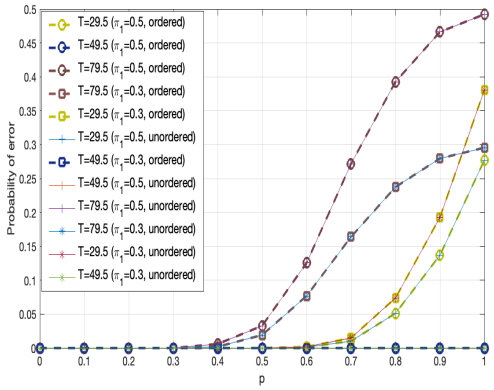


Fig. 4.9: P_e as a function of p with different values of T for the CEOT-based system for $\pi_1 = 0.3$ and $\pi_1 = 0.5$.

Fig. 4.10 shows that the UB we obtained is a relatively tight UB compared with the UB obtained in [88] for the average fraction of number of transmissions required as a function of the attacking probability p in the CEOT-based system. Fig. 4.11 presents the average fraction of transmissions required $N_{t,CEOT}/N$ in the CEOT-based system as a function of p for different values of prior probability and T when $\alpha = 0.3$. We observe from Fig. 4.11 that when $T \to T^*$ (T^* here is roughly $N/2$), the system is most likely to have the highest transmissions required in the network if the prior probabilities of both hypotheses are 0.5. However, when the prior probabilities change, the value of T that results in the highest transmissions required might also change. It is clear that a smaller T results in a larger average number of transmissions required to decide $\mathcal{H}_0$ and a smaller average number of transmissions required to decide $\mathcal{H}_1$. For a relatively small $\pi_1 < 0.5$,

[5]The threshold closest to the optimal threshold T^* is 49.5 in Fig. 4.9 when $\pi_1 = 0.5$ and $\pi_1 = 0.3$.

the probability of the FC deciding $\mathcal{H}_0$ is higher. Consequently, the system that uses $T = 29.5$ has higher number of transmissions required when compared to the system that uses $T = 49.5$ given $\pi_1 = 0.3$. Thus, there is a relationship between the average number of transmissions needed and the detection performance of the system. With an appropriately designed threshold at the FC, it is possible to save transmissions while still guaranteeing the quality of the decision. In Fig. 4.12, we plot $N_{t,CEOT}/N$ (i.e., the average number of transmissions required) as a function of s to show that a fairly small value of s is sufficient for the derived result to serve as an UB on the average number of transmissions required.

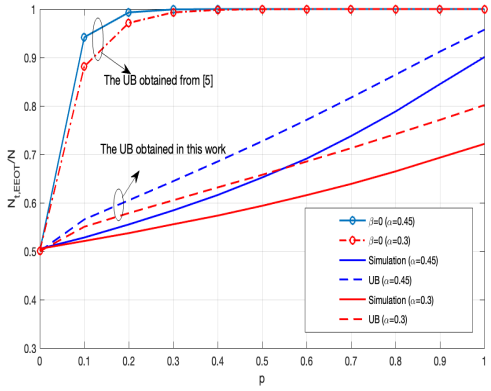


Fig. 4.10: Benchmarking upper bounds for the fraction of the number of transmissions required $N_{t,CEOT}/N$ as a function of p with different values of α when $\pi_1 = 0.5$. The actual average number of transmissions for the system (simulation result in the figure) is obtained via Monte Carlo method given $s = 20$.

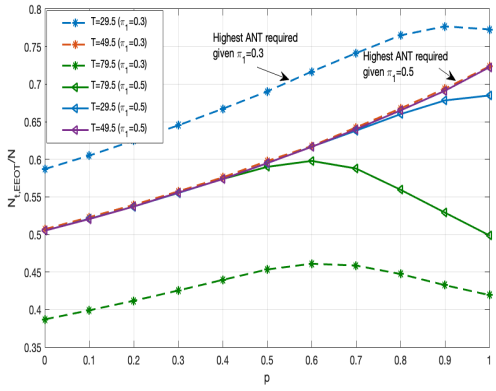


Fig. 4.11: $N_{t,CEOT}/N$ as a function of p with different values of T when $\alpha = 0.3$ and $\pi_1 = 0.3, 0.5$.

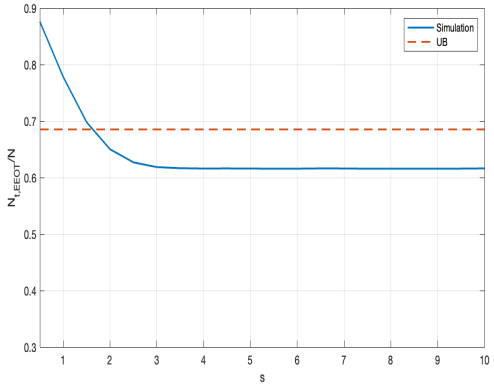


Fig. 4.12: $N_{t,CEOT}/N$ as a function of s given $p = 0.6$, $N = 100$ and $\pi_1 = 0.5$. Red dashed line indicates the UB of average number of transmissions we obtained given $p = 0.6$. Note that for $s > 1.6$, the UB we obtained serves as a valid UB.

4.6 Summary

In this chapter, the effect of Byzantine attacks on the performance of the conventional OT-based system and the CEOT-based system were investigated. We derived the error probability and the number of saved transmissions for OT-based systems under different Byzantine attacks. We also obtained some upper bounds and the lower bounds on the number of transmissions saved for the OT-based systems under different Byzantine attacks. The simulation results showed that the Byzantine nodes can both maximize the probability of error and significantly increase the number of transmissions needed to make the final decision when they adopt the optimal attacking strategy. A comparison of the robustness of CEOT-based and conventional OT-based systems was made, shedding light on how to employ OT-based frameworks in an attack-prone environment. Some possible countermeasures to mitigate the impact of Byzantines on OT-based systems were also discussed.

CHAPTER 5
DISTRIBUTED QUANTIZED DETECTION OF SPARSE SIGNALS UNDER BYZANTINE ATTACKS

In this chapter, we investigate distributed detection of sparse stochastic signals with quantized measurements under Byzantine attacks, where sensors may send falsified data to the FC to degrade system performance. Here, the Bernoulli-Gaussian (BG) distribution is used to model sparse stochastic signals. Several detectors with significantly improved detection performance are proposed by incorporating estimates of attack parameters into the detection process.

5.1 Introduction

With the development of compressive sensing (CS) [15, 20, 24, 118] in recent years, the sensors in sensor networks often send low-dimensional compressed measurements to the FC instead of high-dimensional sparse data, thereby improving bandwidth efficiency and reducing the communication overhead. A high-dimensional signal is sparse when only a few entries in the signal are non-zero, and others are zeros. Under the CS framework, the reconstruction and the detection of sparse

signals have received considerable attention. Here, we are interested in detecting compressed sparse signals.

5.1.1 Related Work

The problem of compressed sparse signal detection in sensor networks has been studied in the literature [21, 32, 55, 68, 100–102, 105, 116]. In these studies, the recovery of sparse signals was not necessarily required. In [21, 105, 116], partly or completely reconstructed sparse signals are required to derive the test statistics for sparse signal detection, while in [32, 55, 68, 100, 101], the test statistics are directly derived from compressed measurements to perform sparse signal detection. In [21] and [105], the authors proposed orthogonal matching pursuit algorithms to detect the presence of a sparse signal based on single measurement vectors and multiple measurement vectors, respectively, by estimating only a fraction of the support set of a sparse signal. In [32], the Bernoulli-Gaussian (BG) distribution was utilized to model the random sparsity of sparse signals, and the generalized likelihood ratio test (GLRT) was proposed to address the unknown degree of sparsity. Note that under the BG model (which is widely used to model the sparsity of signals [32, 52, 53, 86, 115]), the sparse signal has zero sparsity degree if the signal is absent, but a nonzero sparsity degree that approaches zero if the signal is present. Due to this property, parameter testing based on the sparsity degree can be employed for sparse signal detection by formulating the problem as a one-sided and close hypothesis testing problem. In [101], instead of GLRT, a method based on the locally most powerful test (LMPT), which is a popular tool for the problems of one-sided and close hypothesis testing, was proposed for detecting sparse signals in sensor networks. The test statistic of the LMPT detector was directly derived from the compressed measurements without any signal recovery. The detectors proposed in [21, 32, 101, 105] assume that the raw signals are transmitted within the network. However, due to limited bandwidth constraints in practical scenarios, it is necessary to consider the case where only quantized data is transmitted over sensor networks. To satisfy this requirement, many studies have been conducted on the design of sparse signal detectors based on quantized data [27, 55, 56, 68, 100, 102, 116].

A two-stage detector based on the GLRT, where sparse signal recovery is integrated into the detection framework, was proposed in [116] for sparse signal detection from 1-bit CS-based measurements. However, due to substantial information loss caused by 1-bit quantization, there is a noticeable performance gap compared to the clairvoyant detector based on analog measurements [27]. To address this issue, the authors in [102] proposed a quantized LMPT detector that enables the system to achieve detection performance comparable to a clairvoyant LMPT detector by selecting a reasonable number of reference sensors. The work was extended in [100] to consider generalized Gaussian noise. Additionally, the authors of [55] proposed an improved-1-bit LMPT detector that optimizes the quantization process and reduces the required number of sensor nodes to compensate for the performance loss caused by 1-bit quantization. The authors of [68] proposed a computationally-efficient generalized LMPT detector for the detection of distributed sparse signals when non-ideal reporting channels between the sensors and the FC are considered. In [56], the authors proposed an energy-efficient censoring-based LMPT detector in clustered sensor networks to address the excessively high energy consumption caused by data transmission in existing centralized LMPT detectors. In this scheme, the cluster head sensors and the ordinary sensors only transmit data that is sufficiently informative to the FC.

5.1.2 Major Contributions

Unlike previous proposed GLRT-based detectors [32, 116] and LMPT-based detectors [55, 68, 100, 101] that focused on attack-free environments, we investigate the impact of Byzantine attacks [25, 60, 64, 66, 82, 96, 110] on the detection performance, and enhance the resilience of the detectors. More specifically, we consider the GLRT-based and LMPT-based detectors with unknown random sparse signals operating under Byzantine attacks. The random unknown sparse signals are still characterized by the BG distribution as in [32, 53, 55, 68, 86, 100, 101, 116]. When such a system is under Byzantine attacks, two factors need to be taken into account: the unknown sparsity of the signal and the presence of unidentified attacks. We assume that the Byzantines do not have perfect knowledge about the actual state of the phenomenon of interest and attack based on their

local decisions, and we also assume that the system does not have perfect knowledge about the attack strategy. Under such assumptions, we evaluate the performance of the GLRT-based and the LMPT-based detectors. The simulation results show that the detectors are vulnerable to Byzantine attacks because their performance degrades.

To improve the resilience of the system in the presence of Byzantine attacks, it is intuitive that we need more information about the attack parameters. In this work, we develop a framework for estimating unknown parameters that are inspired by the works in [83, 84], where supervised machine learning was utilized as quality of transmission estimator for optical transport networks. In [84] and [83], a fraction of the total data is used to obtain a sufficiently accurate estimate of the unknown underlying parameters. Correspondingly, a subset of the sensors in this work is randomly selected, with their decisions serving as training samples for estimating the unknown attack parameters in the network. We introduce the notion of reference sensors to represent those sensors whose local decisions serve as training samples in our problem and propose the generalized likelihood ratio test with reference sensors (GLRTRS) and the locally most powerful test with reference sensors (LMPTRS) with adaptive thresholds, given that the sparsity degree and the attack parameter are unknown. The proposed detectors allow us to obtain excellent system performance. When the fraction of Byzantines in the networks is assumed to be known, we propose enhanced LMPTRS (E-LMPTRS) and enhanced GLRTRS (E-GLRTRS) detectors which can further improve the detection performance of the system. The main contributions of this work are summarized as follows.

- We perform a comprehensive performance analysis of existing GLRT-based and LMPT-based detectors in the presence of Byzantine attacks. Our analysis and simulation results reveal the degree to which both detectors are vulnerable to attacks.

- We propose a novel approach to design resilient GLRT and LMPT based detectors by considering the potential existence of adversarial Byzantine attacks. Specifically, we integrate the estimation of attack parameters into the detection process.

- Given that the sparsity degree and the attack parameters (i.e., the fraction of Byzantine nodes and the probability that Byzantines flip local decisions) are unknown, we propose GLRTRS and LMPTRS detectors with adaptive thresholds. Our simulation results indicate that both GLRTRS and LMPTRS detectors are resilient to Byzantine attacks. They can achieve detection performance close to that of the benchmark likelihood ratios test (LRT) detector, which has perfect knowledge of the sparsity degree and attack parameters.

- When the fraction of Byzantines in the networks is assumed to be known, we propose E-GLRTRS and E-LMPTRS detectors, which further improve the detection performance of the system by filtering out potential malicious sensors. Our simulation results show that the proposed enhanced detectors outperform LMPTRS and GLRTRS detectors.

5.2 System model

Consider the binary hypothesis testing problem of detecting sparse signals where hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$ indicate the presence and absence of the sparse signal, respectively. We consider a distributed network consisting of one FC and N sensors that observe the signals that share the joint sparsity pattern[1] as shown in Fig. 5.1. Let y_i be the received observation at sensor $i \in \{1, 2, \ldots, N\}$. We assume that all the observations are independent and identically distributed (i.i.d.) conditioned on the hypotheses. For sensor i, the observation y_i is modeled as

$$y_i = \begin{cases} n_i & \text{under } \mathcal{H}_0 \\ \mathbf{h_i}^T \mathbf{x_i} + n_i & \text{under } \mathcal{H}_1, \end{cases} \tag{5.1}$$

where $\mathbf{x_i} \in \Re^{M\times 1}$ is the sparse signal received by sensor i, $\mathbf{h_i} \in \Re^{M\times 1}$ is the channel gain of sensor i, which is modeled as a random vector to account for the variability and uncertainty in the

[1]Joint sparsity pattern indicates that non-zero elements of all the signals occur at the same locations, and the sparsity pattern is the same across all signals. This assumption of joint sparsity pattern can be readily observed in the field of compressed sensing, e.g., [17, 40, 73, 77].

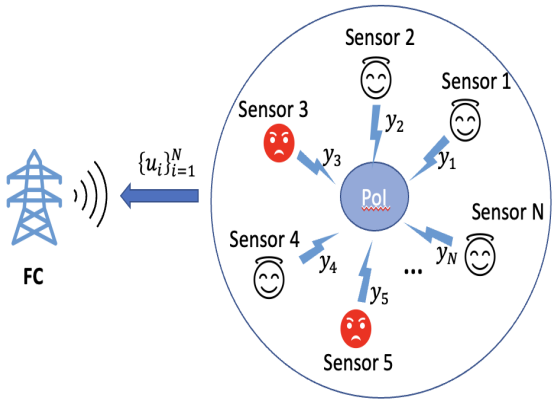


Fig. 5.1: System model of distributed sensor network. The red sensors are malicious.

communication channel, and n_i is Gaussian noise with zero mean and variance σ_n^2. Based on the received compressed measurements $\{y_i\}_{i=1}^N$ from all the sensors, the FC makes a global decision about the absence or presence of the sparse signals.

We adopt the BG distribution introduced in [32,53,55,68,86,100,101,116] to model the sparse signals where the joint sparsity pattern is shared among all the signals observed by the sensors. The locations of nonzero coefficients in x_i are assumed to be the same across all the sensors. Let $\mathbf{s} \in \Re^{M\times 1}$ describe the joint sparsity pattern of $\{\mathbf{x}_i\}_{i=1}^N$, where

$$\begin{cases} s_m = 1, & \text{for } \{x_{i,m} \neq 0, i = 1, 2, \ldots, N\} \\ s_m = 0, & \text{for } \{x_{i,m} = 0, i = 1, 2, \ldots, N\} \end{cases} \tag{5.2}$$

for $m = 1, 2, \ldots, M$. $\{s_m\}_{m=1}^M$ are assumed to be i.i.d. Bernoulli random variables with a common parameter p ($p \to 0^+$), where $P(s_m = 1) = p$ and $P(s_m = 0) = 1-p$. In other words, p represents the sparsity degree of the sparse signal $\mathbf{x}_i$ for $\forall i \in \{1, 2, \ldots, N\}$. Moreover, each element of $\mathbf{x}_i$ is assumed to follow an i.i.d. Gaussian distribution $\mathcal{N}(0, \sigma_x^2)$ [59]. Therefore, the BG distribution is imposed on $x_{i,m}$ as

$$x_{i,m} \sim p\mathcal{N}(0, \sigma_x^2) + (1-p)\delta(x_{i,m}), \tag{5.3}$$

where $\delta(\cdot)$ is the Dirac delta function. Due to the limited bandwidth, the sensors send their quantized observations instead of raw observations $\{y_i\}_{i=1}^{N}$ to the FC. We assume that a fraction α of the total N sensors, namely, αN sensors, are compromised by the Byzantines. We also assume that the compromised sensors are uniformly distributed in the network. In other words, a sensor i can be honest (H) with probability $1-\alpha$ or Byzantine (B) with probability α. The Byzantines may intentionally send falsified local decisions to the FC with an attack probability, i.e., the probability that Byzantines flip their decision. The fraction of Byzantines α and the probability that Byzantines flip their decision, P_A, are considered attack parameters. Note that the fusion rule is assumed not to be altered by Byzantine nodes.[2] Let $\mathbf{z_i}$ denote the actual quantized observation at sensor $i \in \{1, 2, \ldots, N\}$. The q-bit quantizer at the i^{th} sensor is defined as

$$\mathbf{z_i} = \begin{cases} \mathbf{v_1} & \tau_{i,0} \leq y_i \leq \tau_{i,1} \\ \mathbf{v_2} & \tau_{i,1} \leq y_i \leq \tau_{i,2} \\ \vdots & \vdots \\ \mathbf{v_{2^q}} & \tau_{i,2^q-1} \leq y_i \leq \tau_{i,2^q}, \end{cases} \tag{5.4}$$

where $\mathbf{v_k}$ is the binary code word with $\mathbf{v_k} \in \{0,1\}^q$ that represents the quantized observation and $\{\tau_{i,l}, l = 0, 1, 2, \ldots, 2^q\}$ are the quantization thresholds. For example, given $q = 2$, we have $\mathbf{v_1} = 00$, $\mathbf{v_2} = 01$, $\mathbf{v_3} = 10$ and $\mathbf{v_4} = 11$. Let $\mathbf{u_i}$ be the binary vector sent to the FC, which represents one of the possible quantizer observations $\{\mathbf{v_k} : k = 1, ...2^q\}$. $\mathbf{u_i}$ can also be interpreted as a (soft) decision. If sensor i is honest, we have $P(\mathbf{u_i} = \mathbf{z_i}|i = H) = 1$, otherwise we have $P(\mathbf{u_i} \neq \mathbf{z_i}|i = B) = P_A$. Here, the probability density function (PDF) of the local decision $\mathbf{u_i}$ if i is honest is given as

$$P(\mathbf{u_i}|i = H, \mathcal{H}_h) = P(\mathbf{z_i}|i = H, \mathcal{H}_h) = \prod_{j=1}^{2^q} P(\mathbf{z_i} = \mathbf{v_j}|i = H, \mathcal{H}_h)^{I(\mathbf{z_i}, \mathbf{v_i})} \tag{5.5}$$

[2]This assumption aligns with some related works such as [33, 66, 109, 112].

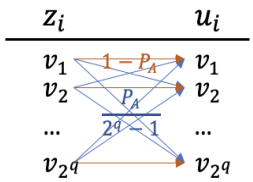


Fig. 5.2: Attack model for a Byzantine node i. With a probability of $P_A/(2^q-1)$, each Byzantine node decides to send a soft decision that differs from the one it believes to be correct. With probability $1-P_A$, the Byzantine nodes send the soft decision that they believe to be correct.

for $h=0,1$, where

$$P(\mathbf{z_i}=\mathbf{v_j}|i=H,\mathcal{H}_h)=P(\tau_{i,j-1}\leq y_i\leq\tau_{i,j}|i=H,H_h) \tag{5.6}$$

based on (5.4) and $I(\mathbf{z_i},\mathbf{v_i})$ is an indicator function that returns 1 if $\mathbf{z_i}$ is element-wise equal equal to $\mathbf{v_i}$ and returns 0 otherwise. In (5.5), we need to know the PDF of y_i, for $i=1,2,\ldots,N$. According to [106], both $y_i|\mathcal{H}_0$ and $y_i|\mathcal{H}_1$ follow Gaussian distributions as shown in (5.7), where $\beta_{i,0}^2=\sigma_n^2$, $\beta_{i,1}^2=\sigma_n^2+p\sigma_x^2||\mathbf{h}_i||_2^2$ and $b\overset{a}{\sim}f(b)$ means variable b asymptotically follows PDF $f(b)$.

$$y_i|\mathcal{H}_0\sim\mathcal{N}(0,\beta_{i,0}^2) \tag{5.7a}$$

$$y_i|\mathcal{H}_1\overset{a}{\sim}\mathcal{N}(0,\beta_{i,1}^2), \tag{5.7b}$$

The proof of (5.7b) is provided in [[106], Appendix B], where the Lyapounov Central Limit Theorem (CLT) is utilized to derive the results. Let $A_{i,j,h}$ represent the probability that y_i falls within the range of $[\tau_{i,j-1},\tau_{i,j}]$ when sensor i is honest under hypothesis $\mathcal{H}_h$, i.e., $P(\tau_{i,j-1}\leq y_i\leq\tau_{i,j}|i=H,\mathcal{H}_h)$. Then $A_{i,j,h}$ is given by

$$A_{i,j,h}=Q(\frac{\tau_{i,j-1}}{\beta_{i,h}})-Q(\frac{\tau_{i,j}}{\beta_{i,h}}) \tag{5.8}$$

for $h=0,1$, where $Q(\cdot)$ denotes the tail distribution function of the standard normal distribution. If sensor i is Byzantine, $\mathbf{u_i}$ does not have to be equal to $\mathbf{z_i}$. The attack model for Byzantine nodes is illustrated in Fig. 5.2. According to the chain rule, the PDF of local decision $\mathbf{u_i}$ is given as (5.11),

where

$$P(\mathbf{u_i} = \mathbf{v_j} | \mathbf{u_i} = \mathbf{z_i}, \mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) = \begin{cases} 1 & j = k \\ 0 & j \neq k, \end{cases} \tag{5.9}$$

$$P(\mathbf{u_i} = \mathbf{v_j} | \mathbf{u_i} \neq \mathbf{z_i}, \mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) = \begin{cases} 0 & j = k \\ \frac{1}{2^q - 1} & j \neq k, \end{cases} \tag{5.10}$$

$$\begin{aligned} P(\mathbf{u_i}|i = B, \mathcal{H}_h) &= \prod_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j}|i = B, \mathcal{H}_h)^{I(\mathbf{u_i}, \mathbf{v_j})} \\ &= \prod_{j=1}^{2^q} [\sum_{k=1}^{2^q} P(\mathbf{z_i} = \mathbf{v_k}|i = B, \mathcal{H}_h) P(\mathbf{u_i} = \mathbf{z_i}|\mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) \\ &\quad \times P(\mathbf{u_i} = \mathbf{v_j}|\mathbf{u_i} = \mathbf{z_i}, \mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) \\ &\quad + P(\mathbf{z_i} = \mathbf{v_k}|i = B, \mathcal{H}_h) P(\mathbf{u_i} \neq \mathbf{z_i}|\mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) \\ &\quad \times P(\mathbf{u_i} = \mathbf{v_j}|\mathbf{u_i} \neq \mathbf{z_i}, \mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h)]^{I(\mathbf{u_i}, \mathbf{v_j})} \end{aligned} \tag{5.11}$$

$P(\mathbf{u_i} \neq \mathbf{z_i}|\mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) = P_A$, $P(\mathbf{u_i} = \mathbf{z_i}|\mathbf{z_i} = \mathbf{v_k}, i = B, \mathcal{H}_h) = 1 - P_A$ and $P(\mathbf{z_i} = \mathbf{v_k}|i = B, \mathcal{H}_h) = Q(\frac{\tau_{i,k-1}}{\beta_{i,h}}) - Q(\frac{\tau_{i,k}}{\beta_{i,h}})$ for $h = 0, 1$. Note that (5.9) and (5.10) are equivalent to $I(i, k)$ and $\frac{1 - I(i,k)}{2^q - 1}$, respectively. Hence, (5.11) can be rewritten as

$$\begin{aligned} &P(\mathbf{u_i}|i = B, \mathcal{H}_h) \\ &= \prod_{j=1}^{2^q} \left\{ \sum_{k=1}^{2^q} A_{i,k,h} \left[(1 - P_A) I(j, k) + \frac{P_A(1 - I(i, k))}{2^q - 1} \right] \right\}^{I(\mathbf{u_i}, \mathbf{v_j})} \\ &= \prod_{j=1}^{2^q} \left\{ \sum_{k=1}^{2^q} A_{i,k,h} \left[(1 - P_A - \frac{P_A}{2^q - 1}) I(j, k) + \frac{P_A}{2^q - 1} \right] \right\}^{I(\mathbf{u_i}, \mathbf{v_j})} \\ &= \prod_{j=1}^{2^q} \left\{ A_{i,j,h}(1 - P_A) + \sum_{k=1, k \neq j}^{2^q} A_{i,k,h} \frac{P_A}{2^q - 1} \right\}^{I(\mathbf{u_i}, \mathbf{v_j})} \end{aligned}$$

$$= \prod_{j=1}^{2^q} \left\{ A_{i,j,h}(1-P_A) + (1-A_{i,j,h})\frac{P_A}{2^q-1} \right\}^{I(\mathbf{u_i},\mathbf{v_j})}$$
$$= \prod_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j} | i = B, \mathcal{H}_h)^{I(\mathbf{u_i},\mathbf{v_j})}. \tag{5.12}$$

Due to the statistical independence of the local decisions $\{u_1, u_2, \ldots, u_N\}$, we have

$$P(\mathbf{U}|\mathcal{H}_h) = \prod_{i=1}^{N}\prod_{j=1}^{2^q}\left[\sum_{X=B,H} P(\mathbf{u_i}=\mathbf{v_j}|i=X,\mathcal{H}_h)P(i=X)\right]^{I(\mathbf{u_i},\mathbf{v_j})} \tag{5.13}$$

for $h = 0, 1$.

