Byzantine Attacks in RIS-Enhanced Cooperative Spectrum Sensing: A Decision Fusion Perspective

Gaoyuan Zhang, Gaolei Song, Boyuan Li*, Zijian Li*, Baofeng Ji, Ruijuan Zheng, Guoqiang Zheng, Tony Q.S. Quek, *Fellow, IEEE*

***Abstract*—From the perspective of hard decision fusion, we investigate Byzantine attacks in Reconfigurable Intelligent Surface (RIS)-enhanced and decode-and-forward relay-assisted Cooperative Spectrum Sensing (CSS) for mobile Cognitive Radio Networks (CRNs) in this paper. Specially, a RIS-enhanced and decode-and-forward relay-assisted CSS configuration is first constructed under dynamic channel scenarios due to user mobility. Subsequently, the channel- and attack-aware hard decision fusion rules are developed, and the optimal channel-aware Byzantine attack strategies are then developed under both small-scale and large-scale attacking scenarios. The corresponding results depict that the optimal attack strategy does not require any a prior knowledge of the global instantaneous Channel State Information (ICSI) (e.g. false alarm probability and detection probability of all the secondary users), although perfect acquisition of ICSI is clearly always not affordable from the attacker perspective, which is further exacerbated by the RIS and decode-and-forward relays involved in CSS and the potential high mobility of secondary users that leads to fast fading channels. Furthermore, our counterintuitive results also indicate that, regardless of the attacker's awareness of the decision fusion rule, the optimal Byzantine attack can be achieved through a unifying framework, the explicit attack strategy may be not unique, and the attacking effectiveness is primarily determined by the fraction of the Byzantine nodes rather than the channel dynamics. That is, to make the channel-aware approach more practical, the challenge that the heavy reliance on the global ICSI and decision fusion rule in obtaining the Byzantine attacks is successfully relaxed. Finally, we empirically validate our theoretical analysis through extensive simulations across a wide range of attacking scenarios.**

***Index Terms*—Cooperative Spectrum Sensing, Byzantine Attack, Reconfigurable Intelligent Surface, Decision Fusion.**

I. Introduction

A. Background

THE continuously exponential growth of lightweight IoT devices in 5G network is leading to increase in spectrum agility [1]–[3]. Traditional static radio spectrum allocation policy is creating "artificial spectrum scarcity", since Radio Frequency (RF) bands are either idle or highly underutilized most of the time and unlicensed users are not allowed to access [4], [5]. Cognitive Radio Networks (CRNs) is considered an enabling technology for efficient utilization of scarce RF spectrum to ameliorating the unavoidable spectrum scarcity by allowing opportunistic dynamic spectrum access to unlicensed secondary users (SUs) without causing any harmful interference to primary users (PUs) with the help of smart protocols and signal processing algorithms, such as spectrum sensing [6]–[8]. Further, to prevail against various channel effects (such as noise, shadowing, multipath fading etc.) and hidden terminal in real CRNs, Collaborative Spectrum Sensing (CSS) is imperative [9]–[11]. CSS can incorporate the spatial diversity to mitigate these problems, wherein different SUs work together with their independent observations and make a collaborative decision to distinguish between two hypotheses, e.g., the absence or presence of a spectrum hole [12]–[14].

An recent emerging technology named Reconfigurable Intelligent Surfaces (RIS) has added a new dimension to CSS design and demonstrated significant potential for augmenting CSS by reconfiguring the sensing and reporting channel characteristics even when CRNs suffer from a low signal-to-noise ratio (SNR) or strong interference [15]–[17]. Unfortunately, this integratation is increasingly vulnerable to Spectrum Sensing Data Falsification (SSDF) or Byzantine attacks, wherein the compromised SUs can independently or collaboratively submit falsified sensing reports to blind the dedicated node called Fusion Center (FC) in infrastructure-based CRNs [18]–[20]. This is attributed to the fact that the unique and powerful channel reconfiguration of RIS can provide active degrees of freedom for attackers to exploit the capabilities of RIS-enhanced CSS to launch more complicated and unpredictable attacks with even greater damage [21]. These intractable security threats are becoming more severe in large-scale CRNs, where massive sensing reports aggregation magnifies the influence of malicious reports [22].

To mitigate Byzantine attacks, a number of countermeasures for traditional CSS, including global decision-based defense [23], reputation based approaches [24], and abnormal statistical-behavior detection based approaches [25], have been proposed. For a detailed discussion and an extensive list of references of Byzantine attack defense, the author is referred to [26]. However, unlike those who were researching ways to use secure decision fusion or anomaly detection to prevail against Byzantine attacks in traditional CSS [27], we restrict our attention towards comprehensive assessment from

Gaoyuan Zhang, Gaolei Song, Baofeng Ji, Ruijuan Zheng and Guoqiang Zheng are with the School of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China. They are also with the Science and Technology Innovation Center of Intelligent system, Longmen Laboratory, Luoyang 471000, China. Gaoyuan Zhang and Baofeng Ji are also with the Institute of Physics, Zhengzhou 451451, Henan Academy of Sciences. Email: gaoyuan.zhang@uestc.edu.cn.

Boyuan Li was with the School of Computer Science and Artificial Intelligence, Zhengzhou University, Zhengzhou, China, 450001. Zijian Li is with Dalian Maritime University, College of Artificial Intelligence. Tony Q.S. Quek is with the Singapore University of Technology and Design, Singapore 487372, and also with Yonsei Frontier Lab, Yonsei University, Seoul 03722, South Korea. Email: tonyquek@sutd.edu.sg.

*Corresponding authors: Boyuan Li (Email: l202311841010602@gs.zzu.edu.cn) and Zijian Li (Email: lizj@dlmu.edu.cn)

an offensive perspective, and demonstrate how attackers can exploit the distributed decision making process to enable more potent attacks by sending false reports in RIS-enhanced CSS. Such adversary-oriented analysis can provide some profound insights of strategic attack-defense interplay for CSS in mobile CRNs given the available resources, which motivates our imperative study.

B. Motivation and Contributions

To the best of our knowledge, there are still some shortcomings in current research on Byzantine attacks in CSS. Firstly, for analysis simplification, classical studies address only the first level of uncertainty accounts for the PU phenomenon as observed by the SUs, the Byzantine attack strategy has pioneered the design under the premise that the sensing report transmission between the SUs and the FC is always reliable [28], and the second level of uncertainty typically due to the report channel between the SUs and the FC has not fully considered. However, the real CRNs are typically affected by receiver noise, channel fading, and interference. Further, the sensing report transmission is important from security point of view since it is one of the most vulnerable phases to attacks. Secondly, from the attacking effectiveness viewpoint, many researchers have extensively arrived a typical conclusion that the channel-aware design is attractive given the available resources considering its inherent adaptivity [29]. In an intuitive sense, the channel-aware approach refers to the integration of the knowledge of the global ICSI, such as the false alarm probability and detection probability for all the SUs, into the design of attacking algorithms. However, wireless communication is naturally uncertain and time-varying due to effects that are not always amenable to modeling, such as high mobility in 5G environment. Thus, perfect ICSI estimation or prediction is clearly complexity intensive and time-consuming, which is computationally challenging that attackers will meet and overcome operating in an adversarial environment and stringent resource constraints, most notably the energy constraint with compromised SUs operating on irreplaceable battery supply. Finally, the channel-aware attacking algorithm is highly sensitive to ICSI estimation and prediction [30]. Most existing attack strategies under this theoretical framework have theoretical significance, but lack practical significance due to their very poor robustness when facing non-stationary communication channels. This paper aims to address the above mentioned issues and provide satisfactory solutions. The main contributions are summarized as follows:

- We propose a RIS-enhanced mobile CRN model subject to Byzantine attacks, where the report channel exhibit non-ideal characteristic.
- We develop the channel- and attack-aware hard decision fusion rules for the FC.
- We analyze and optimize the channel-aware Byzantine attacks from the perspective of our developed hard decision fusion rules.
- We reveal that as long as the fraction of the Byzantine nodes remains unchanged, the optimal attack strategy remains consistent even if the adversary is in the absence of any side information about the global ICSI.
- We develop extensive simulations to verify our theoretical analysis across a wide range of attacking scenarios.

It is worth pointed out that we pay all our attention towards distributed CSS under Byzantine attacks. The extension to hard decision fusion case in distributed wireless sensor networks (WSNs) under Byzantine attacks [31] is straightforward but not pursued here. Moreover, the ensuing analysis is not tailored uniquely for distributed CSS, and thus the results of this paper may be easily extended to include any distributed communication configuration impaired by Byzantine attacks [32]. Finally, most of the ensuing discussions can directly apply to multibit (soft) spectrum sensing output case [33], although it is not conceptually simple. Such analyses, however, will be under investigation in our future work.

C. Paper Organization

To the best of our knowledge, this is the first work investigating the hard decision fusion to design Byzantine attacks in RIS-enhanced CSS for mobile CRNs. The remainder of this work takes the following structure. Section II introduces related research work on attack methods. Section III proposes a CSS network architecture. Section IV develops the decision fusion rule for different scenarios. Section V analyzes the conditions that optimal attacks should satisfy when the optimal decision fusion rule is known. Section VI explores the conditions that optimal attacks should satisfy when the suboptimal decision fusion rule is known. Section VII provides some typical conclusions. Section IX verifies the effectiveness of the theoretical analysis through simulation experiments. Section X provides the conclusions and prospects for future research direction.

II. Related Work

We will discuss Byzantine attacks in Distributed WSNs, CSS, and Intelligent Cooperative Systems separately. Details of some related works are summarized in table I.

1) Byzantine Attacks in Distributed WSNs: In the early studies on Byzantine attacks, it was commonly assumed that the true observation information was completely known. For example, under this assumption, [34] formulated the problem as a Kullback-Leibler divergence (KLD) minimization and, through a "water-filling" procedure, derived the optimal attacking distributions for binary detection with quantized sensor observations. The study concluded that if the fraction of Byzantine nodes is less than half, the attackers cannot fully disrupt the system. However, although KLD is applicable for asymptotic analysis with a fixed number of sensors, it cannot directly provide accurate performance metrics. To overcome this limitation, [35] analyzed the impact of Byzantine nodes on distributed Bayesian detection and employed Chernoff information as the performance metric, systematically investigating the asymptotic behavior of such systems. Unlike [34], which only obtained the optimal blinding strategy, [35] further provided a closed-form expression for the optimal Byzantine attack. To overcome the limitations of the aforementioned asymptotic analyses, Kailkhura et al. directly investigated

TABLE I: SUMMARY OF RELATED WORKS ON BYZANTINE ATTACKS

Works	System Model	Performance Indicators	RIS Enhancement	Antenna Number		Required Priori Information		Type of attack
				PU	SU	P_{D_i}/P_{F_i}	True Sensing Information	
[34]	Binary Hypothesis Testing with M-ary Detection	KLD	✗	Single	Single	✔	✔	Static
[35]	Binary Hypothesis Testing with Hard Detection	Chernoff Information	✗	Single	Single	✔	✗	Static
[33]	Binary Hypothesis Testing with Quantized Detection	Modified Deflection Coefficient	✗	Single	Single	✔	✔	Static
[36]	Binary Hypothesis Testing with M-ary Detection	KLD	✗	Single	Single	✔	✗	Static
[37]	M-ary Hypothesis Testing with Hard Detection	Minimum Fraction of Byzantine Nodes	✗	Single	Single	✔	✔	Static
[38]	Binary Hypothesis Testing with Hard Detection	Global Detection performance	✗	Single	Single	✔	✗	Dynamic
[39]	Binary Hypothesis Testing with Hard Detection	Malicious User Revenue	✗	Single	Single	✔	✗	Dynamic
[40]	Binary Hypothesis Testing with Hard Detection	MI	✗	Single	Single	✗	✗	Dynamic
Ours	Binary Hypothesis Testing with Hard Detection	Reliability of Decision Fusion Statistics	✔	Single	M	✗	✗	Dynamic

the non-asymptotic scenario. Using the more direct probability of error as the performance metric, they derived the minimum fraction of attackers required to completely blind the system and provided closed-form solutions for the optimal attack strategies that maximize performance degradation when blinding is not achievable. In a subsequent study, [36] demonstrated, using the blinding probability, that uniform falsification can blind the FC, but only if the vast majority of nodes are compromised. To overcome this limitation, [37] evaluated attack probabilities at the sensor level and, under the assumptions of symmetric falsification and that Byzantine attackers have access to local output statistics, derived an attack strategy capable of blinding a single sensor with the minimum compromise probability.

2) Byzantine Attacks in CSS: In the context of CSS, the complexity of sensing data and the critical need for accurate spectrum allocation make the system particularly vulnerable to sophisticated attacks. Consequently, research in this area has focused on more advanced attack models, such as those employing M-ary quantized data frameworks. For instance, [33] conducted a more in-depth and systematic study of SSDF attacks under an M-ary quantized data framework based on the Neyman-Pearson criterion. It introduced a more sophisticated probabilistic SSDF attack model for CSS that the negative effect of defined probabilistic SSDF attack for the CSS with M-ary quantized data has been characterized, and the condition of the proposed SSDF attack model to make the FC completely incapable of inferring the status of PU has been derived. From a different perspective, [41] proposes a robust evaluation benchmark from a FC perspective that integrates extended sensing with actual transmission feedback. Practical factors such as channel imperfections and inference errors are incorporated to focus on the detection of inconsistent sensors, providing a complementary perspective to the traditional analysis of blinding attacks. Similarly, [42] examined attack strategies in the absence of any defense mechanisms, analyzing both attack strength and attack probability, and introduced a classical trust-based CSS algorithm to execute the attack strategies and evaluate the security of Byzantine attackers. These findings emphasize that a successful Byzantine strategy must carefully balance disruption and stealth, especially in systems employing trust or reputation mechanisms.

