Video Surveillance Camera Identity Recognition Method Fused With Multi-Dimensional Static and Dynamic Identification Features

Zhijie Fan, Fudan University, China

Zhiwei Cao, The Third Research Institute of Ministry of Public Security, China*

Xin Li, People's Public Security University of China, China

Chunmei Wang, Fudan University, China

Bo Jin, The Third Research Institute of Ministry of Public Security, China

Qianjin Tang, The Third Research Institute of Ministry of Public Security, China

ABSTRACT

With the development of smart cities, video surveillance networks have become an important infrastructure for urban governance. However, by replacing or tampering with surveillance cameras, an important front-end device, attackers are able to access the internal network. In order to identify illegal or suspicious camera identities in advance, a camera identity identification method that incorporates multidimensional identification features is proposed. By extracting the static information of cameras and dynamic traffic information, a camera identity system that incorporates explicit, implicit, and dynamic identifiers is constructed. The experimental results show that the explicit identifiers have the highest contribution, but they are easy to forge; the dynamic identifiers rank second, but the traffic preprocessing is complex; the static identifiers rank last but are indispensable. Experiments on 40 cameras verified the effectiveness and feasibility of the proposed identifier system for camera identification, and the accuracy of identification reached 92.5%.

KEYWORDS

Identification Feature, Identity Recognition, Intrusion Detection, Self-Information Amount, Static and Dynamic Identification Features, Video Surveillance Camera

INTRODUCTION

In recent years, with the rapid development of public security big data applications, video surveillance network has become a powerful tool for security management and crime investigation (Guo et al., 2017; Kim & Park, 2021; Nazaré & Schwartz, 2016). The video surveillance network established

DOI: 10.4018/IJISP.319304

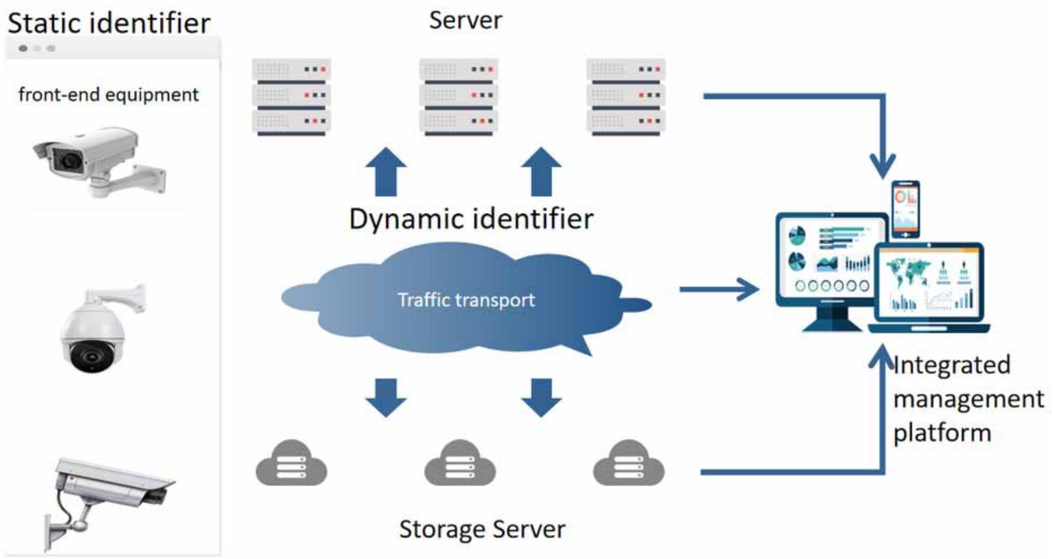
*Corresponding Author

by the public security department is characterized by large scale of video front-end equipment, wide geographical coverage, diverse system application software, rich bearing business functions and strict monitoring requirements, which helps to improve the efficiency of public security and assist in case investigation (Ma et al., 2018; Thiyagarajan et al., 2018). However, video surveillance systems are exposed to complex network security risks, such distributed networks are easy targets for criminals, and in the event of a major network security incident will result in the unavailability of video surveillance resources, which in turn may threaten the security of the state, society, enterprises and individuals (Ma et al., 2018; Thiyagarajan et al., 2018). Industry and academia have invested heavily in the development of prevention tools and methods to reduce potential losses by moving the prevention barrier forward through monitoring and early warning.

As an important front-end device of video surveillance network, video surveillance cameras are characterized by a large number and variety. If a video surveillance camera is illegally replaced, it can easily develop into a security risk for the whole video network. However, as part of the video dedicated network cameras are leased, purchased, installed and replaced by operators, management departments are often unable to clarify all front-end devices, while the lack of overall security management mechanism for front-end devices, without an efficient front-end device identity authenticity detection method. Therefore, even if video surveillance cameras are illegally replaced, it is often difficult for security management to detect it.

Therefore, in this paper, we propose a video surveillance camera identity identification method that incorporates multi-dimensional static and dynamic identification features. As shown in figure 1, firstly, static identifiers of video surveillance cameras are collected and extracted, including explicit and invisible identifiers. Then a series of dynamic traffic feature sets are extracted starting from the video surveillance camera network traffic perspective. The multidimensional identifiers are fused and the amount of self-information for each identifier is analyzed. Finally, the video surveillance camera identification index is determined according to the contribution amount of the identifier in the identification system. The experimental results prove that the method in this paper can accurately and efficiently perform video surveillance camera identification, which has important theoretical and practical significance.

Figure 1.
Video surveillance network workflow



RELATED WORKS

Summary

Video surveillance cameras belong to a special IoT device used for image acquisition. IoT device identification and marking technology refers to the identification of device information (type, brand, series, model, etc.) of the target device by means of network sniffing, data mining, etc., and the access in the subsequent process to be authenticated. In the IoT device side access, the security policy based on the physical characteristics of the device avoids the complex and inefficient computation of traditional security policies, and physical information features are generally difficult to forge, and numerous researchers study defense strategies for the perception layer (Islam et al., 2017; Xie et al., 2017). The low computing power of device hardware in IoT systems, energy consumption limits complex sensing networks bring unique security threats and defense requirements for IoT end devices, such as key verification algorithms and cross-domain authentication under low computation (Yang et al., 2021).

