Hindawi
Security and Communication Networks
Volume 2020, Article ID 6641844, 13 pages
https://doi.org/10.1155/2020/6641844

*Review Article*

# Research and Analysis of Electromagnetic Trojan Detection Based on Deep Learning

**Jiazhong Lu,[1] Xiaolei Liu ,[2] Shibin Zhang,[1] and Yan Chang[1]**

[1] *School of Cybersecurity, Chengdu University of Information Technology, Chengdu 610225, Sichuan, China*
[2] *Institute of Computer Application, China Academy of Engineering Physics, Mianyang, Sichuan 621900, China*

Correspondence should be addressed to Xiaolei Liu; liuxiaolei@caep.cn

Received 22 October 2020; Revised 28 October 2020; Accepted 9 November 2020; Published 25 November 2020

Academic Editor: Ting Chen

The electromagnetic Trojan attack can break through the physical isolation to attack, and the leaked channel does not use the system network resources, which makes the traditional firewall and other intrusion detection devices unable to effectively prevent. Based on the existing research results, this paper proposes an electromagnetic Trojan detection method based on deep learning, which makes the work of electromagnetic Trojan analysis more intelligent. First, the electromagnetic wave signal is captured using software-defined radio technology, and then the signal is initially filtered in combination with a white list, a demodulated signal, and a rate of change in intensity. Secondly, the signal in the frequency domain is divided into blocks in a time-window mode, and the electromagnetic signals are represented by features such as time, information amount, and energy. Finally, the serialized signal feature vector is further extracted using the LSTM algorithm to identify the electromagnetic Trojan. This experiment uses the electromagnetic Trojan data published by Gurion University to test. And it can effectively defend electromagnetic Trojans, improve the participation of computers in electromagnetic Trojan detection, and reduce the cost of manual testing.

## 1. Introduction

With the development of information technology, electronic devices of various functions are continuously designed, such as computers and printers, which generate electromagnetic radiation during use. Electromagnetic radiation can also leak data in the device, threatening information security. Electromagnetic leakage is divided into two categories. The first type is passive leakage, which is an inevitable leakage caused by electronic equipment in normal work. Electromagnetic waves leaking from the display were captured by Hidema Tanaka [1]. The second category is active leakage, such as the experiment by Kuhn and Anderson [2], which conducts electromagnetic leakage in the form of actively transmitting specified information. This type of electronic leakage is malicious hardware or software in an electronic device that regularly leaks a specified signal to the outside world by controlling the electronic device. This type of electromagnetic leakage is called an electromagnetic Trojan [3]. Different from other common computer Trojan viruses, this type of electromagnetic Trojan does not use the system equipment such as the network to exchange information with the outside world, which causes the electromagnetic Trojan to break through physical isolation and be more difficult to detect. Such electromagnetic Trojans not only threaten the information security of ordinary users but may even threaten the physically isolated internal network.

At present, the main defense methods for electromagnetic Trojan attacks focus on passive defense, such as electromagnetic shielding and signal interference. The active detection methods mainly focus on monitoring the electromagnetic signals in the range, establishing a white list of normal signals and capturing and storing electromagnetic signals. The experienced security personnel observed the signal spectrum to find the electromagnetic Trojan attack, which is extremely inefficient [4]. Moreover, due to a large number of electromagnetic waves in the space, the amount of data captured per second can reach 1G to 2G [5]. Therefore, the classification efficiency by manual means is low. At the same time, the black/white list-based electromagnetic wave data processing scheme cannot effectively detect the new electromagnetic Trojan attack.

Although the study of electromagnetic leakage has not stopped so far, there is no effective and active defense method. Our summary study found that the defense methods of electromagnetic Trojans can be divided into the following categories: (1) cutting off the transmission channel; (2) applying protection on the device; (3) electromagnetic wave record analysis. In summary, most of the defense methods for electromagnetic Trojans focus on the protection of electromagnetic wave channels, or the electromagnetic Trojans are defended by software-defined radio-related APIs.

We propose to monitor the electromagnetic wave signals in the monitoring area and characterize the signals. The flow calculation of big data is used to process electromagnetic wave data with huge data volume online. At the same time, we combine the fast classification function of deep learning to detect abnormal signals. Therefore, the protection of information security can greatly improve the detection efficiency of the electromagnetic Trojan signal, improve the security of the physical isolation network, and promote the development and progress of the electromagnetic Trojan detection.

## 2. Related Work

Xu et al. [6] used electromagnetic wave leakage to design a hardware Trojan for information acquisition of specific equipment. Their experiments show that the electromagnetic Trojan can silently obtain the 128 bit AES encryption key stored in the crypto chip.

Xu et al. [7] conducted a theoretical analysis of the electromagnetic leakage of the power line and established a radiation leakage model based on the power line. The electromagnetic leakage of the power line was verified by the measured data.

At the Black Hat Conference in August 2018, the Eurecom research team presented at the conference and demonstrated the results of using the electromagnetic leakage principle to attack mixed-signal wireless chips. By capturing the electromagnetic leakage of the chip, the key information of the encryption chip is obtained, so that the highly secure cryptographic algorithm is no longer safe.

In our previous work [8], considering the features of different power sources in different locations, combined with spark streaming technology and machine learning classification technology, a joint platform for electromagnetic signal anomaly detection based on big data analysis is proposed. The electromagnetic signal is abnormally detected by feature comparison and small-signal analysis, and the position and number between the signal sources are determined by three-point positioning and signal attenuation. The experimental results show that the method can detect abnormal electromagnetic signals and classify abnormal electromagnetic signals well, and the accuracy rate can reach 95%, and the positioning accuracy can reach 89%. However, our previous method can only detect abnormal electromagnetic signals, but it cannot detect electromagnetic Trojans.

Aiming at the above problems, this paper proposes an electromagnetic environment anomaly detection method based on deep learning. The electromagnetic Trojan signal is analyzed and modeled from the perspectives of time, frequency domain, and time domain. A large number of features are proposed to distinguish the normal signal from the electromagnetic Trojan signal, and the electromagnetic signal analysis system in the big data environment is designed. The system supports electromagnetic signal capture, filtering, characterization, and online/offline analysis in a big data environment. The system has a strong ability to expand through a hierarchical division of labor. Finally, the electromagnetic signal detection data set is constructed by using the electromagnetic Trojan disclosed by Ben Gurion University. The algorithm and design prototype system proposed in this paper are tested and evaluated.

### 2.1. Our Contribution

(1) This paper proposes the features of the new electromagnetic Trojan, which can analyze the signal in the frequency domain, time domain, and signal working mode and distinguish between abnormal signals and noise signals.

(2) This paper presents a new signal research perspective. In the traditional electromagnetic wave analysis technology, most of the research angles are the time domain and frequency domain information of the signal. In this paper, the working mode of the electromagnetic Trojan is compared and analyzed, and the sampled signal is compared to the flow data packet. A complete signal flow is analogous to the one-session mode. This method can improve the detection accuracy.

(3) We designed an electromagnetic signal analysis system that satisfies the big data environment. Through strict modular segmentation, the multidevice collector is supported in parallel to collect signals, filter data, and summarize and analyze the signals, so that the system has high ductility. At the same time, the monitoring bandwidth of the signal is increased through multidevice time slice rotation, various signal filtering schemes are designed, and the pressure of system data processing is reduced by signal filtering.