3) Byzantine Attacks in Intelligent Cooperative Systems: More recent research trends have shifted toward developing more realistic and adaptive attack models, particularly for intelligent cooperative systems. For example, [38] proposed a dominant cooperative probability attack (DCPA) model, which aims to reduce detectability and enhance stealthiness. This DCPA model contains auxiliary cooperative attackers (ACAs) that launch attacks and a dominant attacker (DA) that determines the ACAs' attack probability intervals according to their respective credibility. By introducing dynamic probability cooperation (i.e., a non-fixed attack pattern), the model can guide the network's decision-making process. Compared with traditional static attack strategies, DCPA exhibits both higher stealth and greater destructiveness. In another advanced application, [43] proposed a probabilistic attack model for federated learning: given a target distribution, the model precisely derives the minimal amount of data contamination required to reduce model performance to the desired level. This work is the first to quantify the attack cost-performance degradation-stealthiness trade-off in a computable form. As a result, attackers can perform local, targeted, and hardly noticeable poisoning at extremely low cost. These developments

TABLE II: DEFINITION OF THE KEY MATHEMATICAL NOTATIONS

Symbols	Meaning	Symbols	Meaning
$\mathbf{h}_{d_i}$	The channel matrices from the PU to the i-th SU.	$\mathbf{h}_r$	The channel matrices from the PU to the RIS.
$\mathbf{H}_i$	The channel matrices from the RIS to the i-th SU.	x_i	Received signal at the i-th SU.
$\mathbf{\Theta}$	Adjustment phase matrix for RIS systems.	p	The transmit power by the PU.
$\mathbf{s}$	The PU transmission symbol.	$\mathbf{n}_i$	Gaussian noise at the i-th SU.
$\mathbf{w}$	The receive beamformer to equalize the received signal.	f_s	The sampling frequency of the received signal.
τ_s	The available sensing time.	T	The signal sampling number ($T = \tau_s f_s$).
λ_i	The sensing threshold of the i-th SU.	γ	The average SNR of the PU measured at the i-th SU.
P_{D_i}	Detection probability of the i-th SU.	P_{F_i}	False alarm probability of the i-th SU.
α	The intensity of the attack.	Q	Q-function.
$P_{1,0}$	The probability that an honest (Byzantine) node sends 1 to the FC when its actual spectrum sensing result is 0.	$P_{0,1}$	The probability that an honest (Byzantine) node sends 0 to the FC when its actual spectrum sensing result is 1.

represent a broader trend towards an adversarial model that is both strategically complex and attuned to the dynamics of the environment. This trend is exemplified by [39], which applied an evolutionary game theory framework allowing attackers to adaptively choose their flipping probability based on payoff optimization. Further advancing this line of inquiry, [40] derived necessary and sufficient conditions for optimal attacks, showing that attackers can remain efficient and stealthy without CSI, thereby reducing energy consumption.

4) Novelty of Our Work: Compared with the previous works, the main novelty of this work is that the reliance on the global ICSI and decision fusion rules [34], [36], [44] in obtaining the Byzantine attacks is fully relaxed, and robust Byzantine attacks can be successfully implemented under stringent resource constraints over time-varying channels. Note here that, similar to the one considered here, a comprehensive coverage on Byzantine attacks in distributed WSNs can be found in [40]. The distinction lies in that [40] is presented from Mutual Information (MI) perspective without considering the RIS enhancement, whereas this work is developed based on decision fusion rules for RIS-enhanced CSS.

III. System Model

As depicted in Fig. 1, we consider a more prevailing model, as not studied extensively in the CSS literature but arguably more relevant to engineering applications. In our mobile CRN, there exists one PU with single antenna, I SUs with M antennas, one FC and an passive RIS with N reflecting elements, where $N > I$. A RIS deployed within the cell is responsible for assisting SUs' spectrum sensing [45]. Normal SUs perform a detection approach for spectrum sensing, then report sensing results to the FC by employing multiple relay nodes. The FC then make global decision based on information gathered from local SUs to infer the absence or presence of a spectrum hole. Some SUs may be compromised by attackers and report falsified sensing results. In addition, the key variables used in this paper is listed in Tables II.

A. Sensing Channel Model

The received signal for the i-th SU can be given by

$$\mathbf{x}_i = \left(\mathbf{h}_{d_i} + \mathbf{H}_i \mathbf{\Theta} \mathbf{h}_r\right) \sqrt{p}\mathbf{s} + \mathbf{n}_i, \tag{1}$$

where $\mathbf{h}_{d_i} \in \mathbb{C}^{M\times 1}$, $\mathbf{h}_r \in \mathbb{C}^{N\times 1}$ and $\mathbf{H}_i \in \mathbb{C}^{M\times N}$ denote the channel matrices from the PU to the i-th SU, PU to the RIS, and RIS to the i-th SU, respectively, where $i = 1, 2, \ldots, I$. p and $\mathbf{s}$ denote the PU's transmit power and information signal. The diagonal phase shift matrix for the RIS is defined as $\mathbf{\Theta} \triangleq \text{diag}\left[e^{j\theta_1}, \cdots, e^{j\theta_N}\right]$, where $\theta_n \in [0, 2\pi)$. $\mathbf{n}_i \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$ represents AWGN at the i-th SU. The SU employs a receive beamformer $\mathbf{w} \in \mathbb{C}^{1\times M}$ to equalize the received signal, where $\|\mathbf{w}\|^2 = 1$, as

$$\mathbf{x}_i = \mathbf{w}^H \mathbf{x}_i = \mathbf{w}^H \left(\mathbf{h}_{d_i} + \mathbf{H}_i \Theta \mathbf{h}_r\right) \sqrt{p}\mathbf{s} + \mathbf{w}^H \mathbf{n}_i. \tag{2}$$

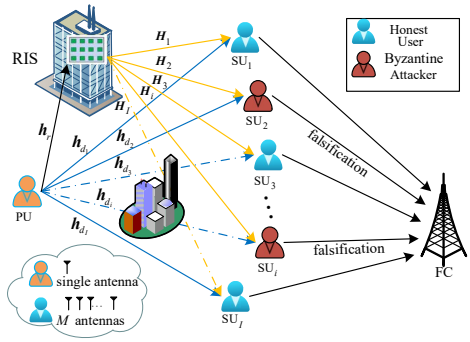


Fig. 1: RIS Enhanced Mobile CRNs under Byzantine Attacks.

Each SU has the ability to collect signal samples to determine whether the PU is active or not during the sensing interval. Therefore, we can rewrite the signal received by the i-th SU at the t-th sampling instant as

$$x_i(t) = \begin{cases} \mathbf{w}^H \mathbf{n}_i(t), & H_0 \\ \mathbf{w}^H [\left(\mathbf{h}_{d_i} + \mathbf{H}_i \mathbf{\Theta} \mathbf{h}_r\right) \sqrt{p}\mathbf{s}(t) + \mathbf{n}_i(t)], & H_1 \end{cases} \tag{3}$$

Energy detection is widely used in CRNs due to its simplicity and effectiveness [46]. It does not require any a priori knowledge of PU's signal and strict phase synchronization, so we only consider energy detection. Let f_s be the sampling rate and τ_s the sensing time, so the number of samples is $T = \tau_s f_s$. Now the test statistic for energy detection is

$$G(x) = \frac{1}{T} \sum_{t=1}^{T} |x_i(t)|^2, \quad t = 1, 2, \ldots, T, \tag{4}$$

with the decision rule

$$G(x) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_i, \quad i = 1, 2, \ldots, I. \tag{5}$$

where λ_i is the pre-determined decision threshold for the i-th SU. For passive RIS, according to energy detection, the test

statistic $G(x)$ can be written as follows

$$G(x)=\begin{cases}\frac{1}{T}\sum\limits_{t=1}^{T}\left|\mathbf{w}^{H}\mathbf{n}_i\left(t\right)\right|^2, & H_0\\ \frac{1}{T}\sum\limits_{t=1}^{T}\left|\mathbf{w}^{H}[(\mathbf{h}_{d_i}+\mathbf{H}_i\Theta\mathbf{h}_r)\sqrt{p}\mathbf{s}(t)+\mathbf{n}_i\left(t\right)]\right|^2, & H_1\end{cases} \tag{6}$$

For larger values of T, the multidimensional Central Limit Theorem [47] can be used to approximate $G(x)$, and modeled with a Gaussian distribution as follows

$$G(x)\sim\begin{cases}N\left(\sigma^2,\frac{\sigma^4}{T}\right), & H_0\\ N\left[p\left|\mathbf{w}^{H}\left(\mathbf{h}_{d_i}+\mathbf{H}_i\Theta\mathbf{h}_r\right)\right|+\sigma^2,\frac{\sigma^4}{T}(1+\gamma)^2\right], & H_1\end{cases} \tag{7}$$

where $\gamma=\frac{p\left|\mathbf{w}^{H}\left(\mathbf{h}_{d_i}+\mathbf{H}_i\Theta\mathbf{h}_r\right)\right|^2}{\sigma^2}$ is the instantaneous received SNR. Since $G(x)$ follows a Gaussian distribution, thus the local false alarm probability P_{F_i} and the local detection probability P_{D_i} can be respectively expressed as

$$P_{F_i}=Q\left[\left(\frac{\lambda_i}{\sigma^2}-1\right)\sqrt{T}\right], \tag{8}$$

and

$$P_{D_i}=Q\left[\left(\frac{\lambda_i}{\sigma^2}-\gamma-1\right)\sqrt{\frac{T}{(1+\gamma)^2}}\right]. \tag{9}$$

where $Q\left(x\right)=\frac{1}{\sqrt{2\pi}}\int_x^{\infty}\exp\left(-\frac{t^2}{2}\right)dt$ denotes the complementary cumulative distribution function of the standard Gaussian distribution.

B. Byzantine Attacks Model

We assume that each Byzantine attacker independently distort the local spectrum sensing reports of SUs to fulfill their malicious objectives. Specifically, we denote the probability that an honest (Byzantine) node sends 0 to the FC when its actual spectrum sensing result is 1 as $P_{0,1}$, while $P_{1,0}$ represents the probability that an honest (Byzantine) node sends 1 to the FC when its actual spectrum sensing result is 0. Denoting the manipulated data as u_i, we have

$$P(u_i=0\mid x_i=1)\triangleq\pi_{0,1}=\alpha P_{0,1} \tag{10}$$

and

$$P(u_i=1\mid x_i=0)\triangleq\pi_{1,0}=\alpha P_{1,0}, \tag{11}$$

where α denotes the fraction of the Byzantine nodes, i.e., the fraction of of SUs affected by the Byzantine attacks. Note here that we assume that each compromised node has the same attacking probabilities $P_{0,1}$ and $P_{1,0}$. However, our analytical results can be extended to the case where compromised nodes have different attacking probabilities.

C. Report Channel Model

We assume that each relay node adopts the Decode-and-Forward (DF) strategy. In this context, each relay transmission process can be equivalently modeled as a Binary Channel (BC). Let $\varepsilon_{i,j,0}$ and $\varepsilon_{i,j,1}$ denote the crossover probabilities of the j-th BC in the i-th path. Multiple BCs in serial configuration can be equivalent to a BC with crossover probabilities $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$ [48].

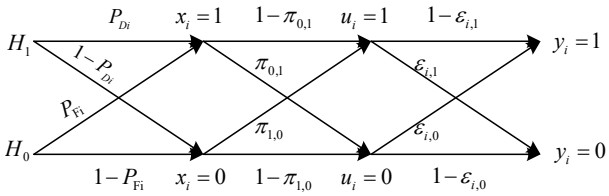


Fig. 2: System Model with Byzantine Attacks.

It is worth pointed out that, based on above discussion, the equivalent transmission model can be represented in Fig. 2. Note that we do not specify the sensing and report channel model or statistics of $\mathbf{h}_{d_i}$, $\mathbf{h}_r$, and $\mathbf{H}_i$. This need may arise due to the fact that we aim to provide a unifying framework in treating many of the unreliable communication channels in CSS of mobile CRNS. Also, we assume that the SUS are potential mobile. Corresponding, the spectrum sensing channel and report channel may involve nonstationary fading, and the channel matrices $\mathbf{h}_{d_i}$, $\mathbf{h}_r$, and $\mathbf{H}_i$ are fast time-varying. This renders the previous approach assuming stationary fading channel statistics impractical and useless. However, the notations have been removed for simplicity.

IV. Fusion Rules Under Byzantine Attacks

We assume the FC is aware of Byzantine attacks and knows the attack parameters, including α, $P_{1,0}$ and $P_{0,1}$. However, the FC cannot identify specific compromised nodes and thus considers each SU to be Byzantine node with the certain probability α. That is, a non-adversarial scenario is considered where Byzantine nodes do not actively confront the FC.