General IoT Devices

For the research of identification methods of IoT devices, the current research hotspots mainly focus on device fingerprint extraction, including device hardware fingerprint, device installation software and traffic fingerprint (Vastel et al., 2018; Thangavelu et al., 2018), protocol fingerprint (Chen et al., 2021), etc. Biometric fingerprint has stability and uniqueness which can uniquely identify objects (Lai et al., 2019). Analogous to the role of biometric fingerprint for human identification, in the IoT, the information with stability and independence can be obtained by extracting the hardware characteristics of devices, communication characteristics between devices, IoT application characteristics, etc. to generate fingerprints for application in IoT device authentication. Device fingerprint is a device inherent property and information feature that can be used for IoT device identity verification, a tool to determine whether it is legal to access the cloud and upload cloud data reliability (Radhakrishnan et al., 2014). The IoT device fingerprinting technology is mainly classified into active device fingerprinting technology, passive device fingerprinting technology and mixed fingerprinting technology according to whether it can be obtained from the propagated data messages.

Active Device Fingerprinting Technology: Active device fingerprinting technology refers to the acquisition of IoT device characteristics through platform APIs (Koo et al., 2019) and other means to actively collect end-user device information or relevant privacy information, such as GPS address-related, IP-related, network type (Wi-Fi, 4G, 5G), cell phone usage, etc. The features collected through JS and flash can realize the identification of smart home IoT devices through software and hardware parameters or plug-in information features, which can distinguish more than 40 types of devices at 97% accuracy (Jose et al., 2017; Li et al., 2020). Analyze the mobile fingerprint of the user's behavior within the APP (Xue et al., 2018), the latest Web page fingerprint, by some website specific image files, js files, etc. By crawling these files and comparing whether the MD5 value is consistent with the rule base to determine the same content management system (CMS).

Passive Device Fingerprinting Technology: Passive device fingerprinting technology refers to extracting features from different data messages sent by devices, such as extracting features from collected data, sensor content noise, Wi-Fi signal variations (Liu et al., 2019), channel status of LoRa packets (Chen et al., 2021; Shen et al., 2021), traffic data features, etc., which can be used as features to identify different devices. Further, the communication fingerprint obtained from the data messages in communication, due to the heterogeneity of IoT, the diversity of IoT device protocols, Hamad et al. proposed a method to extract the connected device fingerprint method that can be used with a variety of connectivity technologies, including Wi-fi, Zigbee, Z-Wave, Ethernet. . Fingerprint extraction from the physical layer of the IoT, such as PUF (Mahmood et al., 2021), exploits the hardware characteristics of integrated circuits to extract fingerprints to perform device authentication by exploiting variations in transmission signals due to hardware and manufacturing inconsistencies.

Mixed Device Fingerprinting Technology: Mixed device fingerprinting technology refers to the fusion of active and passive device fingerprint technology in the same device identification and authentication system, which corresponds the device identifier generated by active device fingerprint technology on the client side with the feature information collected by passive device fingerprint technology on the server side and related to the protocol stack, so that all devices have unique device identification IDs. IoT application scenarios are becoming increasingly diverse, authentication with multiple factors and constraints will become a trend to address support multi-category device authentication, secure communication and data reliability.

Video Surveillance Cameras

In order to strengthen the construction of security protection of public security video surveillance network, a large number of scholars have also conducted in-depth research on video surveillance network security (Akbanov et al., 2019; Andresini et al., 2021; Kilincer et al., 2021). Li et al. constructed a network security posture indicator system for video private networks and gave a method to quantify the indicators. Duan et al. implemented the extraction and evaluation of cyber security posture elements using, for example, a hidden Markov model. The above study mainly focuses on the macro security posture of video private networks and does not focus on the identity security risk of front-end devices. At the same time the quantification of indicators is based on expert experience and is relatively rough. The video surveillance cameras fingerprinting technology is mainly classified into noise fingerprinting technology, image fingerprinting technology and fingerprinting technology for general IoT-like devices according to physical and applied features of the video surveillance camera.

Noise Fingerprinting Technology: A large of works already exist on sensor content noise, such as Das et al. used the frequency responses of speakers and microphones from two wireless IoT devices as acoustic hardware fingerprints, energy change features extracted from CAN frames captured by Telematics sensors as fingerprints, and fingerprints using camera frame noise (Khan et al., 2019). Camera Sensor Pattern Noise (SPN) in Photo-Response Non-Uniformity (PRNU) is a type of noise signal, which is the sensor pattern noise caused by the sensor defect of the imaging capture device that changes the nature of the digital image. PRNU is widely used in extracting camera hardware characteristics because of its intrinsic robustness, for example, for PRNU in pictures The similarity of two discrete signals is determined by operation, and a wavelet denoising algorithm is used to extract the specific fingerprint of the camera presentation sensor (Mandelli et al., 2020).

Image Fingerprinting Technology: Image fingerprinting technology refers to the PRNU features extracted from two images to do inter-correlation, if two images from the same camera, will see correlation images appear multiple spikes, find out the location of the most obvious peak in, calculate the peak correlation energy ratio height, determine its value is greater than a certain threshold, to decide belong to the camera for authentication. Bondi et al. proposed a random projection that approximately preserves the geometry of the point cloud composed of fingerprints, and effectively reduces the dimensionality of the space where the fingerprints are located by quantizing the PRNU and compressing the pipeline of image residuals to improve the application of video real-time verification. On top of this, Taspinar et al. optimized the efficiency of random projection to reduce the dimensionality of fingerprints by pipeline compression function.

Fingerprinting Technology for General IoT-like Devices: Like general IoT devices, hardware authentication technology for video surveillance cameras based on device fingerprint is a hot issue in cyberspace security. Device fingerprinting refers to the use of characteristic information of a device to generate a unique identifier. Based on the device identifier, users can perform device authentication and device anomaly detection. Currently, there are three main techniques for device fingerprint authentication: transient feature-based, modulated signal-based, and internal sensor-based. Lukas et al. identified video surveillance cameras by sensor pattern noise and validated the effectiveness of the method on 320 images from 9 different video surveillance cameras. Zhang et al. used a two-way authentication protocol to study RFID tag authentication technology in order to solve the authentication problem of IoT system, and the experimental results show that the technology can quickly confirm the identity information and ensure that the IoT is completely. However, the paper is a two-way authentication from the perspective of traditional cryptography and the method is difficult to use in the field of video surveillance camera identification.

However, all the fingerprinting technologies above (such as active device fingerprinting, passive device fingerprinting and mixed fingerprinting) are used in the filed of traditional sensor IoT devices, they are not suitable for video surveillance cameras that have different applications and features from the general IoT devices. For the noise fingerprinting technology and image fingerprinting technology, they all are intrusive identity identification methods that need to complete the operational deployment on the video surveillance cameras before they are activated, they are not suitable for the existing been used and deployed video surveillance cameras. Our method is a non-invasive identity identification method for video surveillance cameras according to their applications and features, it can not only solve the video surveillance cameras identity identification and recognition problem well, but also can reduce the overall video surveillance network construction cost, improve the efficiency of the security use of existing video surveillance cameras.