5.3 GLRT and Quantized LMPT Detectors

In this section, we start with a brief review of the GLRT and the quantized LMPT detectors where all the sensors are assumed to be honest so that they send uncorrupted decisions to the FC, i.e., $\mathbf{u_i} = \mathbf{z_i}$. Then, the performance of the GLRT and the quantized LMPT detectors under Byzantine attacks is evaluated. The sparse signals here are characterized by the BG model. Under the BG model, the problem of distributed detection of sparse stochastic signals can be formulated as a problem of one-sided and close hypothesis testing which is given as

$$\begin{cases} \mathcal{H}_0: & p = 0 \\ \mathcal{H}_1: & p \to 0^+. \end{cases} \tag{5.14}$$

5.3.1 Fusion Rule for GLRT and Quantized LMPT Detectors with Honest Sensors

GLRT Detector

The fusion rule of the GLRT detector is given by

$$\frac{\max_p P(\mathbf{U}|\mathcal{H}_1;p)}{P(\mathbf{U}|\mathcal{H}_0;p=0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda', \tag{5.15}$$

We can obtain the estimated sparsity degree $\hat{p}$ via maximum-likelihood estimation (MLE) which is given as $\hat{p} = \arg\max_p P(\mathbf{U}|\mathcal{H}_1;p)$. By replacing p by $\hat{p}$ in (5.15) and taking the logarithm of both sides of (5.15), the fusion rule can be expressed as

$$\Gamma_{GLRT} = \sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{z_i}=\mathbf{v_j})G_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_1, \tag{5.16}$$

where $G_{i,j} = \hat{A}_{i,j,1} - \hat{A}_{i,j,0}$, $\hat{A}_{i,j,1} = Q(\frac{\tau_{i,j-1}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}}) - Q(\frac{\tau_{i,j}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}})$ and $\hat{A}_{i,j,0} = A_{i,j,0}$.

Quantized LMPT Detector

Since the sparsity degree p is positive and close to zero under $\mathcal{H}_1$, and $p = 0$ under $\mathcal{H}_0$, the problem of distributed detection of sparse stochastic signals can be performed via locally most powerful tests as shown in [100]. Firstly, the logarithm form of the LRT, which is given by

$$lnP(\mathbf{U}|\mathcal{H}_1;p) - lnP(\mathbf{U}|\mathcal{H}_0) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} ln(p_0/p_1), \tag{5.17}$$

is considered for decision-making at the FC, where $P(\mathbf{U}|\mathcal{H}_h) = \prod_{i=1}^{N} P(\mathbf{u_i}|\mathcal{H}_h, i = H)$ and $P(\mathcal{H}_h) = p_h$ for $h = 0, 1$. Due to the fact that the sparsity degree p is close to zero, the first-order Taylor's series expansion of $lnP(\mathbf{U}|\mathcal{H}_1;p)$ around zero is given as

$$lnP(\mathbf{U}|\mathcal{H}_1;p) = lnP(\mathbf{U}|\mathcal{H}_1;p=0) + p\left(\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p}\right)_{p=0}. \tag{5.18}$$

By substituting (5.18) in (5.17), the test statistic of the quantized LMPT detector is given by

$$\left(\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p}\right)_{p=0} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{ln(p_0/p_1)}{p} = \lambda_2, \tag{5.19}$$

where

$$\begin{aligned}\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1;p)}{\partial p} &= \sum_{i=1}^{N} \frac{\partial lnP(\mathbf{u_i}|\mathcal{H}_1, i=H;p)}{\partial p} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{2^q} w_{i,j} I(\mathbf{u_i} = \mathbf{v_j})\end{aligned} \tag{5.20}$$

and $w_{i,j} = \frac{\sigma_x^2||h_i||_2^2}{2\beta_{i,1}^3}\left[\tau_{i,j-1}\Phi(\frac{\tau_{i,j-1}}{\beta_{i,1}}) - \tau_{i,j}\Phi(\frac{\tau_{i,j}}{\beta_{i,1}})\right] A_{i,j,1}^{-1}$. Here, $\Phi(\cdot)$ denotes the CDF of the standard normal distribution. Hence, the decision rule is given as

$$\Gamma_{LMPT} = \sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j})\widetilde{w}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_2, \tag{5.21}$$

where $\widetilde{w}_{i,j} = (w_{i,j})_{p=0}$. Next, we evaluate the detection performance of the GLRT and the quantized LMPT detectors in the presence of Byzantines.

5.3.2 Performance Analysis of the GLRT and the Quantized LMPT Detectors in the Presence of Byzantines

Let $L = \sum_{i=1}^{N} L_i$ denote the global statistic for the fusion rule given in (5.16) or (5.21), where $L_i = \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j})d_{i,j}$ and $d_{i,j} \in \{\widetilde{w}_{i,j}, g_{i,j}\}$. According to the Lyapunov CLT, L approximately follows a Gaussian distribution with mean $E(\sum_{i=1}^{N} L_i)$ and variance $Var(\sum_{i=1}^{N} L_i)$ when N is sufficiently large. Under both hypotheses, $E(L)$ and $Var(L)$ are given as

$$\begin{aligned}E(L|\mathcal{H}_h) &= \sum_{i=1}^{N} E(L_i|\mathcal{H}_h) = \sum_{i=1}^{N} E\left(\sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j})d_{i,j}\right) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h)d_{i,j}\end{aligned}$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{2^q} [P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, i = H)(1-\alpha) + P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, i = B)\alpha] d_{i,j} \tag{5.22}$$

and

$$\begin{aligned} Var(L|\mathcal{H}_h) &= \sum_{i=1}^{N} Var(L_i|\mathcal{H}_h) = \sum_{i=1}^{N} \left[E\left(L_i^2|\mathcal{H}_h\right) - E(L_i|\mathcal{H}_h)^2\right] \\ &= \sum_{i=1}^{N} E\left[\left(\sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) d_{i,j}\right)^2\right] - E(L|\mathcal{H}_h)^2 \\ &= \sum_{i=1}^{N} \sum_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h) d_{i,j}^2 - E(L|\mathcal{H}_h)^2 \\ &= \sum_{i=1}^{N} \sum_{j=1}^{2^q} [P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, i = H)(1-\alpha) + P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, i = B)\alpha] d_{i,j}^2 - E(L|\mathcal{H}_h)^2, \end{aligned} \tag{5.23}$$

respectively. Using the expression in (5.22) and (5.23), the probabilities of detection and false alarm can be calculated as

$$P_d = P(L > \lambda|\mathcal{H}_1) = Q\left(\frac{\lambda - E(L|\mathcal{H}_1)}{\sqrt{Var(L|\mathcal{H}_1)}}\right) \tag{5.24}$$

and

$$P_f = P(L > \lambda|\mathcal{H}_0) = Q\left(\frac{\lambda - E(L|\mathcal{H}_0)}{\sqrt{Var(L|\mathcal{H}_0)}}\right), \tag{5.25}$$

respectively, where $\lambda \in \{\lambda_1, \lambda_2\}$.

Next, we investigate the optimal attack strategy that can be adopted by Byzantines. From the attackers' perspective, the optimal strategy is to render the system blind, aiming to achieve a probability of detection equal to 1/2. To determine the optimal attack strategy, we utilize the deflection coefficient, which provides a simple and yet effective measure of the global probability of detection. The deflection coefficient is defined as $D_f = \frac{(E(L|\mathcal{H}_1) - E(L|\mathcal{H}_0))^2}{Var(L|\mathcal{H}_1)}$. Thus, to blind the

FC, Byzantines need to strategically design the attack parameters so that $D_f = 0$, i.e., $E(L|\mathcal{H}_1) = E(L|\mathcal{H}_0)$. By utilizing (5.22), we can obtain

$$\alpha P_A = \frac{\sum_{i=1}^{N}\sum_{j=1}^{2^q}(A_{i,j,1}-A_{i,j,0})d_{i,j}}{\sum_{i=1}^{N}\sum_{j=1}^{2^q}\left[\frac{1}{2^q-1}+(1-\frac{1}{2^q-1})(A_{i,j,1}-A_{i,j,0})\right]d_{i,j}}. \tag{5.26}$$

When αP_A equals the right-hand side of (5.26), the attackers are able to blind the FC. From the simulation results presented later in this chapter, both the GLRT and the quantized LMPT detectors are very vulnerable to Byzantine attacks, even if the attack parameter P_A is very small. A possible explanation is that, since detectors make their decisions based on observations with the same mean and slightly different variances under the two hypotheses, it is easy for them to make incorrect decisions in the presence of Byzantines.

5.4 Resilient Detector under Byzantine Attack

In order to improve the resilience of the detector, we attempt to elicit some additional information regarding the attack parameters from the local decisions of some sensors and incorporate it into the design of the fusion rule. In general, a detector's performance improves as additional information is obtained, e.g., sparsity degree p, the fraction of Byzantines α, and attack probability P_A. Intuitively, a GLRT detector can be designed, which takes both the unknown sparsity degree and the unknown attack parameters into consideration, as shown in (5.27).

$$\frac{\max_{p,P_A,\alpha} P(\mathbf{U}|\mathcal{H}_1;p)}{\max_{P_A,\alpha} P(\mathbf{U}|\mathcal{H}_0;p=0)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda''. \tag{5.27}$$

If we assume that the sparse signals are weak and the number of sensors is large, the MLE attains its asymptotic PDF, and an appropriate threshold λ'' can be found based on the asymptotic detection performance of the GLRT detectors (see Sec. 6.5 in [51]). However, sparse signals need not be weak. In that case, it is not tractable to obtain an appropriate threshold value λ''. Moreover, the presence of nuisance parameters P_A and α results in a degradation of the detection performance of

GLRT detectors.

To overcome these problems, as alluded to earlier, we randomly select a fraction of the sensors as reference sensors from the set of all sensors and estimate unknown parameters (i.e., α, P_A and p) in two steps. In the first step, nuisance attack parameters are estimated based on the local decisions coming from the reference sensors. In the second step, the estimated attack parameters are utilized to estimate the unknown sparsity degree p based on the local decisions from the remaining sensors. The proposed GLRTRS detector is based on the above parameter estimates. As the LMPT-based detector does not require the knowledge of the sparsity degree p, the only estimation occurs in the first step, which is the estimation of the nuisance attack parameters. Later in this section, we will provide details about the proposed GLRTRS and LMPTRS detectors.

Since we carry out the entire estimation process in two steps, we would like to minimize the performance loss caused by partitioning the estimation process. Let us take the GLRT detector presented in (5.27) as an example. Suppose we want to partition the entire estimation process into two steps, as described above. In that case, we want to ensure that the performance degradation caused by the unknown sparsity degree p is negligible while estimating the attack parameters. In other words, the two pairs of estimated attack parameters we obtain, which are $\{\alpha_{H_1}, P_{A,H_1}\} = \arg\max_{\alpha,P_A} P(\mathbf{U}|\mathcal{H}_1, p, \alpha, P_A)$ and $\{\alpha_{H_0}, P_{A,H_0}\} = \arg\max_{\alpha,P_A} P(\mathbf{U}|\mathcal{H}_0, p = 0, \alpha, P_A)$, should be very close to each other. To complete this task, we introduce reference sensors to assist us. We randomly select a set of reference sensors from the set of all the sensors to estimate the unknown nuisance attack parameters P_A and α.[3] At the reference sensors, we employ different predefined thresholds so that the decisions of the reference sensors satisfy Assumption 5.1 below.

Assumption 5.1. *The probability $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_h)$ (or $Pr(\mathbf{z_i} = v_1|\mathcal{H}_h)$) is approximately equal to 1 for $h = 0, 1$.*

Note that the condition in Assumption 5.1 can be attained when reference sensors send v_{2^q} (or v_{2^1}) with a probability that is close to 1, regardless of the underlying true hypothesis $\mathcal{H}_h$. To satisfy

[3]Since we have assumed that α fraction of Byzantine nodes are uniformly distributed in the network, there are α fraction of Byzantine nodes within both the set of reference sensors and remaining sensors.

Assumption 5.1, one of the simplest methods is to either set $\widetilde{\tau}_{j,2^q-1} \ll \tau_{i,1}$ or $\tau_{i,2^q} \ll \widetilde{\tau}_{j,1}$. This is because the limit $\lim_{\widetilde{\tau}_{j,2^q-1}\to-\infty} Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_h) = 1$ (or $\lim_{\widetilde{\tau}_{j,1}\to+\infty} Pr(\mathbf{z_i} = v_1|\mathcal{H}_h) = 1$) always holds.[4] It allows us to ensure that the performance degradation caused by the unknown sparsity degree p is negligible while the attack parameters are being estimated.

In the following subsections, we consider two cases: (i) The sparsity degree p and the attack parameters $\{\alpha, P_A\}$ are all unknown; (ii) α is known, but sparsity degree p and attack probability P_A are unknown.

5.4.1 Networks with Unknown p, α and P_A

Two detectors are proposed in this subsection: the GLRTRS detector that requires the estimation of unknown parameter p, and the LMPTRS detector that does not require the estimation of p.

GLRTRS Detector

According to (5.13), we are able to obtain

$$P(\mathbf{U}|\mathcal{H}_h) = \prod_{i=1}^{N}\prod_{j=1}^{2^q}\left[A_{i,j,h} + x\left(\frac{1}{2^q-1} - A_{i,j,h} - \frac{A_{i,j,h}}{2^q-1}\right)\right]^{I(\mathbf{u_i},\mathbf{v_j})} \tag{5.28}$$

where $x = \alpha P_A$. For convenience, instead of considering the two attack parameters α and P_A separately, we consider a single attack parameter x. The problem of distributed detection of a sparse stochastic signal can be formulated as

$$\begin{cases} \mathcal{H}_0: & p = 0, 0 \le x \le 1 \\ \mathcal{H}_1: & p \to 0^+, 0 \le x \le 1 \end{cases}. \tag{5.29}$$

[4]Based upon (5.7), the observation y_i for $i \in \{1, 2, \ldots, N\}$ has zero mean and different variances that are related to sparsity degree p given different hypotheses. Since a sparse signal is considered for which the sparsity degree p tends to 0, it is possible to design reasonable quantizer thresholds for reference nodes. A reasonable quantizer threshold refers to a quantizer threshold that is not excessively large or small. From experiments, it has been shown that $\tau_{i,1} - \widetilde{\tau}_{j,2^q-1} = 6$ (or $\widetilde{\tau}_{j,1} - \tau_{i,2^q} = 6$) is sufficient to satisfy Assumption 5.1 for the reference sensors. Therefore, if Assumption 5.1 is satisfied, it is highly likely that the reference sensors will continue to send the same decision regardless of the true underlying hypothesis.

The fusion rule of the GLRTRS detector is given by

$$\frac{\max_p \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, \hat{x})}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p=0, \hat{x})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda, \tag{5.30}$$

where N_{ref} is the number of reference sensors and they are labelled as $1, 2, 3 \ldots, N_{ref}$. The estimate of the unknown attack parameter x, i.e., $\hat{x}$ is made via MLE based on the reference sensors data. Here, the estimated attack parameter x is given as

$$x_{H_h} = \arg\max_x P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) \tag{5.31}$$

for $h = 0, 1$. $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ in (5.31) is the joint pmf of local decisions coming from the reference sensors and it is given as

$$\begin{aligned} P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) &= \prod_{i=1}^{N_{ref}} \prod_{j=1}^{2^q} \left[\sum_{X=B,H} P(\mathbf{u_i}=\mathbf{v_j}|i=X, \mathcal{H}_h) P(i=X) \right]^{I(\mathbf{u_i},\mathbf{v_j})} \\ &= \prod_{i=1}^{N_{ref}} \prod_{j=1}^{2^q} \left[C_{i,j,h} + x \left(\frac{1}{2^q - 1} - C_{i,j,h} - \frac{1}{2^q - 1} C_{i,j,h} \right) \right]^{I(\mathbf{u_i}=\mathbf{v_j})} \end{aligned} \tag{5.32}$$

for $h = 0, 1$, where $C_{i,j,h} = Q(\frac{\widetilde{\tau}_{i,j-1}}{\beta_{i,h}}) - Q(\frac{\widetilde{\tau}_{i,j}}{\beta_{i,h}})$.

Note that if Assumption 5.1 holds and is employed at any q-bit quantizer of reference sensors, i.e., $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1) \approx Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0) \approx 1$ for any reference sensor i, the absolute value of $\widetilde{\tau}_{j,2^q-1}$ will be sufficiently large, and thus, the difference between the probabilities $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1)$ and $Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0)$ will be really small. Let $E_i = Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_1) - Pr(\mathbf{z_i} = v_{2^q}|\mathcal{H}_0)$ denote the difference between the probabilities of local decisions under $\mathcal{H}_1$ and $\mathcal{H}_0$ for any reference sensor i. According to Eq. (5.7), we have $E_i = Q(\frac{\widetilde{\tau}_{j,2^q-1}}{\beta_{i,1}}) - Q(\frac{\widetilde{\tau}_{j,2^q-1}}{\beta_{i,0}})$. The values of E_i as a function of $\widetilde{\tau}_{j,2^q-1}$ are shown in Fig. 5.3. We can observe that for a sufficiently large (or small) value of $\widetilde{\tau}_{j,2^q-1}$, for example, $\widetilde{\tau}_{j,2^q-1} = -6$, E becomes significantly small, with $E < 10^{-6}$.

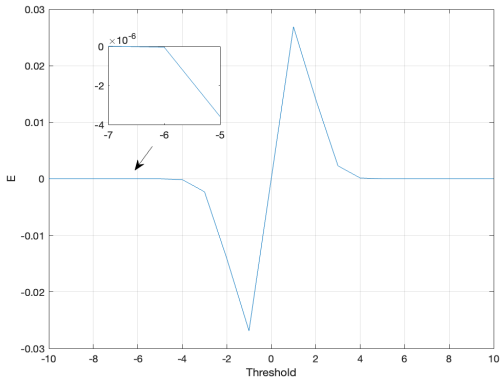


Fig. 5.3: E versus $\widetilde{\tau}_{j,2^q-1}$ given $p = 0.05$, $\sigma_x^2 = 5$, $\sigma_n^2 = 5$, $q = 1$ and $||\mathbf{h_i}||_2 = 1$ for all i.

Based on the above discussion, we can easily derive

$$P(u_i|\mathcal{H}_h, p, x) \approx (1-x)^{I(\mathbf{u_i}=\mathbf{v}_{2^q})} \prod_{j=1}^{2^q-1} \left(\frac{x}{2^q-1}\right)^{I(\mathbf{u_i}=\mathbf{v_j})} \tag{5.33}$$

for any reference sensor i. So the difference between the estimated x under different hypotheses will be significantly small and can be assumed negligible, i.e., $x_{H_0} \approx x_{H_1}$. This result is employed in the following theorem stating that the estimator considered in (5.31) is an efficient MLE when Assumption 5.1 is satisfied.

Theorem 5.1. *The MLE of the unknown attack parameter x based on the data from the reference sensors is unbiased, and it attains the Cramér–Rao lower bound (CRLB) of the problem, which equals $\frac{(1-x)x}{N_{ref}}$.*

PROOF: Please see Appendix A.8. ■

By replacing $\hat{x}$ by x_{H_1} in $P(\mathbf{u_i}|\mathcal{H}_1, p, x_{H_1})$ and $\hat{x}$ by x_{H_0} in $P(\mathbf{u_i}|\mathcal{H}_1, p=0, x_{H_0})$ in (5.30), the fusion rule can be reformulated as

$$\frac{\max_p \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_{H_1})}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p=0, x_{H_0})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa, \tag{5.34}$$

where $P(\mathbf{u_i}|\mathcal{H}_h, p, x_{H_h}) = \prod_{j=1}^{2^q} P(\mathbf{u_i} = \mathbf{v_j}|\mathcal{H}_h, p, x_{H_h})$. Since x_{H_0} is approximately the same as

x_{H_1}, i.e., $x_{H_0} \approx x_{H_1}$, choosing x_{H_0} or x_{H_1} as the estimated x under both hypotheses, or choosing the average of x_{H_0} and x_{H_1} as the estimated x under both hypotheses are all acceptable options. Here, we opt to replace both x_{H_1} and x_{H_0} in (5.34) with their averaged estimate $x_H = \frac{x_{H_1}+x_{H_0}}{2}$. The fusion rule then can be simplified as follows:

$$\frac{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_H)}{\prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_0, p=0, x_H)} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa, \tag{5.35}$$

where κ is the threshold to be set in order to ensure the desired probability of false alarm PFA. Next, we calculate the estimated sparsity degree $\hat{p}$, which is given as

$$\hat{p} = \arg\max_{p} \prod_{i=N_{ref}+1}^{N} P(\mathbf{u_i}|\mathcal{H}_1, p, x_H). \tag{5.36}$$

Upon taking the logarithm of both sides of (5.35), the simplified fusion rule is given as

$$\Gamma_{GLRTRS} = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) F_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa', \tag{5.37}$$

where $\kappa' = \log(\kappa)$, $F_{i,j} = f_{i,j,1} - f_{i,j,0}$, $f_{i,j,h} = \hat{A}_{i,j,h} + x_H \left(\frac{1}{2^q-1} - \hat{A}_{i,j,h} - \frac{1}{2^q-1}\hat{A}_{i,j,h} \right)$, $\hat{A}_{i,j,1} = Q(\frac{\tau_{i,j-1}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}}) - Q(\frac{\tau_{i,j}}{\sqrt{\sigma_n^2+\hat{p}\sigma_x^2}})$ and $\hat{A}_{i,j,0} = A_{i,j,0}$. Assume that $N - N_{Nef}$ is sufficiently large, the global statistic Γ_{GLRTRS} then follows a Gaussian distribution with mean

$$E(\Gamma_{GLRTRS}|H_h) = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} F_{i,j} P(\mathbf{u_i} = \mathbf{v_j}|H_h, x_H, p) \tag{5.38}$$

and variance

$$\begin{aligned} Var(\Gamma_{GLRTRS}|H_h) = &\sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} F_{i,j}^2 P(\mathbf{u_i} = \mathbf{v_j}|H_h, x_H, p) \\ &- E^2(\Gamma_{GLRTRS}|H_h) \end{aligned} \tag{5.39}$$

for $h = 0, 1$. With (5.38) and (5.39), the probabilities of detection and false alarm are respectively

given as

$$P_d = Q\left(\frac{\kappa' - E(\Gamma_{GLRTRS}|H_1)}{\sqrt{Var(\Gamma_{GLRTRS}|H_1)}}\right), \tag{5.40}$$

$$P_f = Q\left(\frac{\kappa' - E(\Gamma_{GLRTRS}|H_0)}{\sqrt{Var(\Gamma_{GLRTRS}|H_0)}}\right). \tag{5.41}$$

For a given false alarm PFA, we can obtain the suboptimal adaptive threshold used by the FC as shown in (5.42).[5]

$$\kappa' = Q^{-1}(PFA)\sqrt{Var(\Gamma_{GLRTRS}|H_0)} + E(\Gamma_{GLRTRS}|H_0) \tag{5.42}$$

LMPTRS Detector

Similarly, after we obtain the estimated attack parameter x_H, the test statistic of the proposed LMPTRS detector can be expressed as

$$\left(\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1, p, x_H)}{\partial p}\right)_{p=0} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{ln(p_0/p_1)}{p}, \tag{5.43}$$

where

$$\begin{aligned}
\frac{\partial lnP(\mathbf{U}|\mathcal{H}_1, p, x_H)}{\partial p} &= \sum_{i=1}^{N} \frac{\partial lnP(\mathbf{u_i}|\mathcal{H}_1, p, x_H)}{\partial p} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{2^q} \frac{\sigma_x^2||h_i||_2^2 I(\mathbf{u_i} = \mathbf{v_j})}{2(p\sigma_x^2||h_i||_2^2 + \sigma_n^2)^{\frac{3}{2}}} \left[\tau_{i,j-1}\Phi\left(\frac{\tau_{i,j-1}}{\sqrt{p\sigma_x^2||h_i||_2^2 + \sigma_n^2}}\right)\right. \\
&\quad \left. -\tau_{i,j}\Phi\left(\frac{\tau_{i,j}}{\sqrt{p\sigma_x^2||h_i||_2^2 + \sigma_n^2}}\right)\right] \frac{1 - x_H - x_H A_{i,j,1}}{A_{i,j,1} + x_H(1 - x_H - x_H A_{i,j,1})} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) g_{i,j}.
\end{aligned} \tag{5.44}$$

[5] Since we obtain the adaptive threshold based on the estimated attack parameter, it is a suboptimal threshold that approximately satisfies a desired false alarm.