A. Optimal Decision Fusion Rule

Following from the ML criterion [49], we can have that the decision fusion statistic for the FC can be given as

$$\Lambda\left(\mathbf{y}\right)=\ln\frac{P(\mathbf{y}|H_1)}{P(\mathbf{y}|H_0)}=\sum_{i=1}^{I}\ln\frac{P(y_i|H_1)}{P(y_i|H_0)}=\sum_{i=1}^{I}\Lambda\left(y_i\right), \tag{12}$$

where $\mathbf{y}=[y_1,y_2,...,y_I]$ is the binary observation vector received at the FC from the I branches after $(J-1)$ hops. The contribution of y_i to the decision fusion statistic $\Lambda\left(\mathbf{y}\right)$ is denoted as $\Lambda\left(y_i\right)$, which can be expressed as (13) as shown at the bottom of the next page. When the output of the report channel $y_i=1$, $\Lambda\left(y_i\right)$ can be expressed as

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\left(1-\pi_{0,1}\right)\left(1-\varepsilon_{i,1}\right)+P_{D_i}\pi_{0,1}\varepsilon_{i,0}\\ +\left(1-P_{D_i}\right)\pi_{1,0}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)\varepsilon_{i,0}\end{array}}{\begin{array}{l}P_{F_i}\left(1-\pi_{0,1}\right)\left(1-\varepsilon_{i,1}\right)+P_{F_i}\pi_{0,1}\varepsilon_{i,0}\\ +\left(1-P_{F_i}\right)\pi_{1,0}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)\varepsilon_{i,0}\end{array}}. \tag{14}$$

When the output of the report channel $y_i=0$, $\Lambda\left(y_i\right)$ can be expressed as

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\left(1-\pi_{0,1}\right)\varepsilon_{i,1}+P_{D_i}\pi_{0,1}\left(1-\varepsilon_{i,0}\right)\\ +\left(1-P_{D_i}\right)\pi_{1,0}\varepsilon_{i,1}+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)\left(1-\varepsilon_{i,0}\right)\end{array}}{\begin{array}{l}P_{F_i}\left(1-\pi_{0,1}\right)\varepsilon_{i,1}+P_{F_i}\pi_{0,1}\left(1-\varepsilon_{i,0}\right)\\ +\left(1-P_{F_i}\right)\pi_{1,0}\varepsilon_{i,1}+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)\left(1-\varepsilon_{i,0}\right)\end{array}}. \tag{15}$$

Bringing (14) and (15) into (13) yields the optimal decision fusion statistic $\Lambda\left(\mathbf{y}\right)$, which can be expressed as (16), shown at the bottom of the next page.

B. Suboptimal Decision Fusion Rules under Ideal Local Sensing Channels

Under the ideal condition of the local sensing channel, we have $P_{D_i}=1$ and $P_{F_i}=0$, and (14) can be rewritten as

$$\Lambda\left(y_i\right)=\ln\frac{\left(1-\varepsilon_{i,1}\right)-\pi_{0,1}\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)}{\varepsilon_{i,0}+\pi_{1,0}\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)}. \tag{17}$$

Similarly, (15) can be rewritten as

$$\Lambda\left(y_i\right)=\ln\frac{\varepsilon_{i,1}+\pi_{0,1}\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)}{\left(1-\varepsilon_{i,0}\right)-\pi_{1,0}\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)}. \tag{18}$$

C. Suboptimal Fusion Rules when the SNR for the Multihop Relay Network is High

At this point, there is $\varepsilon_{i,0}\to 0$ and $\varepsilon_{i,1}\to 0$, then (14) can be rewritten as

$$\Lambda\left(y_i\right)=\ln\frac{P_{D_i}\left(1-\pi_{0,1}\right)+\left(1-P_{D_i}\right)\pi_{1,0}}{P_{F_i}\left(1-\pi_{0,1}\right)+\left(1-P_{F_i}\right)\pi_{1,0}}, \tag{19}$$

and (15) can be rewritten as

$$\Lambda\left(y_i\right)=\ln\frac{P_{D_i}\pi_{0,1}+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)}{P_{F_i}\pi_{0,1}+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)}. \tag{20}$$

D. Suboptimal Decision Fusion Rules when the SNR for the Multihop Relay Network is Low

When the SNR for each relay channel is low, each relay transmission process can be approximately modeled as a Binary Symmetric Channel (BSC) according to the fact that $\varepsilon_{i,j,0}\to 0.5$ and $\varepsilon_{i,j,1}\to 0.5$. In this context, we have that $\varepsilon_{i,0}=\varepsilon_{i,1}\approx 0.5$. Let $\delta_{i,j}=\log\frac{1-\varepsilon_{i,1}}{\varepsilon_{i,0}}$, following from the fact that [50]

$$\begin{aligned}\frac{1-\varepsilon_{i,1}}{\varepsilon_{i,0}}&=\frac{\frac{1}{2}\left[1+\prod_{j=1}^{J}\left(1-2\varepsilon_{i,j,1}\right)\right]}{1-\frac{1}{2}\left[1+\prod_{j=1}^{J}\left(1-2\varepsilon_{i,j,0}\right)\right]}\\&\approx\exp\left[\min_{1\le j\le J}\left(\delta_{i,j}\right)\right],\end{aligned} \tag{21}$$

we can directly rewrite (14) as (22), which is shown at the bottom of the next page. Under low channel SNR conditions, the exponential term in (22) can be accurately approximated by its first-order Taylor expansion without introducing unnecessary errors [51], we can immediately simplify (22) as:

$$\Lambda\left(y_i\right)\approx\ln\frac{1+\left[P_{D_i}\left(1-\pi_{0,1}\right)+\left(1-P_{D_i}\right)\pi_{1,0}\right]\min_{1\le j\le J}\left(\delta_{i,j}\right)}{1+\left[P_{F_i}\left(1-\pi_{0,1}\right)+\left(1-P_{F_i}\right)\pi_{1,0}\right]\min_{1\le j\le J}\left(\delta_{i,j}\right)}. \tag{23}$$

Using the fact that, $x\to 0$, $\log\left(1+x\right)\to x$, (23) can be further simplified as

$$\Lambda\left(y_i\right)\approx\left(1-\pi_{0,1}-\pi_{1,0}\right)\left(P_{D_i}-P_{F_i}\right)\min_{1\le j\le J}\left(\delta_{i,j}\right). \tag{24}$$

Similarly, when $y_i=0$, we can easily have that

$$\Lambda\left(y_i\right)\approx-\left(1-\pi_{0,1}-\pi_{1,0}\right)\left(P_{D_i}-P_{F_i}\right)\min_{1\le j\le J}\left(\delta_{i,j}\right). \tag{25}$$

V. Optimal Byzantine Attacks when the Optimal Decision Fusion Rule is Known

A. Optimal Small-Scale Byzantine Attack

For the convenience of analysis, we start by respectively modifying the optimal LLR-based decision fusion statistics given in (14) and (15) as

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{D_i}\right)\varepsilon_{i,0}\\-\left[P_{D_i}\pi_{0,1}-\left(1-P_{D_i}\right)\pi_{1,0}\right]\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}{\begin{array}{l}P_{F_i}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{F_i}\right)\varepsilon_{i,0}\\-\left[P_{F_i}\pi_{0,1}-\left(1-P_{F_i}\right)\pi_{1,0}\right]\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}, \tag{26}$$

and

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\varepsilon_{i,1}+\left(1-P_{D_i}\right)\left(1-\varepsilon_{i,0}\right)\\+\left[P_{D_i}\pi_{0,1}-\left(1-P_{D_i}\right)\pi_{1,0}\right]\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}{\begin{array}{l}P_{F_i}\varepsilon_{i,1}+\left(1-P_{F_i}\right)\left(1-\varepsilon_{i,0}\right)\\+\left[P_{F_i}\pi_{0,1}-\left(1-P_{F_i}\right)\pi_{1,0}\right]\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}. \tag{27}$$

We first pay our attention towards (26). Here, the fraction of Byzantine nodes satisfies $\alpha\le 0.5$, thus for any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, provided that $\pi_{0,1}=\pi_{1,0}$, then (26) can be reduced to

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{D_i}\right)\varepsilon_{i,0}\\-\pi_{0,1}\left(2P_{D_i}-1\right)\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}{\begin{array}{l}P_{F_i}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{F_i}\right)\varepsilon_{i,0}\\+\pi_{0,1}\left(1-2P_{F_i}\right)\left(1-\varepsilon_{i,0}-\varepsilon_{i,1}\right)\end{array}}. \tag{28}$$

Further, it is evident from (28) that $2P_{D_i}-1>0$ and $1-2P_{F_i}>0$. If both $\pi_{0,1}$ and $\pi_{1,0}$ reach their maximum value α, i.e., $P_{0,1}=P_{1,0}=1$, the numerator of (28) reaches its minimum while the denominator attains its maximum. Considering that the numerator is larger than the denominator, then the magnitude of $\Lambda(y_i)$ achieves the minimum at this point, and thus results in the lowest reliability of $\Lambda(y_i)$.

$$\Lambda\left(y_i\right)=\ln\frac{\begin{array}{l}P_{D_i}\left(1-\pi_{0,1}\right)P\left(y_i|u_i=1\right)+P_{D_i}\pi_{0,1}P\left(y_i|u_i=0\right)\\+\left(1-P_{D_i}\right)\pi_{1,0}P\left(y_i|u_i=1\right)+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)P\left(y_i|u_i=0\right)\end{array}}{\begin{array}{l}P_{F_i}\left(1-\pi_{0,1}\right)P\left(y_i|u_i=1\right)+P_{F_i}\pi_{0,1}P\left(y_i|u_i=0\right)\\+\left(1-P_{F_i}\right)\pi_{1,0}P\left(y_i|u_i=1\right)+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)P\left(y_i|u_i=0\right)\end{array}}. \tag{13}$$

$$\begin{aligned}\Lambda\left(\mathbf{y}\right)=&\sum_{i:y_i=1}\ln\frac{P_{D_i}\left(1-\pi_{0,1}\right)\left(1-\varepsilon_{i,1}\right)+P_{D_i}\pi_{0,1}\varepsilon_{i,0}+\left(1-P_{D_i}\right)\pi_{1,0}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)\varepsilon_{i,0}}{P_{F_i}\left(1-\pi_{0,1}\right)\left(1-\varepsilon_{i,1}\right)+P_{F_i}\pi_{0,1}\varepsilon_{i,0}+\left(1-P_{F_i}\right)\pi_{1,0}\left(1-\varepsilon_{i,1}\right)+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)\varepsilon_{i,0}}\\&+\sum_{i:y_i=0}\ln\frac{P_{D_i}\left(1-\pi_{0,1}\right)\varepsilon_{i,1}+P_{D_i}\pi_{0,1}\left(1-\varepsilon_{i,0}\right)+\left(1-P_{D_i}\right)\pi_{1,0}\varepsilon_{i,1}+\left(1-P_{D_i}\right)\left(1-\pi_{1,0}\right)\left(1-\varepsilon_{i,0}\right)}{P_{F_i}\left(1-\pi_{0,1}\right)\varepsilon_{i,1}+P_{F_i}\pi_{0,1}\left(1-\varepsilon_{i,0}\right)+\left(1-P_{F_i}\right)\pi_{1,0}\varepsilon_{i,1}+\left(1-P_{F_i}\right)\left(1-\pi_{1,0}\right)\left(1-\varepsilon_{i,0}\right)}.\end{aligned} \tag{16}$$

Similarly, substituting $\pi_{0,1} = \pi_{1,0}$ into (27) yields

$$\Lambda(y_i) = \ln \frac{\begin{array}{c} P_{D_i}\varepsilon_{i,1} + (1-P_{D_i})(1-\varepsilon_{i,0}) \\ +\pi_{0,1}(2P_{D_i}-1)(1-\varepsilon_{i,0}-\varepsilon_{i,1}) \end{array}}{\begin{array}{c} P_{F_i}\varepsilon_{i,1} + (1-P_{F_i})(1-\varepsilon_{i,0}) \\ -\pi_{0,1}(1-2P_{F_i})(1-\varepsilon_{i,0}-\varepsilon_{i,1}) \end{array}}. \tag{29}$$

If both $\pi_{0,1}$ and $\pi_{1,0}$ reach their maximum value α, i.e., $P_{0,1} = P_{1,0} = 1$, the numerator of (29) reaches its maximum while the denominator attains its minimum. Considering that the numerator is smaller than the denominator, then the magnitude of $\Lambda(y_i)$ achieves the minimum, and also results in the lowest reliability of $\Lambda(y_i)$.

Finally, it is crucial to note from (28) and (29) that the fraction of Byzantine nodes α, the probabilities $P_{0,1}$ and $P_{1,0}$, do not independently affect the decision fusion statistic $\Lambda(y_i)$. In fact, they are intertwined with the detection probability P_{D_i}, the false alarm probability P_{F_i}, and the report channel ICSI $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, and affect the magnitude of $\Lambda(y_i)$ in an interdependent manner. For example, in the case of the Always-Yes (AY) attack, we have $P_{0,1} = 0$ and $P_{1,0} = 1$. Correspondingly, we have $\pi_{0,1} = 0$ and $\pi_{1,0} = \alpha$, leading to

$$\begin{cases} P_{D_i}\pi_{0,1} - (1-P_{D_i})\pi_{1,0} = (P_{D_i}-1)\alpha < 0 \\ P_{F_i}\pi_{0,1} - (1-P_{F_i})\pi_{1,0} = (P_{F_i}-1)\alpha < 0 \\ |P_{D_i}\pi_{0,1} - (1-P_{D_i})\pi_{1,0}| < |P_{F_i}\pi_{0,1} - (1-P_{F_i})\pi_{1,0}|. \end{cases} \tag{30}$$

Following the result in (30), analysis of (26) and (27) shows that, compared to the no-attack case ($P_{0,1} = 0$, $P_{1,0} = 0$), the increment of the numerator in (26) is less than that of the denominator, while in (27) the decreasment of the numerator is less than that of the denominator, and the numerator and denominator merge and gradually become identical. Consequently, both the magnitude and reliability of $\Lambda(y_i)$ are degraded. Similarly, for the Always-No (AN) attack, $P_{0,1} = 1$ and $P_{1,0} = 0$, there are $\pi_{0,1} = \alpha$ and $\pi_{1,0} = 0$, leading to

$$\begin{cases} P_{D_i}\pi_{0,1} - (1-P_{D_i})\pi_{1,0} = \alpha P_{D_i} > 0, \\ P_{F_i}\pi_{0,1} - (1-P_{F_i})\pi_{1,0} = \alpha P_{F_i} > 0, \\ |P_{D_i}\pi_{0,1} - (1-P_{D_i})\pi_{1,0}| > |P_{F_i}\pi_{0,1} - (1-P_{F_i})\pi_{1,0}|. \end{cases} \tag{31}$$

The same observation can be easily drawn as we depict for AY attack. Following the results reported in (28) and (29), we can immediately achieve an argument that for small-scale attacks, the optimal strategy is the **Always-False (AF) attack**. This is achieved by setting the flip probabilities to their maximum, i.e., $\pi_{0,1} = \pi_{1,0} = \alpha$ (corresponding to $P_{0,1} = P_{1,0} = 1$).