Compared with Android devices and tablets, video surveillance cameras have a single function, and based on device fingerprints, there is still less research on identity recognition for video surveillance cameras. Yin et al proposed the use of a decision tree approach to classify device fingerprints, followed by the implementation of network entry detection for dedicated IP video networks. Among them, device fingerprinting includes several dimensions, such as IP address, services running on the device, memory usage information, etc. But the study ignored the video surveillance camera traffic information.

The dynamic set of traffic features of each video surveillance camera can comprehensively reflect the traffic characteristics of the front-end video surveillance camera, which is the basis for the unique identification and subsequent accurate identification and network intrusion detection of video surveillance cameras in the front-end network. If someone were to replace a video surveillance camera with a malicious one, most of the explicit identifiers are easily forged and it would be difficult for the public safety department to detect an anomaly. However, the traffic characteristics of the video surveillance camera are difficult to tamper with and can therefore be used as an important indicator for identification (Jo & Kim, 2018). Regarding the study of traffic features, Wei-Chao Yang et al proposed an IoT device identification method based on traffic fingerprinting, which has better detection effect and stability. Meidan et al proposed to use machine learning methods to analyze network traffic of IoT devices and complete device identification with an overall classification accuracy of 99%. Currently, there is still a gap in the research on video surveillance camera identification by traffic features of video surveillance cameras, and in-depth research is needed.

In the video surveillance network, if the identifier of a front-end access device is not consistent with the local database, it can be considered as an abnormal device. Device anomaly access intrusion detection algorithms, on the other hand, are more likely to use algorithms such as association rules, support vector machines, random forests, and decision trees to identify anomalous data in a dataset.

A review of the literature revealed that:

1. There is less research on video surveillance camera identification, and there is an urgent need to carry out research on identification for video surveillance front-end devices.
2. Relevant departments still rely on explicit identifiers to determine the identity of the device, and the use of implicit identification features with video surveillance camera traffic, a dynamic feature information, is not considered in the identity.
3. No valid method is used to assess the extent to which different identifiers contribute to device identification.

In order to solve the above problems, a suitable method is needed to capture, analyze, converge, and fuse video surveillance camera features. To this end, this paper proposes a video surveillance camera identification method that incorporates multidimensional identification features to provide security management with an effective means of identifying illegal access to front-end devices.

The main contributions of this paper are as follows:

1. A video surveillance camera identification system containing both static and dynamic identifiers is designed. Multi-dimensional feature acquisition and extraction from video surveillance camera explicit and implicit features and traffic features.
2. A video surveillance camera multidimensional feature analysis method based on self-information quantity and information entropy is proposed. In this paper, we propose to identify video surveillance cameras based on the decision tree model in machine learning, and experimentally calculate the contribution ranking of each identifier.
3. Experiments are conducted on 41 terminal device feature datasets, and the results show that the accuracy of this paper's method in correct video surveillance camera identification reaches 93.1%, which can effectively detect video surveillance camera identity forgery and tampering behavior.

The structure of this paper is as follows, Section 3 details the video surveillance camera identity recognition method incorporating multidimensional identification features, Section 4 verifies the feasibility and superiority of this paper's method through experiments, and finally concludes the whole paper by pointing out the shortcomings and looking forward to future research directions.

VIDEO SURVEILLANCE CAMERA IDENTIFICATION METHOD

In this paper, we propose a video surveillance camera identity identification method that incorporates multidimensional static and dynamic identification features, and the main research method is shown in Figure 2. Firstly, collect the data of video surveillance cameras, then extract the static and dynamic features of each video surveillance camera. Secondly, construct feature vector, label and generate initial data set, and is divided into training set and test set. Thirdly, with the training set, machine learning models such as decision trees and artificial neural networks are trained. Finally, with the test set, the accuracy of video surveillance camera identification capability is verified.

We divide the original identifier system into a static identifier system and a dynamic identifier system. The identifiers included in each system are shown in Figure 3, and each identifier will be described in detail in subsequent sections.

Construction of identifier system

Static identifier system

The static identifier system includes both explicit and invisible identifiers, which are extracted and fused on the basis of the original identifier system to generate the final identifier system. Traditional front-end video surveillance cameras are uniquely identified using a method similar to explicit identifiers, which can be easily forged and tampered with. Implicit identifiers alone do not have the ability to uniquely identify, and the ability to identify is improved by combining multiple implicit identifiers. In this paper, we combine their advantages and disadvantages, and choose a combination of explicit and implicit identifiers to build and generate a static identifier system.

Explicit identifiers are symbols that clearly identify video surveillance cameras. The explicit identifiers used in this paper for video surveillance camera identification include IP addresses and MAC addresses, as shown in Table 1.

The five types of implicit identifiers to be extracted in this paper are shown in Table 2, which collect video surveillance camera implicit identifiers from device parameter information, channel

Figure 2.
Flow chart of our method

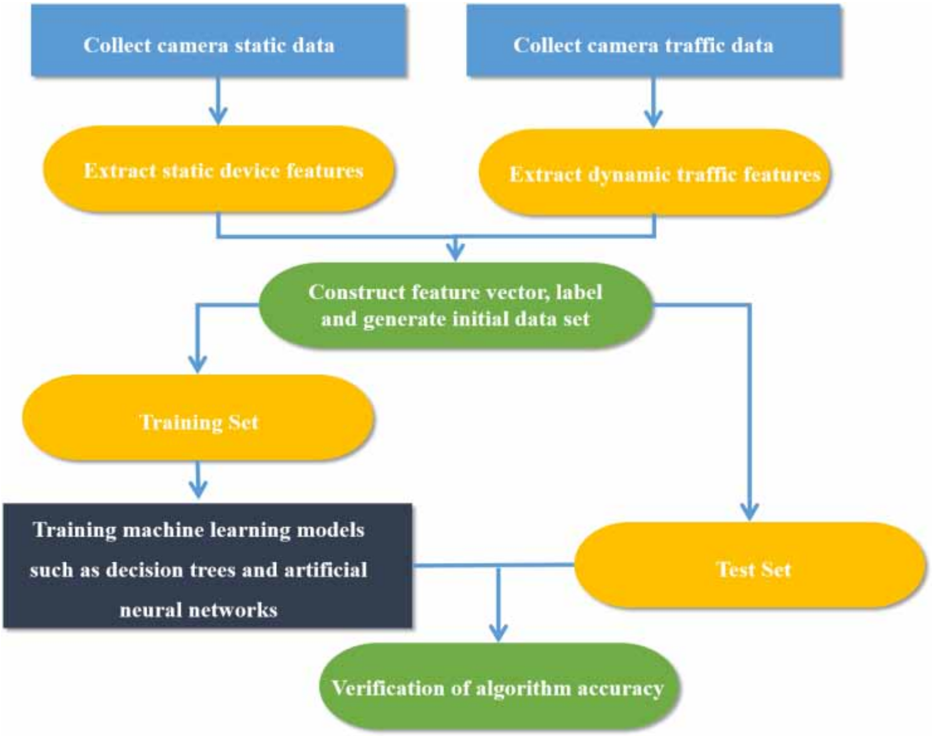

Collect camera static data
Collect camera traffic data
Extract static device features
Extract dynamic traffic features
Construct feature vector, label and generate initial data set
Training Set
Training machine learning models such as decision trees and artificial neural networks
Test Set
Verification of algorithm accuracy

