Sticky Fingers: Resilience of Satellite Fingerprinting against Jamming Attacks

Joshua Smailes
University of Oxford
joshua.smailes@cs.ox.ac.uk

Edd Salkield
University of Oxford
edd.salkield@cs.ox.ac.uk

Sebastian Köhler
University of Oxford
sebastian.kohler@cs.ox.ac.uk

Simon Birnbach
University of Oxford
simon.birnbach@cs.ox.ac.uk

Martin Strohmeier
armasuisse Science + Technology
martin.strohmeier@armasuisse.ch

Ivan Martinovic
University of Oxford
ivan.martinovic@cs.ox.ac.uk

Abstract—In the wake of increasing numbers of attacks on radio communication systems, a range of techniques are being deployed to increase the security of these systems. One such technique is radio fingerprinting, in which the transmitter can be identified and authenticated by observing small hardware differences expressed in the signal. Fingerprinting has been explored in particular in the defense of satellite systems, many of which are insecure and cannot be retrofitted with cryptographic security.

In this paper, we evaluate the effectiveness of radio fingerprinting techniques under interference and jamming attacks, usually intended to deny service. By taking a pre-trained fingerprinting model and gathering a new dataset in which different levels of Gaussian noise and tone jamming have been added to the legitimate signal, we assess the attacker power required in order to disrupt the transmitter fingerprint such that it can no longer be recognized. We compare this to Gaussian jamming on the data portion of the signal, obtaining the remarkable result that transmitter fingerprints are still recognizable even in the presence of moderate levels of noise. Through deeper analysis of the results, we conclude that it takes a similar amount of jamming power in order to disrupt the fingerprint as it does to jam the message contents itself, so it is safe to include a fingerprinting system to authenticate satellite communication without opening up the system to easier denial-of-service attacks.

I. MOTIVATION

Due to the growing availability of cheap Software-Defined Radio (SDR) hardware combined with increasing global reliance upon satellite systems and their data, attacks on space systems are a real threat to critical infrastructure. Furthermore, the great cost of launching and operating satellites means there are a large number of legacy satellite systems lacking proper cryptographic security. A number of techniques have been proposed to secure satellite communications in the absence of cryptography, relying upon analysis of the signal and other factors to authenticate messages. Fingerprinting is one of these techniques, and involves looking at the signal impairments caused by small differences in the transmitter hardware to authenticate the transmitter and distinguish it from attacker-controlled devices. This has been well-explored in terrestrial radio systems, and has seen use more recently in satellite systems, which have the additional difficulty of high levels of atmospheric noise.

However, an adversary's goal is not always to spoof communication; sometimes simple denial of service is sufficient. Traditionally this has been achieved through jamming techniques, involving the use of targeted noise or other signals to stop the legitimate signal from being properly decoded. This has been observed widely in the real world, particularly in the recent jamming attacks on Starlink in the Ukraine war [1, 2].

If fingerprinting techniques are used to secure a ground system, incoming messages may also be rejected if the transmitter fingerprint does not match the expected value. Therefore, an attacker may achieve easier denial of service by simply disrupting the fingerprint. With satellite systems, this is a particular concern – the high levels of atmospheric noise have already distorted the signal such that any identifiable transmitter characteristics are difficult to extract, which may make the attacker's goal even easier to accomplish.

A. Contributions

In this paper we show that physical-layer satellite fingerprinting techniques are resilient against jamming attacks, so their use does not increase the threat of denial-of-service attacks. We build upon our previous work in satellite fingerprinting, using the SATIQ system as a case study [3]. We focus on the Iridium satellite constellation used for communication, first performing an analysis of the level of Gaussian noise necessary to disrupt messages and the hardware required to achieve this. This is followed by data collection, in which we collect $540\,066$ messages with varying levels of noise added to the incoming signal, and the creation of further datasets by adding synthetic jamming to clean signals in software.[1] We next perform an evaluation of this data, showing that in order to disrupt the transmitter fingerprint through jamming, a similar or greater transmit power is required than to disrupt the message contents via traditional jamming techniques, with a

[1]The dataset has been made openly available at https://zenodo.org/records/10678124, and the code at https://github.com/ssloxford/SatIQ-noise.
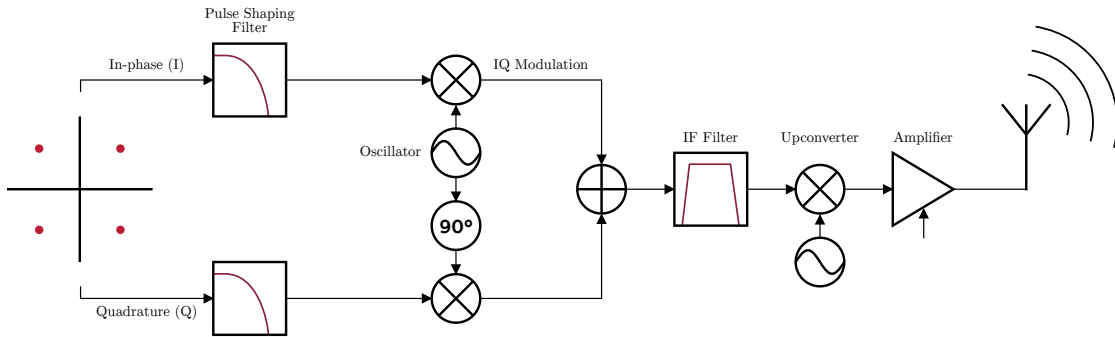
Fig. 1. An illustration of the hardware components involved in QPSK signal modulation. Each hardware component can introduce its own impairments upon the signal, which can be used to identify the transmitter through fingerprinting.

difference of between 0.3 and $2\,\mathrm{dB}$. Finally, we conclude with a discussion of these results and their potential root causes.

II. BACKGROUND

In this section we provide an overview of radio fingerprinting techniques, the jamming approaches that we evaluate, and the message structure and error correcting code attributes of the target protocol.

A. Fingerprinting

Fingerprinting techniques are used to identify radio transmitters by looking solely at the received signal. This is possible due to small variations in the analog components of the transmitter hardware (illustrated in Figure 1) introducing different impairments onto the signal, even between identical copies of the same hardware [4]. These impairments can be used to identify or authenticate transmitters, protecting against spoofing and replay attacks by attacker-controlled devices. This is the main use case for fingerprinting in satellites, many of which are not secured with cryptography for various reasons. Examples include NASA's MODIS data which is open by design [5], GEO-KOMPSAT-2A with leaked keys [6], and COMS-1 which has broken cryptography [7].