B. Optimal Large-Scale Byzantine Attack

When the fraction of Byzantine nodes satisfies $\alpha > 0.5$, for any given values of P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, substituting $\pi_{0,1} + \pi_{1,0} = 1$ into (26) and (27) yields the following results. When the output of the report channel $y_i = 1$

$$\Lambda(y_i) = \ln \frac{\varepsilon_{i,0} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})}{\varepsilon_{i,0} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})} = 0, \tag{32}$$

when the output of the report channel $y_i = 0$

$$\Lambda(y_i) = \ln \frac{1-\varepsilon_{i,0}-\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})}{1-\varepsilon_{i,0}-\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})} = 0. \tag{33}$$

Following the results reported in (32) and (33), we have that when $\alpha > 0.5$, for any P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, any attack strategy satisfying $\pi_{0,1} + \pi_{1,0} = 1$ forces $\Lambda(y_i)$ to 0, making it completely unreliable. In this context, the optimal attack is the **Random (RD) attack**, which satisfies $\pi_{0,1} + \pi_{1,0} = 1$. The explicit attack strategy is implicitly not unique.

VI. Optimal Byzantine Attacks when the suboptimal Decision fusion rule is known

A. Optimal Small-Scale Byzantine Attack under Ideal Local Sensing Channels

For the suboptimal fusion rule under ideal sensing channel, when the fraction of Byzantine attack satisfies $\alpha \leq 0.5$, for any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, provided that $\pi_{0,1} = \pi_{1,0}$, (17) can be changed to

$$\Lambda(y_i) = \ln \frac{1-[\varepsilon_{i,1} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]}{\varepsilon_{i,0} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})}, \tag{34}$$

and (18) can be changed to

$$\Lambda(y_i) = \ln \frac{\varepsilon_{i,1} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})}{1-[\varepsilon_{i,0} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]}. \tag{35}$$

The derivative of (34) with respect to $\pi_{0,1}$ yields

$$\frac{\partial\Lambda(y_i)}{\partial\pi_{0,1}} = \frac{-(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1+\varepsilon_{i,0}-\varepsilon_{i,1})}{[\varepsilon_{i,0}+\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]\{1-[\varepsilon_{i,1}+\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]\}}. \tag{36}$$

For the denominator given in (36), we can have that

$$\begin{cases} \varepsilon_{i,0} < \varepsilon_{i,0} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1}) < \dfrac{1}{2}, \\ \dfrac{1}{2} < 1 - \left[\varepsilon_{i,1} + \pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})\right] < 1-\varepsilon_{i,1}. \end{cases} \tag{37}$$

Then we can immediately get that $\frac{\partial\Lambda(y_i)}{\partial\pi_{0,1}} < 0$. With the above result, we can easily obtain that since the decision fusion statistic $\Lambda(y_i)$ is positive and monotonically decreases as $\pi_{0,1}$ increase, its magnitude $|\Lambda(y_i)|$ is minimized when $\pi_{0,1}$ reach maximum, i.e., $\pi_{0,1} = \pi_{1,0} = \alpha$ (corresponding to $P_{0,1} = P_{1,0} = 1$). This magnitude minimization directly corresponds to the lowest decision reliability and thus the optimal attack. Similarly, the derivative of (35) with respect to $\pi_{0,1}$ yields

$$\frac{\partial\Lambda(y_i)}{\partial\pi_{0,1}} = \frac{(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1-\varepsilon_{i,0}+\varepsilon_{i,1})}{[\varepsilon_{i,1}+\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]\{1-[\varepsilon_{i,0}+\pi_{0,1}(1-\varepsilon_{i,0}-\varepsilon_{i,1})]\}}. \tag{38}$$

$$\begin{aligned} \Lambda(y_i) &= \ln \frac{P_{D_i}(1-\pi_{0,1})(1-\varepsilon_{i,1}) + P_{D_i}\pi_{0,1}\varepsilon_{i,0} + (1-P_{D_i})\pi_{1,0}(1-\varepsilon_{i,1}) + (1-P_{D_i})(1-\pi_{1,0})\varepsilon_{i,0}}{P_{F_i}(1-\pi_{0,1})(1-\varepsilon_{i,1}) + P_{F_i}\pi_{0,1}\varepsilon_{i,0} + (1-P_{F_i})\pi_{1,0}(1-\varepsilon_{i,1}) + (1-P_{F_i})(1-\pi_{1,0})\varepsilon_{i,0}} \\ &\approx \ln \frac{[P_{D_i}\pi_{0,1} + (1-P_{D_i})(1-\pi_{1,0})]\frac{\varepsilon_{i,0}}{\varepsilon_{i,1}} + [P_{D_i}(1-\pi_{0,1}) + (1-P_{D_i})\pi_{1,0}]\exp[\min_{1\leq j\leq J}(\delta_{i,j})]}{[P_{F_i}\pi_{0,1} + (1-P_{F_i})(1-\pi_{1,0})]\frac{\varepsilon_{i,0}}{\varepsilon_{i,1}} + [P_{F_i}(1-\pi_{0,1}) + (1-P_{F_i})\pi_{1,0}]\exp[\min_{1\leq j\leq J}(\delta_{i,j})]}. \end{aligned} \tag{22}$$

Then we can get that $\frac{\partial \Lambda(y_i)}{\partial \pi_{0,1}} > 0$. Considering that the decision fusion statistic $\Lambda(y_i)$ is a negative, monotonically increase as $\pi_{0,1}$ increase, and its magnitude $|\Lambda(y_i)|$ is minimized when when $\pi_{0,1}$ reach maximum. This configuration, i.e., $\pi_{0,1} = \pi_{1,0} = \alpha$, achieves the optimal attack by minimizing the decision reliability.

B. Optimal Small-Scale Byzantine Attack when the SNR for the Multihop Relay Network is High

For any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, provided that $\pi_{0,1} = \pi_{1,0}$, (19) can be rewritten as

$$\Lambda\left(y_i\right) = \ln \frac{P_{D_i} - \left(2P_{D_i} - 1\right)\pi_{0,1}}{P_{F_i} + \left(1 - 2P_{F_i}\right)\pi_{0,1}}, \tag{39}$$

and (20) can be changed to

$$\Lambda\left(y_i\right) = \ln \frac{1 - \left[P_{D_i} - \left(2P_{D_i} - 1\right)\pi_{0,1}\right]}{1 - \left[P_{F_i} - \left(2P_{F_i} - 1\right)\pi_{0,1}\right]}. \tag{40}$$

The derivative of (39) with respect to $\pi_{0,1}$ yields

$$\frac{\partial \Lambda\left(y_i\right)}{\partial \pi_{0,1}} = -\left(\frac{2P_{D_i} - 1}{P_{D_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1}} + \frac{1 - 2P_{F_i}}{P_{F_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1}}\right). \tag{41}$$

Considering that

$$\begin{cases} 2P_{D_i} - 1 > 0, \\ 1 - 2P_{F_i} > 0, \\ P_{D_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1} > 0, \\ P_{F_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1} > 0, \end{cases} \tag{42}$$

then we can have that $\frac{\partial \Lambda(y_i)}{\partial \pi_{0,1}} < 0$. For small-scale attacks, analysis shows that the fusion statistic $\Lambda(y_i)$ is a positive and monotonically decreasing function with respect to $\pi_{0,1}$. Therefore, its magnitude $|\Lambda(y_i)|$ is minimized when $\pi_{0,1}$ reach maximum, i.e., $\pi_{0,1} = \pi_{1,0} = \alpha$. Similarly, taking the derivative of (40) with respect to $\pi_{0,1}$ yields

$$\frac{\partial \Lambda\left(y_i\right)}{\partial \pi_{0,1}} = \frac{2P_{D_i} - 1}{P_{D_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1}} + \frac{1 - 2P_{F_i}}{P_{F_i}\left(1 - 2\pi_{0,1}\right) + \pi_{0,1}}. \tag{43}$$

As indicated in (43), we can still easily derive that when $\pi_{0,1} = \pi_{1,0} = \alpha$, the magnitude $|\Lambda(y_i)|$ is minimized, leading to the lowest reliability of $\Lambda(y_i)$.

C. Optimal Small-Scale Byzantine Attack when the SNR for the Multihop Relay Network is Low

In this context, if substituting $\pi_{0,1} = \pi_{1,0}$ into (24) and (25), we can respectively have

$$\Lambda\left(y_i\right) = \left(1 - 2\pi_{0,1}\right)\left(P_{D_i} - P_{F_i}\right) \min_{1 \le j \le J}\left(\delta_{i,j}\right), \tag{44}$$

and

$$\Lambda\left(y_i\right) \approx -\left(1 - 2\pi_{0,1}\right)\left(P_{D_i} - P_{F_i}\right) \min_{1 \le j \le J}\left(\delta_{i,j}\right). \tag{45}$$

Further, if both $\pi_{0,1}$ and $\pi_{1,0}$ reach their maximum α, it is straightforward to observe that $(1-2\pi_{0,1})$ attains its minimum. Consequently, (44) and (45) respectively reach their minimum and maximum. This implies that the magnitude of the fusion statistic $\Lambda(y_i)$ is minimized, thereby leading to the lowest possible reliability of $\Lambda(y_i)$.

D. Optimal Large-Scale Byzantine Attack with Ideal Local Channel

For any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$, if $\pi_{0,1}$ and $\pi_{1,0}$ satisfy $\pi_{0,1} + \pi_{1,0} = 1$, then (17) can be revised as

$$\Lambda\left(y_i\right) = \ln \frac{1 - \varepsilon_{i,1} - \pi_{0,1}\left(1 - \varepsilon_{i,0} - \varepsilon_{i,1}\right)}{\varepsilon_{i,0} + \left(1 - \pi_{0,1}\right)\left(1 - \varepsilon_{i,0} - \varepsilon_{i,1}\right)} = 0, \tag{46}$$

and (18) can be revised as

$$\Lambda\left(y_i\right) = \ln \frac{\varepsilon_{i,0} + \pi_{0,1}\left(1 - \varepsilon_{i,0} - \varepsilon_{i,1}\right)}{1 - \varepsilon_{i,1} - \left(1 - \pi_{0,1}\right)\left(1 - \varepsilon_{i,0} - \varepsilon_{i,1}\right)} = 0. \tag{47}$$

It is evident that for any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$, and $\varepsilon_{i,1}$, the magnitude of the fusion statistic $\Lambda(y_i)$ is always 0, indicating its reliability is minimized.

E. Optimal Large-Scale Byzantine Attack when the SNR for the Multihop Relay Network is High

If we bring $\pi_{0,1} + \pi_{1,0} = 1$ into the suboptimal fusion rule (19) and (20), we can respectively have

$$\Lambda\left(y_i\right) = \ln \frac{P_{D_i}\left(1 - \pi_{0,1}\right) + \left(1 - P_{D_i}\right)\pi_{1,0}}{P_{F_i}\left(1 - \pi_{0,1}\right) + \left(1 - P_{F_i}\right)\pi_{1,0}} = 0, \tag{48}$$

and

$$\Lambda\left(y_i\right) = \ln \frac{P_{D_i}\pi_{0,1} + \left(1 - P_{D_i}\right)\left(1 - \pi_{1,0}\right)}{P_{F_i}\pi_{0,1} + \left(1 - P_{F_i}\right)\left(1 - \pi_{1,0}\right)} = 0. \tag{49}$$

It is evident that for any given P_{D_i}, P_{F_i} $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$ the magnitude of the fusion statistic $\Lambda\left(y_i\right)$ is always 0, indicating that the reliability of the fusion statistic is minimized.

F. Optimal Large-Scale Byzantine Attack when the SNR for the Multihop Relay Network is Low

If we bring $\pi_{0,1} + \pi_{1,0} = 1$ into (24) and (25), we can respectively have

$$\Lambda\left(y_i\right) = \left(1 - \pi_{0,1} - \pi_{1,0}\right)\left(P_{D_i} - P_{F_i}\right) \min_{1 \le j \le J}\left(\delta_{i,j}\right) = 0, \tag{50}$$

and

$$\Lambda\left(y_i\right) \approx -\left(1 - \pi_{0,1} - \pi_{1,0}\right)\left(P_{D_i} - P_{F_i}\right) \min_{1 \le j \le J}\left(\delta_{i,j}\right) = 0. \tag{51}$$

Similarly, at this point, the decision fusion statistic $\Lambda\left(y_i\right)$ is 0 for any given P_{D_i}, P_{F_i}, $\varepsilon_{i,0}$ and $\varepsilon_{i,1}$.

The above interpretation cognizant of the fact that the principle of the Byzantine attack is to minimize the magnitude of the decision fusion statistic $\Lambda(y_i)$, and this magnitude minimization could significantly reduce the reliability of the information gathered from local SUs by the FC. This is not a coincidence: the attackers is to make the data that the FC receives from the SUs such that no information is conveyed. Thus, the FC becomes blind if the conditional probabilities given in (12) are identical. In such a scenario, the best that the FC can do is to make decisions solely based on the priors, resulting in the most degraded performance at the FC. Based on the above discussion, we derive optimal Byzantine attack strategies for both large-scale and small-scale scenarios, as summarized in Algorithm 1.