## 3. Electromagnetic Trojan Signal Detection Method

This paper proposes a joint platform for electromagnetic signal detection based on big data analysis. The electromagnetic Trojan in hardware is detected by the features of electromagnetic signals in time, information quantity, and energy.

*3.1. System Design.* This paper proposes a scheme for capturing electromagnetic waves using software-defined radio technology and proposes several new features for characterizing the electromagnetic wave signals. The method proposed in this paper overcomes the determination of the existing scheme and combines the advantages of the existing scheme to perform electromagnetic Trojan detection. With the help of software-defined radio technology, the scope of electromagnetic wave monitoring is increased, and the problem of a large amount of data caused by increasing the signal acquisition range is overcome by proposing a large number of new features into the machine learning model for classification.

In the detection method of this chapter, the electromagnetic signal is first collected, the main acquisition range is 10 Hz–3000 MHz signal data, and the abnormality and background noise are distinguished; secondly, the electromagnetic signal is analyzed and filtered; the purpose is to screen out the meaningful signal, reduce storage and detection expenses, characterize the filtered signal from the perspective of energy, information volume, time, etc. in the time domain and the frequency domain, convert it into a time-dependent feature sequence depicting the state of the signal corresponding to the time state, and finally use LSTM. The method performs classification detection, selects a suspicious abnormal signal, and locates and alarms the signal according to the intensity information of the signal.

Our approach procedures are shown in Figure 1, which contains the key stages and necessary operations of the Trojan signal detection method.

The first stage is the electromagnetic signal capture phase, and the signal capturer and signal analysis plug-in are connected to the laboratory's server. In this way, the server can directly control the arrester for real-time signal capture and analyze the electromagnetic signal to find out the signal data being transmitted within the monitored frequency range. This automated work is a huge improvement in detection efficiency.

The second stage is the electromagnetic signal filtering stage. For large data electromagnetic signals, the signal must be filtered to save detection cost and shorten the detection time. Therefore, most of the normal, white, background, and noise signals need to be excluded. Three filtering methods are used in this method to effectively do this.

The third stage is the feature extraction stage. The filtered signal is divided into blocks in the form of the time window, and the filtered electromagnetic wave signals are extracted from the angles of energy, time, and information entropy, and a serialized electromagnetic wave feature matrix is generated.

The fourth stage is anomaly detection, and there is no effective characterization of the working mode and periodic features of the signal in the serialized signal feature matrix. Therefore, the LSTM algorithm is used to further mine the time regularity of the signal and feature extraction. Then, softmax is used to classify the features of LSTM mining and detect abnormal signals.

*3.2. Electromagnetic Signal Processing.* Signal processing includes signal acquisition and signal filtering.

*3.2.1. Signal Acquisition.* This paper uses the software-defined radio (SDR) technology to capture radio signals in space. This technology can easily capture, analyze, and process signals in each frequency band in the current location. Software-defined ratios can acquire electromagnetic wave signals within a specified range by dynamically configuring the operating frequency band and sampling range. Such a collection method does not need to modify the hardware device and can conveniently collect the signals in the monitoring range by using the time slice rotation mode.

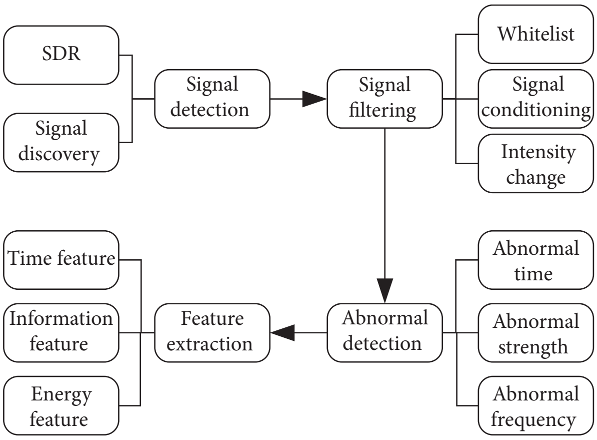


Figure 1: Approach procedures.

Since the modulation modes of the signals in the respective frequency bands are different, the specific bearer content of the signal cannot be obtained in the data acquisition phase. Therefore, the signals we acquire at this stage are complex signal data in the frequency domain, and then these signals are further processed according to the frequency band and modulation protocol.

Unlike other signal receiving devices, the normal signal acceptance process requires the sender and the receiver to agree on the transmission time, the transmission band, and the coding mode. In our signal acquisition process, we need to monitor the electromagnetic signals in the range, which makes it impossible to specify the working frequency range of the signal. Therefore, we need one more signal discovery process than other signal receiving devices in the working process. This paper proposes a signal discovery method based on the signal processing scheme in random signal analysis. The source of the background noise signal is mainly the low-energy signal after the band-pass filtering of the signal outside the sampling range. Therefore, although the background signal is indeterminate in intensity, the intensity value follows a uniform distribution in the distribution, the intensity mean difference is small, and the intensity variance is also smaller; the signaled channel and the unsignaled edge zone will suddenly increase the intensity of the sampling point, so the intensity mean and intensity variance will increase compared to the background signal; the sampling point is the signaled channel. When the intensity of the sampling point reaches the maximum value, the intensity variance will be larger. In this paper, by sampling the number of sampling points, the sampling points with signals are searched in a rolling manner, and the corresponding information is sent to the signal filtering module for filtering, which reduces the load caused by the useless data on the system. So, in addition to capturing the signal, the module needs to pick the true signal from the background noise. The collection procedures are as shown in Figure 2. Set the center frequency *f* and the sampling rate *s* for each device and then control the device for sampling. The variance and the mean are calculated every *k* samples. If the mean is large, the sample point is retained; otherwise, the sample point variance is calculated. If the variance is greater than the defined variance threshold, discard these points; otherwise, it is considered that there is a signal at these sampling points to save the data.

Figure 3(a) shows the signal power distribution of the corresponding frequency range. The ordinate unit is dBm. Figure 3(b) is the variance distribution of the intensity of the corresponding sampling point. In Figure 3, we sample 4096 units and perform variance calculation. We can see that there are two obvious signals in Figure 3(a). In the corresponding position in Figure 3(b), you can see two peaks with large variances (the first half of Figure 3(b)). Some data have no obvious variance fluctuations, and the reason why the data in the first half of Figure 3(a) are significantly different is that the power and intensity are logarithmically processed during the conversion, so the small difference between the intensities is logarithmic. The process is enlarged, so the variance changes during the calculation process are not obvious. The signal can be effectively found by the variance of the power of the sampling points in the corresponding frequency range [9].

(1) Whitelist Signal: the whitelist mechanism is added, the signal frequency is excluded in the whitelist, and the whitelist library supports user customization, such as broadcast signals and special frequency signals. The whitelist is shown in Table 1.

(2) It can mediate the signal: this type of signal can be filtered by signal protocol unpacking. For example, the FM broadcast signal is unpacked by the FM protocol to filter out the FM signal, as shown in Table 2.

(3) Intensity change signal: the filtering scheme adopts multiple collection points to collect at the same time, then calculates the intensity of the collected signal and filters according to the intensity of the collected signal, and filters out the signal transmission source are not within the system scope.