Figure 3.
Video surveillance camera identifier system

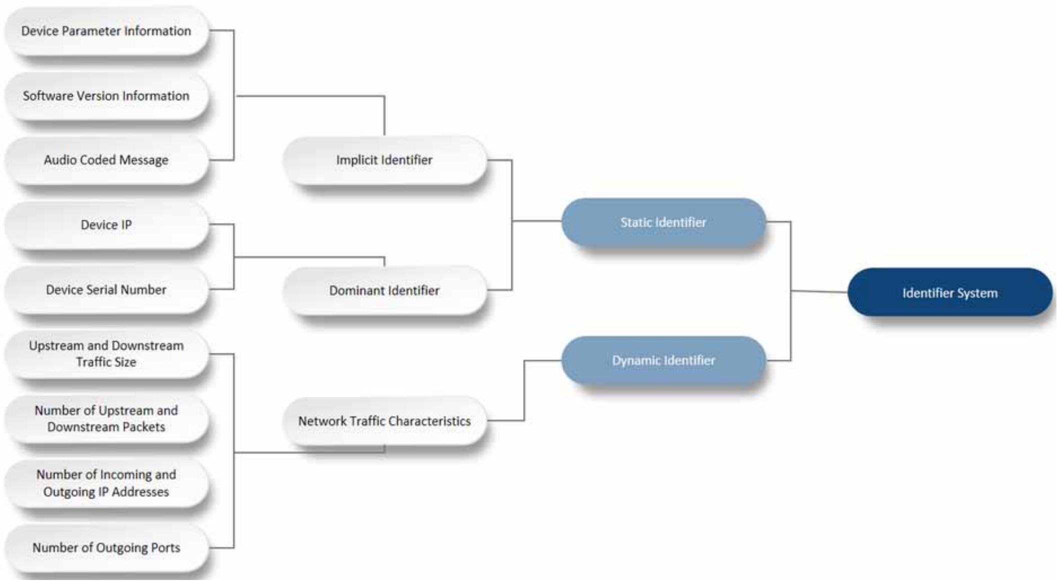

Device Parameter Information
Software Version Information
Audio Coded Message
Device IP
Device Serial Number
Upstream and Downstream Traffic Size
Number of Upstream and Downstream Packets
Number of Incoming and Outgoing IP Addresses
Number of Outgoing Ports
Implicit Identifier
Dominant Identifier
Network Traffic Characteristics
Static Identifier
Dynamic Identifier
Identifier System

Table 1.
Examples of partial video surveillance camera explicit identifiers

Device IP	MAC Address
192.168.1.65	28:57:be:29:7b:4a
192.168.1.63	28:57:be:b1:2f:a9
192.168.1.61	28:57:be:59:d6:5c
192.168.1.62	28:57:be:39:d7:56

information, audio coding information, video coding information and software information, respectively. The names of the implicit identifiers included in each class of identifiers are indicated in the remarks of Table 2. Table 3 lists some of the implicit identifier information.

Dynamic identifier system

In terms of dynamic identifier system, for the application scenario characteristics of front-end video surveillance cameras, this paper defines and extracts from the perspective of video surveillance camera traffic features. The dynamic traffic feature set is composed of a series of metrics that comprehensively reflect the traffic characteristics of the front-end video surveillance cameras, and is the basis for the unique identification of video surveillance cameras in the front-end network and subsequent accurate identification and network intrusion detection.

The dynamic identifier system of video surveillance cameras initially constructed in this paper is shown in Table 4. Network traffic features include four main categories, which are upstream and downstream traffic features, upstream and downstream packet features, inflow and outflow IP features, and outflow destination port number features, containing a total of 10 features.

Table 2.
Examples of categories of hidden video surveillance camera identifiers

Hidden Identifier Categories	Instructions
Device parameter information	Includes the device brand, model, serial number, and operating system
Channel information	Includes the channel name, channel size, open service, signaling and protocol type, and audio and video transmission protocol type
Audio coding information	Includes the audio coding type, audio sampling rate and other information
Video coding information	Includes the video encoding type, video frame rate, bit rate type and other information
Software information	Includes the main software version and DSP software version

Table 3.
Examples of hidden video surveillance camera identifiers

OS	Equipment Brand	Equipment Model	Open Cervice	Audio Encoding Type
Linux	dahua	IPC-HF	554: rtsp	G711_A
OpenWrt	hikvision	DS-2CD	9010: sdr	AAC
Windows	huawei	IPC622	20000: dnp	G711_U
Android	Lize intelligent	IDS-2CD	8080: http	G711_A

In this paper, we extract the content of a particular video surveillance camera at a specific time of day to get information such as traffic size, packet size, and IP characteristics. This information is determined by the video streaming protocol of a particular video surveillance camera, making it difficult for offenders to falsify such features, which tend to change accordingly once the front-end video surveillance camera is maliciously replaced. If managers can count and compare device traffic characteristics in real time, they can quickly detect changes in front-end video surveillance cameras.

Quantization of Identifier Contribution Value

Based on the initial construction of static and dynamic raw identifier systems, the contribution values of different identifiers need to be quantified to facilitate the selection and extraction of appropriate identifiers by users and management, while the contribution values also demonstrate that the identifiers can assist in video surveillance camera identification. In this paper, the self-information quantity and information entropy in information theory are used to analyze and judge the information quantity that an identifier has and construct the final identifier feature vector (Harte & Newman, 2014; Li et al., 2018; Majumdar et al., 2018; Zheng & Wang, 2018).