The fusion rule can be reformulated as

$$\Gamma_{LMPTRS} = \sum_{i=1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) \widetilde{g}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma', \tag{5.45}$$

where $\gamma' = \frac{ln(p_0/p_1)}{p}$ and $\widetilde{g}_{i,j} = (g_{i,j})_{p=0}$. Like the one employed earlier, we can derive the threshold γ' in (5.45) for a given false alarm PFA. We can obtain that

$$\gamma' = Q^{-1}\left(PFA\right) \sqrt{Var(\Gamma_{LMPTRS}|H_0)} + E(\Gamma_{LMPTRS}|H_0), \tag{5.46}$$

where

$$E(\Gamma_{LMPTRS}|H_0) = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} \widetilde{w}_{i,j} P(\mathbf{u_i} = \mathbf{v_j}|H_0, x_H, p = 0) \tag{5.47}$$

and

$$Var(\Gamma_{LMPTRS}|H_0) = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} \widetilde{w}_{i,j}^2 P(\mathbf{u_i} = \mathbf{v_j}|H_0, x_H, p = 0) - E^2(\Gamma_{LMPTRS}|H_0). \tag{5.48}$$

5.4.2 Networks with Known α, Unknown p and Unknown P_A

When it is assumed that we know the fraction of Byzantine nodes α in the network, we can obtain more accurate information and achieve better detection performance. In this subsection, the GLRTRS and the LMPTRS detectors are further enhanced by introducing a local decision filter at the FC, which allows us to select sensors that are more likely to be honest. The proposed enhanced detectors are referred to as the E-GLRTRS and the E-LMPTRS detectors.

Upon receiving local decisions $\{\mathbf{U}(1), \ldots, \mathbf{U}(t)\}$ until time step t, where $\mathbf{U}(t) = \{\mathbf{u_1}(t), \ldots, \mathbf{u_N}(\mathbf{t})\}$, each sensor's statistical behavior is used to filter local decisions. The local decision filter distinguishes malicious nodes from honest nodes at time t by the following

$$\sum_{j=1}^{2^q} |R_j - \widetilde{p}_t(\mathbf{u_i} = \mathbf{v_j})| \underset{b_i(t)=0}{\overset{b_i(t)=1}{\gtrless}} \tau, \forall i \in \{N_{ref}+1, \ldots, N\}, \tag{5.49}$$

where $R_j = min(P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_1), P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_0))$ is a benchmark value to filter out the potential malicious sensors[6] and $b_i(t)$ represents the behavioral identity of sensor i at time t. If $b_i(t) = 1$, the sensor i is regarded as an honest node; otherwise, it is regarded as a potential Byzantine node. $\widetilde{p}_t(\mathbf{u_i} = \mathbf{v_j})$ is the empirical probability of $\mathbf{u_i} = \mathbf{v_j}$ until time step t according to the history of local decisions and it is given as

$$\widetilde{p}_t(\mathbf{u_i} = \mathbf{v_j}) = \frac{\sum_{q=1}^{t} I(\mathbf{u_i}(q), \mathbf{v_j})}{t}, \tag{5.50}$$

where $\mathbf{u_i}(q)$ is the $\mathbf{u_i}$ at time step q. The left side of (5.49) measures the deviation of the empirical probability of $\mathbf{u_i} = \mathbf{v_j}$ from the benchmark value R_j. Sensors are potential Byzantine nodes if the deviation exceeds a predefined threshold τ. Based on the behavioral identity of all the sensors $\{b_i(t)\}_{i=1}^{N}$ at time step t, we can obtain the fusion rules of enhanced detectors. Note that both GLRTRS and LMPTRS have the form

$$\sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} I(\mathbf{u_i} = \mathbf{v_j}) W_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta, \tag{5.51}$$

where $(W_{i,j}, \eta) \in \{(\widetilde{g}_{i,j}, \gamma'), (F_{i,j}, \kappa')\}$. Hence, the enhanced fusion rule at time step t is given by

$$\Gamma_E(t) = \sum_{i=N_{ref}+1}^{N} \sum_{j=1}^{2^q} b_i(t) I(\mathbf{u_i}(t) = \mathbf{v_j}) W_{i,j}(t) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \eta(t). \tag{5.52}$$

Let $\alpha_t(t)$ and $P_A(t)$ denote the probability that a sensor is a Byzantine node and the probability that a Byzantine node attacks at time step t, respectively, and let α be the initial value of α_t. We first obtain the estimated attack probability $\hat{p}_A(0) = x_H(0)/\alpha$ at time $t = 0$ as initial value of $\hat{P}_A$, where $x_H(0) = \frac{x_{H_1}(0) + x_{H_0}(0)}{2}$ and $x_{H_h}(0)$ is given in (5.31) for $h = 0, 1$. After filtering the possible Byzantine nodes, the value of α_t at time step $t = 0$ is updated according to $\{b_i(0)\}_{i=N_{ref}+1}^{N}$. The

[6]Note that based upon (5.7), the observation $y_i, \forall i \in \{1, 2, \ldots, N\}$ has zero mean and different variances that are related to the sparsity degree p given different hypotheses. Regardless of the quantizer thresholds that have been chosen, sensors tend to transmit the same decisions with slightly different probabilities based upon different hypotheses, i.e, $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_1)$ and $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_0)$ are slightly different. The simplest method of choosing R_j is to take the minimum value between $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_1)$ and $P(\mathbf{u_i} = \mathbf{v_j}|i = H, \mathcal{H}_0)$.

Table 5.1: Summary of GLRT-based and LMPT-based detectors under different scenarios.

unknown $\{P_A, \alpha, p\}$: GLRTRS: $\sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})F_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa'$ LMPTRS: $\sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})\widetilde{g}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma'$
known α and unknown $\{P_A, p\}$: E-GLRTRS: $\sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q} b_i(t)I(\mathbf{u_i}(t)=\mathbf{v_j})F_{i,j}(t) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa'(t)$ E-LMPTRS: $\sum_{i=N_{ref}+1}^{N}\sum_{j=1}^{2^q} b_i(t)I(\mathbf{u_i}(t)=\mathbf{v_j})\widetilde{g}_{i,j}(t) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \gamma'(t)$
commonly used GLRT-based detector: $\sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})G_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_1$
LMPT-based detector [100]: $\sum_{i=1}^{N}\sum_{j=1}^{2^q} I(\mathbf{u_i}=\mathbf{v_j})\widetilde{w}_{i,j} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \lambda_2$

updating rule is given as

$$\alpha_t(0) = \alpha - \frac{\sum_{i=N_{ref}+1}^{N} b_i(0)}{N - N_{ref}}. \tag{5.53}$$

At the next time step, the updated $\alpha_t(0)$ is employed as the new prior to estimate $\hat{p}_A(2)$ and $\hat{P}_A(1) = \frac{\sum_{i=0}^{1}\hat{p}_A(i)}{2}$. The value of α_t is also updated at time step $t = 1$ according to $\{b_i(1)\}_{i=N_{ref}+1}^{N}$ in the same manner as (5.53), i.e., $\alpha(1) = \alpha(0) - \frac{\sum_{i=N_{ref}+1}^{N} b_i(1)}{N-N_{ref}}$, and becomes the new prior at the next time step. Thus, at time step t, $\alpha_t(t-1) = \alpha_t(t-2) - \frac{\sum_{i=N_{ref}+1}^{N} b_i(t-1)}{N-N_{ref}}$ is utilized to obtain $\hat{P}_A(t) = \frac{\sum_{i=0}^{t}\hat{p}_A(i)}{t+1}$. By replacing x_H and $F_{i,j}$ with $X_H(t) = \hat{P}_A(t)\alpha_t(t-1)$ and $b_i(t)W_{i,j}$, respectively, in (5.38) and (5.39), we can obtain $E(\Gamma_E(t)|H_h)$ and $Var(\Gamma_E(t)|H_h)$. Similarly, for a given false alarm PFA, we can obtain the threshold used by the FC at time step t, which is given as $\eta(t) = Q^{-1}(PFA)\sqrt{Var(\Gamma_E(t)|H_0)} + E(\Gamma_E(t)|H_0)$. To compare the detectors over all of the scenarios we consider, we provide a summary table shown in Table 5.1.

5.5 Simulation results and Discussion

In this section, we present the simulation results to evaluate the performance of the proposed detectors in the presence of Byzantine attacks and compare them with the quantized LMPT-based detector (proposed in [100]) and the commonly used GLRT-based detector. Via simulations, we analyze the performance of the proposed schemes in terms of the probability of error in the system. The channel gains $\{\mathbf{h_i}\}_{i=1}^{N}$ are all assumed to be sampled from normal distribution with a homogeneous scenario so that $||\mathbf{h_i}||_2 = 1, \forall i$ as described in [100]. Table 5.2 presents the parameter settings for reference. Unless otherwise noted, we assume the number of sensors N to be 280. When reference sensors are employed, we employ $N_{ref} = 80$ out of 280 sensors as reference sensors, except when we evaluate system performance as a function of N_{ref}.

Table 5.2: Summary of parameter settings.

	N	N_{ref}	σ_n^2	σ_x^2	$\|\|\|\mathbf{h_i}\|\|_2$
value	280	80	1	5	1
	α	PFA	π_1	μ_w	p
value	0.3	0.4	0.5	0	0.05

In Fig. 5.4, we demonstrate the error probabilities of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$, the GLRT detector, and the proposed GLRTRS detector. Two different quantizers are employed, i.e., $q = 1$ and $q = 2$. The error probability of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$ shown in Fig. 5.4 is used as the benchmark to assess the performance of the proposed detectors. It can be observed that the GLRT detector is extremely vulnerable to attacks for both one-bit quantization and multilevel quantization, and a small fraction of Byzantine nodes α with a small attack parameter P_A are sufficient to break down the entire system. However, the proposed GLRTRS detector can obtain an error probability close to that of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$. We can observe from Fig. 5.4 that in the cases of $q = 1$ and $q = 2$, the GLRTRS detector outperforms the commonly used GLRT-based detector in the presence of attacks, with a performance close to the benchmark LRT detector. Note that the

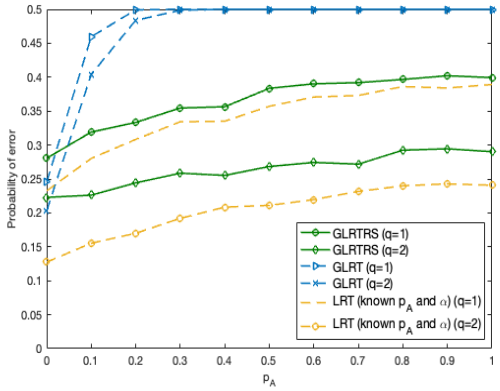


Fig. 5.4: Comparison of Pe for the GLRTRS, LRT and GLRT detectors.

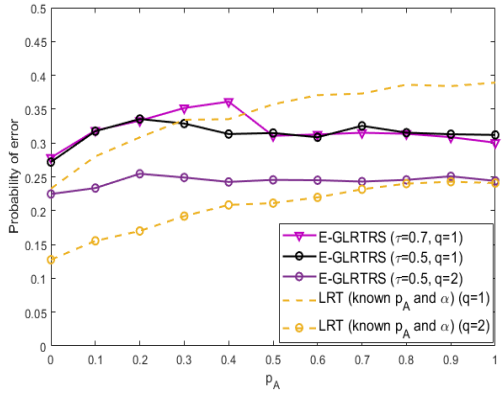


Fig. 5.5: Pe versus P_A when different values of q and the different values of threshold τ are utilized for the E-GLRTRS detectors.

GLRTRS detector uses only 200 sensors for detection purposes and exhibits performance close to the benchmark detector that uses 280 sensors for detection purposes. Hence, when no attacks are present, the commonly used GLRT-based detector performs slightly better. The number of quantization levels also affects the performance of the GLRTRS detector. As shown in Fig. 5.4, with an increase in q, the error probability of the proposed GLRTRS detector further decreases due to the reduction of performance losses caused by quantization. From Fig. 5.4, we can also observe that the difference between the benchmark error probability and the error probability of the proposed GLRTRS detector is larger when the value of q increases. It is because the GLRTRS detector is a sub-optimal detector, while the benchmark LRT detector is an optimal one.

If we assume that the fraction of Byzantine nodes α is known to the system, The error prob-

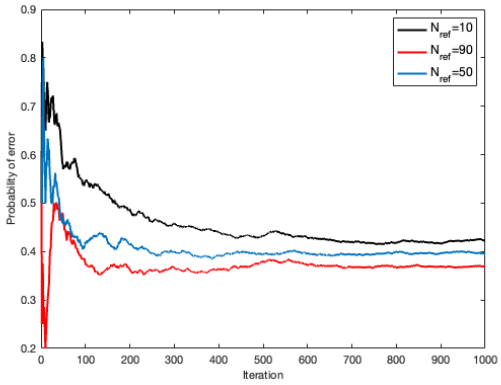


Fig. 5.6: Pe versus the number of iterations when different values of N_{ref} are utilized for the GLRTRS detector.

ability of the system can be further reduced by employing the E-GLRTRS detector. As shown in Fig. 5.5, the error probability of the E-GLRTRS detector decreases with an appropriately designed threshold τ compared to the GLRTRS detector. We can filter out different numbers of potential Byzantine nodes with different values of the threshold τ in (5.49). A potential Byzantine node can be either an actual Byzantine or a falsely identified one. It is obvious that a smaller threshold results in greater false filtering, while a larger threshold results in greater miss filtering. False filtering implies that honest nodes are falsely filtered out, whereas miss filtering implies that malicious nodes remain unfiltered. Both false filtering and miss filtering result in degrading the system's performance. Therefore, the system will likely perform better if the threshold τ is set appropriately. As shown in Fig. 5.5, $\tau = 0.5$ is more appropriate than $\tau = 0.7$. It can be observed that when $\tau = 0.5$, $q = 1$ and $P_A > 0.3$, the E-GLRTRS detector outperforms the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$. This is because the E-GLRTRS detector filters out potential Byzantine nodes and utilizes the rest of the sensors for detection. In contrast, the benchmark LRT detector utilizes all the sensors for detection purposes. Although the E-GLRTRS detector is inferior to the benchmark LRT detector when $q = 1$ and $P_A < 0.3$, the difference in error probabilities is not too significant.

In Fig. 5.6, the error probability and the convergence rate of the GLRTRS detector with different number of reference nodes are presented. The number of sensors used for detection purposes

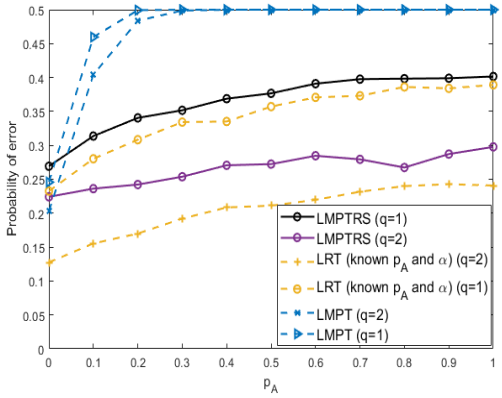


Fig. 5.7: Comparison of Pe for the LMPTRS, LRT and quantized LMPT detectors.

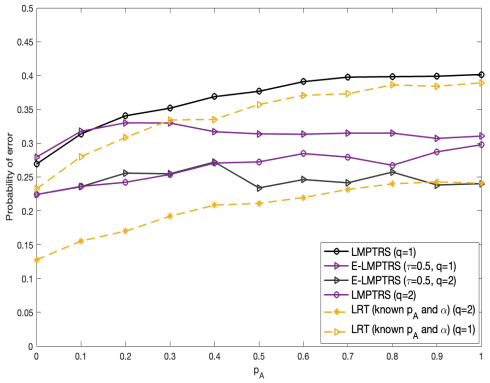


Fig. 5.8: Pe versus P_A when different values of q are utilized for the LMPTRS and the E-LMPTRS detectors.

in the GLRTRS detectors with different values of N_{ref} are equal to 200, i.e., $N - N_{ref} = 200$. It can be observed that the convergence rate is faster, and the error probability is lower when more reference nodes are used.

Fig. 5.7 shows the error probabilities of the LRT detector with perfect knowledge of $\{P_A, \alpha, p\}$, the quantized LMPT detector (proposed in [100]) and the proposed LMPTRS detector for $q = 1$ and $q = 2$, respectively. We can observe that the quantized LMPT detector proposed in [100] is also extremely vulnerable to attacks for both one-bit and multilevel quantization when all the p, P_A and α are unknown. However, it can be observed that when $q = 1$, the proposed LMPTRS detector is capable of obtaining an error probability close to the benchmark error probability that is obtained by employing the LRT detector with perfect knowledge of the attack parameters $\{P_A, \alpha, p\}$. Simi-

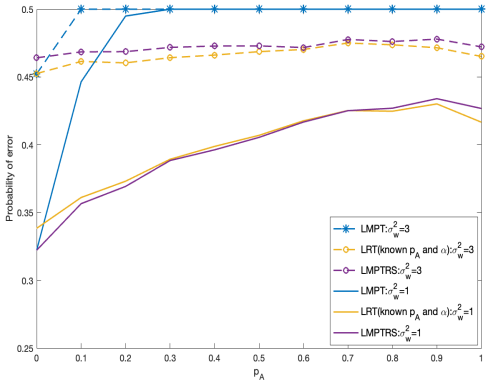


Fig. 5.9: P_e versus P_A for benchmark LRT, LMPT and LMPTRS detectors under Laplace distributed noise. The noise has a mean of $\mu_w = 0$ and a variance of σ_w^2 with probability of false alarm ($PFA = 0.4$). The sparse signals are assumed to asymptotically follow Gaussian distribution with mean 0 and variance $p\sigma_x^2||\mathbf{h}_i||_2^2$.

lar to the conclusion we obtained from Fig. 5.4, the LMPTRS detector outperforms the quantized LMPT detector proposed in [100] in the presence of attacks. The error probability of the proposed LMPTRS detector decreases with increasing q, and a higher value of q increases the difference between the benchmark error probability and the proposed LMPTRS detector error probability. It is also possible to further reduce the error probability of the system by assuming that the fraction of Byzantine nodes α is known to the system. As shown in Fig. 5.8, the E-LMPTRS detector outperforms both the quantized LMPT detector and the benchmark LRT detector with perfect knowledge of the attack parameters by filtering potential Byzantine nodes when $q = 1$. When q increases (e.g., $q = 2$), the E-LMPTRS detector still outperforms the quantized LMPT detector. In Fig. 5.9, we demonstrate the performance of our proposed detectors, which were originally designed for the simple Gaussian case, in the presence of one realization of generalized Gaussian noise. The noise here is assumed to follow the Laplace distribution, which is a special case of the generalized Gaussian distribution with parameter $\beta = 1$. We also note that according to [100], all types of generalized Gaussian distributed high-dimensional sparse signals asymptotically follow Gaussian distributions. We can observe that our proposed detector exhibits a certain level of resilience to the Byzantine attack when the tail of the distribution is not heavy.

5.6 Summary

The distributed detection problem of sparse stochastic signals with quantized measurements in the presence of Byzantine attacks was investigated. The sparse stochastic signals were characterized by their sparsity degrees, and the BG distribution was utilized to model sparsity. We proposed the LMPTRS and GLRTRS detectors with adaptive thresholds, given that the sparsity degree p and the attack parameters, i.e., α and P_A are unknown. The simulation results showed that the LMPTRS and GLRTRS detectors outperformed the LMPT detector under attack and achieved detection performance close to the benchmark LRT detector with perfect knowledge of the attack parameters and sparsity degree p. When the fraction of Byzantines α in the networks is assumed to be known, the E-LMPTRS and E-GLRTRS detectors were proposed to further improve the detection performance of the system by filtering out potential malicious sensors. Simulation results showed that the proposed enhanced detectors outperform LMPTRS and GLRTRS detectors.

CHAPTER 6
HUMAN-MACHINE HIERARCHICAL NETWORKS FOR DECISION MAKING UNDER BYZANTINE ATTACKS

In this chapter, we consider the human-machine collaborative decision-making networks in the present of Byzantine attacks. A belief-updating algorithm is proposed based on a hierarchical framework where local decisions from physical sensors act as reference decisions to improve the quality of human sensor decisions. The proposed algorithm effectively defends against Byzantine attacks, even when most physical sensors are malicious, significantly enhancing the performance of the human-machine collaborative system. The effect of available side information on the decision quality of individual human is also investigated.

6.1 Introduction

In high stake scenarios where human lives and assets are at risk, automatic physical sensor-only decision-making may not be sufficient [79, 87, 107]. Further, in some circumstances, such as remote sensing and emergency access systems, humans may possess additional side information

in addition to the common observations available from both physical sensors and humans. Thus, it may be necessary to incorporate humans in decision-making, intelligence gathering, and decision control. The emerging human-machine inference networks aim to combine humans' cognitive strength and sensors' sensing capabilities to improve system performance and enhance situational awareness.

Unlike physical sensors that can be programmed to operate with fixed parameters, human behavior and decisions are governed by psychological processes which are quite complex and uncertain. Hence, traditional signal processing and fusion schemes can not be adopted directly for integrating sensor measurements with human inputs. It is imperative to construct a framework to capture attributes associated with human-based sources of information so that they can be fused with data from physical sensors.

6.1.1 Related Work

There have been studies that employ statistical signal processing to address human-related factors in human-machine collaborative decision making. For instance, the authors of [104] studied decision fusion performance when the individual human agents use different thresholds modeled as random variables to make local decisions regarding a given phenomenon of interest (PoI). The authors in [79] proposed a hybrid system that consists of multiple human sub-populations, with the thresholds of each sub-population characterized by non-identically distributed random variables and a limited number of machines (physical sensors) whose exact values of thresholds are known. For such a hybrid system, they derived the asymptotic performance at the fusion center in terms of Chernoff information. The authors in [87, 107] showed that adding human inputs may or may not improve the overall performance of human-sensor networks, and they derived the conditions under which performance is improved. Furthermore, collaborative decision-making in multi-agent systems was investigated when the rationality of participating humans is modeled using prospect theory [28, 29]. To a large extent, the literature on human-machine collaborative networks has not considered the distributed nature and the openness of wireless networks in which the physical

sensors deployed in the network are low-cost, insecure, and vulnerable to various attacks, e.g., jamming, wiretap, spoofing [14, 26, 39] and Byzantine attacks [54, 119]. Here, we are interested in Byzantine attacks, where physical sensors in the network might be compromised and send falsified data to the FC.