*3.2.2. Signal Filtering.* Before data analysis, the data need to be preprocessed; the purpose is to delete duplicate information and redundant information and reduce the data processing load of the system under the premise of preserving abnormal data. In the electromagnetic Trojan signal analysis platform of this paper, three filtering modes are adopted:

In filtering electromagnetic signals, some fixed frequencies may be used inside the system as a frequency band for secure communication or a frequency band not suitable for monitoring. Therefore, for these frequency bands, we have established a whitelist list as shown in Table 1. First, we use whitelist signal filtering to compare the input signal frequency with the frequency in the whitelist and directly filter it into a normal signal. If it is not in the turn, it can be used to mediate signal filtering.

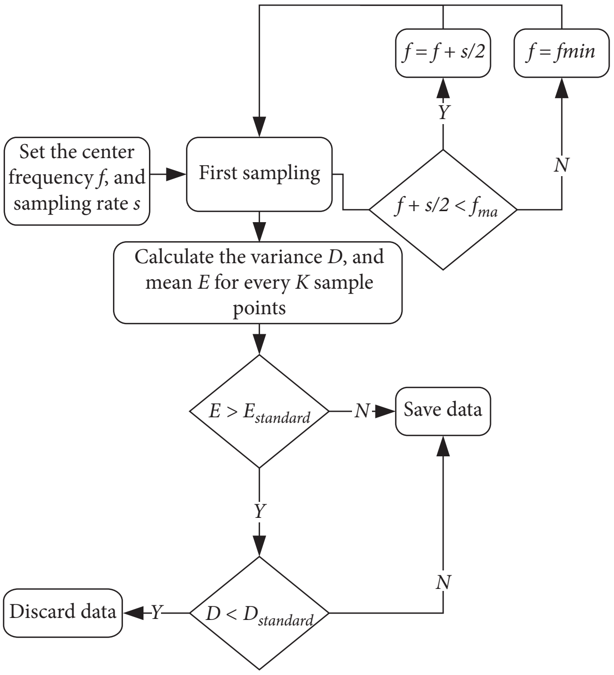

$f = f + s/2$
$f = fmin$
Y
Set the center frequency $f$, and sampling rate $s$
First sampling
N
$f + s/2 < f_{ma}$
Calculate the variance $D$, and mean $E$ for every $K$ sample points
$E > E_{standard}$
N
Save data
Y
N
Discard data
Y
$D < D_{standard}$

Figure 2: Collection procedures.

Among the received electromagnetic signals, there are a large number of electromagnetic signals that are normal electromagnetic signals conforming to the specifications, such as broadcast signals and mobile communication signals. Therefore, a demodulate signal filtering scheme is added after the whitelist signal. In the mediation signal filtering, the signal demodulation plug-in is used to demodulate and filter the electromagnetic signal and combined with the distribution of the signal frequency band in the analysis region, the signal that can be normally demodulated and conforms to the specification is used as a normal signal, and the protocol cannot be parsed or not. The signal of the local signal list is output to the intensity change signal for filtering. The demodulate electromagnetic signal protocol in the partially monitored frequency band is shown in Table 2.

In the intensity change signal, the greater the intensity change value, the closer the signal source is to the collection point, otherwise, the farther the signal source is from the collection point. The following relationship exists between the wireless signal transmission power *PR* and the received power PT:

$$\mathrm{PR} = \frac{\mathrm{PT}}{(r * n)}, \tag{1}$$

where *r* is the distance between the transmitting point and the receiving point and *n* is the environmental constant. From this, we can deduct the relationship between the fixed point of the launching point and the fixed distance of the two receiving points and the distance between the transmitting point and the intensity:

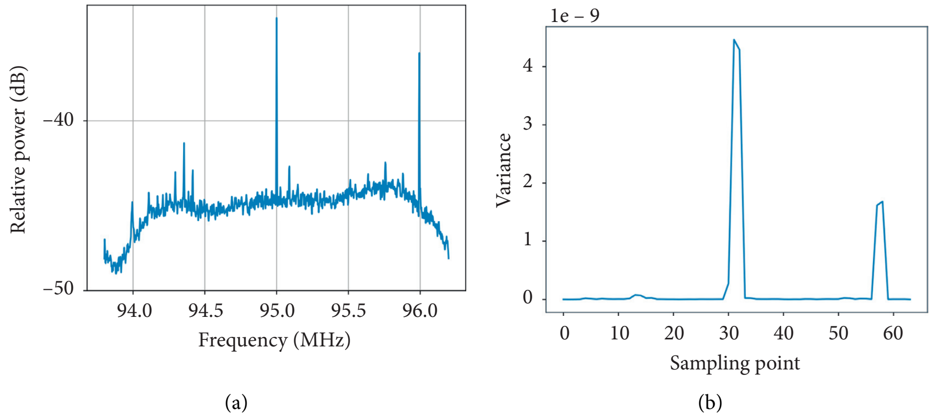


Figure 3: Performance of the frequency domain signal on the variance. (a) Signal frequency domain diagram from 94 to 96 MHz; (b) Sample point variance distribution from 94 to 96 MHz.

Table 1: Partial whitelist signals.

| Frequency band (MHz) | |
|---|---|
| 1.7 | 50 |
| 1.9 | 144.100 |
| 3.5/3.8 | 144.200 |
| 10.7 | 1000 |
| 2.4 | 2412 |

Table 2: 10 Hz – 3000 MHz partial signal protocol.

| Frequency band (MHz) | Protocol |
|---|---|
| 1710–1725<br>885–909 | Mobile 2G signal protocol |
| 2010–2025 | Mobile 3G signal protocol |
| 1880–1890<br>2320–2370<br>2575–2635 | Mobile 4G signal protocol |
| 909–915<br>1745–1755 | Unicom 2G signal protocol |
| 1940–1955 | Unicom 3G signal protocol |
| 2300–2320<br>2555–2575<br>1755–1765 | Unicom 4G signal protocol |
| 825–840 | Telecom 2G signal protocol |
| 1920–1935 | Telecom 3G signal protocol |

$$\frac{1}{r} = \frac{\left(\mathrm{PT}_1/\mathrm{PT}_2 - 1\right)}{x_{12}}, \tag{2}$$

where $r$ is the distance from the signal transmission point to the receiving point 2, $\mathrm{PT}_1$ is the power received at the receiving point 1, $\mathrm{PT}_2$ is the power received at the receiving point 2, and $x_{12}$ is the distance between the two receiving points. The ratio between the distance and the signal strength can be obtained from the distance between the two points, as shown in Figure 4. The signal strength of the sample point is inversely proportional to the distance. The greater the signal strength, the closer the distance.

### 3.3. *Electromagnetic Trojan Signal Feature Extraction.*

Based on the working mode of electromagnetic Trojan, combined with the experience of security schemes in APT traffic detection [10], this paper quantifies the electromagnetic waves in terms of time angle and information transmission. During the characterization of electromagnetic waves, our classification is based on the center frequency of the channel occupied by the signal, and then the time-window division method is used to perform dicing extraction on the signal. This feature extraction method can greatly compress our data volume while keeping the behavior of the signal as much as possible. Because of the limitation of electromagnetic leakage in one-directional propagation, electromagnetic Trojans usually use channels that are not commonly used or used at a lower frequency in the communication process. Therefore, it is a good choice to distinguish the center frequency of the signal.