Self-information based assessment of identifier contribution has been effectively applied in several domains (Majumdar et al., 2018; Santos, 2019). In this paper, we first get the information about the contribution of the identifier in that front-end video surveillance camera identifier system by calculating the self-information of different identifiers in the identifier system; and for the overall identifier system, the expectation of the self-information of the identifier is calculated to get the information represented by the identifier system in the whole. In other words, self-information amount and information entropy quantify the information contribution of identifiers from individual and whole respectively (Rincón & Cheng, 2018; Zhang et al., 2017).

Assuming that each front-end video surveillance camera has n identifiers, use $\mathrm{P_i}$ to denote the probability that the ith identifier can correctly identify the video surveillance camera, and event $\mathrm{X_i}$ to denote that the ith identifier can correctly identify the video surveillance camera. The formula for calculating the self-information amount of event $\mathrm{X_i}$ is:

$$\mathrm{I}\left(\mathrm{Xi}\right) = -\mathrm{lnP_i} \tag{1}$$

Table 4.
Examples of video surveillance camera flow characteristics

Number	Traffic Characteristic Category	Traffic Characteristic Name
1	Upstream and downstream flow	Upstream traffic per hour
2		Downstream flow per hour
3		Total traffic per hour
4	Upstream and downstream packet	Uplink data packets per hour
5		Downlink packets per hour
6		Total packets per hour
7	IP Inbound outbound IP	Number of outgoing IP addresses per day
8		Number of inbound IP addresses per day
9		Daily total number of connected IP addresses
10	Number of outbound destination ports	Number of outbound destination ports in a day

Information entropy is designed to quantify and measure information, and from the formula, information entropy is the mathematical expectation of the amount of self-information, i.e., it is the calculation of the size of the information brought by the whole system of identifiers. If the uncertainty brought by a certain identifier system is large, the information content of that identifier system is considered to be large, and its contribution is also considered to be correspondingly large. Therefore, the identifier system with the most information can be compared by calculating the information entropy. The formula for calculating information entropy is as follows.

$$H(X) = -\sum_{i=1}^{n} P_i \ln P_i \tag{2}$$

The amount of information will be used as one of the important indicators for the selection of the final identifier system of the front-end video surveillance camera, while the amount of self-information can be used as one of the preliminary bases for the identification of the device identity, which will provide a good research basis for the subsequent research of the identification algorithm of the front-end video surveillance camera identity and the intrusion detection of the front-end network.

Extraction of Identifier Feature Vector

On the basis of identifying the selected identifiers, this section characterizes the identifiers and integrates them into feature vectors for fast and efficient identifier matching.

The feature processing in this paper draws on the discretization methods (Cui et al., 2018; Ho et al., 2021; Li et al., 2018; Lv et al., 2017; Shokoohi-Yekta et al., 2017; Zou et al., 2018). The discretization operation is divided into two main steps:

1. collecting historical data for each identifier.
2. using expert experience to organize and discretize the historical data.

Representative discrete methods include (Chormale & Ghatule, 2020; Rundo et al., 2018; Schmidhuber, 2015):

1. For some identifiers containing multiple fields, assign values to each field separately. For example, assign values [1,2,3,...,n] to different software versions.
2. For identifiers with relatively stable values, floating intervals are set based on the highest and lowest records of historical data. If the existing data is in the floating interval, assign a value of 1, otherwise assign a value of 0. For example, a video surveillance camera daily outflow ip number between [4, 5, 6, 7] floating, if a day the video surveillance camera outflow ip number is 8, then assign a value of 0.
3. For identifiers containing continuous type values, the values can be mapped to a smaller space without changing the relative size of the values. For example, if the number of video surveillance camera traffic packs at a location fluctuates between 1000-90,000, discrete rules can be developed based on expert experience to discrete the traffic packs into [1,2,3,4,..., n].

Based on the identifier system of 7 static and 10 dynamic identifiers listed in Table 1, Table 2, and Table 3. The decision to retain all identifiers after quantification of contribution values, the feature vector of the ith video surveillance camera can be expressed as:

$$P_i = \{f_{1,i}, f_{2,i}, f_{3,i}, \ldots, f_{7,i}\} \tag{3}$$

Table 5.
Identifier feature transformation method

Identifier Name	Feature Transformation Method
Packet length	Perform discrete operations and convert them to numbers
Channel name	Discrete numbers 1-4
Software version	Different software versions correspond to different numbers
OS Category	Different OS correspond to different numbers

EXPERIMENTS

Experimental Procedure

1. Experimental environment

The software environment used in this paper: Python 3.9.2, Scikit-Learn 0.24.2, Numpy 1.20.3, Graphviz 0.16, Pydotplus 2.0.2, Pandas 1.2.4. Hardware environment: Inter Core i5-8250U 8G RAM.In the experiment section of this paper, the CART decision tree algorithm was used to build the training model. The maximum depth of the decision tree was 8, the loss function adopted the information gain, the splitting strategy of each node adopted the Best strategy, and the default values of other parameters were adopted.

2. Experimental data set

In this experiment, the characteristic data of 25 video surveillance cameras, 5 servers, 8 ordinary PCs and 3 cell phone devices from different vendors are simulated and collected by referring to the real traffic characteristics of video surveillance cameras, servers, ordinary PCs and cell phone devices in video Internet and LAN.

With reference to the defined video surveillance camera identifier system (Du et al., 2019; Kim & Lee et al., 2020; Kim & Park, 2021; Rashwan et al., 2016), this experiment selected some of the more representative static and dynamic identifiers for this experimental device feature portrayal, as shown in Table 6.

Table 6.
Identifier selected in the experimental environment

Identifier Category	Identifier Name
Static identifier	Device IP
	OS Type
Dynamic identifier	Number of IP addresses exchanged with the video surveillance camera
	Upstream traffic (into the video surveillance camera)
	Downflow (out of video surveillance camera)
	Signaling protocol type
	Audio and video transmission protocol type
	Type of service provided by the video surveillance camera

The operating system types involved in this experiment include Linux, Windows, Android and IOS. Video surveillance cameras cover domestic mainstream manufacturers including Dahua, Haikang, Yuvision, Fluorite, etc. The network data message protocols supported by these video surveillance cameras include RTSP, SIP, HTTP, ONVIF and other signaling protocols and audio and video transmission protocols such as RTP and TCP. For the type of services provided by the video surveillance camera and the greater relevance of each application manufacturer, this experiment mainly selected two standard audio and video services, RTSP and ONVIF, and set their attribute set to {include, not include}.

For the static identifier of the video surveillance camera, this experiment uses the video surveillance camera parameter setting tool to obtain the video surveillance camera parameter information.For upstream and downstream traffic, this experiment simulates the data with a traffic count interval of 10 minutes and the traffic unit is Mbit.