There are two main approaches to fingerprinting. Transient fingerprinting techniques look at the signal transient: the portion of the signal where the transmitter first powers up. This requires precise timing, but has been demonstrated to contain sufficient information for authenticating transmitters in security contexts [8]. On the other hand, steady-state fingerprinting looks at the signal as a whole, pulling identifying information out of the modulated portion of the signal [9].

In both steady-state and transient fingerprinting, the fingerprint is also affected by properties of the wireless channel: background noise (including interference from legitimate sources), free space path loss, and multipath distortion all add additional impairments onto the signal, making fingerprint extraction more difficult. This is particularly prominent in satellite systems, in which signals travel through hundreds of kilometers of atmosphere before reaching the receiver. A number of approaches have been explored to mitigate this issue, including averaging multiple messages to eliminate noise [10, 11], adding noise to clean signals during training [12], and looking at high sample rate signals to capture more detail and a wider range of features [3].

B. Jamming

In a jamming scenario, the attacker emits an interfering signal in the same frequency band as the victim signal, within the vicinity of a victim receiver. Denial of service is achieved once the received power of the attacker's signal causes sufficient bit errors, making a certain proportion of packets undecodable. The received power of the interfering signal emitted by the attacker can be measured relative to the victim signal received power, giving the attacker-to-victim power ratio.

Against a fingerprinting system, the attacker's objective is to cause victim messages to be rejected as illegitimate, by inducing fingerprinter error on a certain proportion of packets. The robustness of the fingerprinter against a jammer is therefore seen by comparing the decoder error rate to the fingerprinter error rate, as the attacker-to-victim power ratio increases.

In this work, we consider two main types of jamming which are widely used in recent analyses of space systems: noise jamming and tone jamming [13, 14]. It is known that these forms of jamming cause denial of service at different attacker-to-victim power ratios; we evaluate both as we are interested to observe the different effects that they cause in the fingerprinting system.

In particular, we focus on the Iridium satellite constellation, used primarily for communication and composed of 66 satellites in Low Earth Orbit (LEO). We proceed to describe the structure of Iridium's messages, paying particular attention to the error correcting code which must be overcome to render the received packets undecodable. In particular, we focus on Iridium Ring Alert (IRA) messages, which are the ones collected for the fingerprinting system being evaluated [3].

C. Structure of Iridium Ring Alert Messages

In order to establish the power required to disrupt communication through jamming attacks, we must first understand the structure of the Iridium messages and their error correcting codes. The Iridium Ring Alert (IRA) messages are designed to be received by all Iridium user terminals. Each message contains information about the transmitting satellite, which includes its position, altitude, and a unique identifier for the satellite and its onboard transmitter. These messages are preceded by a synchronization header, which is always

identical for every transmission; the fingerprinter identifies the transmitter through small variations in this header. The remainder of the message is 93 bits in length, followed by a variable length set of data pages. The 93 bit section is protected by an interleaved "BCH" Error Correcting Code (ECC), which performs error detection and correction on the received bits – we focus on this fixed section as we analyze the error correcting capacity of the code. An overview of the precise operation of BCH codes was given by *Walters et al.* [15].

Specifically, the ECC is applied to the message in three interleaved blocks of 21 bits each. Each 21 bit block is protected by an additional 10 parity bits to form an encoded 31 bit block; the encoded message is given by multiplying the input message by the polynomial:[2]

$$g(x) = x^{10} + x^7 + x^5 + x^4 + x^2 + x + 1$$

The resulting code has minimum Hamming distance 5; a hard decoder can therefore correct any 2 bit error per block. Therefore the probability of a block decode error for $p = \mathbb{P}(bit\ error)$, is given by:

$$\mathbb{P}(block\ error) = 1 - (1-p)^{31} - 31p(1-p)^{30} - \binom{31}{2}p^2 \cdot (1-p)^{29}$$

The overall message is therefore decodable under this strategy if and only if there are no more than 2 bit errors for every block within the message. The probability of a message decode error, given input bit error p, is therefore given by:

$$\mathbb{P}(message\ error) = 1 - (1 - \mathbb{P}(block\ error))^3$$

Under this model, a 50% message error rate is achieved at a bit error rate of $p \approx 0.08$.

D. Related Work

Some existing works look at attacks on fingerprinting systems, but the field has not yet been widely explored. The primary focus of most of these works has been on impersonating device fingerprints: for example, the authors of [16] use arbitrary waveform generators to replay messages with sufficient precision that the fingerprint is duplicated. In [17] a similar approach is attempted using SDR hardware at lower sample rates, with limited success. It has also been demonstrated that location fingerprinting techniques (using radio characteristics to determine the location of transmitters) can be disrupted through Gaussian noise [18]. The authors of [19] provide an overview of research that targets Specific Emitter Identification (SEI) systems (an alternative terminology for fingerprinting); all of these works focus on replay, impersonation, and circumvention of the system. Finally, [20–22] investigate attacks on SEI/fingerprinting systems in which an attacker causes misclassification of malicious messages through the use of adversarial machine learning. To the best of our knowledge, denial-of-service attacks through jamming and disruption of transmitter fingerprinting systems have not yet been investigated.

TABLE I. TABLE OF COMPONENTS REQUIRED TO EXECUTE A JAMMING ATTACK ON IRIDIUM SYSTEMS.[3]

Component		Cost (USD)	Power/Gain
SDR	HackRF One	340	—
Module amplifier	Mini Circuits ZHL-5W-2GX+	1195	7 dB W
RF IC amplifier	Qorvo QPA2237	155+100[4]	10 dB W
GPS patch antenna	Pulse 673-GPSGMSMA	16	0 dB
High gain antenna	Chelton FPA21-16L/1258	—[5]	21 dB

III. THREAT MODEL

The objective of the attacker is to transmit radio interference, in the form of Additive White Gaussian Noise or a single tone, to cause errors in the fingerprinter which results in misclassification of the transmitter. The desired outcome, from the attacker's perspective, is to cause misclassification even though the original signal remains decodable. Therefore, as discussed in Section II, we measure the robustness of the fingerprinting system by considering the attacker-to-victim power ratio required to achieve a given error rate on the decoder and fingerprinter systems.

We assume that the attacker has access to off-the-shelf software-defined radio hardware and a suitable amplifier to transmit in the correct frequency band. The attacker's capabilities include maintaining presence within the vicinity of the receiver, such that the transmitted signal is received by the victim. We further assume that the victim antenna is omnidirectional, the type of antenna used in real-world Iridium transceivers [23].