*3.3.1. Time Feature.* Electromagnetic Trojans work by using electromagnetic radiation for information leakage, but software-defined radio technology does not allow the use of electromagnetic waves from the circuit for the reception. Therefore, the working mode of the electromagnetic Trojan can only be a one-way transmission and a connectionless working mode. Through the study of the electromagnetic Trojan, we found that, to work properly in this mode, the electromagnetic Trojan will take a certain heartbeat mechanism to determine whether the Trojan host is active. And before sending important information, it will send a signal to synchronize. Therefore, the electromagnetic Trojan signal has a strong feature in time. So, in time, we extracted the features as shown in Table 3.

Interval features include minimum interval, maximum interval, average interval, and interval variance. The electromagnetic signal sent by the electromagnetic Trojan attacker is automatically sent with the programming. The transmission of most electromagnetic Trojan signals has a

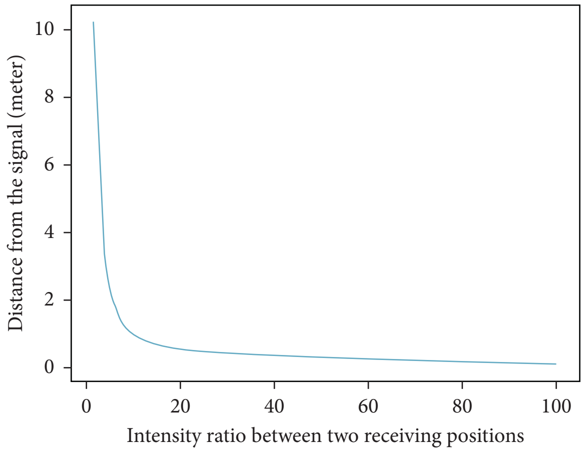


Figure 4: Relationship between two receiving position intensity ratios and signal source distance.

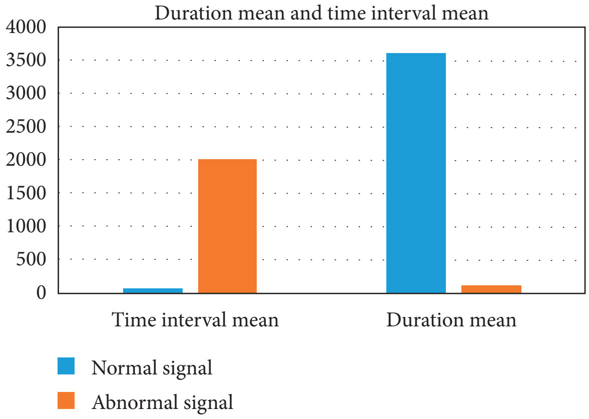


Figure 5: The difference between the normal signal and the abnormal signal in time.

Table 3: Time-based electromagnetic wave features.

| Features | Description |
|---|---|
| Minimum interval | Minimum time for signal interval |
| Maximum interval | Maximum time of signal interval |
| Average interval | Average time of signal interval |
| Interval variance | Variance of signal interval |
| Maximum duration | Minimum time for uninterrupted signal |
| Maximum duration | Maximum time for uninterrupted signal |
| Duration variance | Variance of the uninterrupted signal |
| Average duration | Average time of the uninterrupted signal |

time interval. The average interval is mostly within a few seconds. We analyze nearly 2000 signals of 6 kinds of electromagnetic Trojans. It was found that less than 10% of the signals had an average transmission time of more than 10 seconds. In the normal signal, this average interval is shorter than that of the Trojan signal, and the transmission time interval of most normal signals is relatively small. For example, the WIFI signal appears as a normal distribution in the time interval, and the minimum interval and the maximum interval are not large, so the signal variance values are small; the Mobile 2G broadcast signal has a heartbeat feature in time, showing a certain regularity, so the signal interval variance used by the channel is small; while the Trojan signal has obvious periodicity and lacks synchronization mechanism, it shows strong regularity in time interval, but the use rate of the electromagnetic Trojan is lower. Therefore, the variance value is generally larger. Therefore, the features of the time interval can be used to distinguish the Trojan signal from the normal signal, as shown in Figure 5.

*3.3.2. Information Volume Features.* We extracted the information volume features of the signal as shown in Table 4. Compared with the features of the normal signal on the channel, the Trojan signal also has obvious differences. Since the channel for normal signal transmission is well-divided, there are multiple devices sharing the channel, so the volume of information on the channel is generally stable and the duty ratio of the channel is high.

We have found that the electromagnetic Trojan cannot accurately control the power during the process of transmitting signals. Compared with the normal signal transmitting equipment, there will be large fluctuations in signal power, occupied channel bandwidth and strength. We block the time window and channel of the signal and treat each block as a packet in the network transmission. All the packets in the window are studied as one session. Each block extracts power data separately and finds the maximum power block, the minimum power block, and the variance between all power blocks in one session. At the same time, the background power in the current environment and the parameter conditions is extracted.

In the working process of the electromagnetic Trojan, to be able to transmit the secret data, the electromagnetic Trojan generally selects the channel with more idle time; otherwise, the normal device communication will be affected during the transmission process or the interference may cause the stealing signal to be unsuccessful. We have to send it out. When a normal device communicates, as long as it meets the communication specifications of the current frequency, it will have its blocking error correction system, so we propose the channel duty cycle (the ratio of the total duration of the signal to the total time of the block). There is a clear difference between the electromagnetic Trojan signal and the normal signal at the channel duty cycle.

The bandwidth refers to the frequency range occupied by the signal. In this paper, the bandwidth decision scheme is the first zero-point method [9]. When the signal strength drops to 0 for the first time, the frequency range is the bandwidth. According to Shannon's theorem, it can be known that, in the noise-and signal-independent Gaussian white noise channels, assuming the signal power S and the noise power N are constant, the relationship between the channel capacity $C$ (b/s) and the channel bandwidth W (Hz) [11] are as shown in equation (3). It can be seen that the channel capacity is proportional to the bandwidth. Therefore, by taking the value of the bandwidth used in the actual transmission of the signal, it can represent the amount of information transmitted by the current signal:

Table 4: Information volume feature.

| Features | Description |
|---|---|
| The number of blocks | The number of blocks that the signal is divided into at a time |
| Average total power per block | Average power per block |
| Single block maximum power | Maximum power per signal |
| Single block minimum power | Minimum power per signal |
| Block power variance | The amount of noise interference in the environment where the signal is located |
| Minimum duty cycle | The minimum amount of data for a single transmission of a signal |
| Average bandwidth | The bandwidth of each frame of the current signal divided by the number of frames |
| Bandwidth variance | Bandwidth change rate, measuring the stability of data transmission |
| Maximum bandwidth | The maximum bandwidth of the signal to be sent, used to calculate the channel width |
| Minimum bandwidth | The minimum amount of data sent by a single signal |

$$C = W \log_2\left(1 + \frac{S}{N}\right). \quad (3)$$

*3.3.3. Energy Features.* The collected signals are stored in the form of complex signals, so we define the energy according to the definition of the complex signal. The energy is defined as shown in the following equation:

$$E = \sqrt{I^2 + Q^2}. \quad (4)$$

There is a close relationship between the magnitude of the energy $E$ and the strength of the transmitted signal and the amount of information transmitted.