3. Experimental steps

The procedure can be roughly divided into four stages, as follows.

(1) Each video surveillance camera corresponds to the construction of a feature vector. A total of 286 sample data were simulated and collected in this experiment, and their classification results were manually labeled. For operating system type, signaling protocol type, audio/video transmission protocol type and service type, this paper adopt the proposed discretization method to process.
(2) In this experiment, 80% of the sample set is randomly selected as the training set, and the remaining 20% is used as the test set.
(3) With the training set, this paper adopts a decision tree algorithm to construct a video surveillance camera identity recognition model.
(4) With the test set, the accuracy of video surveillance camera identification capability is verified, and the contribution of the selected identifier in the experiment is obtained.

Experimental Results and Analysis

1. Experimental evaluation index

In this paper, the evaluation metrics for video surveillance camera identification are built based on a confusion matrix that visualizes the true and predicted values of each class. In multi-categorization problems, the confusion matrix can more intuitively detect how many categories are misclassified and into which categories they are misclassified (table 7).

Video surveillance camera identity recognition can be considered as a classification problem, and the evaluation metrics for the prediction effect of the classification problem are mainly Train_Precision for training accuracy and Predict_Precision for prediction accuracy. Training accuracy is how many of the training datasets are correctly classified after modeling; prediction accuracy calculates how many of the test sets are correctly predicted to be classified.

Table 7.
Classification problem confusion matrix structure

	The Prediction Normal	The Prediction Abnormal
The Actual Normal	TP	FN
The Actual Abnormal	FP	TN

The accuracy evaluation metrics applied in this paper are defined as:

$$\Pr ecision = \frac{TP + TN}{TP + TN + FN + FP} \tag{4}$$

2. Results analysis

In this paper, we propose to identify the video surveillance camera based on decision tree model in machine learning (Du et al., 2020; Guo et al., 2017; Kalbo et al., 2020; Liu et al., 2018; Nanda & Patra, 2021). The process of generating a decision tree is to define some division criterion such that the features satisfying the criterion continuously divide the data set into subsets with higher purity and lower uncertainty. Among them, the feature selection criteria to measure the purity and uncertainty of the dataset before and after dividing the dataset include: information gain, information gain rate, and Gini index. Usually, after dividing the dataset using a certain feature selection criterion during the iterative process, the purity of each data subset is higher and the uncertainty is lower than that of the dataset before the division. The decision tree generated in this experiment is shown in Figure 4:

The leaf node represents the final classification result, and the path from the root node to the leaf node represents its step-by-step classification process. The training data set confusion matrix and test data set confusion matrix of this experiment are shown in Table 8 and Table 9 respectively:

```
According to Formula (1), it can be calculated as follows:
Train_Precision=0.961
Predict_Precision=0.931
```

The experimental results show that the generated decision tree can correctly classify 96.1% of the training sets and correctly identify 93.1% of the video surveillance cameras in the test sets.

The ranking of the contribution degree of each identifier selected in this experiment is shown in Table 10:

As seen in Table 10, the total contribution of the four features - number of IPs exchanged with the video surveillance camera, uplink traffic, downlink traffic and signaling protocol type - reaches 81% in this experiment, and they play a key role in identifying the video surveillance camera.

In summary, this experiment uses a decision tree algorithm to construct a video surveillance camera identity recognition model by collecting the feature values of selected static identifiers and

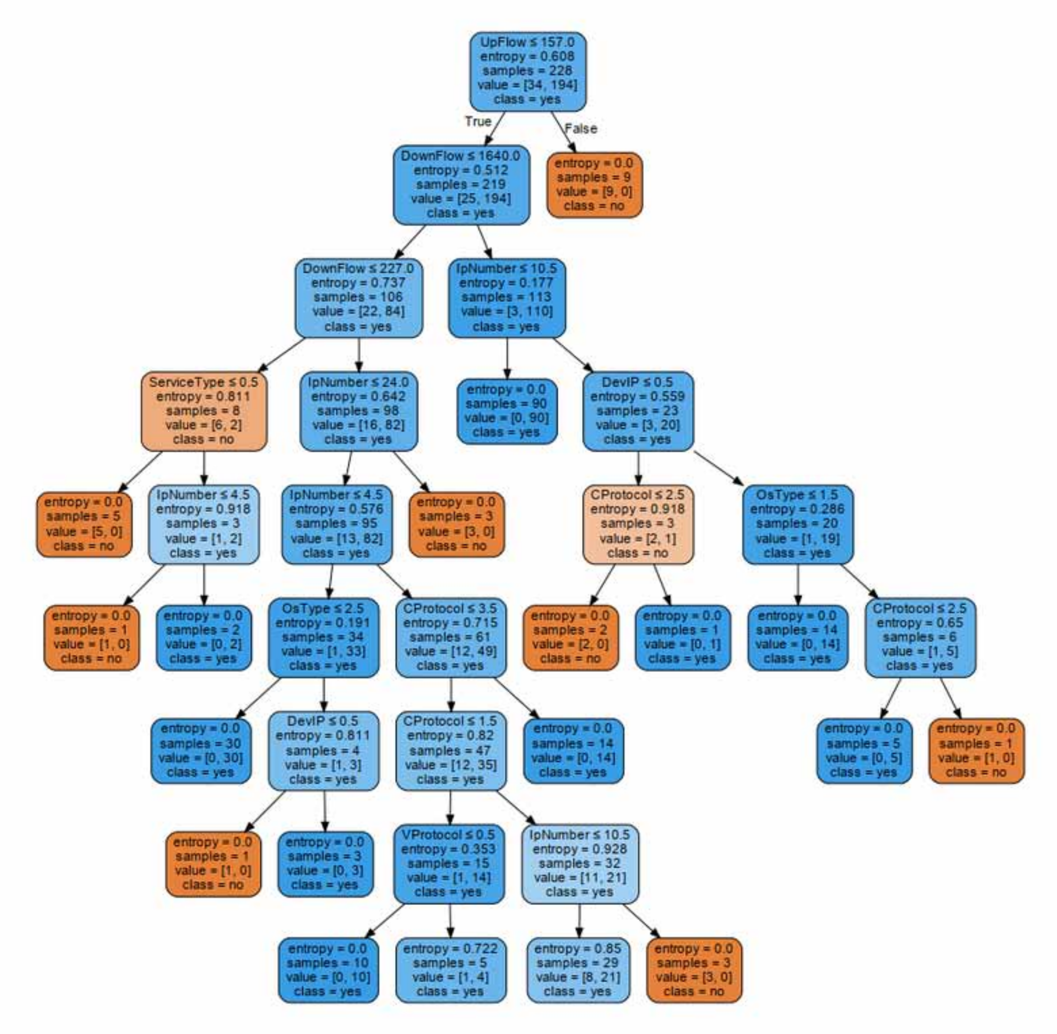
Table 8.
Train data set confusion matrix structure

	The Prediction Normal	The Prediction Abnormal
The Actual Normal	25	9
The Actual Abnormal	0	194

Table 9.
Test data set confusion matrix structure

	The Prediction Normal	The Prediction Abnormal
The Actual Normal	3	1
The Actual Abnormal	3	5 1

Figure 4.
Decision tree generated by experiment



dynamic identifiers, and the recognition accuracy reaches 93.1%. Although only eight of the defined identifiers were selected in the experimental session, the proposed method is also suitable for selecting more identifiers and constructing the corresponding video surveillance camera identification model.