Algorithm 1 Optimal Byzantine Attack Algorithm

Input: I (number of SUs); x (PU status after SU decision, $x_i \in \{0,1\}$);
  α (attack intensity); $P_{0,1}$; $P_{1,0}$.
Output: u (attack result)
if $\alpha \leq 0.5$ **then**
  (AF attack)
  for $i = 1$ **to** I **do**
    if $i \in$ Byzantine_nodes **then**
      $u_i \leftarrow x_i \oplus 1$
    else
      $u_i \leftarrow x_i$
    end if
  end for
else
  (satisfy $\alpha(P_{0,1} + P_{1,0}) = 1$*)*
  for $i = 1$ **to** I **do**
    if $i \in$ Byzantine_nodes **then**
      $u_i \leftarrow \mathrm{BC}(P_{0,1}, P_{1,0}, x_i)$
    else
      $u_i \leftarrow x_i$
    end if
  end for
end if
return u

VII. Discussion

As shown in the above analysis of the optimal Byzantine attack under both optimal and suboptimal decision fusion rules, we find that the attacker can achieve the optimal attack effect through a specific attack strategy regardless of whether the decision fusion rule is known or not. The key observations are summarized as follows:

1) For small-scale Byzantine attacks, the optimal attack strategy is the AF attack, i.e., flipping the local spectrum sensing report so as to minimize the magnitude of the decision fusion statistic $\Lambda(y_i)$ from the attacker perspective. This magnitude reduction degrades the global decision reliability of the FC.
2) For large-scale Byzantine attacks, the optimal attack strategy is to flip local spectrum sensing report extensively such that the decision fusion statistic $\Lambda(y_i)$ is driven to 0. In this context, the FC becomes incapable of recovering any meaningful information from the corrupted received spectrum sensing reports, resulting in a complete failure of its decision-making.
3) The attacker can achieve optimal impact even when no tangible ICSI is available, simply by selecting an appropriate strategy based on the fraction of the Byzantine nodes. Thus, the reliance on the global ICSI and decision fusion rules in obtaining the Byzantine attacks can be relaxed, and the Byzantine attacks can be successfully implemented over time-varying mobile channels without any a prior knowledge of decision fusion rules or ICSI.

VIII. Attack performance evaluation

According to our previous analysis of optimal Byzantine attack strategies and their impact on decision fusion, this section conducts a systematic evaluation of typical attack strategies. Based on their characteristics, these attacks can be categorized into the following four types [40]:

1) AN attack: data 0 remains unchanged, data 1 is flipped to 0;
2) AY attack: data 1 remains unchanged, data 0 is flipped to 1;
3) AF attack: data 0 is flipped to 1, data 1 is flipped to 0;
4) RD attack: data 0 is randomly flipped to 1, data 1 is randomly flipped to 0.

A. Performance Evaluation when $y_i = 1$

In order to evaluate the overall attack performance we first evaluate the performance of individual branches, if the attack performance of each branch is consistent, it can also indicate the strength of the attack performance of different attack methods on the overall system.

When $y_i = 1$, (26) is reduced to

$$\Lambda(y_i) = \ln \frac{A}{B}, \tag{52}$$

where A and B are

$$\begin{cases} A = P_{D_i}(1-\varepsilon_{i,1}) + (1-P_{D_i})\varepsilon_{i,0} \\ \quad - [P_{D_i}\pi_{0,1} - (1-P_{D_i})\pi_{1,0}](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ B = P_{F_i}(1-\varepsilon_{i,1}) + (1-P_{F_i})\varepsilon_{i,0} \\ \quad - [P_{F_i}\pi_{0,1} - (1-P_{F_i})\pi_{1,0}](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{53}$$

The four attack modes can be represented as follows.

AN attack:

$$\begin{cases} A_{AN} = \varepsilon_{i,0} + P_{D_i}(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1-\alpha) \\ B_{AN} = \varepsilon_{i,0} + P_{F_i}(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1-\alpha). \end{cases} \tag{54}$$

AY attack:

$$\begin{cases} A_{AY} = \varepsilon_{i,0} + [\alpha + (1-\alpha)P_{D_i}](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ B_{AY} = \varepsilon_{i,0} + [\alpha + (1-\alpha)P_{F_i}](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{55}$$

AF Attack:

$$\begin{cases} A_{AF} = \varepsilon_{i,0} + \{\alpha - (2\alpha-1)P_{D_i}\}(1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ B_{AF} = \varepsilon_{i,0} + \{\alpha - (2\alpha-1)P_{F_i}\}(1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{56}$$

RF Attack:

$$\begin{cases} A_{RD} = \varepsilon_{i,0} + [P_{D_i} + P_{1,0} - P_{D_i}\alpha(P_{0,1}+P_{1,0})](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ B_{RD} = \varepsilon_{i,0} + [P_{D_i} + P_{1,0} - P_{F_i}\alpha(P_{0,1}+P_{1,0})](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{57}$$

1) AN Attacks and AY Attacks: For the AN attack, the numerator is approximately proportional to: $P_{D_i}(1-\alpha)$. The denominator is approximately proportional to: $P_{F_i}(1-\alpha)$. Therefore, the overall decision statistic is proportional to: $(P_{D_i} - P_{F_i})(1-\alpha)$.

Similarly, for the AY attack, the numerator is approximately proportional to: $\alpha + (1-\alpha)P_{D_i}$. The denominator is approximately proportional to: $\alpha + (1-\alpha)P_{F_i}$. Therefore, the overall decision statistic is proportional to: $(P_{D_i} - P_{F_i})(1-\alpha)$. Thus, both the AN and AY attacks have the same overall decision statistic, proportional to: $(P_{D_i} - P_{F_i})(1-\alpha)$. Therefore, whether small or large-scale, the attack performance of AN and AY attacks is considered identical.

2) AF Attacks and AN Attacks: According to (56), for the AF attack, the numerator is approximately proportional to: $[\alpha-(2\alpha-1)P_{D_i}]$. The denominator is approximately proportional to: $[\alpha-(2\alpha-1)P_{F_i}]$. Therefore, the overall decision statistic is proportional to: $(P_{D_i}-P_{F_i})(2\alpha-1)$.

Since it is known that $P_{D_i}-P_{F_i}>0$, comparing $(1-\alpha)$ and $|2\alpha-1|$ can reveal which attack results in a smaller magnitude for $\Lambda(y_i)$.

- **Small-scale attack**: When $\alpha<0.5$, it can be concluded that under small-scale attacks, the AF attack outperforms the AN attack.
- **Large-scale attack**: When $\frac{1}{2}\leq\alpha\leq\frac{2}{3}$, the AF attack performs better. When $\frac{2}{3}\leq\alpha\leq 1$, the AN or AY attacks perform better.

3) RF Attack and AN Attack: For the RF attack, the numerator is approximately proportional to:

$[P_{D_i}+\alpha P_{1,0}-P_{D_i}\alpha(P_{0,1}+P_{1,0})]$.

The denominator is approximately proportional to:

$[P_{F_i}+\alpha P_{1,0}-P_{F_i}\alpha(P_{0,1}+P_{1,0})]$.

Therefore, the overall decision statistic is proportional to:

$(P_{D_i}-P_{F_i})[\alpha(P_{0,1}+P_{1,0})-1]$.

Performance comparison of RD and AN Attacks, we can use $|\alpha(P_{0,1}+P_{1,0})-1|$ and $(1-\alpha)$ to determine which attack results in a smaller magnitude for $\Lambda(y_i)$.

- **Small-scale attack**: Since $\alpha<0.5$, comparing $1-\alpha(P_{0,1}+P_{1,0})$ and $(1-\alpha)$, we can conclude that: When $0\leq P_{0,1}+P_{1,0}\leq 1$, the AN attack performs better. When $1\leq P_{0,1}+P_{1,0}\leq 2$, the RD attack performs better.
- **Large-scale attack**: We have $|\alpha(P_{0,1}+P_{1,0})-1|=\alpha(P_{0,1}+P_{1,0})-1$. When $0\leq P_{0,1}+P_{1,0}\leq 1$, it is concluded that $(1-\alpha)\geq|\alpha(P_{0,1}+P_{1,0})-1|$, meaning that the performance of AN or AY attacks is always better than that of the AF attack. When $1\leq P_{0,1}+P_{1,0}\leq 2$, setting $(1-\alpha)=\alpha(P_{0,1}+P_{1,0})-1$, we get: $P_{0,1}+P_{1,0}=\frac{2}{\alpha}-1$. When $1\leq P_{0,1}+P_{1,0}\leq\frac{2}{\alpha}-1$, the RD attack performs better. When $\frac{2}{\alpha}-1\leq P_{0,1}+P_{1,0}\leq 2$, the AN attack performs better.

4) RD Attack and AF Attack: To compare the performance of the RD attack with the AF attack, we can use $|\alpha(P_{0,1}+P_{1,0})-1|$ and $(2\alpha-1)$ to determine which attack has better performance.

- **Small-scale attack**: When $0\leq P_{0,1}+P_{1,0}\leq 1$, since AF attacks require $P_{0,1}=P_{1,0}=1$, the AF attack does not exist at this point; when $1\leq P_{0,1}+P_{1,0}\leq 2$, the AF attack performs better.
- **Large-scale attack**: When $1\leq P_{0,1}+P_{1,0}\leq\frac{2}{\alpha}-2$, the AF attack performs better; when $\frac{2}{\alpha}-2\leq P_{0,1}+P_{1,0}\leq 2$, the RD attack performs better.

B. Performance Evaluation when $y_i=0$

When $y_i=0$, (27) can be reduced to

$$\Lambda(y_i)=\ln\frac{C}{D}, \tag{58}$$

where C and D are

$$\begin{cases} C=P_{D_i}\varepsilon_{i,1}+(1-P_{D_i})(1-\varepsilon_{i,0}) \\ \quad +[P_{D_i}\pi_{0,1}-(1-P_{D_i})\pi_{1,0}](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ D=P_{F_i}\varepsilon_{i,1}+(1-P_{F_i})(1-\varepsilon_{i,0}) \\ \quad +[P_{F_i}\pi_{0,1}-(1-P_{F_i})\pi_{1,0}](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{59}$$

Substituting the four attack patterns at this point yields:

AN attack:

$$\begin{cases} C_{AN}=(1-\varepsilon_{i,0})-P_{D_i}(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1-\alpha) \\ D_{AN}=(1-\varepsilon_{i,0})-P_{F_i}(1-\varepsilon_{i,0}-\varepsilon_{i,1})(1-\alpha). \end{cases} \tag{60}$$

AY attack:

$$\begin{cases} C_{AY}=(1-\varepsilon_{i,0})-[\alpha+(1-\alpha)P_{D_i}](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ D_{AY}=(1-\varepsilon_{i,0})-[\alpha+(1-\alpha)P_F](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{61}$$

AF attack:

$$\begin{cases} C_{AF}=(1-\varepsilon_{i,0})-\{\alpha-(2\alpha-1)P_{D_i}\}(1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ D_{AF}=(1-\varepsilon_{i,0})-\{\alpha-(2\alpha-1)P_{F_i}\}(1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{62}$$

RD attack:

$$\begin{cases} C_{RD}=(1-\varepsilon_{i,0})-[P_{D_i}+\alpha P_{1,0}-P_{D_i}\alpha(P_{0,1}+P_{1,0})](1-\varepsilon_{i,0}-\varepsilon_{i,1}) \\ D_{RD}=(1-\varepsilon_{i,0})-[P_{F_i}+\alpha P_{1,0}-P_{F_i}\alpha(P_{0,1}+P_{1,0})](1-\varepsilon_{i,0}-\varepsilon_{i,1}). \end{cases} \tag{63}$$

Similar to the analysis when $y_i=1$, the strength of the attack performance of different attack methods can also be obtained for different scales of attack intensity when $y_i=0$. However, the conclusions drawn are consistent with those for $y_i=1$, and the specific analyses are shown in the Appendix A.

TABLE III: Parameters Used in Simulations

Parameter	Detailed description
Antenna Number for PU	1
Number of SUs I	8, 10, 12
Antenna Number for each SU M	4, 6, 8
Number of RIS reflective elements N	5, 7, 9
Channel condition $\mathbf{h}_{d_i}$, $\mathbf{h}_r$, and $\mathbf{H}_i$	Rayleigh fading with normalised average power
Noise condition $\mathbf{n}_i(t)$	Complex AWGN
Sample number T	50
Phase shift of RIS $\mathbf{\Theta}$	Randomly generated
Number of hops J	4, 6, 8
Information sequence length	504 bits
Number of simulation cycles	Get at least 3000 frame errors

IX. NUMERICAL RESULTS AND DISCUSSIONS

This study assesses the influence of Byzantine attacks on system performance, quantified by Bit Error Rate (BER), within time-varying channel environments. Our analysis systematically investigates six key factors, such as the proportion of Byzantine attackers, the number of antennas per SU, and the number of reflecting elements in the RIS, to determine optimal attack strategies for both small-scale and large-scale scenarios. The default simulation parameters are as follows: 10 SUs each equipped with 6 antennas, an RIS with 9 elements, 8 relay nodes, and a data sequence length of 504 bits. A complete list of parameters is provided in Table III.

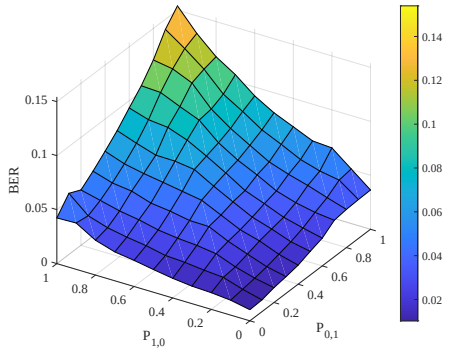


Fig. 3: The BER Performance of Optimal Fusion Rule under Small-Scale Attacks.