As shown in Figure 6, the collected complex signals are plotted in a complex coordinate system. Figure 6(a) shows the step-by-step situation of the sampling point when the channel is idle. Figure 6(b) shows the distribution of sampling points when the distance is constant and there is a small amount of data for transmission. Figure 6(c) shows the same distance. The distribution of sample points sends large amounts of data.

According to the definition of energy in equation (4), the visual display of energy in the graph is the distance from the sampling point to the off-center. As can be seen from Figure 6, as the amount of data transmitted by the channel increases, the energy distribution on the channel also appears. The obvious difference is that energy can be used to distinguish the amount of data transmitted on the current channel.

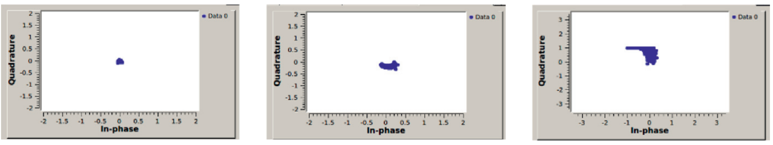
So, we can find the relationship between energy and the amount of transmission at a certain distance. Since the normal signal and the electromagnetic Trojan signal have significant differences in the usage rule and the amount of transmitted information, the stability of signal transmission is poor. Therefore, from the perspective of energy, we propose the features of maximum energy, minimum energy, energy mean, maximum energy change rate, and minimum energy change rate.

*3.3.4. Harmonic Features.* Harmonic signal features are the most prominent signal features of electromagnetic Trojans. Harmonic is a signal component that appears to be an integer multiple of the fundamental frequency when the signal is converted from the time domain to the frequency domain by the Fourier transform. The frequency is several times that of the fundamental wave. The generation of harmonics affects the normal transmission of the signal, which is a noise during data transmission and affects the normal transmission of signals on the surrounding channels. Therefore, the suppression of harmonics is an extremely important part in the signal transmission structure. Since the spectrum shifts before the signal is transmitted, the signal is converted from the low frequency to the high frequency, so the multiple of the frequency may change, but the frequency interval between the harmonics cannot be changed.

The electromagnetic Trojan's transmitting equipment is generally used for displays and wires. Compared with the dedicated signal transmitting equipment, in addition to the relevant features discussed above, there is also a problem with the depth of the harmonic signal. As shown in Figure 7(a), the abscissa is the frequency of the signal, the vertical axis is time, and the depth of the color represents the signal strength here (where the deeper the color, the stronger the intensity, and the stronger the signal is). Figure 7(b) shows the normal signal. It can be seen that the signaled part is a black line and no harmonics are found. To prevent harmonics from causing noise effects, harmonic suppression has been used before the signal is transmitted to minimize the influence of harmonics, so this is a big difference between the electromagnetic Trojan and the normal signal.

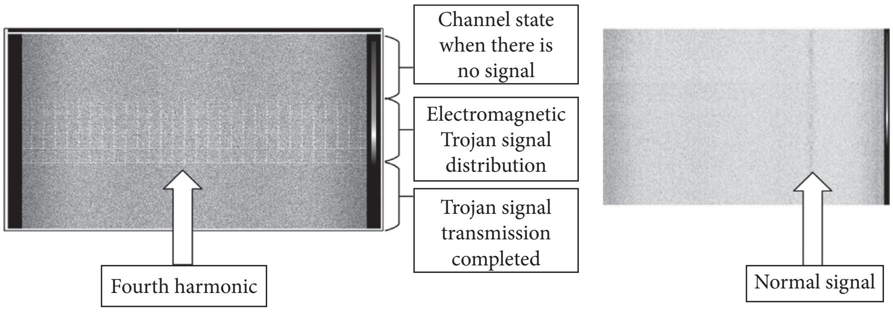
## 4. Electromagnetic Trojan Detection and Signal Classification

*4.1. Experimental Environment.* The acquisition equipment of this experiment is TV stick RTL2832Ux3, HACKrfx1. The acquisition range is TV bar RTL2832U (10 Hz to 1800 MHz, maximum sampling rate 3.2 MHz) and HACKrf (10 Hz to 6 GHz, maximum sampling rate 128 MHz). Rolling monitoring of electromagnetic signals from 10 Hz to 6 GHz [11] is possible by HACKrf one + sdr (software-defined radio). The data compression sensing algorithm and serialized memory are used to compress the data, and the compression effect is greater than 0.5. The hardware environment configuration of the experiment is shown in Table 5.



(a) (b) (c)

Figure 6: Relationship between energy size and data transfer volume. (a) Energy map when the channel is idle; energy maps for (b) a small amount of data transmission and (c) a large amount of data transmission.


Channel state when there is no signal
Electromagnetic Trojan signal distribution
Trojan signal transmission completed
Fourth harmonic
Normal signal

(a) (b)

Figure 7: Signal waterfall diagram. (a) Electromagnetic Trojan harmonic signal waterfall map; (b) normal signal waterfall map.

*4.2. Experimental Data Set.* The GSMem system, which was released in 2015 by Mordechai Guri [12] of Ben Gurion University, was used in this experiment. The system can signal the electromagnetic leakage of the computer CPU. In the transmission process, it only needs 4 KB of CPU cache space, no root or administrator privileges, no need to call the system API, and intel or AMD CPU signal transmission can be made on Windows or Linux system, and for the signal acceptance, only a mobile phone supporting the GSM communication protocol is required. Figure 8 shows a screenshot of a CPU Trojan provided by Mordechai Guri. By running the CPU Trojan on the target computer, you can receive the information to be compromised on the phone next to it.

The system-bus-radio system provided by William Entriken was also used in this experiment. The system is written in C language, and an electromagnetic Trojan is installed on the display to transmit the data to be transmitted in the form of electromagnetic waves.

*4.3. Data Collection.* The experimental data are divided into two parts. One part is the normal electromagnetic signal data, the signal comes from the fixed collection point in the laboratory, the frequency is from 10 MHz to 3000 MHz, and the acquisition time is one week, marked as normal signals. Another part of the Trojan data is generated using the public system. The sampling method, acquisition time, and acquisition point are consistent with the normal data and marked as abnormal signals. In the data acquisition process, we use signal filtering and signal extraction to characterize the data, which greatly reduces the amount of data processing. The data set is shown in Table 6.

Table 5: Hardware environment configuration.

| Project | Configuration |
|---|---|
| Server | Dell PowerEdge R730XD |
| Processor | 2 Intel® Xeon® E5 – 2630 v3,<br>8 cores 16 threads |
| RAM | 8 * 16 GB RDIMMs |
| SDR | HACKrf one, RTL2832U *R*820 T |
| Hard disk | 2 TB * 10 |
| Collector | HACKrf one * 1, RTL2832U * 3 |

*4.4. Experimental Indicators.* In this experiment, the following indicators were selected for evaluation: FN, FP, TN, TP, FPR, FNR, TPR, TNR, precision, recall, and F1-score. False negative (FN) refers to the number of samples that are determined to be electromagnetic Trojan signals but are normal signals. False positive (FP) is judged to be a normal signal but is the number of samples of the electromagnetic Trojan signal. True negative (TN) is that the electromagnetic Trojan signal is correctly determined as the number of samples of the electromagnetic Trojan signal. True positive (TP) is that the normal signal is correctly determined as the number of samples of the normal signal.