6.1.2 Major Contributions

In contrast to most existing work, we aim to construct robust human-machine collaborative decision-making systems. We consider the general scenario where some sensors in the network are compromised by adversaries (Byzantines) so that they send falsified data to human agents. A belief updating and reputation-based scheme, where human agents and physical sensors interact with each other in decision-making, is proposed to mitigate the effect of Byzantine attacks. The proposed scheme consists of three parts: belief updating at human agents, decision-making at the FC, and reputation updating at the FC. In the belief updating part, the human agents make their local decisions based on their observations regarding the PoI and the decisions received from the physical sensors over a short time window. Within this short window, the human agents update their beliefs of the physical sensors' behavioral identities and further update their likelihood ratios (LRs) about the PoI. The belief updating phase involves collecting information from human agents and physical sensors to contribute to the decision-making at the FC. However, the belief-updating processes at the human agents based on short-term information, i.e., local decisions made by the sensors, may only reflect sensors' behavior over a short period. Consequently, the reputations of physical sensors are also updated over time at the FC to assist in the identification of Byzantine sensors and in mitigating their impact during the decision-making process. Moreover, we study under which conditions human agents can improve the quality of their decisions by using their side information if available. Our simulation results show that the proposed scheme can effectively defend against Byzantine attacks and enhance the quality of human agents' decisions.

6.2 System model

In this section, we consider a network model consisting of one FC, M human agents (human sensors), and N physical sensors, all of which make threshold-based binary decisions based on independent observations regarding the PoI. Unlike physical sensors, which employ deterministic thresholds, human sensors are assumed to use random thresholds to make decisions, which account for humans' cognitive biases. We also assume that the human agents have a similar background, e.g., culture, education level, and experience. [1] To account for the similar background they are assumed to have, it is reasonable to assume that a known probability distribution characterizes the random thresholds used by human sensors in this work. The thresholds used by the physical sensors are assumed to be the same and deterministic, which are $\boldsymbol{\tau} = [\tau_1, \ldots, \tau_N]^T$. The thresholds used by the human sensors are denoted by $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_M]^T$ and they are independent identically distributed (i.i.d.) random variables where ξ_i follows a probability density function (pdf) $f(\xi)$ for $i = 1, \ldots, M$. In this work, we assume that all the human sensors are honest and put in their best effort to make decisions. We also assume that a fraction α of the N physical sensors are Byzantine nodes and the FC is unaware of the identity of Byzantine nodes in the network. Hence, each sensor has the probability of α being a Byzantine node. As a result of the cognitive biases present in human sensors, some of them might perform worse than others when detecting the PoI. We utilize all useful information from the decisions coming from all the sensors (including physical and human sensors) in the network by employing a human-machine network that is constructed hierarchically, and a belief updating scheme is proposed.

[1] According to studies on human behavior [36, 41, 79, 108, 108], different backgrounds have profound effect on a person's decision-making process, the quality of decisions, as well as the ability to make decisions. To account for the diversity of human populations, we can assume that humans with different backgrounds use random thresholds to follow different distributions. In contrast, random thresholds used by humans with the same background follow the same distribution.

6.2.1 Belief-updating Scheme

The system model is shown in Fig. 6.1, where a hierarchical framework is established. All human agents are connected to a small set of physical sensors. Each human agent makes local decisions based on its raw observations and then updates its belief regarding the behavioral identity of the connected physical sensors and its LR based on the local decisions coming from the connected physical sensors' during the time interval $(nT, (n+1)T]$ for $n = 0, 1, \ldots$.

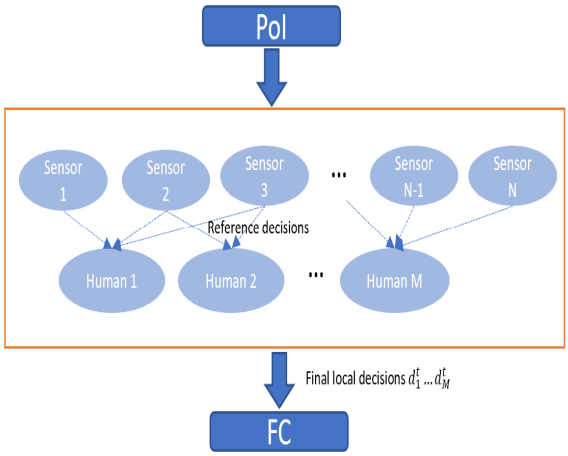


Fig. 6.1: System model

Remark 6.1. *Note that the behavioral identity and the LR are updated at $nT+1$, $nT+2$,... until $(n+1)T$. The human sensor uses this updated information to make a decision at time $(n+1)T$. In general, the human sensors collect information from the connected reference physical sensors to update the LR during $(nT, (n+1)T]$. We assume that T is not large so that the true underlying hypothesis does not change during $(nT, (n+1)T]$. At each time step $t = T, 2T, 3T, \ldots$, the final decisions made by the FC regarding the presence of the PoI are based on the decisions received from the humans.*

Let y_i^t and z_m^t denote the observation of sensor i and human m at time t, respectively.[2] The LR of physical sensor $i \in \{1, \ldots, N\}$ and human agent $m \in \{1, \ldots, M\}$ at time t are given as $L_{S,i}^t = \frac{f(y_i^t|\mathcal{H}_1)}{f(y_i^t|\mathcal{H}_0)}$ and $L_{H,m}^t = \frac{f(z_m^t|\mathcal{H}_1)}{f(z_m^t|\mathcal{H}_0)}$, respectively, where hypothesis $\mathcal{H}_1$ indicates the presence of

[2]The observations of both human and physical sensors are assumed to be of the same type and are i.i.d..

the PoI and hypothesis $\mathcal{H}_0$ indicates the absence of the PoI. Thus, the decision rule of the physical sensor i at time step t is given by

$$v_i^t = \begin{cases} 1 & L_{S,i}^t \geq \tau_i \\ 0 & otherwise \end{cases} \tag{6.1}$$

Here, we assume that identical physical sensors are deployed, and identical thresholds are utilized at the sensors. Therefore, we have $P_{d,i} = P_d$ and $P_{f,i} = P_f$ for $i = 1, 2, \ldots, N$. Let $u_{i,m}^t$ denote the local decision sent by physical sensor i to the connected human agent m at time t, where $m \in \mathcal{M}_m$. If sensor i is malicious, i.e., $i = B$, we assume that $u_{i,m}^t = 1 - v_i^t$; if it is honest, i.e., $i = H$, we assume $u_{i,m}^t = v_i^t$. The decision rule of the human agent m based on its raw observation at time step t is given by

$$b_m^t = \begin{cases} 1 & L_{H,m}^t \geq \xi_m \\ 0 & otherwise \end{cases} \tag{6.2}$$

Some key notations used in this chapter are listed in Table 6.1 for the convenience of readers. The belief-updating and decision-making process at each human sensor during any time interval $(nT, (n+1)T]$ proceeds as follows for $n = 0, 1, \ldots$:

Belief-updating

1. At time $t \in (nT, (n+1)T]$, $q_{m,i,t}$ and $r_{m,i,t}$ are updated, respectively, as

$$q_{m,i,t} = Pr(u_{i,m}^t = 1 | i_{t-1} = H) = \pi_{1,m,t-1} D_{i,H} + \pi_{0,m,t-1} F_{i,H} \tag{6.3}$$

and

$$r_{m,i,t} = Pr(u_{i,m}^t = 1 | i_{t-1} = B) = \pi_{1,m,t-1} D_{i,B} + \pi_{0,m,t-1} F_{i,B}, \tag{6.4}$$

Table 6.1: List of Notations Used

N	number of physical sensors
M	number of human sensors
$\mathcal{M}_m$	set that consists of physical sensors connected to human sensor m
$\mathcal{N}_i$	set that consists of human sensors connected to physical sensor i
α	fraction of Byzantines in the network
v_i^t	the actual local decision made by sensor i at time t
$u_{i,m}^t$	the local decision sent by physical sensor i to human agent m at time t
b_m^t	the local decision made by human sensor m at time t which is only based on its raw observations
d_m^t	the local decision made by human sensor m at time t when both the local decisions coming from connected physical sensors and b_m^t are utilized
$\pi_{h,m,t}$	probability that $\mathcal{H}_h$ is true at time t for human m
$\lambda_{m,t}$	LR at time t at human sensor m
$w_{m,i,t}$	belief that physical sensor i is honest at human sensor m at time t
$w_{m,i,t}(u_{i,m}^t = h)$	belief that physical sensor i is honest given $u_{i,m}^t = h$ at human sensor m at time t
$\delta_{m,i,t}$	LR based on decision coming from physical sensor i at human sensor m at time t
$\delta_{m,i,t}(u_{i,m}^t = h)$	LR given $u_{i,m}^t = h$ at human sensor m at time t
$r_{m,i,t}$	probability of $u_{i,m}^t = 1$ given sensor i is malicious
$q_{m,i,t}$	probability of $u_{i,m}^t = 1$ given sensor i is honest
$D_{i,H}$(or $D_{i,B}$)	probability of $u_{i,m}^t = 1$ given physical sensor i is honest (or malicious) and $\mathcal{H}_1$ is true
$F_{i,H}$(or $F_{i,B}$)	probability of $u_{i,m}^t = 1$ given physical sensor i is honest (or malicious) and $\mathcal{H}_0$ is true

for $i \in \mathcal{M}_m$, where $D_{i,X} = Pr(u_{i,m}^t = 1|\mathcal{H}_1, i = X) = \int_{\tau_i}^{\infty} Pr(y_i^t|\mathcal{H}_1, i = X)\mathrm{d}y_i^t$ and $F_{i,X} = Pr(u_{i,m}^t = 1|\mathcal{H}_0, i = X) = \int_{\tau_i}^{\infty} Pr(y_i^t|\mathcal{H}_0, i = X)\mathrm{d}y_i^t$ for $X = H$ or B. Based on (6.3) and (6.4), the belief that physical sensor i is honest is updated as

$$\begin{aligned} w_{m,i,t}(u_{i,m}^t = 1) =& \frac{Pr(i_{t-1} = H|u_{i,m}^t = 1)}{Pr(i_{t-1} = B|u_{i,m}^t = 1)} \\ =& \frac{Pr(u_{i,m}^t = 1|i_{t-1} = H)Pr(i_{t-1} = H)}{Pr(u_{i,m}^t = 1|i_{t-1} = B)Pr(i_{t-1} = B)} \\ =& w_{m,i,t-1}\frac{q_{m,i,t-1}}{r_{m,i,t-1}} \end{aligned} \tag{6.5}$$

given $u_{i,m}^t = 1$ and

$$w_{m,i,t}(u_{i,m}^t = 0) = \frac{Pr(i_{t-1} = H|u_{i,m}^t = 0)}{Pr(i_{t-1} = B|u_{i,m}^t = 0)} = w_{m,i,t-1}\frac{1 - q_{m,i,t-1}}{1 - r_{m,i,t-1}} \tag{6.6}$$

given $u_{i,m}^t = 0$ for $i \in \mathcal{M}_m$, where $Pr(i_{t-1} = B(\text{or } H))$ denotes the probability of sensor i being malicious (or honest) at time step t. Note that $Pr(i_{nT+1} = B) = \alpha$ and $Pr(i_{nT+1} = H) = 1 - \alpha$ for $n = 0, 1, 2 \ldots$. Hence, the initial belief is $w_{m,i,nT+1} = (1 - \alpha)/\alpha$ for $i = 0, 1, \ldots, N$. Given $u_{i,m}^t$, the belief that sensor i is honest at time t is $w_{m,i,t} = \frac{Pr(i_t = H)}{Pr(i_t = B)} = w_{m,i,t}(u_{i,m}^t = 1)^{u_{i,m}^t} w_{m,i,t}(u_{i,m}^t = 0)^{1-u_{i,m}^t}$.

2. For physical sensor i at time t, $\delta_{m,i,t}(u_{i,m}^t = h)$ is given by

$$\delta_{m,i,t}(u_{i,m}^t = 1) = \frac{D_{i,B} + D_{i,H} w_{m,i,t-1}}{F_{i,B} + F_{i,H} w_{m,i,t-1}} \tag{6.7}$$

for $h = 1$ and

$$\delta_{m,i,t}(u_{i,m}^t = 0) = \frac{(1 - D_{i,B}) + (1 - D_{i,H}) w_{m,i,t-1}}{(1 - F_{i,B}) + (1 - F_{i,H}) w_{m,i,t-1}} \tag{6.8}$$

for $h = 0$. Hence, given $u_{i,m}^t$, the LR at time t is $\delta_{m,i,t} = \delta_{m,i,t}(u_{i,m}^t = 0)^{1-u_{i,m}^t} \delta_{m,i,t}(u_{i,m}^t = 1)^{u_{i,m}^t}$.

3. *Decision-making at human sensor m at time $t = nT + 1$:*

$$\lambda_{m,nT+1} = \frac{\pi_1}{\pi_0} \frac{\beta_m^{b_m^{nT+1}} (1-\beta_m)^{1-b_m^{nT+1}}}{\gamma_m^{b_m^{nT+1}} (1-\gamma_m)^{1-b_m^{nT+1}}} \underset{d_m^{nT+1}=0}{\overset{d_m^{nT+1}=1}{\gtrless}} \kappa', \tag{6.9}$$

where $\beta_m = \int_{\xi_m}^{\infty} f(z_m^t|\mathcal{H}_1)\mathrm{d}z_m^t$, $\gamma_m = \int_{\xi_m}^{\infty} f(z_m^t|\mathcal{H}_0)\mathrm{d}z_m^t$ are the probabilities of detection and false alarm for human agent m.

Decision-making at human sensor m at time $t \in [nT + 2, (n + 1)T]$:

$$\lambda_{m,t} = \lambda_{m,t-1} \frac{\beta_m^{b_m^t} (1-\beta_m)^{1-b_m^t}}{\gamma_m^{b_m^t} (1-\gamma_m)^{1-b_m^t}} \prod_{j \in \mathcal{M}_m} \delta_{m,j,t} \underset{d_m^t=0}{\overset{d_m^t=1}{\gtrless}} \kappa' \tag{6.10}$$

Decision-making at FC

At the time $t = (n + 1)T$, the fusion rule at the FC is given as

$$\sum_{m=1}^{M} d_m^{(n+1)T} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \kappa, \tag{6.11}$$

where κ is the threshold used by the FC.

Reputation-updating at FC

At time $t = T, 2T \ldots$, the reputation of sensor i is given as $r_i^t = r_i^{t-T} + A_i^{t-T}$, where

$$A_i^{t-T} = \begin{cases} \Delta \frac{c_{i,t}}{|\mathcal{N}_i|} & c_{i,t} > c_{i,t}/2 \\ -\Delta(1 - \frac{c_{i,t}}{|\mathcal{N}_i|}) & otherwise. \end{cases} \tag{6.12}$$

$|\mathcal{N}_i|$ is the cardinality of $\mathcal{N}_i$, $c_{i,t}$=$\sum_{m \in \mathcal{N}_i} I(w_{m,i,t})$ where

$$I(w_{m,i,t}) = \begin{cases} 1 & w_{m,i,t} > 1 \\ -1 & otherwise \end{cases} \tag{6.13}$$

and Δ is the step size to update the reputation of each physical sensor. According to (6.12), a sensor's reputation increases if most human agents believe it is honest, and vice versa. The more human agents vote in favor of the same decision, the greater the increment in the reputation of sensors. When r_i^t is smaller than a threshold η, sensor i is identified as Byzantine and initial reputation is $r_i^0 = 1$ for $i = 0, 1, \ldots, N$.

6.2.2 Human Sensors with Side Information

Thus far, we have assumed that human and physical sensors only receive i.i.d. observations. In this subsection, we assume that human sensors may also possess side information[3] about the PoI other than the common features, which both physical and human sensors can observe. Assume that the human sensor m possesses the side information w_m^t related to the PoI in addition to the common attribute z_m^t for $m = 1, 2, \ldots, M$. To emulate the actions humans take to incorporate the data gathered from side information and observations into their decision-making process, two operations are employed in this work which are OR operation and AND operation [79].

OR Operation

The decision rule when using the OR operation to include side information is given by

$$e_m^t = \begin{cases} 1 & b_m^t = 1 \quad \text{or} \quad w_m^t = 1 \\ 0 & otherwise \end{cases} \tag{6.14}$$

where w_m^t is the side information indicating whether $\mathcal{H}_1$ is present or not and is assumed to be binary for human sensor m. The accuracy of side information is denoted as $Pr(w_m^t = 1|\mathcal{H}_1) = \beta_{m,side}$, and $Pr(w_m^t = 1|\mathcal{H}_0) = \gamma_{m,side}$. We assume that the side information $\{w_m^t\}_{m=1}^M$ is independent among different human sensors. Given the side information, the likelihoods of e_m^t given

[3]The side information refers to the additional information owned by human sensors which could come from previous professional experience or other sources.

$\mathcal{H}_1$ and $\mathcal{H}_0$ are shown, respectively, as [79]

$$f(e_m^t|\mathcal{H}_1)=\beta_{m,side}e_m^t+(1-\beta_{m,side})(\bar{\beta}^{e_m^t}(1-\bar{\beta})^{1-e_m^t}) \tag{6.15}$$

$$f(e_m^t|\mathcal{H}_0)=\gamma_{m,side}e_m^t+(1-\gamma_{m,side})(\bar{\gamma}^{e_m^t}(1-\bar{\gamma})^{1-e_m^t}), \tag{6.16}$$

where $\bar{\beta} = \int_{-\infty}^{\infty} f(\xi)Pr(b^t = 1|\mathcal{H}_1,\xi)\mathrm{d}\xi$ and $\bar{\gamma} = \int_{-\infty}^{\infty} f(\xi)Pr(b^t = 1|\mathcal{H}_0,\xi)\mathrm{d}\xi$ are the averaged probabilities of detection and false alarm for all human agents, respectively. Based on (6.15) and (6.16), we can derive the probability of detection $P_{d,m,side}^{OR}$ and probability of false alarm $P_{f,m,side}^{OR}$ for human sensor m that adopts OR operation, which are given, respectively, by $P_{d,m,side}^{OR} = \beta_{m,side} + (1 - \beta_{m,side})\bar{\beta}$ and $P_{f,m,side}^{OR} = \gamma_{m,side} + (1 - \gamma_{m,side})\bar{\gamma}$. Thus, the overall probability of error for human sensor m using OR operation becomes

$$P_{e,m,side}^{OR} = \pi_0 P_{f,m,side}^{OR} + \pi_1(1 - P_{d,m,side}^{OR}) \tag{6.17}$$

AND Operation

The decision rule when employing the AND operation to include the human observation and side information is expressed as

$$e_m^t = \begin{cases} 1 & b_m^t = 1 \text{ and } \quad w_m^t = 1 \\ 0 & otherwise \end{cases} \tag{6.18}$$

Given the side information, the likelihoods of e_m^t given $\mathcal{H}_1$ and $\mathcal{H}_0$ are expressed, respectively, as [79]

$$f(e_m^t|\mathcal{H}_1) = \beta_{m,side}\bar{\beta}^{e_m^t}(1 - \bar{\beta})^{1-e_m^t} + (1 - \beta_{m,side})(1 - e_m^t) \tag{6.19}$$

$$f(e_m^t|\mathcal{H}_0) = \gamma_{m,side}\bar{\gamma}^{e_m^t}(1 - \bar{\gamma})^{1-e_m^t} + (1 - \gamma_{m,side})(1 - e_m^t) \tag{6.20}$$

Based on (6.19) and (6.20), we can derive the probability of detection $P_{d,m,side}^{AND}$ and the probability of false alarm $P_{f,m,side}^{AND}$ for human sensor m that adopts AND operation, which are given,

respectively, by $P^{AND}_{d,m,side} = \beta_{m,side}\bar{\beta}$ and $P^{AND}_{f,m,side} = \gamma_{m,side}\bar{\gamma}$. Thus, the overall error probability for human sensor m using AND operation becomes

$$P^{AND}_{e,m,side} = \pi_0 P^{AND}_{f,m,side} + \pi_1(1 - P^{AND}_{d,m,side}) \tag{6.21}$$

Comparison between OR and AND Operation

As a result of the above analysis, the performance of each human sensor in terms of error probability is obtained when using the OR operation and when using the AND operation, i.e., $P^{OR}_{e,m,side}$ and $P^{AND}_{e,m,side}$. It is also easy to obtain the averaged error probability of each human sensor without obtaining any side information. It is given by

$$P_e = \pi_0\bar{\gamma} + \pi_1(1 - \bar{\beta}), \tag{6.22}$$

where π_h denotes the probability that $\mathcal{H}_h$ is true for $h = 0, 1$. Based on (6.22), (6.21) and (6.17), we can derive the conditions under which the quality of human sensors' decisions is improved by utilizing different operations to utilize side information. The derived conditions are stated in Theorem 6.1.

Theorem 6.1. *When the following conditions are satisfied, the side information could help individual human sensors make better decisions. For a specific human sensor $m \in \{1, \ldots, M\}$, we have*

- *the quality of decisions is improved by utilizing AND operation when $\frac{\bar{\beta}}{\bar{\gamma}} \leq \frac{\pi_0(1-\gamma_{m,side})}{\pi_1(1-\beta_{m,side})}$.*
- *the quality of decisions is improved by utilizing OR operation when $\frac{\pi_0(1-\bar{\gamma})}{\pi_1(1-\bar{\beta})} \leq \frac{\beta_{m,side}}{\gamma_{m,side}}$.*
- *OR operation performs better than AND operation when $\pi_0(1–2\bar{\gamma})\gamma_{m,side}–\pi_1(1–2\bar{\beta})\beta_{m,side} \leq \pi_1\bar{\beta} - \pi_0\bar{\gamma}$.*

PROOF: The above conditions can be derived by comparing the value of $P^{AND}_{e,m,side}$ and P_e, the value of $P^{OR}_{e,m,side}$ and P_e, and the value of $P^{AND}_{e,m,side}$ and $P^{OR}_{e,m,side}$. ■

6.3 Simulation results and Discussion

Some numerical results are presented in this section. Assume $y_i^t|\mathcal{H}_h \sim \mathcal{N}(\mu, \sigma_h^2)$ and $z_m^t|\mathcal{H}_h \sim \mathcal{N}(\mu_h, \sigma_h^2)$ for $h = 0, 1$, where $\mu_1 = 4$, $\mu_0 = 0$ and $\sigma_1^2 = \sigma_0^2 = 2$. The human thresholds are assumed to follow the Gaussian distribution with parameters (μ_τ, σ_τ), where $\mu_\tau = 2$ and $\sigma_\tau = 2$. We set $N = 60$, $M = 20$, $T = 10$, $\Delta = 0.03$, $\kappa' = 1$, $\kappa = M/2$, $\eta = 0.2$, $\tau_i = 2$ for $i = 1, \ldots, N$ and $\pi_1^{nT+1} = \pi_0^{nT+1} = 0.5$ for $n = 0, 1, \ldots$. Note that the term 'iterations' used in the following figures refers to iterations during the belief updating phase.

In Table 6.2, we show the comparison of the error probabilities of the systems that adopt CV (Chair-Varshney rule), MR (Majority rule), and MRH (Majority rule with human sensors) when α is known. MRH only utilizes the decisions from human sensors, while MR and CV utilize both human and physical sensors. The MR system uses decisions from all the sensors (including physical and human sensors) by performing a simple majority vote to make a final decision. In contrast, our proposed scheme employs a hierarchical framework to construct the human-machine collaborative network. As seen in Table 6.2, MR breaks down when most sensors participating in the decision-making process are malicious. However, our proposed scheme can still achieve comparable performance to the optimal CV rule. We can see in Fig. 6.2 that the fraction of humans making correct decisions increases significantly within a small number of iterations in our proposed scheme, which indicates a rapid improvement in the quality of humans' decisions.

Table 6.2: System error probability as a function of α

	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
CV	2.3e-7	3.1e-7	4e-7
MR	7e-5	0.07	0.996
MRH	2.5e-3	2.5e-3	2.5e-3
Proposed	2.7e-7	3.7e-7	4.3e-7

Although the proposed scheme requires the knowledge of α when each human sensor updates the belief regarding the behavioral identity of the corresponding physical sensors, this knowledge is not necessarily needed to guarantee a good performance. Choosing an appropriate predefined α can alleviate the performance degradation caused by the absence of knowledge of α. In Fig.