CONCLUSION

With the development of smart city and public security big data, the video surveillance network is bound to achieve a higher speed of development. While video surveillance cameras are widely placed in urban corners, they also face the risk of being illegally replaced and tampered with, which, if spread, will bring serious challenges to the entire video network and public society. In this paper, we propose a video surveillance camera identification method that fuses multidimensional identification features. By extracting the explicit and implicit static features of video surveillance cameras and the dynamic features of traffic, we construct a video surveillance camera identifier indicator system, followed by

Table 10.
The contribution of the selected identifier in the experiment

Identifier Name	Characteristic Importance Coefficient	Ranking
Number of IP addresses exchanged with the video surveillance camera	0.251392	1
Upstream traffic (into the video surveillance camera)	0.238855	2
Downflow (out of video surveillance camera)	0.206824	3
Signaling protocol type	0.113581	4
OS Type	0.071143	7
Device IP	0.069007	5
Type of service provided by the video surveillance camera	0.033868	6
Audio and video transmission protocol type	0.015330	8

the fusion and screening of identifiers to compose a feature vector, and analyze the level of indicator contribution using mutual information quantity and information entropy to lay the foundation for intrusion detection and device identification. The experimental results show that explicit identifiers have the largest self-information and contribution but are the most easily forged; dynamic identifiers have the second largest contribution but have a larger workload for traffic collection and analysis; implicit identifiers have a smaller contribution but still have some identity identification capability. Based on the identifier system proposed in this paper, the identification accuracy in 41 terminal devices reaches 93.1%, and the features are not easily tampered with, which has high application value. However, the effectiveness in other video network front-end devices still needs experimental proof.

In the future, firstly, more dynamic features of video surveillance cameras will be extracted and feature fusion will be performed to obtain more effective recognition features; secondly, based on the identifier metric system and feature vector, it is a promising direction to achieve video surveillance camera intrusion anomaly detection by collecting the replaced video surveillance camera data; finally, we will use convolutional neural network and long and short term memory neural network for feature extraction and classification of identifier data to improve the intrusion detection effect.

ACKNOWLEDGMENT

This work is supported by the Natural Science Foundation of Shanghai (21ZR1422000), the China Postdoctoral Science Foundation (2020M670998), the Shanghai Talent Development Fund (2020016) and the Science and Technology Program project of the Ministry of Public Security (2019JZX004). Recommended for acceptance by Dr. Zhijie Fan. (Corresponding author: Zhiwei Cao, Co-first Author: Chunmei Wang and Zhiwei Cao). We thank the anonymous reviewers and the editors for their insightful comments and assistance.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

Akbanov, M., Vassilakis, V. G., & Logothetis, M. D. (2019). Ransomware detection and mitigation using software-defined networking: The case of WannaCry. *Knowledge-Based Systems*, *76*, 111–121.

Andresini, G., Appice, A., & Malerba, D. (2021). Nearest cluster-based intrusion detection through convolutional neural networks. *Knowledge-Based Systems*, *216*, 106798. doi:10.1016/j.knosys.2021.106798

Appasha Chormale, A., & Ghatule, A. P. (2020). Cloud Intrusion Detection System Using Fuzzy Clustering and Artificial Neural Network. *Journal of Physics: Conference Series*, *1478*, 012030. doi:10.1088/1742-6596/1478/1/012030

Atia, M. M., Noureldin, A., & Korenberg, M. J. (2013). Dynamic Online-Calibrated Radio Maps for Indoor Positioning in Wireless Local Area Networks. *IEEE Transactions on Mobile Computing*, *12*(9), 1774–1787. doi:10.1109/TMC.2012.143

Badr, S. M. (2013). Adaptive Layered Approach using C5.0 Decision Tree for Intrusion Detection Systems (ALIDS). *International Journal of Computers and Applications*, *66*, 18–22.

Bondi, L., & Bestagini, P. (2018). Improving prnu compression through preprocessing, quantization and coding. *IEEE Transactions on Information Forensics and Security*, *3*, 1–1.

Chen, N., Hu, A., & Fu, H. (2021). Lora radio frequency fingerprint identification based on frequency offset characteristics and optimized lorawan access technology. In *The 5th IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. IEEE.

Cui, J., Zhang, Y., Cai, Z., Liu, A., & Li, Y. (2018). Securing display path for security-sensitive applications on mobile devices. *Cmc-computers Materials & Continua*, *55*, 17–35.

Das, A., Borisov, N., & Caesar, M. (2014). Do you hear what i hear?: fingerprinting smart devices through embedded acoustic components. In *The 2014 ACM SIGSAC Conference on Computer and Communications Security*. ACM.

Du, H., Chen, L., Qian, J., Hou, J., Jung, T., & Li, X. (2020). PatronuS: A System for Privacy-Preserving Cloud Video Surveillance. *IEEE Journal on Selected Areas in Communications*, *38*(6), 1252–1261. doi:10.1109/JSAC.2020.2986665

Du, L., Zhang, W., Fu, H., Ren, W., & Zhang, X. (2019). An efficient privacy protection scheme for data security in video surveillance. *Journal of Visual Communication and Image Representation*, *59*, 347–362. doi:10.1016/j.jvcir.2019.01.027

Guo, J., Zheng, P., & Huang, J. (2017). An Efficient Motion Detection and Tracking Scheme for Encrypted Surveillance Videos. *ACM Transactions on Multimedia Computing Communications and Applications*, *13*, 1–23.

Hamad, S. A., Zhang, W. E., & Sheng, Q. Z. (2019). IoT device identification via network-flow based fingerprinting and learning. In *The 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*. IEEE.

Harte, J., & Newman, E. A. (2014). Maximum information entropy: A foundation for ecological theory. *Trends in Ecology & Evolution*, *29*(7), 384–389.