FP rate (FPR) is the ratio of the number of electromagnetic Trojan signal samples predicted to be normal signals to the actual number of negative samples, the ratio of FP to FP + TN. FN rate (FNR), which is predicted as the ratio of the number of normal signals of the electromagnetic Trojan signal to the actual number of normal signals, that is, the ratio of FN to FN + TP. TP rate (TPR) is the ratio of the number of normal signals correctly predicted to the actual number of normal samples, the ratio of TP to FN + TP. TN rate (TNR) is correctly predicted as the ratio of the number of electromagnetic Trojan signals to the actual number of

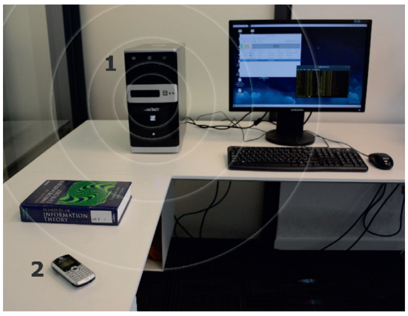


Figure 8: Display electromagnetic trojan.

Table 6: Experimental data set.

| Data set | Feature size (GB) |
|---|---|
| Normal signal | 531 |
| CPU signal | 52 |
| Display signal | 34 |

electromagnetic Trojans, the ratio of TN to FP + TN. The precision accuracy value is the ratio of the number of correctly classified samples to the total number of samples. The recall is the ratio of the number of samples correctly classified by the classifier to the actual number of positive samples. F1-scoreis the weighted harmonic average of accuracy and recall that is used to measure the effectiveness of the test method. The specific formula is as follows:

$$\begin{aligned} A\,(\text{accuracy}) &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},\\ P\,(\text{precision}) &= \frac{\text{TP}}{\text{TP} + \text{FP}},\\ R\,(\text{recall}) &= \frac{\text{TP}}{\text{TP} + \text{FN}},\\ F1\ \text{score} &= \frac{2 * P * R}{P + R}. \end{aligned} \tag{5}$$

## 5. Anomaly Detection and Comparative Experiment

We used two evaluation methods to evaluate the three deep learning algorithms of LSTM, RNN, and CNN, which are the cross method and percentage segmentation method. LSTM is the classification algorithm used in this article, and RNN is the method used in the latest hardware Trojan detection patent in 2018. The electromagnetic waves are characterized according to the features extracted in the patent, and then the electromagnetic signals are classified using the RNN algorithm used in the patent.

In the experiment, the collected and filtered signals contain a large number of noise signals and normal signals in addition to the Trojan signal. Therefore, we divide the classification model into three parts: the input layer, the LSTM layer, and the classifier. The input layer is responsible for converting the characterized vector into a feature sequence for input to the LSTM layer for analysis. The LSTM layer receives the serialized signal data input from the input layer, selectively memorizes and learns the current input data, and then retains the important information and generates a 128-dimensional feature vector for characterizing the current model state. The classifier layer accepts the 128-dimensional vector of the LSTM layer for electromagnetic Trojan signal recognition.

The cross method uses randomization of data into 10 equal parts, 9 of which are used as training sets, and the remaining one is used as a test set for accuracy and accuracy. The detection process needs to be repeated 10 times. The advantage of this is that the data are completely randomized and scrambled, resulting in higher accuracy and lower false positives, eliminating the problem of false negatives caused by duplicate data.

In this experiment, the three deep learning algorithms of LSTM, CNN, and RNN are verified by the cross method. The experimental results are shown in Figure 9. It can be seen from the figure that the accuracy rate of LSTM is 96.41%, the accuracy rate is 98.65%, the performance of RNN is relatively poor, the accuracy rate is 93.51%, and the accuracy rate is 95.56%; the performance of CNN algorithm is even worse, the accuracy rate only 85.62%, and the accuracy rate is only 91.75%. The accuracy in this experiment is higher than that of other algorithms. The features extracted by this experiment are better for the recognition of electromagnetic Trojan signals. The false-negative rate of the system is lower, but the false alarm rate is slightly higher. The reason may be that the features extracted in this paper are not characterized by time series, and the regularity of time series is the main difference between electromagnetic Trojan signal and normal signal, so the performance of circulating neural network is better. For feature classification with a long duration, LSTM performs much better than RNN, so LSTM has higher accuracy for electromagnetic Trojan recognition.

The percentage division method divides the data into two parts, one for 75% and the other for 25%, with 75% of the data for training and the rest for testing, as shown in Figure 10. The advantage is that the algorithm detection operation is not complicated, saving time and high efficiency. However, its accuracy and false-negative rate are slightly inferior to those of the cross, which is suitable for classification with a small amount of data and obvious features. In this experiment, because the electromagnetic signal data is very large and has many features, the overall test results are not satisfactory, but it can be seen that LSTM has higher accuracy than CNN and RNN and CNN has the worst performance; the reason may be, in the detection process, the detection system mainly focuses on the working mode of the electromagnetic Trojan signal. The discrimination between the electromagnetic Trojan signal and the normal signal is mainly in the period, so the deep neural network has a better performance. However, the accuracy and accuracy are consistent with those in the cross-certification, indicating that the system's false-negative rate is still low and the false positive rate needs to be improved.

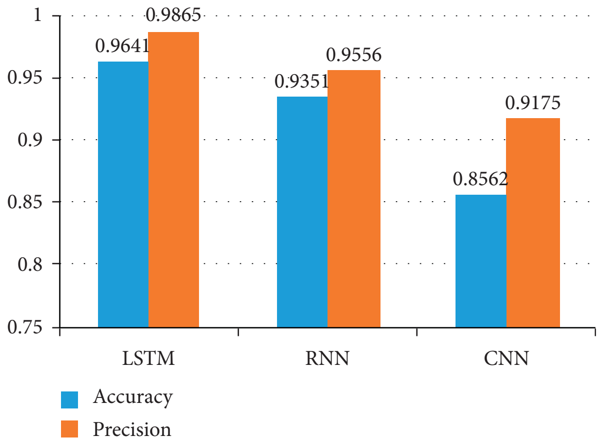


Figure 9: 10-fold cross-validation results.

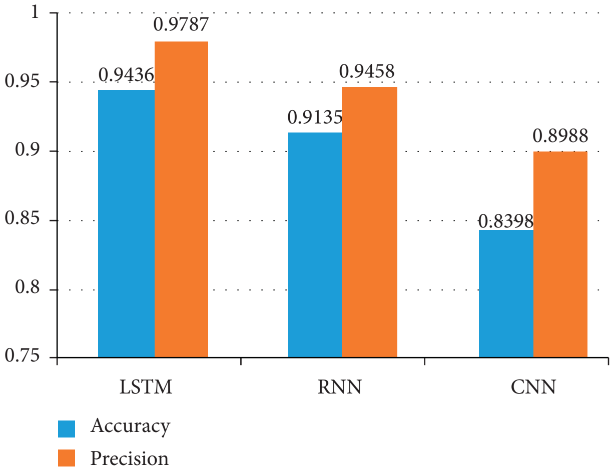


Figure 10: Percentage division method results.