A. Required Hardware Budget

We proceed to analyze the requirements of performing this attack with respect to two example transmitter systems; this allows us to estimate the hardware budget of the attacker, and understand the range over which the attacker can operate.

Firstly, the attacker uses a software-defined radio to emit either Gaussian noise or a single tone; no upconverter is required since Iridium operates at $1600\,\mathrm{MHz}$, which is well within the frequency range of most SDRs. Secondly, the attacker requires a suitable amplifier, which can either be bought off-the-shelf, or made from components to reduce cost. Finally, the attacker requires an antenna: we consider both an omnidirectional GPS patch antenna, which can deny service to receivers within a given area, and a high gain antenna which can target one particular receiver. The full hardware components required to run the attack are tabulated in Table I.

It can be seen that, including the cost of the SDR, the equipment can cost as little as $611\,\mathrm{USD}$; therefore deploying this attack is well within realm of a motivated hobbyist.

B. Effective Attack Range

To better understand the threat caused by attackers of this class, we consider the distances over which jamming can

---

[2]This ECC is implemented in the *iridium-toolkit* software decoder, which can be found at the following URL: https://github.com/muccc/iridium-toolkit.
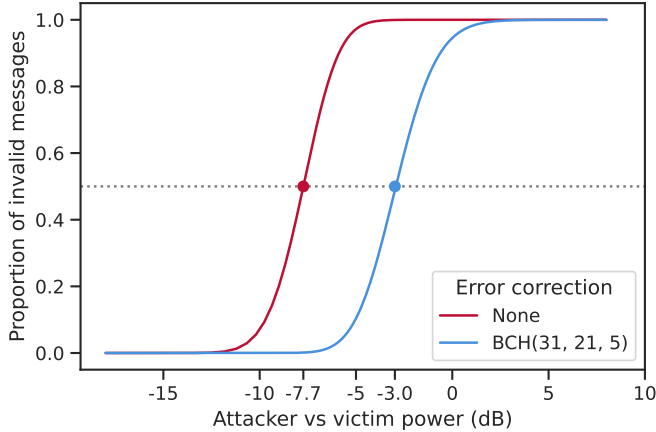
Fig. 2. The proportion of Iridium messages which fail to decode as the jammer power increases, with and without the use of Iridium's built-in error correcting codes. A dashed line represents the point at which half of all messages fail to decode.
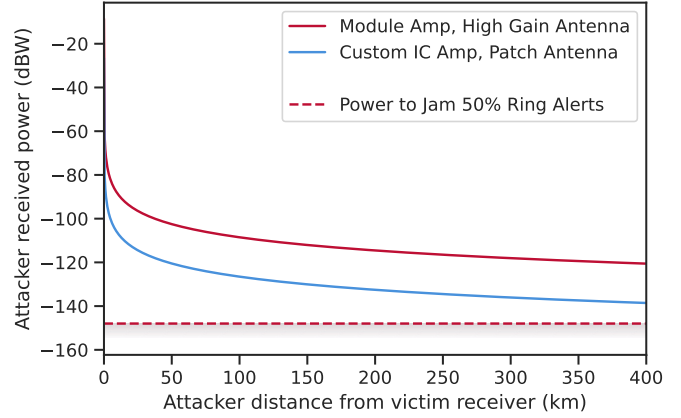

Fig. 3. The received power of a noise jammer as distance to the victim antenna varies, under free space path loss. The dashed line represents the power required to cause a 50% loss rate of Iridium Ring Alert messages; the region where any message loss is caused is shaded. It can be seen that with either set of equipment, line-of-sight attackers deny service over long distances.

be conducted on IRA messages with the hardware from the previous section.

Since IRA messages use Quadrature Phase Shift Keying (QPSK), it is established that the bit error probability corresponds to the *energy per bit to noise power spectral density ratio*, E_b/N_0 by:

$$\mathbb{P}(bit\ error) = 0.5 \cdot erfc(\sqrt{E_b/N_0})$$

where $erfc$ is the complementary error function derived from the cumulative Gaussian distribution.

Since QPSK encodes two bits per symbol, the jammer-to-signal power ratio is related to E_b/N_0 by:

$$E_b/N_0 = \frac{1}{2} \cdot \frac{1}{P_a/P_v}$$

Applying the ECC relationship between message error rate and bit error rate described in Section II, in which $\mathbb{P}(message\ error) = 0.5$ when $\mathbb{P}(bit\ error) \approx 0.08$, the attacker-to-victim power ratio to cause a 50% message error rate is given by:

$$P_a/P_v = \frac{1}{2}\frac{1}{erfc^{-1}(2 \cdot 0.08)} = 0.503 = -2.98\,\text{dB}$$

This is illustrated in Figure 2, which also shows the lower power required to deny service if error correction is not used.

We can finally calculate the distances at which this P_a/P_v can be achieved as distance varies, for our satellite system. It has been previously determined in the context of radio over-shadowing attacks that the peak received power of the Iridium-NEXT constellation is $-145\,\text{dB W}$ [24]; therefore service is denied when the attacker's incident power is $-145 - 2.98 = -147.98$ dB W. Assuming that the attacker has a clear line of sight to the victim, the attenuation of the attacker's signal over distance is given by the free space path loss formula:

$$fspl = 20\log_{10}(d) + 20\log_{10}(f) - 27.55$$

Where d is distance in meters, and f the frequency in MHz, $f \approx 1600\,\text{MHz}$ for Iridium.

By applying this attenuation to the attacker's transmit power using the hardware given in Table I, we can determine the distances over which service can be denied. The results are shown in Figure 3, where it can be seen that either set of equipment is capable of denying services at very long distances.

It may seem surprising that communications can be jammed even at distances of hundreds of kilometers, but it should be considered that the attacker's $10\,\text{dB W}$ transmitter is competing with a $9.5\,\text{dB W}$ victim transmitter that itself undergoes $781\,\text{km}$ of path loss [24]. Although we still expect these values to be lower in a real-world setting where the attacker's signal is subject to multipath propagation and line-of-sight restrictions, this analysis nonetheless shows that cheaply available hardware is more than sufficient to deny service at long distances.