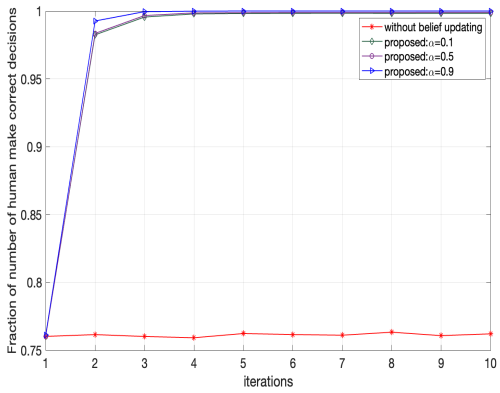


Fig. 6.2: Fraction of number of humans that make correct decisions versus the number of iterations when the system is aware of α

6.3 and Table 6.3, we compare the performance of the different systems when α is replaced with a predefined value α_e in (6.5) and (6.6). Fig. 6.3 shows the fraction of the number of humans that make correct decisions given different values of α_e. We can observe that the system with $\alpha = 0.9$ performs worse when $\alpha_e = 0.3$ and the system with $\alpha = 0.1$ performs worse when $\alpha_e = 0.7$. However, $\alpha_e = 0.5$ works well for the system with any fraction of Byzantine nodes. This is because when $\alpha_e \gg \alpha$ (or $\alpha_e \ll \alpha$), α_e significantly overestimates (or underestimates) α which results in performance degradation. Thus, $\alpha_e = 0.5$ is a good choice when we do not know α. Table 6.3 show that the system that adopts the proposed scheme can outperform the systems that adopt CV, MR, and MRH when α is unknown. Although there is a performance degradation compared to the systems that are aware of α, i.e., the performance shown in Fig. 6.2 and Table 6.2, the performance degradation is negligible for the proposed scheme. Thus, whether we know the actual α or not, the proposed scheme can always achieve a good performance. In Fig. 6.4, we show that the proposed scheme performs well in identifying Byzantine nodes in both cases, i.e., the system is aware/unaware of α.

Table 6.3: System error probability as a function of α given $\alpha_e = 0.5$

	$\alpha = 0.1$	$\alpha = 0.5$	$\alpha = 0.9$
CV	8e-4	3.1e-7	1.8e-3
MR	7e-5	0.07	0.996
MRH	2.5e-3	2.5e-3	2.5e-3
Proposed	1.7e-5	4.1e-7	3.3e-5

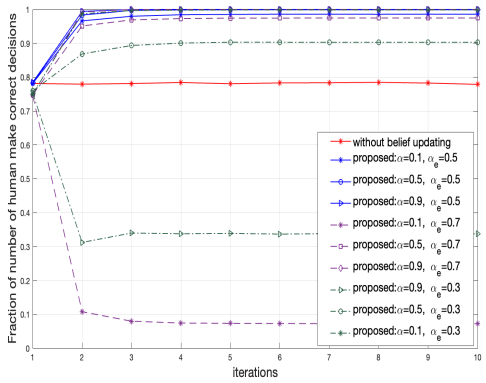


Fig. 6.3: Fraction of the number of humans that make correct decisions versus the number of iterations when the system does not know α.

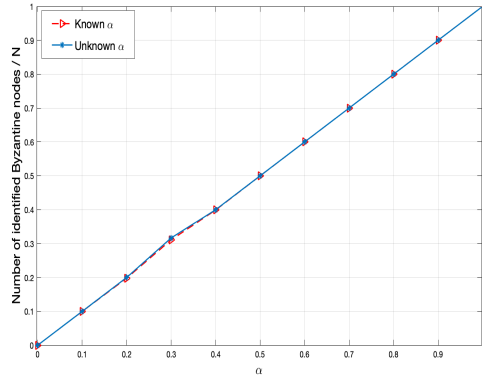


Fig. 6.4: The ratio of identified Byzantine nodes to the total number of sensors versus α for the proposed scheme when α is known and unknown. If α is unknown, we set $\alpha_e = 0.5$.

The impact of different operations while incorporating the individual performance of a human is illustrated in Fig. 6.5. The relationships among the error probability, the detection probability of the side information β_{side}, and the false alarm probability of the side information γ_{side} are shown. It can be observed that given certain values of γ_{side}, the error probability of a human sensor decreases as $\beta side$ increases for both OR and AND operations. Given certain parameters $(\beta_{side}, \gamma_{side})$, we can make a better choice among OR operation, AND operation, and no operation (i.e., no side information is utilized). For example, no operation is a better choice given $\beta_{side} \leq 0.81$ and $\gamma_{side} = 0.1$ and AND operation is a better choice given $\beta_{side} \geq 0.9$ and $\gamma_{side} = 0.3$. Our results shown in Fig. 6.5 are also consistent with Theorem 6.1 we obtained earlier.

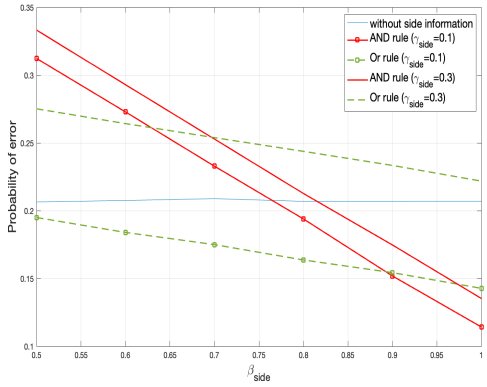


Fig. 6.5: The average probability of error versus β_{side} given different values of γ_{side} for any human sensor m without side information, as well as for any human sensor m that uses OR or AND operations.

6.4 Summary

In this chapter, we have proposed a belief-updating scheme in a human-machine hierarchical network. The local decisions from physical sensors served as reference decisions to improve the quality of human sensor decisions. At the same time, the belief that each physical sensor is malicious was updated during the decision-making process. The impact of side information from an individual human sensor and comparing different operations used to incorporate the side information were also analyzed. Simulation results showed that the quality of human sensors' decisions could be improved by employing the proposed scheme even when most physical sensors in the system are malicious. Moreover, our proposed scheme did not require the knowledge of the actual fraction of malicious physical sensors to guarantee the performance of our proposed scheme. Hence, the proposed scheme can successfully defend against Byzantine attacks and improve the quality of human sensors' decisions.

CHAPTER 7
CONCLUSION AND FUTURE DIRECTIONS

In this dissertation, we aimed to enhance the resilience of various energy-efficient WSNs for the inference task under Byzantine attacks. In Chapters 2 and 3, we conducted in-depth exploration, analysis, and enhancements to the audit bit-based mechanism. In Chapter 2, we evaluated the performance of the traditional audit bit-based mechanism under a more general and practical attack strategy. The results we obtained indicated that the attacker could blind the FC by adopting a very simple attack strategy. To overcome this problem, we proposed an enhanced audit bit-based mechanism, which relaxes the hard constraints on the attack strategies it can withstand. Our results showed that the proposed enhanced audit bit-based scheme outperforms traditional audit bit-based scheme. Building upon the enhanced audit bit framework, we proposed an advanced audit bit-based scheme that not only enhances system robustness but also significantly reduces the redundancy associated with audit bits.

In Chapter 3, we extended the work in Chapter 2 to tackle challenges in scenarios where prior knowledge of the attack strategies is unavailable. We proposed two algorithms to defend against Byzantine attacks in WSNs when the FC is not aware of the attacking strategy. The history of local decisions and idea of audit bit-based mechanism was utilized to update the reputation of sensors and help the system accurately identify Byzantine nodes. Our simulation results showed that we are able to achieve superior detection performance and the enhanced ability of identifying

Byzantine nodes by employing anchor nodes even when the Byzantines exceed half of the total number of sensors in the network.

In Chapter 4, we investigated the impact of Byzantine attacks on the performance of both the conventional OT-based system and the CEOT-based system. We derived the error probability and the number of saved transmissions for OT-based systems under different Byzantine attack scenarios. Additionally, we derived upper and lower bounds on the number of transmissions saved for OT-based systems under various Byzantine attack strategies. The simulation results revealed that Byzantine nodes, when employing optimal attack strategies, could both maximize the probability of error and significantly increase the number of transmissions required to reach a final decision. We also conducted a comparison of the robustness of CEOT-based and conventional OT-based systems, shedding light on how to implement OT-based frameworks in environments susceptible to attacks. Some possible countermeasures to mitigate the impact of Byzantines on OT-based systems were also discussed.

In Chapter 5, we investigated the distributed detection problem of sparse stochastic signals with quantized measurements in the presence of Byzantine attacks. We proposed two robust detectors based on traditional GLRT and LMPT detectors with adaptive thresholds, given that the sparsity degree and the attack strategy are unknown. The simulation results showed that the proposed detectors outperformed both LMPT and GLRT detectors under attack and achieved detection performance close to the benchmark LRT detector with perfect knowledge of the attack strategy and sparsity degree. When the fraction of Byzantines in the networks is assumed to be known, two enhanced detectors building on the previous proposed robust detectors were proposed to further improve the detection performance of the system by filtering out potential malicious sensors. Simulation results showed a good resilience of our proposed detectors.

In Chapter 6, we proposed a belief-updating scheme in a human-machine hierarchical network. The local decisions from physical sensors served as reference decisions to improve the quality of human sensor decisions. At the same time, the belief that each physical sensor is malicious was updated during the decision-making process. The simulation results showed that the proposed

scheme could enhance the performance of the system, even in scenarios where a majority of the physical sensors in the system are malicious, and where there is a lack of knowledge regarding the actual fraction of malicious physical sensors. Moreover, the impact of side information from an individual human sensor and comparing different operations used to incorporate the side information were also analyzed.

Next, we discuss some promising future directions of the work presented in this dissertation.

7.1 Suggestions for Future Research

Next we discuss some potential work and the future directions that should be pursued. The future works mainly focus on the design of resilient human-machine collaborative networks and resilient decentralized networks with quantized decisions.

7.1.1 Resilient Decentralized Networks with Quantized Decisions

In decentralized networks, there is no central server responsible for fusing data from each sensor. Instead, sensors transmit their measurements only to neighboring sensors, and a final decision for a detection problem is reached once a consensus is achieved. Decentralized networks find applications in various fields, including blockchain technology, IoT (Internet of Things) systems, distributed computing, and mesh networks. Previous studies, such as [44, 62, 69], have introduced consensus-based algorithms for addressing detection problems in decentralized networks, where raw observations are broadcasted.

Inspired by the works mentioned in [1, 49], we plan to design resilient decentralized networks where only binary decisions need to be broadcasted among sensors. We will evaluate such system's performance under a fixed network topology. Additionally, we aim to extend our research to address scenarios where sensors transmit M-ary decisions to their neighboring sensors. Furthermore, we will explore the design of resilient decentralized networks that can adapt to dynamically changing network topologies.

7.1.2 Resilient Human-machine Collaborative Networks

Building upon our previous work, which involved modeling human decision-making and decision fusion using random thresholds and Bayesian hierarchical models in Chapter 6, we now aim to expand our research in the following directions:

Human Decision's Uncertainty and Reliability in Human-machine Collaborative Networks

In the previous work, which was discussed in Chapter 6, we focused exclusively on scenarios where human agents were assumed to be honest while the physical sensors were considered potentially malicious. However, our future work will aim to expand this research to encompass situations where both human agents and physical sensors may be unreliable. The unreliable human agents refer to some lazy humans who are greedy and seek to gain monetary rewards without exerting any effort on their part. This type of behavior in humans is referred to as "no-effort attack" here, in which they earn money by making random guesses as part of their decision-making process. We will propose resilient algorithms for this kind of scenario. Our future work will focus on designing resilient algorithms to mitigate and adapt to these complex scenarios. Moreover, we plan to extend our previous work (6), which is a binary detection problem, to an M-ary detection problem.

Human limited Memory and Behavior Uncertainty

One characteristic of human agents is their limited processing capability due to their limited memory. When making decisions, these individuals perceive a quantized version of observations or provide a quantized evaluation of the information they process. We intend to explore the performance of such human-machine collaborative networks, taking into account this inherent feature of humans.

Furthermore, drawing inspiration from [93] which employed quantized priors to address bounded rationality among different human sub-populations in the decision-making process, we aim to

construct human-machine collaborative networks that involve human agents from various sub-populations by incorporating the concept of quantized priors.

7.1.3 Resilient Energy-efficient Networks

Another future work will involve an extension of our prior work. In the previous work, which was discussed in Chapter 4, our primary emphasis was on evaluating the performance of OT-based systems. However, there remains a need for resilient OT-based schemes. We are planning to design some resilient OT-based schemes. Furthermore, the OT-based framework discussed in Chapter 4 focused on the detection problem. The main idea is that only informative sensors are required to transmit their data to the FC. We are planning to leverage the idea of this framework to address estimation problems in energy-efficient WSNs.

APPENDIX A
APPENDIX: PROOFS OF VARIOUS RESULTS

A.1 Proof of Theorem 2.1, Chapter 2

Instead of directly analyzing the property of P_e^I in terms of p_2, we utilize Bhattacharyya distance $\mathcal{BD}$ as a surrogate to asymptotically characterize the detection performance of the system for simplicity. The relationship between Bhattacharyya distance and the probability of error P_e^I is $\lim_{N\to\infty} \frac{ln(P_e^I)}{N} \leq \mathcal{BD}$. For discrete probability distribution, $\mathcal{BD} = \sum_{\mathbf{u}\in\mathcal{U}} -ln\sqrt{P(\mathbf{u}|\mathcal{H}_1)P(\mathbf{u}|\mathcal{H}_0)}$, where $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{2^N}\}$ is the set of all the possible realizations of vector $\mathbf{u} = [u_1, u_2, \ldots, u_N]$. Let $f_i(u_i|i \in \underline{S}) = P(u_i|\mathcal{H}_1, i \in \underline{S})P(u_i|\mathcal{H}_0, i \in \underline{S})$ and $f_i(u_i|i \in \overline{S}) = P(u_i|\mathcal{H}_1, i \in \overline{S})P(u_i|\mathcal{H}_0, i \in \overline{S})$. Due to the fact that sensors independently send their local decisions, $\mathcal{BD}$

is given as

$$\begin{aligned}\mathcal{BD} &= \sum_{\mathbf{u}\in\mathcal{U}} -ln\sqrt{\prod_{i\in\underline{S}} f_i(u_i|i\in\underline{S})\prod_{i\in\overline{S}} f_i(u_i|i\in\overline{S})} \\ &= \sum_{\mathbf{u}\in\mathcal{U}} -ln\sqrt{\prod_{i=1}^{N}\mathcal{F}_i(u_i)} \\ &= \sum_{\mathbf{u}\in\mathcal{U}} -ln\sqrt{\prod_{i=1}^{N}\left(\sum_{d_i\in\mathcal{Q}}\mathcal{F}_i(u_i|d_i)P(d_i)\right)} \\ &= \sum_{\mathbf{u}\in\mathcal{U}} -ln\sqrt{\prod_{i=1}^{N}E_{d_i}\{\mathcal{F}_i(u_i|d_i)\}}\end{aligned} \tag{A.1}$$

where $\mathcal{Q}=\{0,1\}$, $\mathbf{d}=[d_1,d_2,\ldots,d_N]$ and $d_i\in\mathcal{Q}$. $\mathcal{F}_i(u_i|d_i)=(\overline{\pi}_{11}^{u_i}(1-\overline{\pi}_{11})^{1-u_i}\overline{\pi}_{10}^{u_i}(1-\overline{\pi}_{10})^{1-u_i})^{1-d_i}(\underline{\pi}_{11}^{u_i}(1-\underline{\pi}_{11})^{1-u_i}\underline{\pi}_{10}^{u_i}(1-\underline{\pi}_{10})^{1-u_i})^{d_i}$. $d_i=1$ indicates that the sensor i is placed in Set $\underline{S}$, otherwise, it is placed in Set $\overline{S}$. For sensor i, $E_{d_i}\{\mathcal{F}(u_i|d_i)\}$ is given as

$$\begin{aligned}&E_{d_i}\{\mathcal{F}(u_i|d_i)\} \\ &\quad= \sum_{q=0,1}\mathcal{F}(u_i|d_i=q)P(d_i=q) \\ &\quad= \overline{\pi}_{11}^{u_i}(1-\overline{\pi}_{11})^{1-u_i}\overline{\pi}_{10}^{u_i}(1-\overline{\pi}_{10})^{1-u_i}P(d_i=1) \\ &\qquad+\underline{\pi}_{11}^{u_i}(1-\underline{\pi}_{11})^{1-u_i}\underline{\pi}_{10}^{u_i}(1-\underline{\pi}_{10})^{1-u_i}P(d_i=0).\end{aligned} \tag{A.2}$$

We now have following two cases:

$u_i=1$ In this case, $E_{d_i}\{\mathcal{F}(u_i|d_i)\}=\overline{\pi}_{11}\overline{\pi}_{10}P(d_i=1)+\underline{\pi}_{11}\underline{\pi}_{10}P(d_i=0)$. We know that $P(d_i=1)+P(d_i=0)=1$ and $\underline{\alpha}^I\le\alpha_0\le\overline{\alpha}^I$. Let $h(t)=\pi_{11}\pi_{10}$ where $t=\alpha p_1$ is the random variable here. We can obtain $\frac{\partial^2 h(t)}{t^2}=2(1-2P_d)(1-2P_f)<0$. Hence, $h(t)$ is a concave function and has the property as following.

$$P(d_i=1)h(t_1)+P(d_i=0)h(t_2)\le h(P(d_i=1)t_1+P(d_i=0)t_2)=h(t_0) \tag{A.3}$$

where $t_1 = \overline{\alpha}^I p_1$, $t_2 = \underline{\alpha}^I p_1$ and $t_0 = \alpha_0 p_1$.

$u_i = 0$ In this case, $E_{d_i}\{\mathcal{F}(u_i|d_i)\} = (1-\overline{\pi}_{11})(1-\overline{\pi}_{10})P(d_i=1) + (1-\underline{\pi}_{11})(1-\underline{\pi}_{10})P(d_i = 0)$. Let $g(t) = (1-\pi_{11})(1-\pi_{10})$ where $t = \alpha p_1$ is the random variable here. We can obtain $\frac{\partial^2 g(t)}{t^2} = 2(1-2P_d)(1-2P_f) < 0$. Hence, $g(t)$ is also a concave function and follows the similar property as (A.3).

Note that we have $\underline{\pi}_{11} = \overline{\pi}_{11} = \pi_{11}$ and $\underline{\pi}_{10} = \overline{\pi}_{10} = \pi_{10}$ when $p_2 = 0$ according to Lemma 2.1. We can conclude that $E_{d_i}\{\mathcal{F}(u_i|d_i)\} \leq \mathcal{F}^0(u_i)$, where $\mathcal{F}^0(u_i) = \pi_{11}^{u_i}(1-\pi_{11})^{1-u_i}\pi_{10}^{u_i}(1-\pi_{10})^{1-u_i}$. We call the grouping in TAS with $p_2 = 0$ as non-effective grouping which is the same as the direct scheme, i.e., $\underline{\alpha}^I = \alpha_0 = \overline{\alpha}^I$, and the grouping in TAS with $p_2 \neq 0$ as effective grouping. According to (A.3), We show that the Bhattacharyya distance of the effective grouping is always larger than that of the non-effective grouping. According to the analysis above, the detection error $P_e^{(I)}$ can achieve the maximum value when $p_2 = 0$ given specific α_0, P_d, P_f and p_1. The probability of error for the system with direct scheme is

$$P_e^{(D)} = \pi_0 Q\left(\gamma_f^{(D)}\right) + \pi_1 Q\left(\gamma_m^{(D)}\right), \tag{A.4}$$

where $\gamma_f^{(D)}$ and $\gamma_m^{(D)}$ are expressed, respectively, as

$$\gamma_f^{(D)} = Q\left(\frac{\log(\frac{\pi_0}{\pi_1})/\sqrt{N} + \sqrt{N}D_0(\alpha_0, p)}{\sqrt{\pi_{10}(1-\pi_{10})W_d^2}}\right) \tag{A.5a}$$

$$\gamma_m^{(D)} = Q\left(\frac{\log(\frac{\pi_0}{\pi_1})/\sqrt{N} + \sqrt{N}D_1(\alpha_0, p)}{\sqrt{\pi_{11}(1-\pi_{11})W_d^2}}\right), \tag{A.5b}$$

and, $D_0(\alpha_0, p) = \pi_{10}\log(\frac{\pi_{10}}{\pi_{11}}) + (1-\pi_{10})\log(\frac{1-\pi_{10}}{1-\pi_{11}})$, $D_1(\alpha_0, p) = \pi_{11}\log(\frac{\pi_{11}}{\pi_{10}}) + (1-\pi_{11})\log(\frac{1-\pi_{11}}{1-\pi_{10}})$ and $W_d = \log(\frac{\pi_{11}(1-\pi_{10})}{\pi_{10}(1-\pi_{11})})$. Thus, for the non-effective grouping, according to (A.5), $D_0(\alpha_0, p) = 0$ can make the system be totally blind when N is large enough. We can easily obtain that $D_0(\alpha_0, p) = 0$ when $\alpha_0 p = \frac{1}{2}$.

A.2 Proof of Lemma 3.1, Chapter 3

We have the following four cases when we consider the MMS results of the sensors in the same group in Chapter 2. Let i and j represent the sensors in the same group.

- If $u_i = z_i$ and $u_j = z_j$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_1 =& P(i = B|u_i = z_i, u_j = z_j) \\ =& \frac{\alpha_0^2 f_{BB}^{(1)} + \alpha_0(1-\alpha_0) f_{BH}^{(1)}}{\alpha_0^2 f_{BB}^{(1)} + \alpha_0(1-\alpha_0)(f_{HB}^{(1)} + f_{BH}^{(1)}) + (1-\alpha_0)^2 f_{HH}^{(1)}},\end{aligned} \tag{A.6}$$

  where $f_{BB}^{(1)} = [2p_1p_2(1-p_1)+(1-2p_1+2p_1^2)(1-p_2)]^2$, $f_{HB}^{(1)} = f_{BH}^{(1)} = (1-p_2)(1-2p_1+2p_1^2)$ and $f_{HH}^{(1)} = 1$.

- If $u_i \neq z_i$ and $u_j = z_j$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_2 =& P(i = B|u_i \neq z_i, u_j = z_j) \\ =& \frac{\alpha_0^2 f_{BB}^{(2)} + \alpha_0(1-\alpha_0) f_{BH}^{(2)}}{\alpha_0^2 f_{BB}^{(2)} + \alpha_0(1-\alpha_0)(f_{HB}^{(2)} + f_{BH}^{(2)}) + (1-\alpha_0)^2 f_{HH}^{(2)}},\end{aligned} \tag{A.7}$$

  where $f_{BB}^{(2)} = [2p_1p_2(1-p_1) + (1-2p_1+2p_1^2)(1-p_2)][1-2p_1p_2(1-p_1) - (1-2p_1+2p_1^2)(1-p_2)]$, $f_{HB}^{(2)} = p_2(1-2p_1+2p_1^2)$,$f_{BH}^{(2)} = 2p_1(1-p_2)(1-p_1)$ and $f_{HH}^{(2)} = 0$.

- If $u_i = z_i$ and $u_j \neq z_j$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_3 =& P(i = B|u_i = z_i, u_j \neq z_j) \\ =& \frac{\alpha_0^2 f_{BB}^{(3)} + \alpha_0(1-\alpha_0) f_{BH}^{(3)}}{\alpha_0^2 f_{BB}^{(3)} + \alpha_0(1-\alpha_0)(f_{HB}^{(3)} + f_{BH}^{(3)}) + (1-\alpha_0)^2 f_{HH}^{(3)}},\end{aligned} \tag{A.8}$$

  where $f_{BB}^{(3)} = [2p_1p_2(1-p_1) + (1-2p_1+2p_1^2)(1-p_2)][1-2p_1p_2(1-p_1) - (1-2p_1+2p_1^2)(1-p_2)]$, $f_{HB}^{(3)} = 2p_1(1-p_2)(1-p_1)$,$f_{BH}^{(3)} = p_2(1-2p_1+2p_1^2)$ and $f_{HH}^{(3)} = 0$.