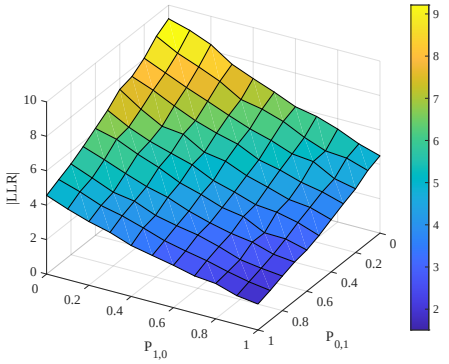


Fig. 4: The Magnitude of LLR for Optimal Fusion Rules under Small-scale Attacks.

A. The Strategy Analysis of Small-Scale Byzantine Attacks under the Optimal Fusion Rule

To substantiate the previously derived optimal strategy for small-scale Byzantine attacks under the optimal fusion rule, we conducted targeted simulations using the default parameter configuration for the optimal small-scale scenario. This verification aims to confirm whether the theoretical predictions hold under practical settings and to quantify the performance degradation caused by such attacks. The BER and the Magnitude of the LLR were adopted as key performance metrics, as they directly reflect decision accuracy and detection confidence. Simulation results reveal that, under optimal attack conditions, the BER attains its maximum value, which is shown in Fig.3, while the Magnitude of LLR reaches its minimum, which is shown in Fig.4, indicating severe impairment of detection reliability. In contrast, without attacks, the BER is minimized and the Magnitude of LLR is maximized, reflecting stable and reliable system operation. These outcomes not only validate the theoretical findings but also underscore the critical vulnerability of the system to optimally configured small-scale Byzantine attacks.

B. The Robustness Evaluation of Small-Scale Byzantine Attack Strategies under Varying Parameters

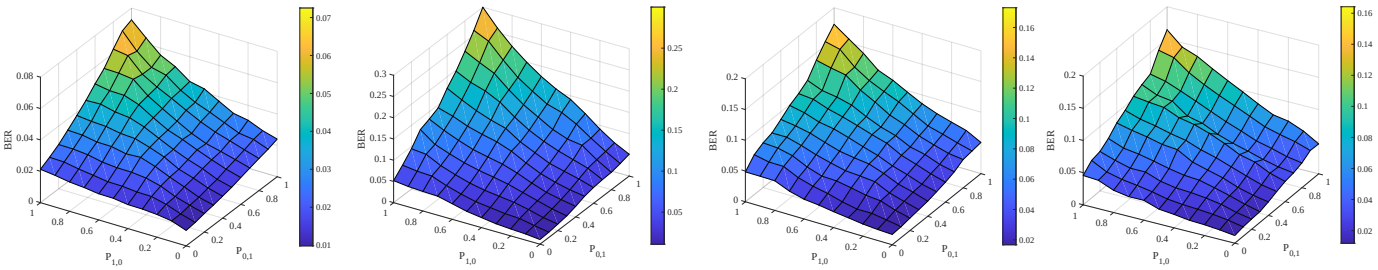
Based on theoretical derivations, this study concludes that in wireless networks, the effectiveness of the optimal small-scale Byzantine attack depends only on the proportion of Byzantine nodes and the signal flipping probability, and is independent of CSI. Accordingly, a multidimensional evaluation framework was established to analyze system performance from three aspects: the number of antennas per SU, the total number of SUs, and the number of RIS reflecting elements. The BER under various parameter settings was used to quantify the effectiveness of different attack strategies. Building upon the previous analysis, we further simulated the system by varying four key channel-related parameters. As shown in Fig. 5, increasing the number of antennas significantly improves overall performance, with smoother BER curves indicating enhanced anti-interference capability.

Nevertheless, even under favorable conditions, the AF attack remains the most destructive strategy, consistently causing the greatest degradation. Fig. 6 shows that without attacks, more SUs reduce BER and improve performance, whereas under a fixed attack ratio, BER degradation stays stable across different SU counts, underscoring the robustness of this attack. Increasing RIS elements lowers BER through enhanced signal manipulation and reception, yet the AF attack still causes the largest performance drop. In summary, across configurations involving attack ratios, SU antennas, SU numbers, and RIS elements, the AF attack consistently proves the most destructive and stable strategy, highlighting its critical role in system security analysis.

We first simulated the optimal large-scale Byzantine attack using default parameters to establish a baseline. As seen in Fig. 7, detection accuracy drops sharply, and Fig. 8 shows the Magnitude of LLR collapsing to its minimum, signifying the attack's success in obscuring the true signal and undermining the FC's reliability. Following this, we analyzed two key channel-related parameters to assess system robustness under fluctuations. Fig. 7 shows that under large-scale Byzantine attacks, the commonly used AF strategy ceases to be optimal. Specifically, when $\alpha(P_{0,1} + P_{1,0}) = 1$, the system's BER locks at 0.5, indicating complete corruption of the received signal and total loss of discriminability. In this state, no internal reconfiguration or adaptive defense can restore reliability, underscoring a fundamental system vulnerability to such powerful attacks.

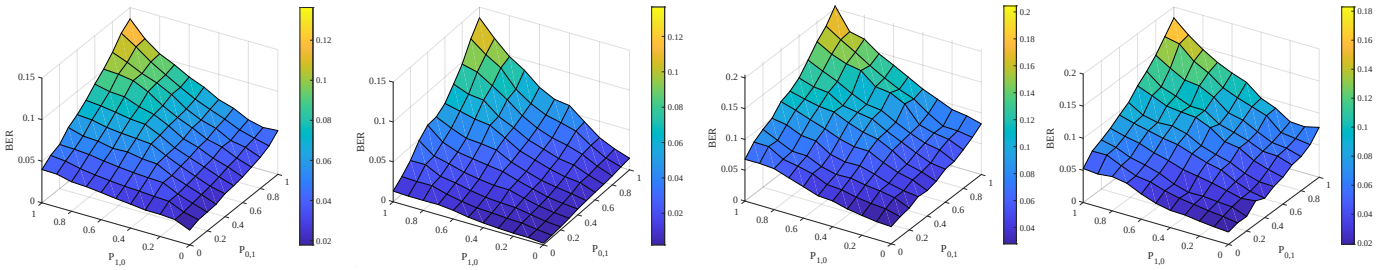
C. The Robustness Evaluation of Large-Scale Byzantine Attack Strategies under Varying Parameters

When the attack ratio α increases from 0.7 to 0.9, the system exhibits significant performance changes. Specifically, while the baseline BER slightly decreases in the absence of attacks, indicating enhanced inherent anti-interference capability under attack conditions, the BER increases markedly. As shown in Fig. 9, when the attack parameters satisfy $\pi_{0,1}+\pi_{1,0} = 1$ (i.e., $\alpha\left(P_{0,1} + P_{1,0}\right) = 1$), the system reaches a worst-case attack scenario. Under this condition, the attacker can continuously force the BER to converge toward 0.5, effectively rendering the decision-making process unreliable. Notably, the destructive power of this attack strategy does not diminish with changes in α. For example, when $\alpha = 0.9$, even partial flipping attacks can substantially degrade system performance. Compared to lower attack intensities, the BER is significantly higher, further demonstrating that the system becomes unrecoverable under large-scale attacks.



(a) $\alpha = 0.2$ (b) $\alpha = 0.4$ (c) $M = 4$ (d) $M = 8$

Fig. 5: The BER performance of Optimal Fusion Rule under different Attack Strengths and Numbers of SUs Antennas.



(a) $I = 8$ (b) $I = 12$ (c) $N = 5$ (d) $N = 7$

Fig. 6: The BER Performance of the Optimal Fusion Rule under Varying Numbers of RIS Reflecting Elements and SUs.

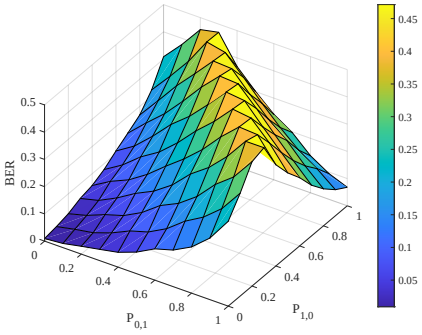


Fig. 7: The BER Performance of Optimal Fusion Rule under Large-Scale Attacks.

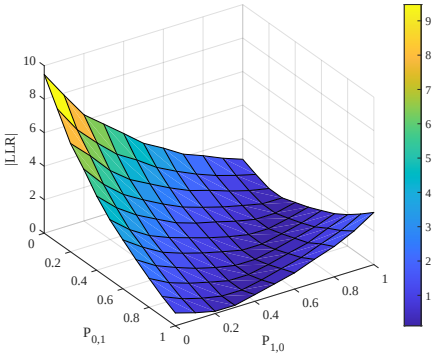


Fig. 8: The Magnitude of LLR for Optimal Fusion Rules under Large-Scale Attacks.

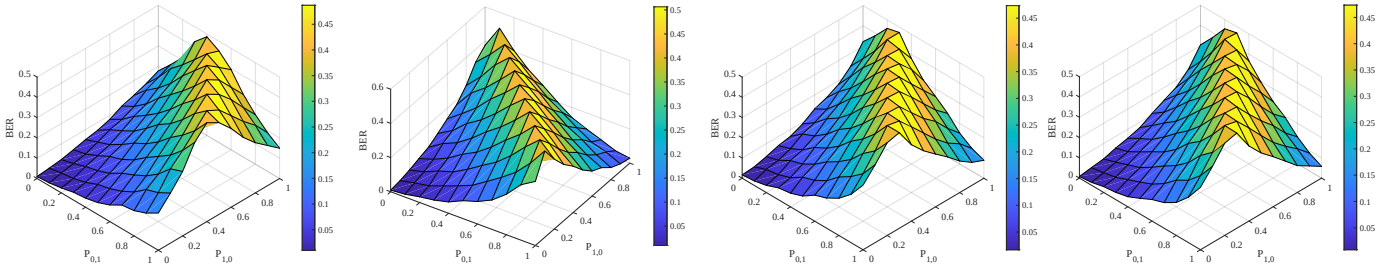
With the attack ratio fixed at $\alpha = 0.8$, the effect of varying the number of antennas M on system performance is studied. Without attacks, increasing antennas significantly improves BER due to spatial diversity gain. However, under optimal attacks, BER quickly converges to 0.5 for all antenna configurations, showing no mitigation of the attack's impact. Similarly, with $\alpha = 0.8$, increasing the number of SUs reduces BER without attacks, reflecting better reliability via cooperative detection. Yet, under optimal attacks, BER remains near 0.5 regardless of SU count, indicating attack resilience to user scale. Fig. 10 shows BER versus transmission SNR for varying RIS elements N. Increasing N greatly improves BER without attacks by enhancing diversity and noise suppression. Under optimal attacks, BER still approaches 0.5, suggesting attack effectiveness is largely unaffected by RIS size.

D. The Strategy Analysis of Large-Scale Byzantine Attacks under the Optimal Fusion Rule

In summary, regardless of system configurations, including attack ratio, antenna count, number of SUs, RIS scale, or relay node deployment, once the attack parameters satisfy $\alpha\left(P_{0,1}+P_{1,0}\right)=1$, the system inevitably enters a decision failure state, with BER steadily converging to 0.5. This attack strategy demonstrates strong consistency and robustness across different network scales and attack intensities, making it one of the most destructive forms of Byzantine attacks.

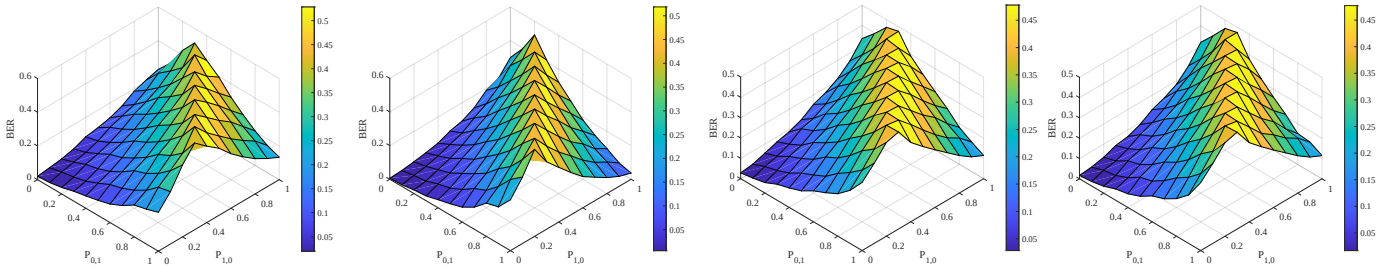
E. Comparison of System Performance under Different Attack Modes

A comparative analysis of system performance under varying adversarial conditions is conducted, with MI serving as the principal performance indicator. The study considers three scenarios: no attack, optimal small-scale attack, and optimal large-scale attack. As illustrated in Fig. 11, the MI under optimal large-scale attack conditions remains nearly zero, signifying a substantial degradation in information exchange capability and overall system reliability. In contrast, although the optimal small-scale attack also leads to a decline in MI, its impact is markedly less severe. These observations



(a) $\alpha = 0.7$ (b) $\alpha = 0.9$ (c) $M = 4$ (d) $M = 8$

Fig. 9: The BER performance of Optimal Fusion Rule under different Attack Strengths and Numbers of SUs Antennas.



(a) $I = 8$ (b) $I = 12$ (c) $N = 5$ (d) $N = 7$

Fig. 10: The BER Performance of the Optimal Fusion Rule under Varying Numbers of RIS Reflecting Elements and SUs.