IV. EXPERIMENT DESIGN

In this paper we test the hypothesis that satellite receiver systems can be more easily disrupted when fingerprinting systems are used to authenticate communication. We achieve this by computing the required power for an attacker to disrupt the fingerprint of a legitimate signal through jamming, and comparing this to the power required to disrupt the contents of the message through conventional jamming techniques. We have already established the attacker's constraints and ability to jam communication in Section III. To assess the robustness of a fingerprinting system in the presence of the same interference, we next gather datasets of Iridium message headers with attacking interference added onto the signal. We collected this data by adding noise onto real Iridium signals using radio hardware, and provide further datasets by adding synthetic interference to clean signals in software.

In this section we provide a summary of our data collection and processing pipeline, our software analysis, and an overview of the data collected.
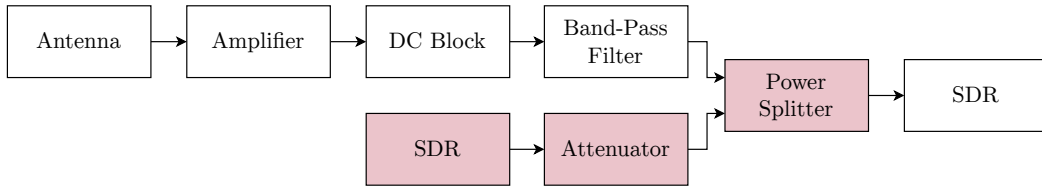
Fig. 4. Overview of the hardware used to collect Iridium signals with additional noise. The hardware that has been added to enable variable noise injection has been highlighted.
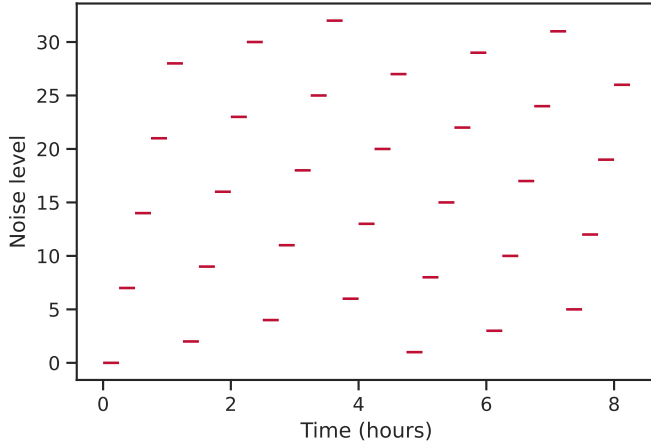

Fig. 5. Level of noise added to the collected data over time. This pattern loops every 8 hours.

A. Data Collection

Our data collection setup is similar to the one used in our original work on SatIQ [3], with one notable difference: an additional transmitting SDR has been connected to the receiver, allowing interference to be added onto the incoming signal. This setup is illustrated in Figure 4, and the hardware used is given in Table II.

In our previous work we note that the accuracy of the system is impacted by the time difference between the "anchor" message and the incoming message to be tested. To remove this as a factor from our experiments, we interleave each of the noise levels – this is illustrated in Figure 5. Every 15 minutes the data collection pipeline is restarted, and the noise level N is incremented according to the formula $N = 7 * i \mod 33$, where i is the number of restarts. This ensures an even spread of noise levels over time, and keeps the time gap between different portions of the dataset consistent. A Gaussian noise function is used, generating signals with amplitude N^2. Note that the amplitude of the noise here is arbitrary, and was selected to ensure a spread of values that covered the range our receiver was able to decode. For our analysis in Section V we look at the attacker-to-victim power ratio P_a/P_v, in decibels, computed from the received data. This is calculated as follows:

$$P_a/P_v = 20 \log_{10}((rms_a - rms_v)/rms_v)$$

Where rms_a and rms_v are the root-mean-squared amplitude of the attacker and victim signals respectively.[6]

TABLE II.   Hardware used for data collection.[7]

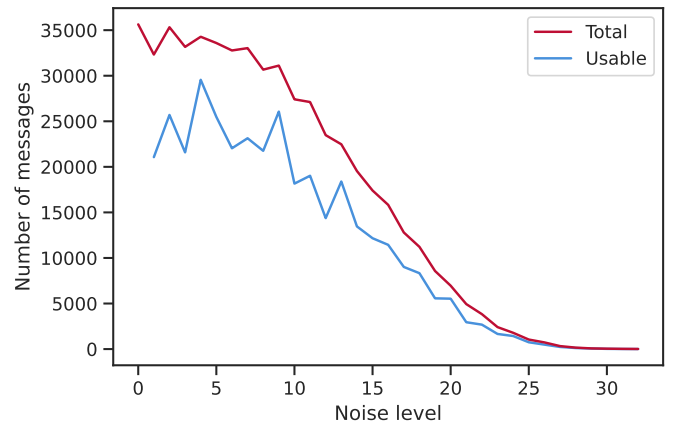| Component	Model number	Cost (USD)
Antenna	Iridium Beam RST740	1245
Low-noise amplifier	Mini-Circuits ZKL-33ULN-S+	209
DC block	NooElec	25
Band pass filter	Mini-Circuits VBF-1560+	45
SDR (receiving)	USRP N210, UBX 40 daughterboard	5086
SDR (transmitting)	BladeRF 2.0 micro xA4	540
Attenuator	Mini-Circuits BW-S20W20+	174
Power splitter	Mini-Circuits ZFRSC-123-S+	95

Fig. 6. Number of messages collected for each noise level, and the number of messages that are useful for analysis (messages from a transmitter also seen in the zero-noise dataset).

B. Data Analysis

We collected data using the above setup for 33 days, during which $540\,066$ messages were received. This dataset has been made openly available to aid future work.[8] For higher noise levels, fewer messages can be decoded; this is exacerbated by the software decoder's lack of error correction. We can see in Figure 6 that the number of received messages drops off smoothly as noise increases, and above a noise level of 32 (corresponding to $P_a/P_v = 7.57\,\mathrm{dB}$) no further messages were received. Also shown in this figure is the number of "usable" messages: the number of messages for which there exists at least one message from the same transmitter in the zero-noise control dataset. Without such a message to compare against, fingerprinting cannot take place. As data collection goes on, this will approach 100% of the dataset.