- If $u_i \neq z_i$ and $u_j \neq z_j$, i is a Byzantine node with probability

$$\begin{aligned}\alpha_4 =& P(i = B|u_i \neq z_i, u_j \neq z_j) \\ =& \frac{\alpha_0^2 f_{BB}^{(4)} + \alpha_0(1-\alpha_0) f_{BH}^{(4)}}{\alpha_0^2 f_{BB}^{(4)} + \alpha_0(1-\alpha_0)(f_{HB}^{(4)} + f_{BH}^{(4)}) + (1-\alpha_0)^2 f_{HH}^{(4)}},\end{aligned} \tag{A.9}$$

where $f_{BB}^{(4)} = [2p_1(1-p_2)(1-p_1) + p_2 p_1^2]^2$, $f_{HB}^{(4)} = f_{BH}^{(4)} = 2p_1p_2(1-p_1)$ and $f_{HH}^{(4)} = 0$.

According to the above results, we have

$$\underline{\alpha} = P(i = B|i \in \underline{\mathcal{T}}) = \alpha_1 \tag{A.10a}$$

$$\begin{aligned}\overline{\alpha} = P(i = B|i \in \overline{\mathcal{T}}) =& \alpha_2 P(u_i \neq z_i, u_j = z_j) + \alpha_3 P(u_i = z_i, u_j \neq z_j) \\ &+ \alpha_4 P(u_i \neq z_i, u_j \neq z_j)\end{aligned} \tag{A.10b}$$

The derivative of $\underline{\alpha}$ with respect to p_1 is given by (note that $f_{HB}^{(1)} = f_{BH}^{(1)}$)

$$\frac{\partial \underline{\alpha}}{\partial p_1} = \frac{\frac{\partial \mathcal{F}_1}{\partial p_1}\mathcal{F}_2 - \frac{\partial \mathcal{F}_1}{\partial p_1}\mathcal{F}_2}{\mathcal{F}_2^2} \tag{A.11}$$

where $\mathcal{F}_1 = \alpha_0^2 f_{BB}^{(1)} + \alpha_0(1-\alpha_0) f_{BH}^{(1)}$, $\mathcal{F}_2 = \alpha_0^2 f_{BB}^{(1)} + \alpha_0(1-\alpha_0)(f_{HB}^{(1)} + f_{BH}^{(1)}) + (1-\alpha_0)^2 f_{HH}^{(1)}$, $\frac{\partial \mathcal{F}_1}{\partial p_1} = \alpha_0^2 \frac{\partial f_{BB}^{(1)}}{\partial p_1} + \alpha_0(1-\alpha_0)\frac{\partial f_{BH}^{(1)}}{\partial p_1}$ and $\frac{\partial \mathcal{F}_2}{\partial p_1} = \alpha_0^2 \frac{\partial f_{BB}^{(1)}}{\partial p_1} + 2\alpha_0(1-\alpha_0)\frac{\partial f_{BH}^{(1)}}{\partial p_1} + (1-\alpha_0)^2 \frac{\partial f_{HH}^{(1)}}{\partial p_1}$. Let $\frac{\partial \underline{\alpha}}{\partial p_1} = 0$, we have

$$\alpha_0^2 \frac{\partial f_{BB}^{(1)}}{\partial p_1} f_{BH}^{(1)} + (1-\alpha_0)\frac{\partial f_{BB}^{(1)}}{\partial p_1} + (1-\alpha_0)^2 \frac{\partial f_{BH}^{(1)}}{\partial p_1} = \alpha_0^2 \frac{\partial f_{BH}^{(1)}}{\partial p_1} f_{BB}^{(1)} \tag{A.12}$$

due to the fact that $\frac{\partial f_{HH}^{(1)}}{\partial p_1} = 0$, where

$$\frac{\partial f_{BB}^{(1)}}{\partial p_1} = 8p_1(2p_2-1)^2(1-2p_1)(1-p_1) + 4p_1(2p_2-1)(1-2p_1)(1-p_2) \tag{A.13}$$

$$\frac{\partial f_{BH}^{(1)}}{\partial p_1} = 2(1-p_2)(2p_1-1). \tag{A.14}$$

We can easily obtain that $p_1 = \frac{1}{2}$ makes the equation (A.12) always hold for any specific $p_2 \in [0, 1]$. In other words, we can obtain that $p_1 = \frac{1}{2}$ could minimize or maximize $\underline{\alpha}$ given a specify p_2. Correspondingly, $p_1 = 0$ or 1 could also maximize or minimize $\underline{\alpha}$. Table A.1 shows all the possible maximum or minimum values of $\underline{\alpha}$ given a specific p_2. According to Table A.1, we have

Table A.1: Possible maximum or minimum values of $\underline{\alpha}$ given a specific p_2.

	$f_{BB}^{(1)}$	$f_{BH}^{(1)}$	$f_{HH}^{(1)}$
$p_1 = 0$	$(1-p_2)^2$	$1-p_2$	1
$p_1 = 1/2$	$1/4$	$(1-p_2)/2$	1
$p_1 = 1$	$(1-p_2)^2$	$1-p_2$	1

$$\underline{\alpha} = \frac{\alpha_0^2/4 + \alpha_0(1-\alpha_0)(1-p_2)/2}{\alpha_0^2/4 + \alpha_0(1-\alpha_0)(1-p_2) + (1-\alpha_0)^2} \tag{A.15}$$

for $p_1 = \frac{1}{2}$ and

$$\underline{\alpha} = \frac{\alpha_0^2(1-p_2)^2 + \alpha_0(1-\alpha_0)(1-p)}{\alpha_0^2(1-p_2)^2 + 2\alpha_0(1-\alpha_0)(1-p_2) + (1-\alpha_0)^2} \tag{A.16}$$

for $p_1 = 0$ or $p_1 = 1$. Let $h = \underline{\alpha} - \alpha_0$ represent the difference between $\underline{\alpha}$ and α_0, and it is given by

$$h = \frac{\alpha_0(1-\alpha_0)(5\alpha_0 - 4 + 2(1-2\alpha_0)(1-p_2))}{4(\alpha_0^2(1-p_2)^2 + 2\alpha_0(1-\alpha_0)(1-p_2) + (1-\alpha_0)^2)} \tag{A.17}$$

for $p_1 = \frac{1}{2}$, and it is given by

$$h = \frac{\alpha_0(1-\alpha_0)(\alpha_0 - 1 + (1-2\alpha_0)(1-p_2))}{\alpha_0^2/4 + \alpha_0(1-\alpha_0)(1-p_2) + (1-\alpha_0)^2} \tag{A.18}$$

for $p_1 = 0$ or $p_1 = 1$. Because we only care about the sign of h and the numerator is always positive. Let h_d denote the numerator of h. We have

$$\frac{\partial h_d}{\partial p_2} = \begin{cases} \frac{\alpha_0(1-\alpha_0)(2\alpha_0-1)}{2} & \text{,if } p_1 = \frac{1}{2} \\ -\alpha_0(1-\alpha_0)(1-2\alpha_0) & \text{,if } p_1 = 0 \text{ or } p_1 = 1 \end{cases} \tag{A.19}$$

We can easily obtained that $\frac{\partial h_d}{\partial p_2} \leq 0$ for $\alpha_0 \leq \frac{1}{2}$ and $\frac{\partial h_d}{\partial p_2} \geq 0$ for $\alpha_0 \geq \frac{1}{2}$ given $\forall p_1 \in \{0, \frac{1}{2}, 1\}$.

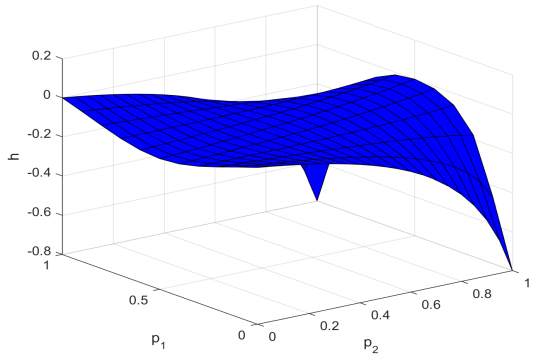


Fig. A.1: h versus p_1 and p_2 given $\alpha_0 = 0.8$.

According to (A.19), We can prove that $h_d \leq 0$ always holds for $\forall p_1 \in \{0, 1\}$. When $p_1 = \frac{1}{2}$, $h \leq 0$ also holds for $\forall \alpha_0 \in [0, \frac{1}{2}]$.

When $\alpha_0 > \frac{1}{2}$ and $p_1 = \frac{1}{2}$, the value of p_2 that guarantees $h \leq 0$ should be smaller than p_2^{max}. p_2^{max} can be obtained from letting $\frac{\partial h}{\partial p_2} = 0$ and it is given by

$$p_2^{max} = \min\left\{\frac{\alpha_0 - 2}{2(1 - 2\alpha_0)}, 1\right\} \tag{A.20}$$

for $\alpha_0 > \frac{1}{2}$ and $p_1 = \frac{1}{2}$. Note that $p_2^{max} = 1$ implies $\frac{\alpha_0 - 2}{2(1-2\alpha_0)} \geq 1$ such that $\alpha_0 \leq 0.8$. Hence, if the fraction of Byzantine nodes in the network is smaller than 0.8, i.e., $\alpha_0 \leq 0.8$, $\underline{\alpha}$ is always smaller than α_0. In other words, we always have a lower probability of existence of Byzantine nodes in set $\underline{\mathcal{T}}$ when $\alpha_0 \leq 0.8$. Fig. A.1 corroborates the results in (A.20).

Since $P(i \in \underline{\mathcal{T}})\underline{\alpha} + P(i \in \overline{\mathcal{T}})\overline{\alpha} = \alpha_0$, we have

$$\begin{aligned} P(i \in \underline{\mathcal{T}})\alpha_0 + P(i \in \overline{\mathcal{T}})\overline{\alpha} &\geq \alpha_0 \\ P(i \in \overline{\mathcal{T}})\overline{\alpha} &\geq \alpha_0(1 - P(i \in \underline{\mathcal{T}}) \\ \overline{\alpha} &\geq \alpha_0 \end{aligned} \tag{A.21}$$

when $p_2 \in [0, p_2^{max}]$, $\alpha_0 \in [\frac{1}{2}, 1]$ and $p_1 \in [0, 1]$. According to the above analysis, (A.21) also holds when $\alpha_0 \in [0, \frac{1}{2}]$, $p_1 \in [0, 1]$ and $p_2 \in [0, 1]$.

A.3 Proof of Lemma 4.1, Chapter 4

According to the fusion rule given in (4.3), we can infer that when inequality $\sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda$ holds, the FC can decide $\mathcal{H}_1$ based on the first k received transmissions. Similarly, when inequality $\sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda$ holds, the FC can decide $\mathcal{H}_0$ based on the first k received transmissions. The minimum value of k that satisfies either of the inequalities in (4.3), i.e., the minimum number of transmissions required to make a decision, is denoted as

$$k_{min} = \begin{cases} k_U^* & \text{when the FC decides } \mathcal{H}_0 \\ k_L^* & \text{when the FC decides } \mathcal{H}_1, \end{cases} \tag{A.22}$$

where

$$k_U^* = \underset{1\le k\le N}{\arg\min} \left\{ \sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda \right\} \tag{A.23}$$

and

$$k_L^* = \underset{1\le k\le N}{\arg\min} \left\{ \sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda \right\} \tag{A.24}$$

denote the minimum number of transmissions required to decide $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively.[1] Under $\mathcal{H}_0$ ($k_{min} = k_U^*$), we have

$$Z_U = \sum_{i=1}^{k_{min}} L_{[i]} + (N-k_{min})|L_{[k_{min}]}| \ge \sum_{i=1}^{N} L_{[i]} = Z, \tag{A.25}$$

and under $\mathcal{H}_1$ ($k_{min} = k_L^*$), we have

$$Z_L = \sum_{i=1}^{k_{min}} L_{[i]} - (N-k_{min})|L_{[k_{min}]}| \le \sum_{i=1}^{N} L_{[i]} = Z. \tag{A.26}$$

[1]Please note that if there is no $k \in \{1,2,\ldots,N\}$ that satisfies the condition $\sum_{i=1}^{k} L_{[i]} + (N-k)|L_{[k]}| < \lambda$ (or if there is no $k \in \{1,2,\ldots,N\}$ that satisfies $\sum_{i=1}^{k} L_{[i]} - (N-k)|L_{[k]}| > \lambda$), we define $k_U^* = \arg\min \emptyset = 0$ (or $k_L^* = \arg\min \emptyset = 0$).

This is because of the fact that $|L_{[1]}| > |L_{[2]}| > \cdots > |L_{[N]}|$). Note that $k_{min} = k_U^*$ is equivalent to $Z_U < \lambda$, and $k_{min} = k_L^*$ is equivalent to $Z_L > \lambda$. Based on (A.25) and (A.26), we can easily infer that $Pr(Z < \lambda | Z_U < \lambda) = 1$ and $Pr(Z > \lambda | Z_L > \lambda) = 1$, respectively. On the other hand, since

$$Z > \lambda \Leftrightarrow \sum_{i=1}^{k} L_{[i]} + \sum_{i=k+1}^{N} L_{[i]} > \lambda \tag{A.27a}$$

$$\Longrightarrow \sum_{i=1}^{k} L_{[i]} > \lambda - \sum_{i=k+1}^{N} L_{[i]} \geq \lambda - (N-k)|L_{[k]}| \tag{A.27b}$$

$$\Longrightarrow \sum_{i=1}^{k} L_{[i]} > \lambda - (N-k)|L_{[k]}| \tag{A.27c}$$

holds $\forall k$, from the definition of k_U^*, it becomes evident that the FC is unable to make a decision $\mathcal{H}_0$ for any value of k. So if $Z > \lambda$, we have $Pr(k_{min} = k_U^*) = 0$ and $Pr(k_{min} = k_L^*) = 1$, i.e., $Pr(Z_U < \lambda) = 0$ and $Pr(Z_L > \lambda) = 1$. It can be concluded that $Pr(Z_L > \lambda | Z > \lambda, \mathcal{H}_j) = 1$. Following a similar procedure, we can also obtain $Pr(Z_U < \lambda | Z < \lambda, \mathcal{H}_j) = 1$.

Based on the above analysis, we can calculate $Pr(Z_L > \lambda | \mathcal{H}_j)$ according to Bayesian rule given as

$$\begin{aligned} Pr(Z_L > \lambda | \mathcal{H}_j) &= \frac{Pr(Z_L > \lambda | Z > \lambda, \mathcal{H}_j) Pr(Z > \lambda | \mathcal{H}_j)}{Pr(Z > \lambda | Z_L > \lambda, \mathcal{H}_j)} \\ &= Pr(Z > \lambda | \mathcal{H}_j). \end{aligned} \tag{A.28}$$

Similarly, we obtain $Pr(Z_U < \lambda | \mathcal{H}_j) = Pr(Z < \lambda | \mathcal{H}_j)$. Hence, the probability of error of the OT-based system is given as

$$\begin{aligned} P_e^{(OT)} &= \pi_0 Pr(Z_L > \lambda | \mathcal{H}_0) + \pi_1 Pr(Z_U < \lambda | \mathcal{H}_1) \\ &= \pi_0 Pr(Z > \lambda | \mathcal{H}_0) + \pi_1 Pr(Z < \lambda | \mathcal{H}_1) = P_e^{(opt)}, \end{aligned} \tag{A.29}$$

where $P_e^{(opt)}$ is the error probability of the unordered system.

A.4 Proof of Theorem 4.1, Chapter 4

Let $\bar{N}_t$ denote the average number of transmissions in the network. $\bar{N}_t$ is given as

$$\bar{N}_t = E(k^*) = \sum_{k=1}^{N} kPr(k^* = k) = \sum_{k=1}^{N} Pr(k^* \geq k) \tag{A.30a}$$

$$= \sum_{k=1}^{N} Pr(k^* \geq k|\mathcal{H}_0)\pi_0 + Pr(k^* \geq k|\mathcal{H}_1)\pi_1, \tag{A.30b}$$

where $Pr(k^* \geq k)$ is the probability that at least k transmissions in the network are needed to make the final decision. Note that k^* is the minimum number of observations/transmissions required to make a decision. k can be considered as the number of observations that have already been received by the FC. The global statistic at the FC is given by $\sum_{i=1}^{k} L_{[i]}$, where $\sum_{i=1}^{k} L_{[i]}$ represents the accumulated LLRs up to the k^{th} transmission at the FC. Next Lemma helps us to obtain the probability of the event that at least k transmissions are required to make the final decision.

Lemma A.1. *The FC can not decide $\mathcal{H}_1$ or $\mathcal{H}_0$ until the FC has received at least k transmissions if $\sum_{i=1}^{k-1} L_{[i]}$ satisfies both $\sum_{i=1}^{k-1} L_{[i]} \leq \lambda+(N-k+1)|L_{[k-1]}|$ and $\sum_{i=1}^{k-1} L_{[i]} \geq \lambda-(N-k+1)|L_{[k-1]}|$.*

PROOF: When the FC received the first $(k-1)$ LLRs, i.e, $[L_{[1]}, L_{[2]}, \ldots, L_{[k-1]}]$, we discuss the cases that the FC can not decide $\mathcal{H}_1$ and the FC can not decide $\mathcal{H}_0$.

Recall that $|L_{[1]}| \geq |L_{[2]}| \cdots \geq |L_{[N]}|$, we have $Z \leq \sum_{i=1}^{k-1} L_{[i]} + (N-k+1)|L_{[k-1]}| = \eta_U$. Obviously, the FC is not able to decide $\mathcal{H}_0$ when $\eta_U > \lambda$. Moreover, (A.31) shows that if the FC doesn't decide $\mathcal{H}_0$ after receiving the first $(k-1)$ LLRs, it can't decide $\mathcal{H}_0$ after receiving the first $(k-2)$ observations.

$$\eta_U = \sum_{i=1}^{k-1} L_{[i]} + (N-k+1)|L_{[k-1]}| \tag{A.31a}$$

$$= \sum_{i=1}^{k-2} L_{[i]} + (N-k+2)|L_{[k-2]}| + (N-k+1)$$

$$\times (|L_{[k-1]}| - |L_{[k-2]}|) + (L_{[k-1]} - |L_{[k-2]}|). \tag{A.31b}$$

As $|L_{[k-1]}| \leq |L_{[k-2]}|$ and $L_{[k-1]} \leq |L_{[k-1]}| \leq |L_{[k-2]}|$, we have $|L_{[k-1]}| - |L_{[k-2]}| \leq 0$ and $L_{[k-1]} - |L_{[k-2]}| \leq 0$ in (A.31b). Hence, we can obtain that (A.31b)$> \lambda$ implies $\sum_{i=1}^{k-2} L_{[i]} + (N-k+2)|L_{[k-2]}| > \lambda$. Following the similar procedure as shown in (A.31), we are able to conclude that if the FC can't decide $\mathcal{H}_0$ after receiving the first $(k-1)$ LLRs, it can't decide $\mathcal{H}_0$ after receiving 0 or 1 or ..., or $(k-2)$ observations.

we can obtain that $\eta_L = \sum_{i=1}^{k-1} L_{[i]} - (N-k+1)|L_{[k-1]}| \leq Z$ after the FC has received the first $(k-1)$ LLRs. Obviously, the FC can not decide $\mathcal{H}_1$ when $\eta_L < \lambda$. Following the similar procedure as shown in (A.31), we can prove that if the FC can't decide $\mathcal{H}_1$ after receiving the first $(k-1)$ largest LLRs, it can't decide $\mathcal{H}_1$ after receiving 0 or 1 or ..., or $(k-2)$ observations. The proof for this is similar as above and is skipped. ■

To evaluate $Pr(k^* \geq k|\mathcal{H}_h)$, we have

$$Pr(k^* \geq k|\mathcal{H}_h) = \int_{\mathbf{l}_{k-1} \in \mathcal{J}} f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h) \mathrm{d}l_1 \ldots \mathrm{d}l_{k-1}, \tag{A.32}$$

where $f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h)$ is the joint pdf of $l_{[1]}, l_{[2]}, \ldots, l_{[k-1]}$ given $\mathcal{H}_h$ for $h = 0, 1$. According to [18], the joint pdf of $l_{[1]}, l_{[2]}, \ldots, l_{[k-1]}$ given $\mathcal{H}_h$ is given as

$$\begin{aligned} & f_{\mathbf{L_{[k-1]}}}(l_{[1]}, \ldots, l_{[k-1]}|\mathcal{H}_h) \\ =& \frac{N!}{(N-k+1)!} \left[\prod_{i=1}^{k-1} f_L(l_i|\mathcal{H}_h)\right] \left[F_{|L|}(l_{k-1}|\mathcal{H}_h)\right]^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}} \end{aligned} \tag{A.33}$$

where $\mathcal{J} = \mathcal{L} \bigcap \mathcal{U} \bigcap \mathcal{D}$ is the intersection of hyperplanes $\mathcal{L}, \mathcal{U}$ and $\mathcal{D}$, and $F_{|L_k|}(l_k|\mathcal{H}_h)$ is the cdf of $|L_k|$ for $h = 0, 1$. By substituting (A.33) in (A.32) and utilizing the law of total expectation, (A.32) can be rewritten as

$$Pr(k^* \geq k|\mathcal{H}_h) = E_{\mathbf{L}_{k-1}} \left[\frac{N!}{(N-k+1)!} \left[F_{|L|}(l_{k-1}|\mathcal{H}_h))\right]^{N-k+1} \mathbf{1}_{\{\mathcal{J}\}}\right] \tag{A.34}$$

for $h = 0, 1$, where $F_{|L|}(l_{k-1}|\mathcal{H}_h))$ is given in (4.12). Note that the Byzantines affect the average

number of transmissions by affecting attack parameters (α, D) in $F_{|L|}(L_{k-1}|\mathcal{H}_h)$.[2]

A.5 Proof of Theorem 4.2, Chapter 4

Let $\bar{N_s}$ denote the average number of transmissions saved in the network given as

$$\bar{N_s} = \sum_{k=1}^{N}(N-k)Pr(k^*=k) = \sum_{k=1}^{N-1} Pr(k^* \leq k) \tag{A.35a}$$

$$= \sum_{k=1}^{N-1} Pr(k^* \leq k|\mathcal{H}_0)\pi_0 + Pr(k^* \leq k|\mathcal{H}_1)\pi_1. \tag{A.35b}$$

Next, we use the following lemma from [18, Chapter 5] to prove Theorem 4.2.

Lemma A.2. *According to Cauchy-Schwarz inequality, we have*

$$|\sum c_i(L_{[i]} - \bar{L})| \leq [\sum (c_i - \bar{c})^2 (N-1)v]^{\frac{1}{2}} \tag{A.36}$$

in terms of empirical mean $\bar{L}$ and empirical variance v for any constants $\{c_i\}_{i=1}^N$. If c_i is non-increasing when i increases, the bound is sharp.

From Lemma A.2, we have $|\sum_{i=1}^{k} L_{[i]} - k\bar{L}| \leq [\sum (c_i - \bar{c})^2 (N-1)v]^{\frac{1}{2}}$ if we let $c_1 = c_2 = \cdots = c_k = 1$ and $c_{k+1} = \cdots = c_N = 0$. Hence, the LB and the UB of $\sum_{i=1}^{k} L_{[i]}$ are given by $g_L \leq \sum_{i=1}^{k} L_{[i]} \leq g_U$, where $g_L = -[\sum (c_i - \bar{c})^2 (N-1)v]^{\frac{1}{2}} + k\bar{L}$ and $g_U = [\sum (c_i - \bar{c})^2 (N-1)v]^{\frac{1}{2}} + k\bar{L}$.