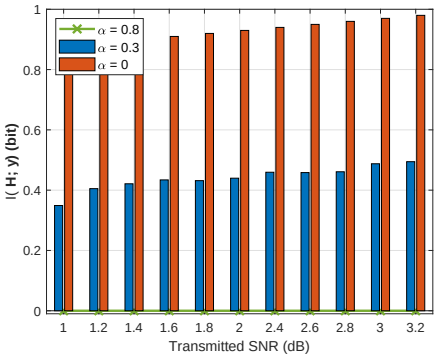


Fig. 11: The MI Performance under Optimal Small-Scale and Large-Scale Attacks across Varying Transmitted SNRs.

demonstrate that the scale of adversarial intervention plays a decisive role in determining the extent of system disruption, underscoring the heightened risk posed by large-scale coordinated attacks.

F. Evaluation of Different Attack Strategies for Small-Scale Byzantine Attacks

In this section, we evaluate the impact of different Byzantine attack strategies under small scale attacks. Fig. 12a and Fig. 12b illustrate Byzantine attacks within the range $0 \leq \alpha \leq \frac{1}{2}$. Fig. 12a corresponds to attack scenarios where the flipping probabilities satisfy $0 \leq P_{0,1} + P_{1,0} \leq 1$. Fig. 12b presents cases where the flipping probabilities satisfy $1 \leq P_{0,1}+P_{1,0} \leq 2$.

From Fig. 12a, we observe that as the attack probability α increases, the system performance deteriorates. Under these conditions, both the AY and AN attacks have a greater impact than random attacks. In contrast, Fig. 12b, where the flipping probabilities satisfy $1 \leq P_{0,1} + P_{1,0} \leq 2$, allows for the possibility of a AF attack. The results indicate that the AF attack is the most detrimental, followed by RF attacks, while the AY and AN attacks have the least impact.

G. Evaluation of Different Attack Strategies for Large Scale Byzantine Attacks

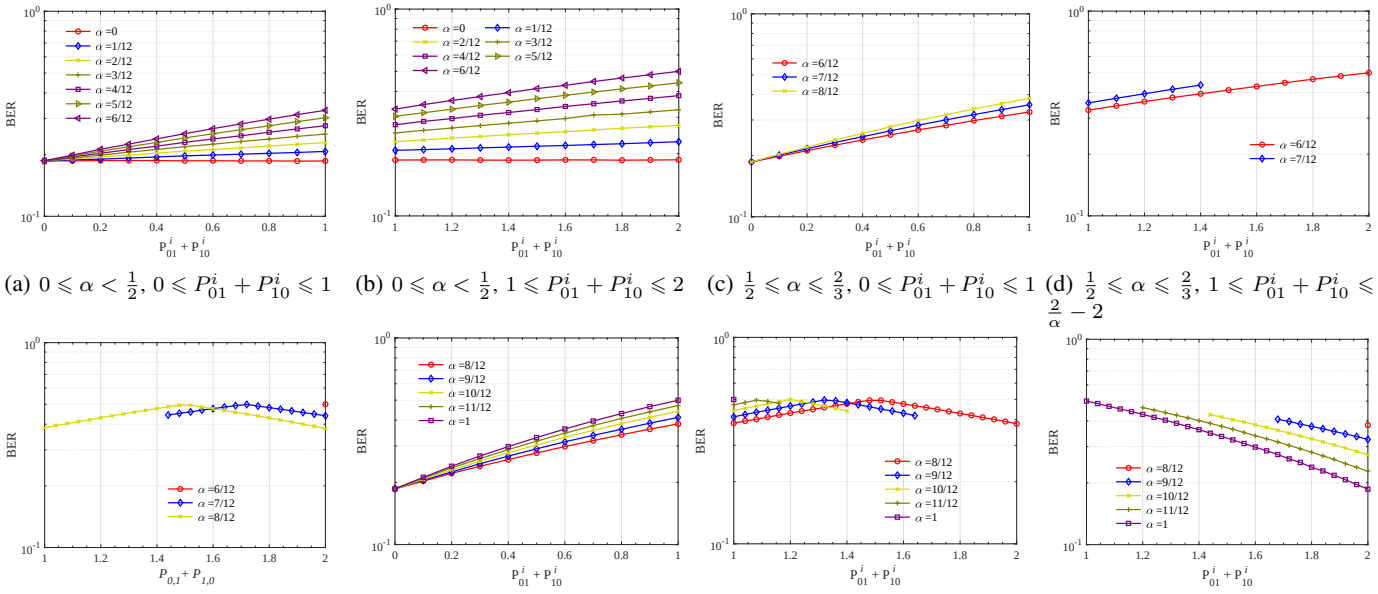
In this section, we evaluate the impact of different Byzantine attack strategies under large scale attacks. Based on the previous analysis, the optimal attack strategy varies depending on the Byzantine node proportion α. Therefore, we divide α into two intervals for analysis: $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$ and $\frac{2}{3} \leq \alpha \leq 1$. In our experiments, we set the number of hops to J=3, the total number of sensors to N=12.

Fig. 12c – Fig. 12e illustrate the Byzantine attack scenarios when $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$, while Fig. 12f-(f) – Fig. 12h correspond to the cases where $\frac{2}{3} \leq \alpha \leq 1$.

Fig. 12c depicts the attack performance when the Byzantine ratio is within $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$ and the flip probability satisfies $0 \leq P_{0,1} + P_{1,0} < 1$. We can observe that for a fixed number of attack nodes, the BER of AY and AN attacks is lower than that of RD attacks, indicating that AY and AN attacks are less disruptive than RD attacks in this scenario.

Fig. 12d examines attack performance for $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$ in the probability range $1 \leq P_{0,1} + P_{1,0} \leq \frac{2}{\alpha} - 2$. Under these conditions, the AF attack is most effective, followed by RD, while AY and AN show the weakest impact.

Fig. 12e complements this analysis for the higher probability range, $\frac{2}{\alpha} - 2 < P_{0,1} + P_{1,0} \leq 2$. Here, the performance ranking shifts: the RD attack becomes the most effective, outperforming the AF attack. The AY and AN attacks remain the least impactful. The case for $\alpha = \frac{1}{2}$ is excluded, as its corresponding range is trivial.



(a) $0 \leqslant \alpha < \frac{1}{2}$, $0 \leqslant P_{01}^i + P_{10}^i \leqslant 1$ (b) $0 \leqslant \alpha < \frac{1}{2}$, $1 \leqslant P_{01}^i + P_{10}^i \leqslant 2$ (c) $\frac{1}{2} \leqslant \alpha \leqslant \frac{2}{3}$, $0 \leqslant P_{01}^i + P_{10}^i \leqslant 1$ (d) $\frac{1}{2} \leqslant \alpha \leqslant \frac{2}{3}$, $1 \leqslant P_{01}^i + P_{10}^i \leqslant \frac{2}{\alpha} - 2$

(e) $\frac{1}{2} \leqslant \alpha \leqslant \frac{2}{3}$, $\frac{2}{\alpha} - 2 \leqslant P_{01}^i + P_{10}^i \leqslant 2$ (f) $\frac{2}{3} \leqslant \alpha \leqslant 1$, $0 \leqslant P_{01}^i + P_{10}^i \leqslant 1$ (g) $\frac{2}{3} \leqslant \alpha \leqslant 1$, $1 \leqslant P_{01}^i + P_{10}^i \leqslant \frac{2}{\alpha} - 1$ (h) $\frac{2}{3} \leqslant \alpha \leqslant 1$, $\frac{2}{\alpha} - 1 \leqslant P_{01}^i + P_{10}^i \leqslant 2$

Fig. 12: Evaluating attack performance with Byzantine proportions

Fig. 12f presents the attack performance under the conditions where the Byzantine ratio satisfies $\frac{2}{3} \leq \alpha \leq 1$ and the flip probability is within the range $0 \leq P_{0,1} + P_{1,0} < 1$. From the figure, it can be observed that when the number of attacking nodes is fixed, the bit error rate of AY and AN attacks is higher than that of the RD attack, indicating that the AY and AN attacks are stronger than the RD attack under these conditions.

Fig. 12g investigates attack performance for Byzantine ratios $\frac{2}{3} \leq \alpha < 1$ within the flip probability range of $1 \leq P_{0,1} + P_{1,0} \leq \frac{2}{\alpha} - 1$. Results show that for a fixed number of attacking nodes, the RD attack is most effective, followed by AY and AN, while AF is least effective. The case $\alpha = 1$ is omitted as its valid probability range collapses to a point.

Fig. 12h extends the analysis to the higher flip probability range of $\frac{2}{\alpha} - 1 < P_{0,1} + P_{1,0} \leq 2$. In this regime, the attack performance ranking changes: the AY and AN attacks become the most effective, surpassing the RD attack. The AF attack remains the weakest.

X. CONCLUSIONS AND FUTURE WORK

We have studied the problem of optimal Byzantine attack strategies in CSS under non-ideal conditions. Assuming that the attackers are aware of the fusion rule, we propose optimal attack strategies for both optimal and suboptimal fusion rules. Our findings indicate that even when Byzantine attackers are unaware of the current fusion strategy, they can still launch optimal attacks. Furthermore, we evaluate the performance of different attack strategies under various conditions. Our future research directions will be in the following areas. We plan to extend the existing optimal attack strategies from the traditional binary hypothesis testing framework to more complex multi-channel spectrum sensing scenarios [52]–[54]. In this extended setting, our focus will be on investigating optimal attack behaviors in multi-dimensional signal spaces [55], particularly how attackers can formulate coordinated interference strategies under varying channel conditions and the coexistence of multiple signal features to maximize disruption of the decision-making process [56]–[58]. In addition, We will further explore defenses against optimal or near-optimal attacks. Based on existing results, lightweight and effective detection and defense schemes will be designed to exploit statistical inconsistencies caused by attackers, improving the robustness and recovery of CSS systems under adversarial conditions.

APPENDIX A

1) AN attacks and AY attacks: The numerator term in an AN attack is roughly proportional to $-P_{D_i}(1-\alpha)$. The denominator, on the other hand, is proportional to $-P_{F_i}(1-\alpha)$, so the overall decision statistic should be proportional to $-(P_{D_i} - P_{F_i})(1-\alpha)$ is proportional.

Similarly the numerator is roughly proportional to $-[\alpha + (1-\alpha) P_{D_i}]$ for an AY attack. The denominator, on the other hand, is proportional to $-[\alpha + (1-\alpha) P_{F_i}]$, so the overall decision statistic should be proportional to $-(P_{D_i} - P_{F_i})(1-\alpha)$ is proportional.

That is, both AN attacks and AY attacks have their overall decision statistics proportional to $-(P_{D_i} - P_{F_i})(1-\alpha)$. This leads to the fact that the attack performance of AN and AY attacks is considered to be the same regardless of whether it is a large-scale attack or a small-scale attack.

2) AF Attack and AN Attack: According to (41), the numerator in the AF attack is approximately proportional to: $-[\alpha - (2\alpha - 1) P_{D_i}]$. The denominator is approximately proportional to:$-[\alpha - (2\alpha - 1) P_{F_i}]$. Therefore, the overall decision statistic is proportional to: $(P_{D_i} - P_{F_i})(2\alpha - 1)$. Since it is known that $P_{D_i} - P_{F_i} > 0$, comparing $(1-\alpha)$

and $|2\alpha - 1|$ can help determine which attack results in a smaller magnitude of the decision statistic.

- **Small-scale attack**: When $\alpha < 0.5$, it can be concluded that under small-scale attacks, the AF attack outperforms the AN attack.
- **Large-scale attack**: when $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$, the AF attack performs better. When $\frac{2}{3} \leq \alpha \leq 1$, the AN or AY attacks perform better.

3) RF Attacks and AN Attacks: For the RF attack, the numerator is approximately proportional to:
$-[P_{D_i} + \alpha P_{1,0} - P_{D_i}\alpha(P_{0,1} + P_{1,0})]$.
The denominator is approximately proportional to:
$-[P_{F_i} + \alpha P_{1,0} - P_{F_i}\alpha(P_{0,1} + P_{1,0})]$.
Therefore, the overall decision statistic is proportional to:
$-(P_{D_i} - P_{F_i})[\alpha(P_{0,1} + P_{1,0}) - 1]$.

To compare the performance of RF and AN attacks, we use $|\alpha(P_{0,1} + P_{1,0}) - 1|$ and $(1 - \alpha)$ to determine which yields a smaller magnitude of the decision statistic

- **Small-scale attack**: since $\alpha < 0.5$, by comparing $1 - \alpha(P_{0,1} + P_{1,0})$ and $(1 - \alpha)$, we can conclude that: when $0 \leq P_{0,1} + P_{1,0} \leq 1$, AN attacks perform better. When $1 \leq P_{0,1} + P_{1,0} \leq 2$, RF attacks perform better.
- **Large-scale attack**: since $\alpha > 0.5$, we have $|\alpha(P_{0,1} + P_{1,0}) - 1| = \alpha(P_{0,1} + P_{1,0}) - 1$. When $0 \leq P_{0,1} + P_{1,0} \leq 1$, it can be concluded that $(1 - \alpha) \geq |\alpha(P_{0,1} + P_{1,0}) - 1|$, meaning AN or AY attacks always perform better than RF attacks. When $1 \leq P_{0,1} + P_{1,0} \leq 2$, by setting $(1 - \alpha) = \alpha(P_{0,1} + P_{1,0}) - 1$, we can find that $P_{0,1} + P_{1,0} = \frac{2}{\alpha} - 1$. When $1 \leq P_{0,1} + P_{1,0} \leq \frac{2}{\alpha} - 1$, RF attacks perform better. When $\frac{2}{\alpha} - 1 \leq P_{0,1} + P_{1,0} \leq 2$, AN attacks perform better.