Of course, evaluation of the merits and demerits of an algorithm is highly accurate. Therefore, we also need to use other indicators to judge the performance of the algorithm. This paper uses the recall rate and F1-measure to evaluate the pros and cons of the detection model. The recall rate and F1-Measure of the three algorithms are shown in Figure 11.

It can be seen from the figure that the detection result of the LSTM algorithm is due to the RNN algorithm. The reason may be that the forgetting gate is added to the LSTM algorithm, which can selectively memorize and forget the data, and the data with a longer period will have better performance, so it has a good classification effect on electromagnetic Trojan signals with long duration.

## 6. Detection Accuracy Factor

Through many experiments on a large amount of data, we found that the setting of the gain parameters and the sampling point distance has a great influence on the accuracy of the electromagnetic Trojan detection system.

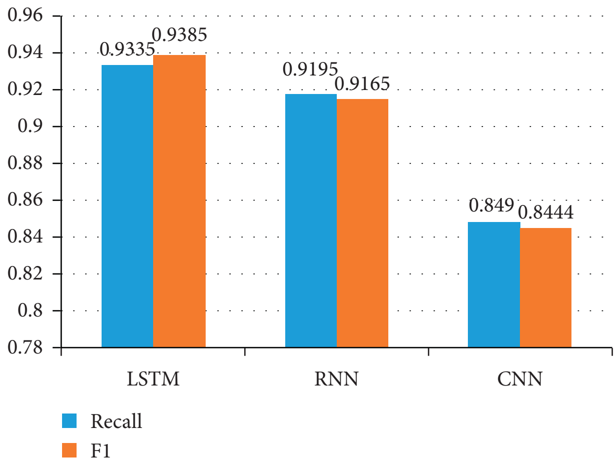


Figure 11: Recall rate and F1 for different algorithms.

### 6.1. Effect of Gain.

*6.1. Effect of Gain.* The gain is the amplification ratio of the acquired signal in dB. The gain calculation formula is shown in the following equation:

$$Z = 10\,\mathrm{lg}\left(\frac{P_1}{P_2}\right). \tag{6}$$

where $P_1$ is the amplified power, $P_2$ is the power before amplification, nd then $P_1$ is the power after $P_2$ is amplified by $Z$ dB. Usually 10 dB represents 10 times magnification, 100 dB represents 20 times magnification, and −10 dB represents 10 times attenuation. Since the gain is multiplied for all signals (whether it is a noise signal or a signal in normal use), all input samples are amplified by a specified multiple, so the noise is amplified and the gain is amplified. The setting of the value will have an impact on the accuracy of the system.

As shown in Figure 12, the gain is set to 50 dB, and the sampled signal is a waterfall graph when the sampling point is 1 m away from the electromagnetic Trojan signal source. The data in the figure are divided into three parts. According to the gradient color in the legend, it can be seen that the deeper the color, the stronger the signal strength, and the lighter the color, the weaker the signal strength. The left part of the figure consists of several random points, which are noise signals (background signals). The cause is that the signal sampling uses band-pass sampling. A band-pass filter is used to filter the signal outside the sampling bandwidth. The signal strength outside the sampling bandwidth after filtering is greatly reduced, resulting in low-power noise. Two obvious signals can be seen in the figure. For the convenience of description, the line on the left side is set to signal A and the line on the right side is set to signal B.

In Figure 12, signal A is the heartbeat broadcast part of the electromagnetic Trojan signal and signal B is an FM broadcast signal at the experimental location. The broadcast signal is selected to compare the intensity change of the Trojan signal and the background noise intensity variation in the control variable. Figure 13 shows the waterfall graph with the sampling gain set to 10 DB and the same set of signals sampled when the sampling point is 1 m away from the

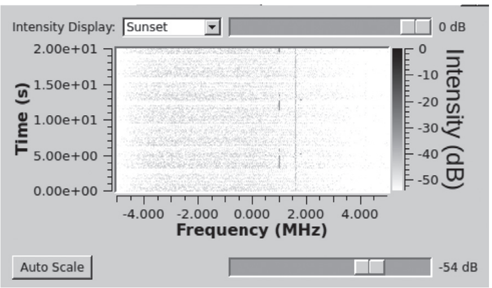

Intensity Display: Sunset

Figure 12: Waterfall diagram with the gain set to 50 and distance of 1 meter.

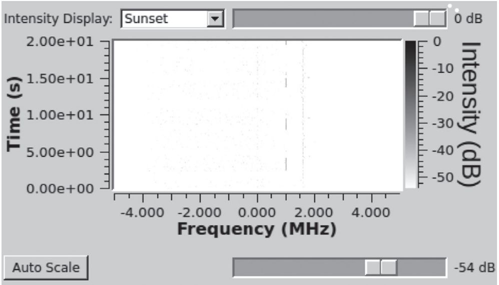

Intensity Display: Sunset

Figure 13: Waterfall diagram with the gain set to 10 and distance of 1 meter.

electromagnetic Trojan source. Comparing Figures 12 and 13, it can be seen that, when the gain is appropriately reduced, the background signal has basically disappeared in the figure, and only the signals A and B are left in the figure. Meanwhile, as compared with Figure 12 with a larger gain, it can be seen that both the signal A and the signal B are greatly weakened. This article sets up a series of experiments for different gains to study the relationship between gain setting and detection accuracy. The test results after the experiment are shown in Figure 14.

It can be seen from Figure 14 that, under the condition of fixed distance, the system detection accuracy increases first and then decreases with the increase of gain, and selecting the appropriate gain value can improve the accuracy of system detection. The reason for this is that the electromagnetic Trojan signal is not transmitted by a professional signal transmitting device. A professional signal transmitting device uses an amplifier to amplify the signal when transmitting the signal, and the electromagnetic Trojan does not have an amplifier when transmitting the signal. Existence: the strength of the signal is slightly higher than the noise signal strength. When the antenna gain is too large, the signal filtering module will filter a part of the electromagnetic Trojan signal as a noise signal, thereby reducing the detection accuracy of the system; meanwhile, when the gain is too small. Electromagnetic Trojan signals and noise signals cannot be distinguished, and data loss can also occur. The difference in gain affects the extraction of the overall features of the signal, but it affects the accuracy of deep neural network modeling and differentiation.

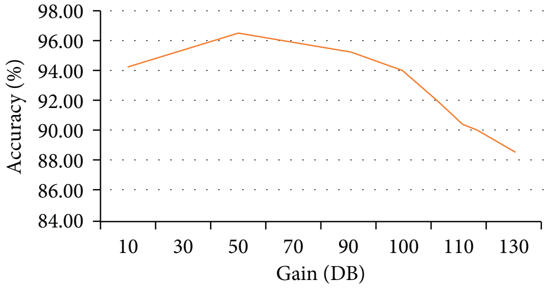


Figure 14: Detection accuracy and gain relationship.

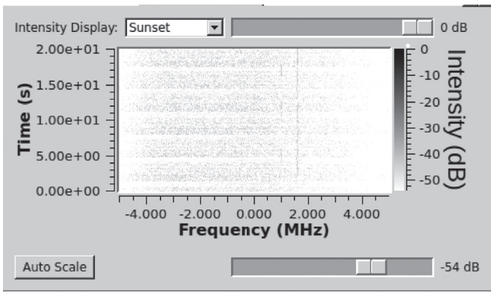

Intensity Display: Sunset

Figure 15: Waterfall diagram with the gain set to 10 and distance of 1 meter.