In Figure 8 we see the effect of adding noise onto the message headers: it is still discernible as a PSK modulated signal,
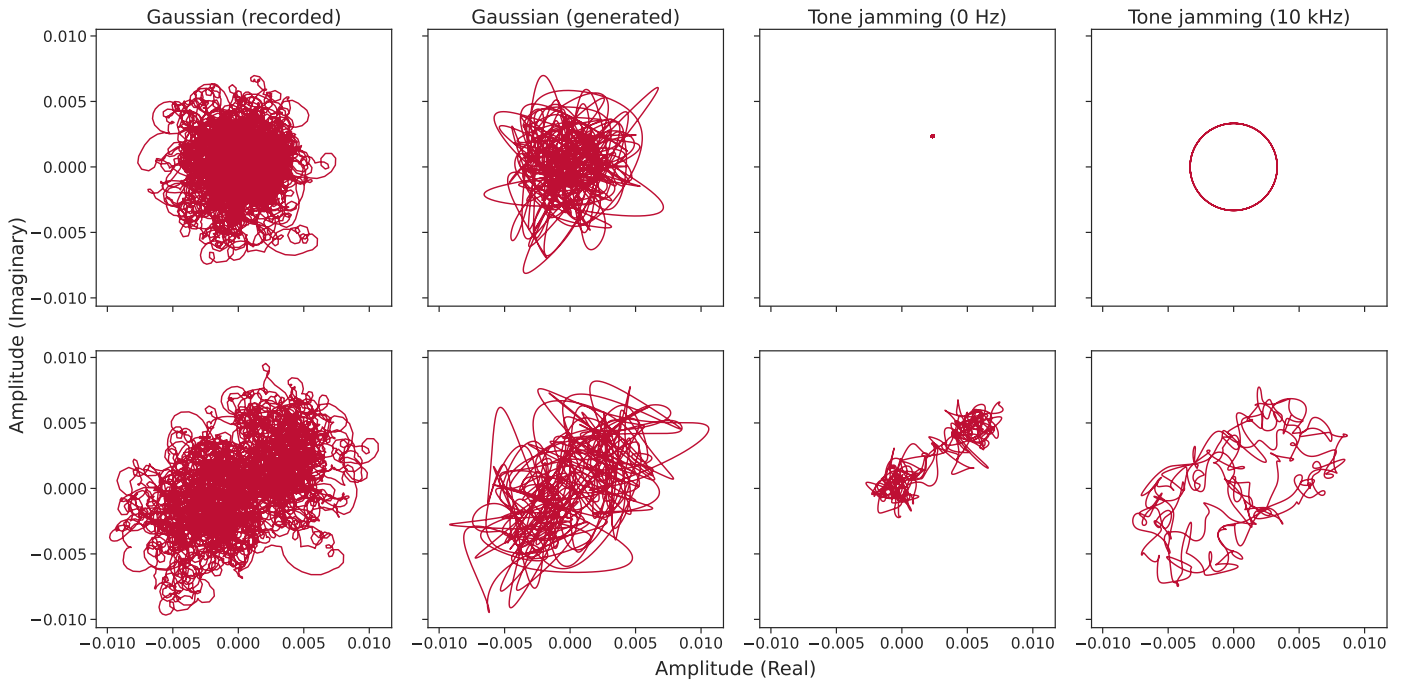
Fig. 7. Comparison of the noise generated for each jamming technique, as constellation diagrams. The top row shows the jamming signal by itself, and the bottom row is the same jamming signal added onto a legitimate Iridium message header.
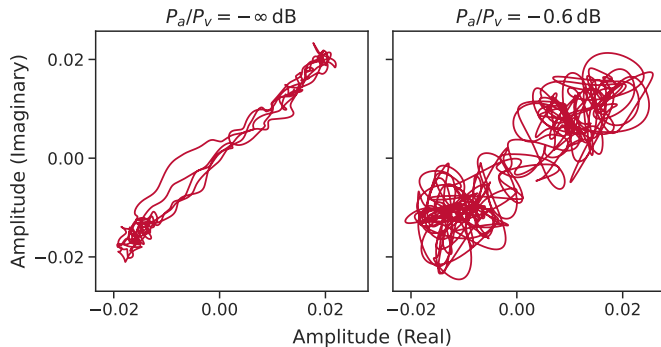

Fig. 8. Two Iridium message headers received during data collection, depicted as constellation plots. The message on the left has no noise added, and the message on the right has Gaussian noise added, with an attacker-to-victim power ratio of -0.6 dB.

but the waveform is (unsurprisingly) significantly noisier. Any of the original impairments on the signal which could have been used to identify the signal have likely disappeared into this noise.

C. Software Jamming

We also performed a software analysis, in which we added different noise and jamming signals to clean signals gathered from our data collection pipeline above. This enabled us to evaluate a wider range of jamming techniques, and removed our reliance upon the decoder pipeline – unlike the data collection above, the dataset does not shrink as we add more noise, since messages have already been demodulated and decoded.

We evaluate Gaussian noise jamming, as in the hardware experiments described above, and tone jamming, in which a constant frequency is added to the incoming signal. These are the two main jamming techniques explored in recent works [25]. For our analysis we use generated Gaussian noise as well as Gaussian noise recorded from our data collection hardware, and tone jamming at different relative frequencies.

We can see each of these jamming techniques illustrated in Figure 7. Each of these is calibrated to an attacker-to-victim ratio of -10 dB. Note that the recorded and generated Gaussian noise look slightly different – it is likely that the recorded Gaussian noise contains some higher frequencies outside the decoder's filter, resulting in a slightly different looking waveform. We also illustrate two different forms of tone jamming: at a relative frequency of 0 Hz, the tone jamming appears as a constant offset, shifting the position of the victim signal by a fixed amount but not changing its shape. At a higher relative frequency (10 kHz) this instead looks like a circle, shifting different parts of the waveform by different amounts.

V. RESULTS

In this section we assess the performance of the fingerprinting model against the noise and jamming signals collected in Section IV. We use the code and model weights provided for SATIQ [3], using the `ae-triplet-final.h5` checkpoint.