4) RF Attacks and AF Attacks: Compare the performance of RF attacks and AF attacks, we can use $|\alpha(P_{0,1} + P_{1,0}) - 1|$ and $(2\alpha - 1)$ to determine which attack performs better.

Following similar analysis method, its easy to conclude that

- **Small-scale attack**: When $0 \leq P_{0,1} + P_{1,0} \leq 1$, AF attack is impossible because they require $P_{0,1} = P_{1,0} = 1$. When $1 \leq P_{0,1} + P_{1,0} \leq 2$, AF attacks perform better.
- **Large-scale attack**: when $1 \leq P_{0,1} + P_{1,0} \leq \frac{2}{\alpha} - 2$, AF attacks perform better. When $\frac{2}{\alpha} - 2 \leq P_{0,1} + P_{1,0} \leq 2$, RF attacks perform better.

REFERENCES

[1] S. Soderi, A. Brighente, S. Xu, and M. Conti, "Multi-ris aided vlc physical layer security for 6g wireless networks," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 15 182–15 195, 2024.

[2] D. Xu, X. Su, G. Premsankar, H. Wang, S. Tarkoma, and P. Hui, "Dynamic hierarchical reinforcement learning framework for energy-efficient 5g base stations in urban environments," *IEEE Transactions on Mobile Computing*, vol. 24, no. 9, pp. 8582–8599, 2025.

[3] M. Sharma and N. Sarma, "Multi-objective optimization for energy efficient cooperative communication in energy-constrained overlay cognitive radio networks," *Physical Communication*, vol. 62, p. 102251, 2024.

[4] M. Xu, X. Song, Y. Zhao, Z. Yin, and Z. Wu, "Deep reinforcement learning-based ris-assisted cooperative spectrum sensing in cognitive radio network," *IEICE Transactions on Communications*, vol. E108-B, no. 4, pp. 404–410, 2025.

[5] G. Zhang, X. He, Y. Mu, W. Wang, Y. Li, B. Ji, and S. Mumtaz, "Serial distributed detection in multihop multirelay wireless sensor networks with end–edge–cloud orchestration under graph-powered computing," *IEEE Internet of Things Journal*, vol. 12, no. 4, pp. 3720–3733, 2025.

[6] Z. Song, Y. She, J. Yang, J. Peng, Y. Gao, and R. Tafazolli, "Nonuniform sampling pattern design for compressed spectrum sensing in mobile cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 23, no. 9, pp. 8680–8693, 2024.

[7] Y. Jiang, X. Li, G. Zhu, H. Li, J. Deng, K. Han, C. Shen, Q. Shi, and R. Zhang, "Integrated sensing and communication for low altitude economy: Opportunities and challenges," *IEEE Communications Magazine*, pp. 1–7, 2025.

[8] B. Cai, J. Ge, and Y.-C. Liang, "Star-ris assisted opportunistic cognitive radio networks," in *2024 IEEE 24th International Conference on Communication Technology (ICCT)*. IEEE, 2024, pp. 947–952.

[9] A. A. Sardar, D. Roy, W. U. Mondal, and G. Das, "Coalition formation for outsourced spectrum sensing in cognitive radio network," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 3, pp. 580–592, 2023.

[10] A. Nasser, H. A. H. Hassan, A. Mansour, K.-C. Yao, and L. Nuaymi, "Intelligent reflecting surfaces and spectrum sensing for cognitive radio networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 3, pp. 1497–1511, 2022.

[11] S. Singh and R. Shrestha, "Ultra-low sensing-time and hardware-efficient spectrum sensor for data fusion-based cooperative cognitive-radio network," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 216–226, 2023.

[12] W. Wu, Z. Wang, Y. Wu, F. Zhou, B. Wang, Q. Wu, and D. W. K. Ng, "Joint sensing and transmission optimization for irs-assisted cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 5941–5956, 2023.

[13] A. Gao, Q. Wang, Y. Wang, C. Du, Y. Hu, W. Liang, and S. X. Ng, "Attention enhanced multi-agent reinforcement learning for cooperative spectrum sensing in cognitive radio networks," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 7, pp. 10 464–10 477, 2024.

[14] J. Sang, M. Zhou, J. Lan, B. Gao, W. Tang, X. Li, S. Jin, E. Basar, C. Li, Q. Cheng, and T. J. Cui, "Multi-scenario broadband channel measurement and modeling for sub-6 ghz ris-assisted wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 23, no. 6, pp. 6312–6329, 2024.

[15] J. Ge, Y.-C. Liang, S. Wang, and C. Sun, "Ris-assisted cooperative spectrum sensing for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 23, no. 9, pp. 12 547–12 562, 2024.

[16] P. Yi, W. Cheng, J. Wang, and W. Zhang, "Ris-assisted seamless connectivity in wireless multi-hop relay networks," *IEEE Transactions on Mobile Computing*, vol. 24, no. 9, pp. 8600–8611, 2025.

[17] A. Albanese, F. Devoti, V. Sciancalepore, M. Di Renzo, A. Banchs, and X. Costa-Pérez, "Ares: Autonomous ris solution with energy harvesting and self-configuration towards 6g," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 12 006–12 019, 2024.

[18] J. Wu, M. Su, X. Xu, L. Qiao, M. Dai, and W. Cao, "Less sample-cooperative spectrum sensing in the presence of large-scale byzantine attack," *IEEE Sensors Letters*, vol. 8, no. 1, pp. 1–4, 2024.

[19] Y. Fu and Z. He, "Massive ssdf attackers identification in cognitive radio networks by using consistent property," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 11 058–11 062, 2023.

[20] A. Parmar, K. Shah, K. M. Captain, M. López-Benítez, and J. R. Patel, "Gaussian mixture model-based anomaly detection for defense against byzantine attack in cooperative spectrum sensing," *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 2, pp. 499–509, 2024.

[21] Z. Luo, S. Zhao, Z. Lu, J. Xu, and Y. E. Sagduyu, "When attackers meet ai: Learning-empowered attacks in cooperative spectrum sensing," *IEEE Transactions on Mobile Computing*, vol. 21, no. 5, pp. 1892–1908, 2020.

[22] A. Abrardo, M. Barni, K. Kallas, B. Tondi *et al.*, *Information Fusion in Distributed Sensor Networks with Byzantines*. Springer, 2021.

[23] B. Kailkhura, S. Brahma, and P. K. Varshney, "Consensus based detection in the presence of data falsification attacks," *arXiv preprint arXiv:1504.03413*, 2015.

[24] L. Chen, X. Shen, X. Zhao, Z. Wang, W. He, G. Xu, and Y. Chen, "Defending dominant cooperative probabilistic attack in crns by js-divergence-based improved reputation algorithm," *Pervasive and Mobile Computing*, vol. 101, p. 101921, 2024.

[25] R. Sarmah, A. Taggu, and N. Marchang, "Detecting byzantine attack in cognitive radio networks using machine learning," *Wirel. Netw.*, vol. 26, no. 8, p. 5939–5950, Nov. 2020.

[26] 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.

[27] S. Zuo, X. Yan, R. Fan, H. Hu, H. Shan, T. Q. S. Quek, and P. Zhao, "Federated learning resilient to byzantine attacks and data heterogeneity," *IEEE Transactions on Mobile Computing*, pp. 1–15, 2025.

[28] 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, 2011.

[29] B. Kim, Y. E. Sagduyu, K. Davaslioglu, T. Erpek, and S. Ulukus, "Channel-aware adversarial attacks against deep learning-based wireless signal classifiers," *IEEE Transactions on Wireless Communications*, vol. 21, no. 6, pp. 3868–3880, 2022.

[30] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 282–310, 2021.

[31] M. Abdelhakim, L. E. Lightfoot, J. Ren, and T. Li, "Distributed detection in mobile access wireless sensor networks under byzantine attacks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 4, pp. 950–959, 2014.

[32] 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.

[33] H. Chen, M. Zhou, L. Xie, and J. Li, "Cooperative spectrum sensing with m-ary quantized data in cognitive radio networks under ssdf attacks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5244–5257, 2017.

[34] 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.

[35] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "Asymptotic analysis of distributed bayesian detection with byzantine data," *IEEE Signal Processing Letters*, vol. 22, no. 5, pp. 608–612, 2014.

[36] 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.

[37] 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.

[38] L. Chen, X. Shen, X. Zhao, Z. Wang, W. He, G. Xu, and Y. Chen, "Defending dominant cooperative probabilistic attack in crns by js-divergence-based improved reputation algorithm," *Pervasive and Mobile Computing*, vol. 101, p. 101921, 2024.

[39] F. Li, R. Lin, J. Wang, J. Hu, F. Shu, and L. Wu, "A fast method to defend against ssdf attacks in the ciov network: Based on dag blockchain and evolutionary game," *IEEE Communications Letters*, vol. 27, no. 12, pp. 3171–3175, 2023.

[40] G. Zhang, P. Wang, W. Wang, Y. Mu, Y. Li, J. Tang, H. Song, H. Wen, and S. Mumtaz, "An information-theoretic approach to distributed detection for mobile wireless sensor networks under byzantine attack in entirely unknown or complicated environment: Design, analysis, and evaluation of the attack strategy," *IEEE Internet of Things Journal*, vol. 12, no. 4, pp. 3689–3706, 2025.

[41] 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, 2019.

[42] 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.

[43] T.-H. Wang, P.-N. Chen, and Y.-C. Huang, "Probabilistic byzantine attack on federated learning," *IEEE Transactions on Signal Processing*, vol. 73, pp. 1823–1838, 2025.

[44] B. Kailkhura, Y. S. Han, S. Brahma, and P. K. Varshney, "Distributed bayesian detection in the presence of byzantine data," *IEEE Transactions on Signal Processing*, vol. 63, no. 19, pp. 5250–5263, 2015.

[45] H. Xie, D. Li, and B. Gu, "Enhancing spectrum sensing via reconfigurable intelligent surfaces: Passive or active sensing and how many reflecting elements are needed?" *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 14 940–14 955, 2024.

[46] J. Wu, J. He, J. Gan, Z. Chen, J. Zhang, Z. Chen, and G. Zhu, "Sequential spectrum sensing against random b yzantine attack in cognitive radio networks," *Transactions on Emerging Telecommunications Technologies*, vol. 34, no. 7, p. e4796, 2023.

[47] W. Feller, *An introduction to probability theory and its applications*. John Wiley & Sons, 1991, vol. 2.

[48] G. Zhang, Y. Li, B. Ji, K. Chen, Y. Mu, and W. Wang, "M-ary distributed decision fusion for multihop relay wireless sensor networks: Decision fusion rule, implementation framework, and performance analysis," *IEEE Internet of Things Journal*, vol. 11, no. 18, pp. 29 569–29 587, 2024.

[49] G. Zhang, C. Ma, K. Chen, Y. Li, H. Li, and C. Han, "Multiple-symbol noncoherent learning detection of coded qam signals in ieee 802.15. 3 wireless multi-media networks," *Physical Communication*, vol. 55, p. 101922, 2022.

[50] G. Zhang, K. Chen, C. Ma, S. K. Reddy, B. Ji, Y. Li, C. Han, X. Zhang, and Z. Fu, "Decision fusion for multi-route and multi-hop wireless sensor networks over the binary symmetric channel," *Computer Communications*, vol. 196, pp. 167–183, 2022.

[51] G. Zhang, K. Chen, J. Tang, H. Song, H. Wen, and S. Mumtaz, "Distributed decision fusion in wireless sensor networks for contextual experience sensing and recommendation forming in e-commerce," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 4, pp. 6817–6829, 2024.

[52] M. A. Mughal, A. Ullah, M. A. Z. Cheema, X. Yu, and N. Jhanjhi, "An intelligent channel assignment algorithm for cognitive radio networks using a tree-centric approach in iot," *Alexandria Engineering Journal*, vol. 91, pp. 152–160, 2024.

[53] Y. Yang, Z. Chen, and R. Zhang, "A cooperative spectrum sensing method based on complex evidence theory and multi-angle weights to defend against ssdf," in *2024 4th International Conference on Electronic Information Engineering and Computer (EIECT)*, 2024, pp. 815–819.

[54] G. Zhu, J. Wu, H. Liang, J. Tang, L. Xia, J. Bao, and W. Cao, "Cost and benefit tradeoff of ssdf attack in cooperative spectrum sensing in cognitive wireless sensor networks," *IEEE Sensors Letters*, vol. 8, no. 5, pp. 1–4, 2024.

[55] A. Chouhan, K. Captain, A. Parmar, and J. Patel, "Defending cooperative spectrum sensing from byzantine attacks: An effective entropy-based weighted algorithm," *IEEE Wireless Communications Letters*, vol. 12, no. 12, pp. 2063–2067, 2023.

[56] H. Liang, J. Wu, T. Liu, H. Wang, and W. Cao, "Efficient cooperative spectrum sensing in uav-assisted cognitive wireless sensor networks," *IEEE Sensors Letters*, vol. 8, no. 10, pp. 1–4, 2024.

[57] X. Fang, M. Jin, Q. Guo, and T. Jiang, "Cnn-transformer based cooperative spectrum sensing in cognitive radio networks," *IEEE Wireless Communications Letters*, pp. 1–1, 2025.

[58] J. Wu, S. Gao, X. Teng, Z. Zhang, M. Dai, H. Ge, and W. Cao, "Beta distribution function-based cooperative spectrum sensing against byzantine attack in cognitive wireless sensor networks," *IEEE Sensors Letters*, vol. 8, no. 5, pp. 1–4, 2024.