LB of $\bar{N_s}$

When the FC decides $\mathcal{H}_1$ in at most k transmissions given hypothesis $\mathcal{H}_h$, we have

$$Pr(k^* \leq k|\mathcal{H}_h) = Pr(\sum_{i=1}^{k} L_{[k]} > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h). \tag{A.37}$$

[2] D affects η_1 and η_0 in $F_{|L|}(L_{k-1}|\mathcal{H}_h)$.

for $h = 0, 1$. It is easy to show that $g_L > \lambda + (N-k)|L_{[k]}|$ implies $\sum_{i=1}^{k} L_{[i]} > \lambda + (N-k)|L_{[k]}|$. Hence, from (A.37), we get

$$Pr(k^* \leq k|\mathcal{H}_h) \geq Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h) \tag{A.38}$$

Similarly, when the FC decides $\mathcal{H}_0$ in at most k transmissions given hypothesis $\mathcal{H}_h$, we get

$$Pr(k^* \leq k|\mathcal{H}_h) \geq Pr(g_U < \lambda - (N-k)|L_{[k]}||\mathcal{H}_h) \tag{A.39}$$

The inequality in (A.39) is true due to the fact that $g_U < \lambda - (N-k)|L_{[k]}|$ implies $\sum_{i=1}^{k} L_{[i]} < \lambda - (N-k)|L_{[k]}|$. Substituting $Pr(k^* \leq k|\mathcal{H}_0)$ and $Pr(k^* \leq k|\mathcal{H}_1)$ in (A.35) with their LBs $Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h)$ and $Pr(g_U < \lambda - (N-k)|L_{[k]}||\mathcal{H}_h)$, respectively, we get

$$\bar{N}_s \geq \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr(g_L > \lambda + n_{UT}|L_{[k]}||\mathcal{H}_h) + Pr(g_U < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h)\right] \tag{A.40}$$

where $n_{UT} = N - k$. A Monte Carlo approach can be utilized to evaluate $Pr(g_L > \lambda + (N-k)|L_{[k]}||\mathcal{H}_h)$ and $Pr(g_U > \lambda - (N-k)|L_{[k]}||\mathcal{H}_h)$. We generate M_2 realizations of $L_{[1]}, L_{[2]}, \ldots, L_{[N]}$ so that the empirical mean $\bar{L}$ and the empirical variance v can be calculated. When M_2 is sufficiently large, $\bar{L}$ approaches the population mean. The population mean and the population variance under $\mathcal{H}_h$ are, respectively, expressed as

$$\delta_h = E[L_i|\mathcal{H}_h] = \alpha\eta_h + (1-\alpha)\mu_h, \;\; \zeta_h^2 = E[L_i^2|\mathcal{H}_h] - \delta_h^2, \tag{A.41}$$

where $E[L_i^2|\mathcal{H}_h] = \alpha E[L_i^2|\mathcal{H}_h, i = B] + (1-\alpha)E[L_i^2|\mathcal{H}_h, i = H] = \beta + \alpha\eta_h^2 + (1-\alpha)\mu_h^2$ for $h = 0, 1$. Substituting the parameters $(\bar{L}, v)$ in (A.40) with parameters $(\delta_h, \frac{N}{N-1}\zeta_h^2)$ under $\mathcal{H}_h$ for $h = 0, 1$ yields

$$\bar{N}_s \geq \sum_{k=1}^{N-1}\sum_{h=0}^{1} \pi_h \left[Pr\left(|L_{[k]}| < \frac{g_L - \lambda}{(N-k)}|\mathcal{H}_h\right) + Pr\left(|L_{[k]}| < \frac{\lambda - g_U}{(N-k)}|\mathcal{H}_h\right)\right], \tag{A.42}$$

where $Pr\left(|L_{[k]}|<r|\mathcal{H}_h\right)=\int_0^r f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h)\mathrm{d}l_{[k]}$ for $r\in\{\frac{g_L-\lambda}{(N-k)},\frac{\lambda-g_U}{(N-k)}\}$. It is given in closed form as [18]

$$f_{L_{[k]}}(l_{[k]}|\mathcal{H}_h)=Nf_L(l_{[k]}|\mathcal{H}_h)\binom{N-1}{k-1}F_L(l_{[k]}|\mathcal{H}_h)^{(N-k)}(1-F_L(l_{[k]}|\mathcal{H}_h))^{(k-1)}. \tag{A.43}$$

Hence, the pdf of $f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h)$ is given by

$$f_{|L_{[k]}|}(l_{[k]}|\mathcal{H}_h)=\frac{\mathrm{d}Pr(|L_{[k]}|\le l_{[k]})}{\mathrm{d}l_{[k]}}=\begin{cases} f_{L_{[k]}}(l_{[k]}|\mathcal{H}_h)-f_{L_{[k]}}(-l_{[k]}|\mathcal{H}_h) & \text{if } l_{[k]}\ge 0 \\ 0 & \text{if } l_{[k]}<0 \end{cases} \tag{A.44}$$

Substituting (A.44) in (A.42), we are able to obtain the lower bound of the number of transmissions saved.

UB of $\bar{N}_s$

It is easy to show that $\sum_{i=1}^{k}L_{[k]}>\lambda+(N-k)|L_{[k]}|$ implies $g_U>\lambda+(N-k)|L_{[k]}|$. Hence, from (A.37), we get

$$Pr(g_U>\lambda+(N-k)|L_{[k]}||\mathcal{H}_h)\ge Pr(k^*\le k|\mathcal{H}_h) \tag{A.45}$$

Similarly, due to the fact that $\sum_{i=1}^{k}L_{[i]}<\lambda-(N-k)|L_{[k]}|$ implies $g_L<\lambda-(N-k)|L_{[k]}|$, we can also get

$$Pr(g_L<\lambda-(N-k)|L_{[k]}||\mathcal{H}_h)\ge Pr(k^*\le k|\mathcal{H}_h). \tag{A.46}$$

Hence, we have

$$\bar{N}_s\le\sum_{k=1}^{N-1}\sum_{h=0}^{1}Pr\big(g_U>\lambda+n_{UT}|L_{[k]}| \text{ or } g_L<\lambda-n_{UT}|L_{[k]}||\mathcal{H}_h\big)\pi_h, \tag{A.47}$$

where $n_{UT} = N - k$ and

$$\begin{aligned}
&Pr(g_U > \lambda + n_{UT}|L_{[k]}| \text{ or } g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\
=&Pr(g_U > \lambda + n_{UT}|L_{[k]}||\mathcal{H}_h) + Pr(g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\
&- Pr(g_U > \lambda + n_{UT}|L_{[k]}| \text{ and } g_L < \lambda - n_{UT}|L_{[k]}||\mathcal{H}_h) \\
=&Pr\left(|L_{[k]}| \leq \frac{g_U - \lambda}{N-k}|\mathcal{H}_h\right) + Pr\left(|L_{[k]}| \leq \frac{\lambda - g_L}{N-k}|\mathcal{H}_h\right) \\
&- Pr\left(|L_{[k]}| \leq \min\left(\frac{g_U - \lambda}{N-k}, \frac{\lambda - g_L}{N-k}\right)|\mathcal{H}_h\right).
\end{aligned} \tag{A.48}$$

Following the similar procedure when we obtain the LB of $\bar{N}_s$, we can get the UB of $\bar{N}_s$. Then, we can obtain the UB and the LB in Theorem 4.2.

A.6 Proof of Theorem 4.4, Chapter 4

According to Equation (4.46), $\bar{N}_{s,CEOT}$ is given as

$$\begin{aligned}
\bar{N}_{s,CEOT} &= \sum_{k=1}^{N-1} Pr(k^* \leq k) \\
&= \sum_{k=1}^{N-1} Pr(k^* \leq k|\Gamma < T)Pr(\Gamma < T) + Pr(k^* \leq k|\Gamma \geq T)Pr(\Gamma \geq T).
\end{aligned} \tag{A.49}$$

LB of $\bar{N}_{s,CEOT}$

Recall that k_0^* and k_1^* denote the minimum number of transmissions needed to make a final decision for descending and ascending ordered local decisions, respectively. It is easy to show that $k_1^* \leq k$ implies $k^* \leq k$ given $\Gamma \geq T$ and $k_0^* \leq k$ implies $k^* \leq k$ given $\Gamma < T$. Hence, we have

$$Pr(k^* \leq k|\Gamma \geq T) \geq Pr(k_1^* \leq k|\Gamma \geq T), \tag{A.50}$$

$$Pr(k^* \leq k|\Gamma < T) \geq Pr(k_0^* \leq k|\Gamma < T). \tag{A.51}$$

Substituting $Pr(k^* \leq k|\Gamma < T)$ and $Pr(k^* \leq k|\Gamma \geq T)$ in (A.49) with their LBs $Pr(k_0^* \leq k|\Gamma < T)$ and $Pr(k_1^* \leq k|\Gamma \geq T)$, respectively, we get

$$\begin{aligned}\bar{N}_{s,CEOT} \geq & \sum_{h=0}^{1}\sum_{k=1}^{N-1} P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) P(\Gamma \geq T|\mathcal{H}_h)\pi_h \\ & + \sum_{k=1}^{N-1} P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right) P(\Gamma < T|\mathcal{H}_h)\pi_h. \end{aligned} \tag{A.52}$$

Since $z_{(i)}$ and $z_{[i]}$ for $\forall i \in \{1, 2, \ldots, N\}$ are non-negative, we have $0 \leq \sum_{i=1}^{k_1^*} z_{(i)} \leq k_1^*$ and $0 \leq \sum_{i=1}^{k_0^*} z_{[i]} \leq k_0^*$. For the fusion rule of equivalent worst case given $\Gamma \geq T$, which is given as

$$\sum_{i=1}^{k_1^*} z_{(i)} \geq T \quad \text{decides } \mathcal{H}_1, \tag{A.53}$$

where $k_1^* \geq T$ is needed to make a decision $\mathcal{H}_1$.

For the fusion rule of equivalent worst case given $\Gamma < T$ given as

$$\sum_{i=1}^{k_0^*} z_{[i]} < T - (N - k_0^*) \quad \text{decides } \mathcal{H}_0, \tag{A.54}$$

where $k_0^* > N - T$ is needed to make a decision $\mathcal{H}_0$. Hence, it is obvious that the FC can not make decision $\mathcal{H}_0$ given $\Gamma < T$ when $k_0^* \leq N - T$ and the FC can not make decision $\mathcal{H}_1$ given $\Gamma \geq T$ when $k < T$. Hence, we have

$$\sum_{k=1}^{T-1} P\left(k_1^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = \sum_{k=1}^{N-T} P\left(k_0^* \leq k|\Gamma < T, \mathcal{H}_h\right) = 0. \tag{A.55}$$

As shown in (4.48), the magnitude of local decisions are ordered in an ascending order, i.e., $|z_{(1)}| \leq |z_{(2)}|, \ldots, \leq |z_{(N)}|$, when we consider the equivalent worst case given $\Gamma \geq T$. It is apparent that $\Gamma \geq T$ implies that the distributed system without ordering would make a decision of $\mathcal{H}_1$. According to Lemma 4.2, the detection performance of the CEOT-based system is the same as that of the distributed system without ordering. We can easily conclude that $k^* \leq k_1^*$ is always

satisfied, which indicates that the minimum number of transmissions required to make a decision for equivalent worst case given $\Gamma \geq T$ is always greater than or equal to the actual minimum number of transmissions required. Since at most $min(N-T, k-T)$ 0s are required when $\Gamma \geq T$ for the unordered distributed system, we have

$$P\left(k_1^* \leq k | \Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{min(N-T,k-T)} \binom{N}{i} \pi_{0,h}^{i} \pi_{1,h}^{N-i}, \tag{A.56}$$

when $k \geq T$.

Similarly, as shown in (4.47), the magnitude of local decisions are ordered in a descending order, i.e., $z_{[1]} \geq z_{[2]}, \ldots, \geq z_{[N]}$, when we consider the equivalent worst case given $\Gamma < T$. Here, $\Gamma < T$ implies that the distributed system without ordering would make a decision of $\mathcal{H}_0$. According to Lemma 4.2, we can also easily conclude that $k^* \leq k_0^*$ is always satisfied, which means that the minimum number of transmissions required to make a decision for equivalent worst case given $\Gamma < T$ is always greater than or equal to the true minimum number of transmissions required. Since at most $min(T-1, k-(N-T+1))$ 1s are required when $\Gamma < T$ for the unordered distributed system, we have

$$P\left(k_0^* \leq k | \Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{min(T-1,k-(N-T+1))} \binom{N}{i} \pi_{1,h}^{i} \pi_{0,h}^{N-i} \tag{A.57}$$

if $k > N - T$.

UB of $\bar{N}_{s,CEOT}$

By substituting k_1^* in (A.53) with k_0^*, we can obtain the fusion rule of equivalent best case given $\Gamma \geq T$ where $k_0^* \geq T$ is needed to make a decision $\mathcal{H}_1$. Similarly, by substituting k_0^* in (A.54) with k_1^*, we can obtain the fusion rule of equivalent best case given $\Gamma < T$ where $k_1^* > N - T$ is needed to make a decision $\mathcal{H}_0$. It is easy to show that $k^* \leq k$ implies $k_1^* \leq k$ given $\Gamma < T$ and

$k^* \leq k$ implies $k_0^* \leq k$ given $\Gamma \geq T$. Hence, we get

$$Pr(k_1^* \leq k|\Gamma < T) \geq Pr(k^* \leq k|\Gamma < T), \tag{A.58}$$

$$Pr(k_0^* \leq k|\Gamma \geq T) \geq Pr(k^* \leq k|\Gamma \geq T). \tag{A.59}$$

Substituting $Pr(k^* \leq k|\Gamma < T)$ and $Pr(k^* \leq k|\Gamma \geq T)$ in (A.49) with their UBs $Pr(k_1^* \leq k|\Gamma < T)$ and $Pr(k_0^* \leq k|\Gamma \geq T)$, respectively, we get

$$\begin{aligned}\bar{N}_{s,CEOT} \leq & \sum_{h=0}^{1}\sum_{k=1}^{N-1} P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right) P(\Gamma < T|\mathcal{H}_h)\pi_h \\ & + \sum_{k=1}^{N-1} P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) P(\Gamma \geq T|\mathcal{H}_h)\pi_h. \end{aligned} \tag{A.60}$$

Following the similar procedure, we have

$$\sum_{k=1}^{N-T} P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right) = \sum_{k=1}^{T-1} P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = 0. \tag{A.61}$$

According to Lemma 4.2, it is apparent that $k_1^* \leq k^*$ is always satisfied if $\Gamma \geq T$, i.e., the minimum number of transmissions required to make a decision for equivalent best case given $\Gamma \geq T$ is always less than or equal to the actual minimum number of transmissions required. Similarly, we can also conclude that $k_0^* \leq k^*$ is always satisfied if $\Gamma < T$. Following a similar procedure to derive the LB, we obtain

$$P\left(k_1^* \leq k|\Gamma < T, \mathcal{H}_h\right) = \sum_{i=0}^{T-1} \binom{N}{i} \pi_{1,h}^i \pi_{0,h}^{N-i}, \tag{A.62}$$

when $k > N - T$ and

$$P\left(k_0^* \leq k|\Gamma \geq T, \mathcal{H}_h\right) = \sum_{i=0}^{N-T} \binom{N}{i} \pi_{0,h}^i \pi_{1,h}^{N-i}, \tag{A.63}$$

when $k \geq T$. Then, we obtain the UB, which is given by substituting (A.61), (A.62) and (A.63) in (A.60), and the LB, which is given by substituting (A.55), (A.56) and (A.57) in (A.52), in

Theorem 4.4.

A.7 Proof of Lemma 4.2, Chapter 4

Recall that k_U^* and k_L^* are the minimum number of transmissions required to decide $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. These are given as follows.

$$k_U^* = \min_{1\leq k\leq N}\left\{\sum_{i=1}^{k} u_{[i]} < T-(N-k)\right\} \tag{A.64}$$

$$k_L^* = \min_{1\leq k\leq N}\left\{\sum_{i=1}^{k} u_{[i]} \geq T\right\}. \tag{A.65}$$

It is obvious that k_U^* and k_L^* can not exist at the same time. If k_U^* is valid, the FC decides hypotheses $\mathcal{H}_0$, and if k_L^* is valid, the FC decides hypotheses $\mathcal{H}_1$. Since only one of the two hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$ can occur at any given time, only one of k_L^* or k_U^* is valid at any given time. Let Z_U, Z_L denote the upper bound and lower bound of $Z=\sum_{i=1}^{N} u_{[i]}$, respectively, if k_U^* or k_L^* is valid. Due to the fact that $u_{[i]}\in\{0,1\}$ for $\forall i\in\{1,2,\ldots,N\}$, we have

$$Z_U = \sum_{i=1}^{k_U^*} u_{[i]} + (N-k_U^*) \geq \sum_{i=1}^{N} u_{[i]} = Z \tag{A.66}$$

if k_U^* is valid and

$$Z_L = \sum_{i=1}^{k_L^*} u_{[i]} \leq \sum_{i=1}^{N} u_{[i]} = Z \tag{A.67}$$

if k_L^* is valid. According to the fusion rule given in (4.20), the FC decides hypothesis $\mathcal{H}_0$ if $Z_U<T$, and hypothesis $\mathcal{H}_1$ if $Z_L\geq T$. Since only one of k_U^* and k_L^* is valid, we have $Pr(k_U^* \text{ is valid}) + Pr(k_L^* \text{ is valid}) = 1$, which is equivalent to $Pr(Z_U<T)+Pr(Z_L\geq T)=1$.

If the FC decides $\mathcal{H}_1$ for an unordered system, we have

$$Z\geq T \overset{implies}{\Longrightarrow} \sum_{i=1}^{k} u_{[i]} + \sum_{i=k+1}^{N} u_{[i]} \geq T \tag{A.68a}$$

$$\implies \sum_{i=1}^{k} u_{[i]} \geq T - \sum_{i=k+1}^{N} u_{[i]} \geq T - (N-k) \tag{A.68b}$$

$$\implies \sum_{i=1}^{k} u_{[i]} \geq T - (N-k) \tag{A.68c}$$

for $\forall k$. We could observe from (A.64) that k_U^* is not valid when $Z \geq T$. So if $Z \geq T$, we have $Pr(Z_U < T) = 0$ and $Pr(Z_L \geq T) = 1$. Hence, we can conclude that $Pr(Z_L \geq T | Z \geq T, \mathcal{H}_j) = 1$. Upon following a similar analysis, we obtain $Pr(Z \geq T | Z_L \geq T, \mathcal{H}_j) = 1$. This allows us to calculate $Pr(Z_L \geq T | \mathcal{H}_j)$ according to Bayes' rule which is given as

$$\begin{aligned} Pr(Z_L \geq T | \mathcal{H}_j) &= \frac{Pr(Z_L \geq T | Z \geq T, \mathcal{H}_j) Pr(Z \geq T | \mathcal{H}_j)}{Pr(Z \geq T | Z_L \geq T, \mathcal{H}_j)} \\ &= Pr(Z \geq T | \mathcal{H}_j). \end{aligned} \tag{A.69}$$

Similarly, we obtain $Pr(Z_U < T | \mathcal{H}_j) = Pr(Z < T | \mathcal{H}_j)$. Hence, the probability of error of the CEOT-based system is given as

$$\begin{aligned} P_e^{(OT)} &= \pi_0 Pr(Z_L \geq T | \mathcal{H}_0) + \pi_1 Pr(Z_U < T | \mathcal{H}_1) \\ &= \pi_0 Pr(Z \geq T | \mathcal{H}_0) + \pi_1 Pr(Z < T | \mathcal{H}_1) = P_e^{(opt)}, \end{aligned} \tag{A.70}$$

where $P_e^{(opt)}$ is the probability of error of the unordered system.

A.8 Proof of Theorem 5.1, Chapter 5

We first consider the scenario where sensors send binary decisions to the FC, i.e., $q = 1$. After that, we consider the system where sensors send q-bit decisions to the FC ($q \geq 2$). Here, we only consider the assumption that $\widetilde{\tau}_{j,2^q} \ll \tau_{i,1}$. Nevertheless, we can reach similar conclusions if we assume $\tau_{i,2^q} \ll \widetilde{\tau}_{j,1}$.

When Sensors Send Binary Decisions (q=1)

The joint pmf of local decisions coming from the reference sensors under hypothesis $\mathcal{H}_h$ is given as $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) = \prod_{i=1}^{N_{ref}} (1-x)^{\mathbf{u_i}} x^{1-\mathbf{u_i}}$ for $h = 0, 1$. Take the logarithm of both sides, we have

$$\begin{aligned}\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) =& \sum_{i=1}^{N_{ref}} [\mathbf{u_i} \log(1-x) + (1-\mathbf{u_i}) \log x] \\ =& Y \log(1-x) + (N_{ref} - Y) \log x, \end{aligned} \tag{A.71}$$

where $Y = \sum_{i=1}^{N_{ref}} \mathbf{u_i}$. Let $\frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x} = 0$, we are able to obtain the estimated attack parameter $\hat{x}_h$ under hypothesis $\mathcal{H}_h$ which maximizes $\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ and the estimated attack parameter $\hat{x}_h$ is given as $\hat{x}_h = 1 - \frac{Y}{N_{ref}}$.

In order to evaluate the estimator performance, it should be noted that it is unbiased since

$$E[\hat{x}_h] = 1 - \frac{1}{N_{ref}} E[Y] = 1 - \frac{1}{N_{ref}} \sum_{i=1}^{N_{ref}} E[\mathbf{u_i}] = x \tag{A.72}$$

The variance of the estimator is given as

$$\begin{aligned} E[\hat{x}_h] =& E[\hat{x}_h^2] - E^2[\hat{x}_h] = E\left[\left(1 - \frac{Y}{N_{ref}}\right)^2\right] - x^2 \\ =& 1 - x^2 - \frac{2}{N_{ref}} E[Y] + \frac{1}{N_{ref}^2} E[Y^2] \\ =& 1 - x^2 - 2(1-x) + \frac{1}{N_{ref}^2} (Var[Y] + E^2[Y]) \\ =& 1 - x^2 - 2(1-x) + \frac{1}{N_{ref}^2} [N_{ref} x(1-x) + N_{ref}^2 (1-x)^2] \\ =& \frac{(1-x)x}{N_{ref}} \end{aligned} \tag{A.73}$$

To evaluate the performance of the estimator, the CRLB can be calculated which is $-\frac{1}{E\left[\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)/\partial x^2\right]}$. Taking the second derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to x, we have $\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x^2} =$

$\sum_{i=1}^{N_{ref}} \left[-\frac{\mathbf{u_i}}{(1-x)^2} - \frac{1-\mathbf{u_i}}{x^2} \right]$. Subsequently, taking the expectation of the above equation, we have

$$\begin{aligned} E\left[\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x^2}\right] &= \sum_{i=1}^{N_{ref}} E\left[\frac{\partial^2 P(\mathbf{u_i}|\mathcal{H}_h, p, x)}{\partial x^2}\right] \\ &= \sum_{i=1}^{N_{ref}} -\frac{1}{(1-x)^2}(1-x) - \frac{1}{x^2}x \\ &= -\frac{N_{ref}}{(1-x)x}. \end{aligned} \tag{A.74}$$

Therefore, the CRLB is $\frac{(1-x)x}{N_{ref}}$ which is the same as (A.73). This indicates that the proposed estimator attains the CRLB; that is, it is an efficient estimator when sensors in the network send binary decisions.

When Sensors Send q-bit Decisions ($q \geq 2$)

The joint pmf of local decisions coming from the reference sensors under hypothesis $\mathcal{H}_h$ is given as $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) = \prod_{i=1}^{N_{ref}} (1-x)^{I(\mathbf{u_i}=\mathbf{v}_{2^q})} \prod_{j=1}^{2^q-1} (\frac{x}{2^q-1})^{I(\mathbf{u_i}=\mathbf{v_j})}$ for $h = 0, 1$. Take the logarithm of both sides, we have

$$\begin{aligned} &\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x) \\ &= \sum_{i=1}^{N_{ref}} I(\mathbf{u_i} = 2^q) \log(1-x) + \sum_{j=1}^{2^q-1} I(\mathbf{u_i} = \mathbf{v_j}) \log(\frac{x}{2^q-1}), \end{aligned} \tag{A.75}$$

Taking the first derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to x, we have

$$\begin{aligned} \frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)}{\partial x} &= \sum_{i=1}^{N_{ref}} \frac{-1}{1-x} I(\mathbf{u_i}=2^q) + \sum_{j=1}^{2^q-1} \frac{1}{x} I(\mathbf{u_i}=\mathbf{v_j}) \\ &= \frac{-Y_1}{1-x} + \frac{Y_2}{x} \qquad (A.76) \\ &= \frac{-Y_1}{1-x} + \frac{N_{ref}-Y_1}{x} \qquad (A.77) \end{aligned}$$

where $Y_1 = \sum_{i=1}^{N_{ref}} I(\mathbf{u_i} = v_{2^q})$ and $Y_2 = \sum_{i=1}^{N_{ref}} \sum_{j=1}^{2^q-1} I(\mathbf{u_i} = \mathbf{v_j})$. In going from (A.76) to

(A.77), the fact that $Y_1 + Y_2 = N_{ref}$ is utilized. Let $\frac{\partial P(\mathbf{U}_{ref}|\mathcal{H}_h,p,x)}{\partial x} = 0$, we are able to obtain the estimated attack parameter $\hat{x}$ which maximizes $\log P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$. The estimated attack parameter $\hat{x}_h$ under hypothesis $\mathcal{H}_h$ is given as $\hat{x}_h = 1 - \frac{Y_1}{N_{ref}}$.