*6.2. Influence of Sampling Distance.* The detection accuracy of the anomaly detection system is also different for the data sampled by different sampling points. Figure 15 shows the signal waterfall diagram when the gain is fixed at 50 dB and the sampling point is 3 meters away from the electromagnetic Trojan signal source. In Figure 12, with the same gain, the sampling point is 1 meter from the Trojan signal source. The intensity of the noise signal and signal B does not change significantly with distance, but the intensity of the electromagnetic Trojan signal is significantly weakened. When the distance increases, the electromagnetic Trojan signal will not be displayed in the waterfall chart. In Figure 4, the electromagnetic wave intensity varies with the propagation distance curve. The electromagnetic wave intensity is inversely proportional to the propagation distance, and the intensity loss is the fastest near the electromagnetic signal emission source. Our experiments used HACKrf one to test the distance and detection accuracy. Set the acceptance gain to 50 dB and test the distance between the acquisition point and the signal source to get the curve, as shown in Figure 15. When the distance is close, the accuracy of electromagnetic Trojan detection is higher. As the distance increases, the accuracy of electromagnetic Trojan detection is lower and lower. When the distance of the electromagnetic Trojan from the sampling point exceeds 4 m, the detection accuracy is greatly reduced, as shown in Figure 16. A study of the propagation distance of CPU leaking Trojans in a paper by Mordechai Guri et al. is as shown in Figure 17. When the electromagnetic Trojan spreads over a certain distance, the attenuation of the Trojan horse will be confused with the signals of other devices around it, failing to properly capture the electromagnetic Trojan signal.

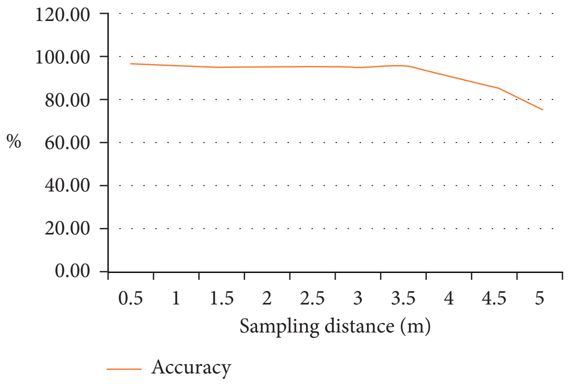


FIGURE 16: Detection accuracy and distance relationship.

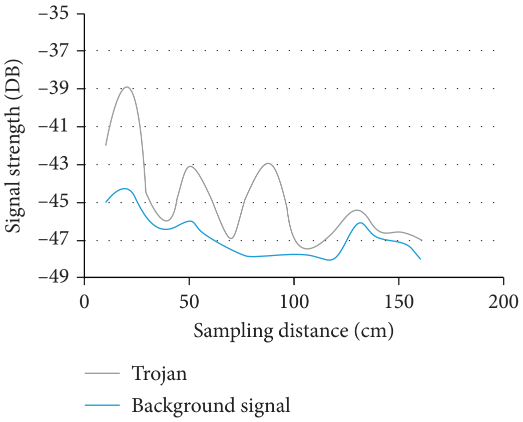


FIGURE 17: Signal strength and sampling distance relationship.

Comprehensive analysis: the gain and sampling point distance have an impact on the detection accuracy of the system. In the process of using HACKrf one, the gain setting is about 50 dB; the closer the electromagnetic Trojan signal source is to the sampling point, the better it is. The sampling density of the sampling points should be improved to improve the detection accuracy of the system.

## 7. Conclusion

In this paper, we characterize electromagnetic signals from the perspective of time, energy, and information and use deep learning to detect and analyze electromagnetic Trojans. Then, the factors affecting the accuracy of the system detection during the Trojan work are quantitatively analyzed, and the gain setting suitable for using HACKrf one and the optimal acquisition radius of the Trojan signal is found. At the same time, this paper also tests the electromagnetic Trojan in the big data environment and compresses the three methods of deep learning. It can be seen that the LSTM algorithm has advantages in detecting electromagnetic Trojan signals.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 62076042 and 61572086), the Key Research and Development Project of Sichuan Province (Nos. 2020YFG0307 and 2018TJPT0012), the Key Research and Development Project of Chengdu (No. 2019-YF05-02028-GX), the Innovation Team of Quantum Security Communication of Sichuan Province (No. 17TD0009), and the Academic and Technical Leaders Training Funding Support Projects of Sichuan Province (No. 2016120080102643).

## References

[1] H. Tanaka, “Information leakage via electromagnetic emanation and effectiveness of averaging technique,” in *Proceedings of the2008 International Conference on Information Security and Assurance (isa 2008)*, pp. 98–101, IEEE, Busan, Korea, April 2008.

[2] M. G. Kuhn and R. J. Anderson, “Soft tempest: hidden data transmission using electromagnetic emanations,” *Information Hiding*, Springer, in *Proceedings of the International Workshop on Information Hiding*, pp. 124–142, April 1998.

[3] J. Balasch, B. Gierlichs, and I. Verbauwhede, “Electromagnetic circuit fingerprints for hardware trojan detection[C].” in *Proceedings of the 2015 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pp. 246–251, IEEE, Dresden, Germany, August 2015.

[4] M. Cozzi, J.-M. Galliere, and P. Maurine, “Thermal scans for detecting hardware Trojans,” *Constructive Side-Channel Analysis and Secure Design*, Springer, in *Proceedings of the International Workshop on Constructive Side-Channel Analysis and Secure Design*, pp. 117–132, April 2018.

[5] W. C. Lu, X. J. Liu, and G. Y. Fang, “Ground Penetrating Radar Data Compression Acquisition Based on Perceptual Compression,” *Journal of Forestry Research*, vol. 3, pp. 1433–1452, 2011.

[6] Q. Xu, X. H Jiang, L. H. Yao et al., “Summary of hardware trojan detection and prevention research,” *Journal of Network and Information Security*, vol. 3, no. 4, pp. 1–13, 2017.

[7] Y. Y. Xu, Q. W. Huang, W. Fan et al., “Modeling and experimental analysis of electromagnetic information leakage based on power line,” *Science China Physics, Mechanics & Astronomy*, vol. 57, no. 57, pp. 22–66, 2014.

[8] J. Z. Lu, W. N. Niu, X. L. Liu et al., “A locable abnormal electromagnetic signal joint detection algorithm,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 33, no. 13, Article ID 1958009, 2019.

[9] Xu Kejun, *Signal Analysis and Processing [M]*, Tsinghua University Press Ltd., Beijing, China, 2006.

[10] J. Z. Lu, K. Chen, Z. L. Zhuo et al., “A temporal correlation and traffic analysis approach for APT attacks detection,” *Cluster Computing*, vol. 201712 pages, 2017.

[11] A. Beitler, A. Caracas, T. Eirich et al., p. 219, 2018 Synchronization in software-defined radio systems: U.S. Patent Application 10/067.

[12] M. Guri, A. Kachlon, O. Hasson et al., “GSMem: data exfiltration from air-gapped computers over {GSM} frequencies,” in *Proceedings of the 24th {USENIX} Security Symposium ({USENIX} Security 15)*, pp. 849–864, Washington, D.C., USA, August 2015.