For each noise level, we compare the fingerprints of received messages to the fingerprints of "anchors" in the zero-noise dataset with the same transmitter ID to provide a distance metric, telling us how similar the two message headers are. If they are sufficiently different, the message will be rejected. We use the distribution of distances at lower noise levels to establish a threshold at which we accept messages – for this analysis we set this to be the 95th percentile.
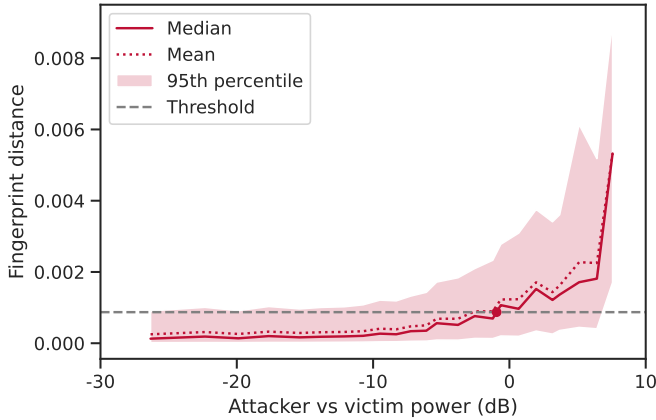
Fig. 9. Distance reported by the SATIQ fingerprinter for incoming messages under different levels of noise as captured by the real-world experimental setup, compared to anchor messages captured with no noise. A dashed line represents the fingerprinter threshold (set at the 95th percentile on the zero-noise dataset), below which messages are accepted by the system.
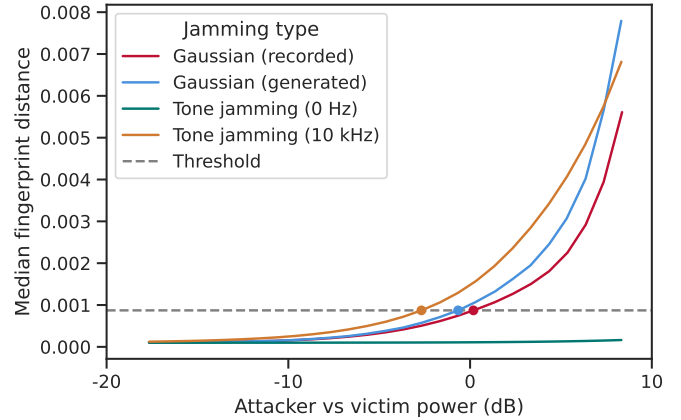

Fig. 10. Distance reported by the SATIQ fingerprinter for incoming messages with different levels of noise added in software, compared to anchors with no noise. The dashed line represents the fingerprinter threshold, accepting 95% of messages when no noise is added.

TABLE III. ATTACKER-TO-VICTIM POWER RATIO (P_a/P_v) REQUIRED TO DISRUPT 50% OF MESSAGES, UNDER EACH JAMMING EXPERIMENT.

Target	Jamming type	P_a/P_v (dB)
Decoder	Gaussian	-2.98
Fingerprinter	Gaussian (hardware)	-0.94
Fingerprinter	Gaussian (recorded)	0.18
Fingerprinter	Gaussian (generated)	-0.66
Fingerprinter	Tone jamming (0 Hz)	16.27
Fingerprinter	Tone jamming (10 kHz)	-2.68

We then find the attacker-to-victim power ratio at which 50% of messages fall above this threshold, and take this to be a successful jamming of the fingerprinting system. We compare this to the power required to disrupt communication using traditional methods, jamming the message through Gaussian noise. This has already been established in Section III to be $-2.98\,\mathrm{dB}$ for Iridium communication.

A. Real-World Experiment

We first look at the data collected using the hardware described in Section IV-A, in which noise is added onto the incoming signals using hardware. These results are summarized in Figure 9. Note the fluctuation at higher levels of attacker power – this is caused by the smaller datasets captured for those levels, as the noise interferes with the decoding pipeline. We can see that, as expected, adding noise to the signal disrupts the fingerprint. The median fingerprint distances crosses the acceptance threshold at an attacker-to-signal ratio of $-0.94\,\mathrm{dB}$; at this point, half of all messages will be rejected by the fingerprinting system. We also plot the mean for comparison purposes – as noise increases, the variance of the distribution of fingerprints increases, bringing the mean distance up faster than the median distance.

This compares very favorably to the jamming analysis in Section III, in which we conclude that an attacker needs $-2.98\,\mathrm{dB}$ to disrupt the message decoder through jamming. Under this configuration, it takes more energy to disrupt the fingerprinter than it does to jam messages.

B. Software Jamming

We also look at the jamming signals added in software in Section IV-C. We can see the effect on the fingerprint as the attacker's transmit power increases in Figure 10. Additionally, the attacker-to-victim power required to disrupt 50% of messages in this way can be seen in Table III. The results for Gaussian jamming (both recorded and generated) are close to our experimental results, although slightly more attacker power is required than in the real-world evaluation. This is likely due

to the additional frequencies present in the Gaussian noise, outside the filter used by the decoder.

We can also see that tone jamming at a relative frequency of $0\,\mathrm{Hz}$ is ineffective; this is likely due to the heavy use of convolutional layers in the fingerprinting model, which are primarily concerned with relative changes in the waveform rather than absolute position. However, we note that tone jamming at a relative frequency of $10\,\mathrm{kHz}$ is highly effective, successfully disrupting the fingerprinter at only $-2.68\,\mathrm{dB}$ relative to the victim signal.

Finally, in Figure 11 we can see the proportion of messages rejected by the system as the attacker power increases, for each jamming technique. We can see from this that tone jamming is significantly more effective at disrupting the fingerprint than Gaussian noise, disrupting over 90% of messages at $0\,\mathrm{dB}$ relative to the victim signal, and reaching 100% at approximately $2.5\,\mathrm{dB}$.

VI. DISCUSSION

In Section III we look at the error rate of the Iridium decoder under Gaussian jamming, as the received attacker power increases. In Figure 12 we compare this to the fingerprinter error under each of the most effective jamming techniques, from both our hardware experiment and software analyses.[9] We can see from this figure that under every form