In order to evaluate the estimator performance, it should be noted that it is unbiased since

$$E[\hat{x}_h] = 1 - \frac{1}{N_{ref}} E[Y_1] = x \tag{A.78}$$

Similarly, the variance of the estimator is given as

$$\begin{aligned} E[\hat{x}_h] =& E[\hat{x}_h^2] - E^2[\hat{x}_h] = E\left[\left(1 - \frac{Y_1}{N_{ref}}\right)^2\right] - x^2 \\ =& 1 - x^2 - \frac{2}{N_{ref}} E[Y_1] + \frac{1}{N_{ref}^2} E[Y_1^2] \\ =& \frac{(1-x)x}{N_{ref}} \end{aligned} \tag{A.79}$$

To evaluate the performance of the estimator, the CRLB can be calculated which is $-\frac{1}{E\left[\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h,p,x)/\partial x^2\right]}$. Taking the second derivative of $P(\mathbf{U}_{ref}|\mathcal{H}_h, p, x)$ with respect to p, we have

$$\begin{aligned} \frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h,p,x)}{\partial x^2} =& \sum_{i=1}^{N_{ref}} -\frac{I(\mathbf{u_i} = 2^q)}{(1-x)^2} - \sum_{i=1}^{2^q-1} \frac{I(\mathbf{u_i} = \mathbf{v_j})}{x^2} \\ =& \sum_{i=1}^{N_{ref}} -\frac{I(\mathbf{u_i} = 2^q)}{(1-x)^2} - \frac{1 - I(\mathbf{u_i} = 2^q)}{x^2} \end{aligned} \tag{A.80}$$

Subsequently, taking the expectation of the above equation, we have

$$\begin{aligned} E\left[\frac{\partial^2 P(\mathbf{U}_{ref}|\mathcal{H}_h,p,x)}{\partial x^2}\right] &= \sum_{i=1}^{N_{ref}} E\left[\frac{\partial^2 P(\mathbf{u_i}|\mathcal{H}_h,p,x)}{\partial x^2}\right] \\ &= \sum_{i=1}^{N_{ref}} -\frac{1}{(1-x)^2}(1-x) - \frac{1}{x^2}x \\ &= -\frac{N_{ref}}{(1-x)x} \end{aligned} \tag{A.81}$$

Therefore, the CRLB is $\frac{(1-x)x}{N_{ref}}$ which is the same as (A.79). This indicates that the proposed estimator attains the CRLB; that is, it is an efficient estimator when sensors in the network send q-bits decisions. This completes our proof.

REFERENCES

[1] S. Alhakeem and P. Varshney, "A unified approach to the design of decentralized detection systems," *IEEE Transactions on Aerospace and Electronic systems*, vol. 31, no. 1, pp. 9–20, 1995. 148

[2] S. Appadwedula, V. V. Veeravalli, and D. L. Jones, "Decentralized detection with censoring sensors," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1362–1373, 2008. 2

[3] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4938–4953, 2011. 65

[4] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, vol. 3. IEEE, 2003, pp. 1713–1723. 65

[5] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007. 2

[6] N. Bessis, "A model to manage smart devices in mobile sensing applications," Ph.D. dissertation, Edge Hill University, UK, 2021. 66, 67

[7] R. S. Blum and B. M. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3229–3235, 2008. 2, 65, 67, 69, 70, 73

[8] P. Braca, S. Marano, and V. Matta, "Single-transmission distributed detection via order statistics," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2042–2048, 2011. 65, 67

[9] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008. 2

[10] W.-N. Chen and I.-H. Wang, "Anonymous heterogeneous distributed detection: Optimal decision rules, error exponents, and the price of anonymity," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7390–7406, 2019. 3

[11] Y. Chen, R. S. Blum, and B. M. Sadler, "Optimal quickest change detection in sensor networks using ordered transmissions," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5. 66, 67

[12] ——, "Ordering for communication-efficient quickest change detection in a decomposable graphical model," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4710–4723, 2021. 66, 67

[13] Y. Chen, B. M. Sadler, and R. S. Blum, "Ordered gradient approach for communication-efficient distributed learning," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2020, pp. 1–5. 66, 67

[14] D. Ciuonzo, A. Aubry, and V. Carotenuto, "Rician MIMO channel-and jamming-aware decision fusion," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866–3880, 2017. 1, 13, 132

[15] D. Ciuonzo, S. H. Javadi, A. Mohammadi, and P. S. Rossi, "Bandwidth-constrained decentralized detection of an unknown vector signal via multisensor fusion," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 6, pp. 744–758, 2020. 101

[16] D. Ciuonzo, P. S. Rossi, and P. K. Varshney, "Distributed detection in wireless sensor networks under multiplicative fading via generalized score tests," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9059–9071, 2021. 13

[17] Y. Cui, S. Li, and W. Zhang, "Jointly sparse signal recovery and support recovery via deep learning with applications in mimo-based grant-free random access," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 788–803, 2021. 105

[18] H. A. David and H. N. Nagaraja, *Order statistics*. John Wiley & Sons, 2004. 161, 162, 164

[19] D. Dolev, "The Byzantine generals strike again," *Journal of algorithms*, vol. 3, no. 1, pp. 14–30, 1982. 67

[20] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006. 101

[21] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse signal detection from incoherent projections," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 3. IEEE, 2006, pp. III–III. 102

[22] G. Fellouris, E. Bayraktar, and L. Lai, "Efficient byzantine sequential change detection," *IEEE Transactions on Information Theory*, vol. 64, no. 5, pp. 3346–3360, 2017. 4

[23] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 167–178, 2012. 65

[24] M. Fornasier and H. Rauhut, "Compressive sensing." *Handbook of mathematical methods in imaging*, vol. 1, pp. 187–229, 2015. 101

[25] Y. Fu and Z. He, "Entropy-based weighted decision combining for collaborative spectrum sensing over Byzantine attack," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1528–1532, 2019. 67, 103

[26] K. Gai, M. Qiu, Z. Ming, H. Zhao, and L. Qiu, "Spoofing-jamming attack strategy using optimal power distributions in wireless smart grid networks," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2431–2439, 2017. 1, 13, 49, 132

[27] F. Gao, L. Guo, H. Li, J. Liu, and J. Fang, "Quantizer design for distributed glrt detection of weak signal in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2032–2042, 2014. 102, 103

[28] B. Geng, S. Brahma, T. Wimalajeewa, P. K. Varshney, and M. Rangaswamy, "Prospect theoretic utility based human decision making in multi-agent systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 1091–1104, 2020. 131

[29] B. Geng, Q. Li, and P. K. Varshney, "Prospect theory based crowdsourcing for classification in the presence of spammers," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4083–4093, 2020. 131

[30] S. S. Gupta and N. B. Mehta, "Ordered transmissions schemes for detection in spatially correlated wireless sensor networks," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1565–1577, 2020. 66, 67

[31] S. S. Gupta, S. K. Pallapothu, and N. B. Mehta, "Ordered transmissions for energy-efficient detection in energy harvesting wireless sensor networks," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2525–2537, 2020. 66, 67

[32] A. Hariri and M. Babaie-Zadeh, "Compressive detection of sparse signals in additive white gaussian noise without signal reconstruction," *Signal Processing*, vol. 131, pp. 376–385, 2017. 8, 102, 103, 106

[33] W. Hashlamoun, S. Brahma, and P. K. Varshney, "Mitigation of byzantine attacks on distributed detection systems using audit bits," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 18–32, 2017. 5, 7, 12, 14, 15, 16, 18, 20, 49, 50, 51, 59, 107

[34] ——, "Audit bit based distributed bayesian detection in the presence of byzantines," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 4, pp. 643–655, 2018. 5, 7, 12, 14, 15, 16, 18, 20, 21, 23, 27, 49, 50, 51, 59

[35] L. Hesham, A. Sultan, M. Nafie, and F. Digham, "Distributed spectrum sensing with sequential ordered transmissions to a cognitive fusion center," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2524–2538, 2012. 65, 67

[36] P. Hodkinson and A. C. Sparkes, "Careership: a sociological theory of career decision making," *British journal of sociology of education*, vol. 18, no. 1, pp. 29–44, 1997. 133

[37] Y.-C. Huang, Y.-J. Huang, and S.-C. Lin, "Asymptotic optimality in byzantine distributed quickest change detection," *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 5942–5962, 2021. 4, 5

[38] C. S. Hyder, B. Grebur, L. Xiao, and M. Ellison, "Arc: Adaptive reputation based clustering against spectrum sensing data falsification attacks," *IEEE Transactions on mobile computing*, vol. 13, no. 8, pp. 1707–1719, 2014. 4

[39] R. P. Jover, J. Lackey, and A. Raghavan, "Enhancing the security of lte networks against jamming attacks," *EURASIP Journal on Information Security*, vol. 2014, no. 1, pp. 1–14, 2014. 1, 13, 49, 132

[40] S. Kafle, B. Kailkhura, T. Wimalajeewa, and P. K. Varshney, "Decentralized joint sparsity pattern recovery using 1-bit compressive sensing," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2016, pp. 1354–1358. 105

[41] D. Kahneman and A. Tversky, "Choices, values, and frames." *American psychologist*, vol. 39, no. 4, p. 341, 1984. 133

[42] B. Kailkhura, S. Brahma, Y. S. Han, and P. K. Varshney, "Optimal distributed detection in the presence of byzantines," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 2925–2929. 13

[43] B. Kailkhura, S. Brahma, and P. K. Varshney, "Data falsification attacks on consensus-based detection systems," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 145–158, 2016. 72

[44] ——, "Data falsification attacks on consensus-based detection systems," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 145–158, 2017. 4, 148

[45] B. Kailkhura, S. Brahma, and P. Varshney, "On the performance analysis of data fusion schemes with byzantines," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 7411–7415. 59, 72, 74

[46] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "On covert data falsification attacks on distributed detection systems," in *2013 13th International Symposium on Communications and Information Technologies (ISCIT)*. IEEE, 2013, pp. 412–417. 14, 49

[47] ——, "Distributed bayesian detection in the presence of byzantine data," *IEEE transactions on signal processing*, vol. 63, no. 19, pp. 5250–5263, 2015. 14, 49

[48] B. Kailkhura, A. Vempaty, and P. K. Varshney, "Collaborative spectrum sensing in the presence of byzantine attacks," in *Cooperative and Graph Signal Processing*. Elsevier, 2018, pp. 505–522. 2

[49] M. Kam, C. Rorres, W. Chang, and X. Zhu, "Performance and geometric interpretation for decision fusion with memory," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 29, no. 1, pp. 52–62, 1999. 148

[50] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009, vol. 344. 53

[51] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993. 114

[52] M. Korki, J. Zhang, C. Zhang, and H. Zayyani, "Block-sparse impulsive noise reduction in ofdm systems—a novel iterative bayesian approach," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 271–284, 2015. 102

[53] ——, "Iterative bayesian reconstruction of non-iid block-sparse signals," *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3297–3307, 2016. 102, 103, 106

[54] L. Lamport, R. Shostak, and M. Pease, "The Byzantine generals problem," in *Concurrency: the Works of Leslie Lamport*, 2019, pp. 203–226. 1, 13, 49, 132

[55] C. Li, Y. He, X. Wang, G. Li, and P. K. Varshney, "Distributed detection of sparse stochastic signals via fusion of 1-bit local likelihood ratios," *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1738–1742, 2019. 102, 103, 106

[56] C. Li, G. Li, and P. K. Varshney, "Distributed detection of sparse signals with censoring sensors in clustered sensor networks," *Information Fusion*, vol. 83, pp. 1–18, 2022. 102, 103

[57] H. Li and Z. Han, "Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3554–3565, 2010. 14

[58] Z. Li, Y. Mo, and F. Hao, "Distributed sequential hypothesis testing with byzantine sensors," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3044–3058, 2021. 5

[59] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE transactions on Wireless Communications*, vol. 7, no. 4, pp. 1326–1337, 2008. 106

[60] H.-Y. Lin, P.-N. Chen, Y. S. Han, and P. K. Varshney, "Minimum Byzantine effort for blinding distributed detection in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 68, pp. 647–661, 2020. 67, 103

[61] S. Lindsey and C. S. Raghavendra, "Pegasis: Power-efficient gathering in sensor information systems," in *Proceedings, IEEE aerospace conference*, vol. 3. IEEE, 2002, pp. 3–3. 15

[62] S. Liu, H. Zhu, S. Li, X. Li, C. Chen, and X. Guan, "An adaptive deviation-tolerant secure scheme for distributed cooperative spectrum sensing," in *2012 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2012, pp. 603–608. 4, 148

[63] X. Liu, T. J. Lim, and J. Huang, "Optimal byzantine attacker identification based on game theory in network coding enabled wireless ad hoc networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2570–2583, 2020. 4

[64] Y. Liu and C. Li, "Secure distributed estimation over wireless sensor networks under attacks," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1815–1831, 2018. 67, 103

[65] A. Manjeshwar and D. P. Agrawal, "Teen: Arouting protocol for enhanced efficiency in wireless sensor networks." in *ipdps*, vol. 1, no. 2001, 2001, p. 189. 15

[66] S. Marano, V. Matta, and L. Tong, "Distributed detection in the presence of Byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 16–29, 2008. 67, 72, 103, 107

[67] E. Masazade, R. Niu, and P. K. Varshney, "Dynamic bit allocation for object tracking in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5048–5063, 2012. 15

[68] A. Mohammadi, D. Ciuonzo, A. Khazaee, and P. S. Rossi, "Generalized locally most powerful tests for distributed sparse signal detection," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 528–542, 2022. 102, 103, 106

[69] A. Mustafa, M. N. U. Islam, and S. Ahmed, "Dynamic spectrum sensing under crash and byzantine failure environments for distributed convergence in cognitive radio networks," *IEEE Access*, vol. 9, pp. 23 153–23 167, 2021. 4, 148

[70] V. S. S. Nadendla, Y. S. Han, and P. K. Varshney, "Distributed inference with m-ary quantized data in the presence of byzantine attacks," *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2681–2695, 2014. 4

[71] N. Nguyen-Thanh and I. Koo, "An enhanced cooperative spectrum sensing scheme based on evidence theory and reliability source evaluation in cognitive radio context," *IEEE Communications Letters*, vol. 13, no. 7, pp. 492–494, 2009. 72

[72] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 4, pp. 1–11, 2005. 15, 80

[73] H. Palangi, R. Ward, and L. Deng, "Distributed compressive sensing: A deep learning approach," *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4504–4518, 2016. 105

[74] M. Pease, R. Shostak, and L. Lamport, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, 1982. 2

[75] F. Penna, Y. Sun, L. Dolecek, and D. Cabric, "Detecting and counteracting statistical attacks in cooperative spectrum sensing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1806–1822, 2011. 13

[76] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, 2004. 16

[77] Z. Qin, J. Fan, Y. Liu, Y. Gao, and G. Y. Li, "Sparse representation for wireless communications: A compressive sensing approach," *IEEE Signal Processing Magazine*, vol. 35, no. 3, pp. 40–58, 2018. 105

[78] C. Quan, B. Geng, Y. Han, and P. K. Varshney, "Enhanced audit bit based distributed Bayesian detection in the presence of strategic attacks," *IEEE Transactions on Signal and Information Processing over Networks*, 2022. 51, 52

[79] C. Quan, B. Geng, and P. K. Varshney, "Asymptotic performance in heterogeneous human-machine inference networks," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, 2020, pp. 584–588. 130, 131, 133, 139, 140

[80] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, no. 2, pp. 554–568, 1996. 65

[81] Z. N. Rawas, Q. He, and R. S. Blum, "Energy-efficient noncoherent signal detection for networked sensors using ordered transmissions," in *2011 45th Annual Conference on Information Sciences and Systems*. IEEE, 2011, pp. 1–5. 65, 67

[82] A. S. Rawat, P. Anand, H. Chen, and P. K. Varshney, "Collaborative spectrum sensing in the presence of byzantine attacks in cognitive radio networks," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 774–786, 2010. 14, 43, 49, 59, 61, 103

[83] I. Sartzetakis, K. Christodoulopoulos, and E. Varvarigos, "Improving qot estimation accuracy through active monitoring," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2017, pp. 1–4. 104

[84] I. Sartzetakis, K. K. Christodoulopoulos, and E. M. Varvarigos, "Accurate quality of transmission estimation with machine learning," *Journal of Optical Communications and Networking*, vol. 11, no. 3, pp. 140–150, 2019. 104

[85] J. Sonnek, A. Chandra, and J. Weissman, "Adaptive reputation-based scheduling on unreliable distributed infrastructures," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 11, pp. 1551–1564, 2007. 49

[86] C. Soussen, J. Idier, D. Brie, and J. Duan, "From bernoulli–gaussian deconvolution to sparse signal restoration," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4572–4584, 2011. 102, 103, 106

[87] N. Sriranga, B. Geng, and P. K. Varshney, "On human assisted decision making for machines using correlated observations," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2020, pp. 1502–1506. 8, 130, 131

[88] N. Sriranga, K. G. Nagananda, R. S. Blum, A. Saucan, and P. K. Varshney, "Energy-efficient decision fusion for distributed detection in wireless sensor networks," in *2018 21st International conference on information fusion (FUSION)*. IEEE, 2018, pp. 1541–1547. 65, 66, 67, 69, 70, 73, 78, 81, 84, 98

[89] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Communications Surveys & Tutorials*, vol. 13, no. 3, pp. 443–461, 2010. 16

[90] L. Tong, Q. Zhao, and S. Adireddy, "Sensor networks with mobile agents," in *IEEE Military Communications Conference, 2003. MILCOM 2003.*, vol. 1. IEEE, 2003, pp. 688–693. 16

[91] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals and Systems*, vol. 1, no. 2, pp. 167–182, 1988. 18, 43, 70

[92] H. L. Van Trees, *Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory*. John Wiley & Sons, 2004. 75

[93] L. R. Varshney and K. R. Varshney, "Decision making with quantized priors leads to discrimination," *Proceedings of the IEEE*, vol. 105, no. 2, pp. 241–255, 2017. 149

[94] P. K. Varshney, *Distributed detection and data fusion*. Springer Science & Business Media, 2012. 13

[95] V. V. Veeravalli and P. K. Varshney, "Distributed inference in wireless sensor networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, no. 1958, pp. 100–117, 2012. 13

[96] A. Vempaty, B. Kailkhura, and P. K. Varshney, *Secure Networked Inference with Unreliable Data Sources*. Springer, 2018. 67, 103

[97] A. Vempaty, P. Ray, and P. K. Varshney, "False discovery rate based distributed detection in the presence of byzantines," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 1826–1840, 2014. 14, 49, 59, 61

[98] J. P. Walters, Z. Liang, W. Shi, and V. Chaudhary, "Wireless sensor network security: A survey," *Security in distributed, grid, mobile, and pervasive computing*, vol. 1, no. 367, p. 6, 2007. 16

[99] W. Wang, L. Chen, K. G. Shin, and L. Duan, "Secure cooperative spectrum sensing and access against intelligent malicious behaviors," in *IEEE INFOCOM 2014-IEEE Conference on Computer Communications*. IEEE, 2014, pp. 1267–1275. 13

[100] X. Wang, G. Li, C. Quan, and P. K. Varshney, "Distributed detection of sparse stochastic signals with quantized measurements: The generalized gaussian case," *IEEE Transactions*

on Signal Processing, vol. 67, no. 18, pp. 4886–4898, 2019. 102, 103, 106, 111, 123, 124, 127, 128

[101] X. Wang, G. Li, and P. K. Varshney, "Detection of sparse signals in sensor networks via locally most powerful tests," *IEEE Signal Processing Letters*, vol. 25, no. 9, pp. 1418–1422, 2018. 8, 102, 103, 106

[102] ——, "Detection of sparse stochastic signals with quantized measurements in sensor networks," *IEEE Transactions on Signal Processing*, vol. 67, no. 8, pp. 2210–2220, 2019. 8, 102, 103

[103] C.-Y. Wei, P.-N. Chen, Y. S. Han, and P. K. Varshney, "Local threshold design for target localization using error correcting codes in wireless sensor networks in the presence of Byzantine attacks," *IEEE transactions on information forensics and security*, vol. 12, no. 7, pp. 1571–1584, 2017. 67

[104] T. Wimalajeewa and P. K. Varshney, "Collaborative human decision making with random local thresholds," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2975–2989, 2013. 8, 131

[105] ——, "Sparse signal detection with compressive measurements via partial support set estimation," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 46–60, 2016. 102

[106] ——, "Compressive sensing-based detection with multimodal dependent data," *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 627–640, 2017. 108

[107] T. Wimalajeewa, P. K. Varshney, and M. Rangaswamy, "On integrating human decisions with physical sensors for binary decision making," in *2018 21st International Conference on Information Fusion (FUSION)*. IEEE, 2018, pp. 1–5. 8, 130, 131

[108] J. R. Winquist and J. R. Larson Jr, "Information pooling: When it impacts group decision making." *Journal of personality and social psychology*, vol. 74, no. 2, p. 371, 1998. 133

[109] J. Wu, P. Li, Y. Chen, J. Tang, C. Wei, L. Xia, and T. Song, "Analysis of byzantine attack strategy for cooperative spectrum sensing," *IEEE Communications Letters*, vol. 24, no. 8, pp. 1631–1635, 2020. 107

[110] J. Wu, T. Song, Y. Yu, C. Wang, and J. Hu, "Generalized Byzantine attack and defense in cooperative spectrum sensing for cognitive radio networks," *IEEE Access*, vol. 6, pp. 53 272–53 286, 2018. 67, 103

[111] ——, "Sequential cooperative spectrum sensing in the presence of dynamic byzantine attack for mobile networks," *PloS one*, vol. 13, no. 7, p. e0199546, 2018. 5

[112] J. Wu, Y. Yu, H. Zhu, T. Song, and J. Hu, "Cost-benefit tradeoff of byzantine attack in cooperative spectrum sensing," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2532–2543, 2020. 107

[113] L. Xiao, D. Xu, N. B. Mandayam, and H. V. Poor, "Attacker-centric view of a detection game against advanced persistent threats," *IEEE Transactions on Mobile Computing*, vol. 17, no. 11, pp. 2512–2523, 2018. 1, 13

[114] Z. Yang, A. Gang, and W. U. Bajwa, "Adversary-resilient distributed and decentralized statistical inference and machine learning: An overview of recent advances under the byzantine threat model," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 146–159, 2020. 4

[115] H. Zayyani, M. Babaie-Zadeh, and C. Jutten, "An iterative bayesian algorithm for sparse component analysis in presence of noise," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4378–4390, 2009. 102

[116] H. Zayyani, F. Haddadi, and M. Korki, "Double detector for sparse signal detection from one-bit compressed sensing measurements," *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1637–1641, 2016. 102, 103, 106

[117] K. Zeng, P. Pawelczak, and D. Cabric, "Reputation-based cooperative spectrum sensing with trusted nodes assistance," *IEEE communications letters*, vol. 14, no. 3, pp. 226–228, 2010. 3, 49

[118] D.-g. Zhang, T. Zhang, J. Zhang, Y. Dong, and X.-d. Zhang, "A kind of effective data aggregating method based on compressive sensing for wireless sensor network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, pp. 1–15, 2018. 101

[119] L. Zhang, G. Ding, Q. Wu, Y. Zou, Z. Han, and J. Wang, "Byzantine attack and defense in cognitive radio networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1342–1363, 2015. 1, 13, 132

[120] L. Zhang, G. Nie, G. Ding, Q. Wu, Z. Zhang, and Z. Han, "Byzantine attacker identification in collaborative spectrum sensing: A robust defense framework," *IEEE Transactions on Mobile Computing*, vol. 18, no. 9, pp. 1992–2004, 2018. 5, 14

[121] L. Zhang, Q. Wu, G. Ding, S. Feng, and J. Wang, "Performance analysis of probabilistic soft SSDF attack in cooperative spectrum sensing," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–9, 2014. 72

VITA

NAME OF AUTHOR: Chen Quan

MAJOR: Electrical and Computer Engineering

EDUCATION:

M.E. 2016 Syracuse University, USA

B.E. 2012 Nanjing University of Science and Technology, China

AWARDS AND HONORS:

Summer Dissertation Fellowship, Syracuse University, 2022