---

[9]Note that the baseline error rate for the fingerprinter depends upon the threshold at which messages are accepted. In this case we set the threshold at the 95th percentile, so the base error rate is 0.05.
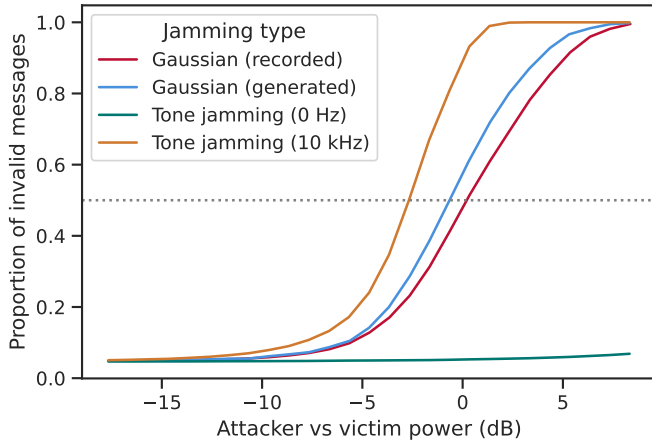
Fig. 11. The proportion of victim messages rejected by the fingerprinting system as the attacker signal strength increases, under different jamming techniques.
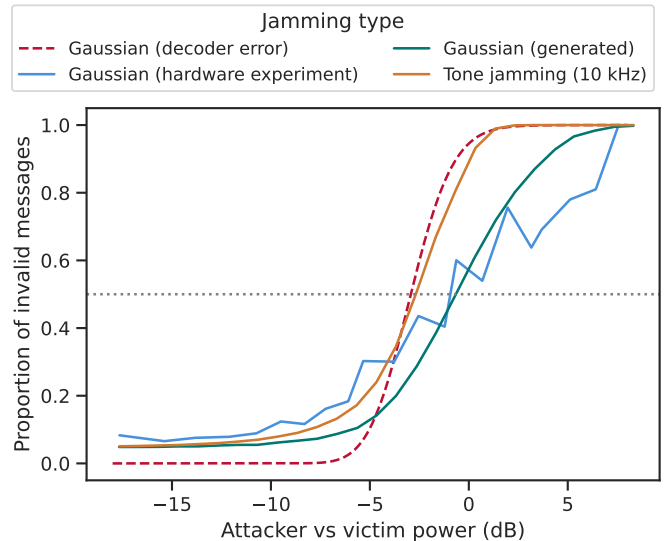

Fig. 12. The proportion of invalid messages as received attacker power increases, comparing the theoretical decoder error (dashed line) to the fingerprinter error under the best-performing experimental configurations.

of jamming tested, it takes a greater amount of attacker power to disrupt the fingerprint than to disrupt the message through jamming. In all cases, the difference in required attacker power is quite small, within $2.5\,\mathrm{dB}$. Tone jamming also requires much less power to disrupt the fingerprint than Gaussian jamming, getting much closer to the power level required for jamming. under Gaussian jamming, the decoder error rate is greater than the fingerprinting system's message rejection rate as the attack power increases. We have therefore refuted our initial hypothesis, showing that the use of fingerprinting to authenticate satellite communication does not significantly weaken the system against jamming attacks.

There are a range of potential explanations for why this is the case. First and foremost, fingerprinting in the context of satellite systems is much more difficult than terrestrial systems, due to the high levels of free space path loss and atmospheric distortion; any identifiable information present in the signal will be subject to significant atmospheric attenuation. In order to achieve good performance, the fingerprinting model must be able to overcome this, and it is likely that in the process of doing so it becomes good at ignoring Gaussian noise – this would be unsurprising, as convolutional autoencoders (the architecture used by this model) have been demonstrated in other works to be effective at removing noise in a range of applications [26, 27]. The fingerprint therefore only fails to be identified when the injected noise is so high that the message itself can no longer be decoded. It may be the case that fingerprinting systems trained on terrestrial data are less robust against noise – they do not need to perform nearly as much noise reduction as satellite fingerprinters, so a smaller amount of noise could be required to disrupt them.

It is also important to note that the decoder and fingerprinter are fundamentally operating on different parts of the signal. The fingerprinter operates on the message header only – this portion of the message is unchanging, and designed to be robust and easy to correlate over. On the other hand, the decoder primarily operates on the message body, which changes between messages and is generally expected to fail first in the presence of noise – enough noise needs to be added to introduce only a small number of bit errors, which may

not destroy all identifying information present in the message header.

It is possible that an attacker aiming to specifically disrupt the fingerprinting system may be able to craft a waveform that disrupts the fingerprinter even more effectively, with lower transmit power. There is a large body of existing research that has demonstrated the effectiveness of adversarial machine learning techniques to execute these types of attacks, in systems ranging from image classifiers [28] to network intrusion detection systems [29]. There has also been some work investigating terrestrial RF fingerprinting systems, with theoretical analyses [21] and small-scale experimentation [22]. Future work may consider evaluating this type of attack on satellite fingerprinting systems. Such techniques usually require black-box access to the target system, if not access to the underlying model weights, and are highly targeted to a single system. We therefore do not consider them in this work, instead considering an attacker employing the use of standard jamming techniques, and concluding that the use of fingerprinting does not open up significant new weaknesses in this case.

VII. Conclusion

In this paper we have provided a novel analysis of the vulnerability of satellite fingerprinting systems against Gaussian and tone jamming interference. We show that the SATIQ fingerprinting system is not significantly more vulnerable to this interference than the Iridium message decoder itself, so fingerprinting can be used without opening the system up to easy denial-of-service attacks.

Existing legacy satellite systems are likely to remain in use for years or decades to come, and in the absence of cryptographic authentication, fingerprinting techniques can provide confidence that downlinked data is legitimate and has not been spoofed or tampered with. This provides an increased degree

of trust in the systems, enabling them to remain useful for the remainder of their operational period.

REFERENCES

[1] Bingyin Ren, Hailong Ge, Guangfei Xu, and Yongxin Zhang. "Anti-Jamming Analysis and Application of Starlink System". In: *International Conference on Networking, Informatics and Computing (ICNETIC)*. 2023.

[2] Valerie Insinna. "SpaceX Beating Russian Jamming Attack was 'Eyewatering': DoD Official". In: *Breaking Defense* (2022). https://breakingdefense.com/2022/04/spacex-beating-russian-jamming-attack-was-eyewatering-dod-official/.

[3] Joshua Smailes, Sebastian Köhler, Simon Birnbach, Martin Strohmeier, and Ivan Martinovic. "Watch This Space: Securing Satellite Communication through Resilient Transmitter Fingerprinting". In: *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*. 2023, pp. 608–621.

[4] Naeimeh Soltanieh, Yaser Norouzi, Yang Yang, and Nemai Chandra Karmakar. "A Review of Radio Frequency Fingerprinting Techniques". In: *IEEE Journal of Radio Frequency Identification* 4.3 (2020), pp. 222–233. ISSN: 2469-7281.

[5] NASA. *Fire Information for Resource Management System (FIRMS)*. 2022. URL: https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms (visited on 05/03/2022).

[6] sam210723. *Receiving Images from Geostationary Weather Satellite GEO-KOMPSAT-2A*. 2020. URL: https://vksdr.com/xrit-rx (visited on 05/09/2022).

[7] sam210723. *COMS-1 LRIT Key Decryption*. 2018. URL: https://vksdr.com/lrit-key-dec (visited on 05/09/2022).

[8] Kasper Bonne Rasmussen and Srdjan Capkun. "Implications of Radio Fingerprinting on the Security of Sensor Networks". In: *2007 Third International Conference on Security and Privacy in Communications Networks and the Workshops-SecureComm 2007*. IEEE, 2007, pp. 331–340. ISBN: 1-4244-0974-8.

[9] Mahsa Foruhandeh, Abdullah Z. Mohammed, Gregor Kildow, Paul Berges, and Ryan Gerdes. "Spotr: GPS Spoofing Detection via Device Fingerprinting". In: *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. WiSec '20. Association for Computing Machinery, July 2020, pp. 242–253. DOI: 10.1145/3395351.3399353.

[10] Weidong Wang and Lu Gan. "Radio Frequency Fingerprinting Improved by Statistical Noise Reduction". In: *IEEE Transactions on Cognitive Communications and Networking* (2022). ISSN: 2332-7731.

[11] Gabriele Oligeri, Simone Raponi, Savio Sciancalepore, and Roberto Di Pietro. "PAST-AI: Physical-layer Authentication of Satellite Transmitters via Deep Learning". In: *arXiv:2010.05470 [cs]* (Oct. 2020). eprint: 2010.05470.

[12] ÖH Tekbaş, Oktay Üreten, and Nur Serinken. "Improvement of Transmitter Identification System for Low SNR Transients". In: *Electronics Letters* 40.3 (2004), pp. 182–183. ISSN: 1350-911X.

[13] Frederick Rawlins, Richard Baker, and Ivan Martinovic. "Death By A Thousand COTS: Disrupting Satellite Communications using Low Earth Orbit Constellations". In: SpaceSec '23 (2023).

[14] Samuel Lefcourt, Nathaniel Gordon, Hanting Wong, and Gregory Falco. "Space cognitive communications: Characterizing radiofrequency interference to improve digital space domain awareness". In: *2022 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE. 2022, pp. 1–7.

[15] Matthew Walters and Sujoy Sinha Roy. "Constant-time BCH error-correcting code". In: *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE. 2020, pp. 1–5.

[16] Boris Danev, Heinrich Luecken, Srdjan Capkun, and Karim El Defrawy. "Attacks on Physical-Layer Identification". In: *Proceedings of the Third ACM Conference on Wireless Network Security*. 2010, pp. 89–98.

[17] Saeed Ur Rehman, Kevin W Sowerby, and Colin Coghill. "Analysis of Impersonation Attacks on Systems Using RF Fingerprinting and Low-End Receivers". In: *Journal of Computer and System Sciences* 80.3 (2014), pp. 591–601. ISSN: 0022-0000.

[18] Philipp Richter, Mikko Valkama, and Elena Simona Lohan. "Attack Tolerance of RSS-Based Fingerprinting". In: *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. Barcelona: IEEE, Apr. 2018, pp. 1–6. ISBN: 978-1-5386-1734-2. DOI: 10.1109/WCNC.2018.8376977. (Visited on 01/04/2024).

[19] Joshua H Tyler, Mohamed KM Fadul, and Donald R Reising. "Considerations, Advances, and Challenges Associated with the Use of Specific Emitter Identification in the Security of Internet of Things Deployments: A Survey". In: *Information* 14.9 (2023), p. 479. ISSN: 2078-2489.

[20] Liting Sun, Da Ke, Xiang Wang, Zhitao Huang, and Kaizhu Huang. "Robustness of Deep Learning-Based Specific Emitter Identification under Adversarial Attacks". In: *Remote Sensing* 14.19 (2022), p. 4996. ISSN: 2072-4292.

[21] Boyang Liu, Haoran Zhang, Yiyao Wan, Fuhui Zhou, Qihui Wu, and Derrick Wing Kwan Ng. "Robust Adversarial Attacks on Deep Learning Based RF Fingerprint Identification". In: *IEEE Wireless Communications Letters* (2023). ISSN: 2162-2337.

[22] Samurdhi Karunaratne, Enes Krijestorac, and Danijela Cabric. "Penetrating RF Fingerprinting-based Authentication with a Generative Adversarial Attack". In: *ICC 2021-IEEE International Conference on Communications*. IEEE, 2021, pp. 1–6. ISBN: 1-72817-122-9.

[23] Beam Communications. *Iridium Beam Active Antenna RST740, Technical Specifications*. Tech. rep. 2017. URL: https://www.beamcommunications.com/document/160

- brochure - rst740 - iridium - active - antenna (visited on 01/11/2024).

[24] Edd Salkield, Marcell Szakály, Joshua Smailes, Sebastian Köhler, Simon Birnbach, Martin Strohmeier, and Ivan Martinovic. "Satellite Spoofing from A to Z: On the Requirements of Satellite Downlink Overshadowing Attacks". In: *Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. 2023, pp. 341–352.

[25] Paola Martinelli, Ernestina Cianca, Mauro De Sanctis, Lanfranco Di Paolo, Alessandro Pisano, and Lorenzo Simone. "Robustness of Satellite Telecommand Links to Jamming Attacks". In: *2012 IEEE First AESS European Conference on Satellite Telecommunications (ESTEL)*. IEEE, 2012, pp. 1–6. ISBN: 1-4673-4688-8.

[26] Aya Saleh Ahmed, Wessam H El-Behaidy, and Aliaa AA Youssif. "Medical Image Denoising System Based on Stacked Convolutional Autoencoder for Enhancing 2-Dimensional Gel Electrophoresis Noise Reduction". In: *Biomedical Signal Processing and Control* 69 (2021), p. 102842. ISSN: 1746-8094.

[27] Debjani Bhowick, Deepak K Gupta, Saumen Maiti, and Uma Shankar. "Stacked Autoencoders Based Machine Learning for Noise Reduction and Signal Reconstruction in Geophysical Data". In: *arXiv preprint arXiv:1907.03278* (2019). arXiv: 1907.03278.

[28] Gabriel Resende Machado, Eugênio Silva, and Ronaldo Ribeiro Goldschmidt. "Adversarial Machine Learning in Image Classification: A Survey Toward the Defender's Perspective". In: *ACM Computing Surveys (CSUR)* 55.1 (2021), pp. 1–38. ISSN: 0360-0300.

[29] Huda Ali Alatwi and Charles Morisset. "Adversarial Machine Learning in Network Intrusion Detection Domain: A Systematic Review". In: *arXiv preprint arXiv:2112.03315* (2021). arXiv: 2112.03315.