Ho, S., Jufout, S. A., Dajani, K., & Mozumdar, M. M. (2021). A Novel Intrusion Detection Model for Detecting Known and Innovative Cyberattacks Using Convolutional Neural Network. *IEEE Open Journal of the Computer Society*, *2*, 14–25.

Islam, M. T., Islam, B., & Nirjon, S. (2017). Soundsifter: Mitigating overhearing of continuous listening device. In *The 15th ACM Annual International Conference on Mobile Systems, Applications, and Services.* ACM.

Jo, H., & Kim, S. (2018). Indoor Smartphone Localization Based on LOS and NLOS Identification. *Sensors (Basel)*, *18*, 3987.

Jose, A. C., Malekian, R., & Ning, Y. (2017). Improving home automation security: Integrating device fingerprinting into smart home. *IEEE Access: Practical Innovations, Open Solutions*, *4*(99), 5776–5787.

Kaemarungsi, K., & Krishnamurthy, P. (2012). Analysis of WLAN's received signal strength indication for indoor location fingerprinting. *Pervasive and Mobile Computing*, *8*, 292–316.

Kalbo, N., Mirsky, Y., Shabtai, A., & Elovici, Y. (2020). The Security of IP-Based Video Surveillance Systems. *Sensors (Basel)*, *20*, 4806.

Khan, S., & Bianchi, T. (2019). Fast image clustering based on camera fingerprint ordering. In *The 2019 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE.

Kilincer, I.F., & Ertam, F., & Engür, A. (2021). Machine learning methods for cyber security intrusion detection: Datasets and comparative study. *Computer Networks*, *188*, 107840.

Kim, J., Lee, D., & Park, N. (2020). CCTV-RFID enabled multifactor authentication model for secure differential level video access control. *Multimedia Tools and Applications*, *79*, 23461–23481.

Kim, J., & Park, N. (2021). Role-based Access Control Video Surveillance Mechanism Modeling in Smart Contract Environment. *Transactions on Emerging Telecommunications Technologies*, *33*, e4227.

Koo, J., Oh, S. R., & Kim, Y. G. (2019). Device identification interoperability in heterogeneous IoT platforms. *Sensors (Basel)*, *19*(6), 1433.

Lai, Y., Qi, Y., He, Y., & Mu, N. (2019). A survey of research on smartphone fingerprinting identification techniques. *Journal of Information Security Research*, *5*(10), 856–878.

Li, A., Xue, S., & Li, X. (2020). Appdna: Profiling app behavior via deep-leaning on function call graphs. *IEEE Transactions on Emerging Topics in Computing*.

Li, D., Wang, Z., Cao, C., & Liu, Y. (2018). Information entropy based sample reduction for support vector data description. *Applied Soft Computing*, *71*, 1153–1160.

Li, H., He, P., Wang, S., Rocha, A., Jiang, X., & Kot, A. C. (2018). Learning Generalized Deep Feature Representation for Face Anti-Spoofing. *IEEE Transactions on Information Forensics and Security*, *13*, 2639–2652.

Liu, P., Yang, P., & Song, W. Z. (2019). Real-time identification of rogue wifi connections using environment-independent physical features. In *The 2019 IEEE INFOCOM Conference on Computer Communications*. IEEE.

Liu, Y., Cao, J. J., Diao, X. C., & Zhou, X. (2018). Survey on Stability of Feature Selection. *Journal of Software*, *29*, 2559–2579.

Liu, Y. W., & Jan, S. S. (2014). Assessment of indoor positioning system using kriging fingerprinting method and IEEE 802.11v standard. *27th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2014, 2*, 1622-1628.

Lv, J., Yang, W., & Man, D. (2017). Device-Free Passive Identity Identification via WiFi Signals. *Sensors (Basel)*, *17*, 2520.

Ma, X., Zhu, B. B., Zhang, T., Cao, S., Jin, H., & Zou, D. (2018). Efficient privacy-preserving motion detection for HEVC compressed video in cloud video surveillance. *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),* 813-818.

Mahmood, K., Shamshad, S., & Rana, M. (2021). Puf enable lightweight key-exchange and mutual authentication protocol for muti-server based d2d communication. *Journal of Information Security and Applications*, *61*(3), 102900.

Majumdar, S., Mukherjee, N., & Roy, A. K. (2018). Information entropy and complexity measure in generalized Kratzer potential. *Chemical Physics Letters*, *716*, 257–264.

Mandelli, S., Argenti, F., & Bestagini, P. (2020). A modified fourier-mellin approach for source device identification on stabilized videos. In *The 2020 IEEE International Conference on Image Processing (ICIP)*. IEEE.

Nanda, M. K., & Patra, M. R. (2020). Intrusion Detection and Classification using Decision Tree Based Key Feature Selection Classifiers. *Intelligent and Cloud Computing*, *153*, 157–170.

Nazare, A. C., & Schwartz, W. R. (2016). A scalable and flexible framework for smart video surveillance. *Computer Vision and Image Understanding*, *144*, 258–275.

Radhakrishnan, S. V., Uluagac, A. S., & Beyah, R. (2014). Gtid: A technique for physical device and device type fingerprinting. *IEEE Transactions on Dependable and Secure Computing*, *12*(5), 519–532.

Rashwan, H. A., Solanas, A., Puig, D., & Martínez-Ballesté, A. (2015). Understanding trust in privacy-aware video surveillance systems. *International Journal of Information Security*, *15*, 225–234.

Rincón, C. A. C., & Cheng, A. M. (2018). SITSA-RT: An Information Theory Inspired Real-Time Multiprocessor Scheduler. *2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC),* 156-163.

Rundo, F., Conoci, S., Ortis, A., & Battiato, S. (2018). An Advanced Bio-Inspired Photo Plethysmo Graphy (PPG) and ECG Pattern Recognition System for Medical Assessment. *Sensors (Basel)*, *18*, 405.

Santos, W. D. (2019). The entropic and symbolic components of information. *Bio Systems*, *182*, 17–20.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks: The Official Journal of the International Neural Network Society, 61*, 85-117.

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815-823.

Shen, G., Zhang, J., & Marshall, A. (2021). Radio frequency fingerprint identification for lora using deep learning. *IEEE Journal on Selected Areas in Communications*, *39*(8), 2604–2616.

Shokoohi-Yekta, M., Hu, B., Jin, H., Wang, J., & Keogh, E. J. (2017). Generalizing DTW to the multi-dimensional case requires an adaptive approach. *Data Mining and Knowledge Discovery*, *31*, 1–31.

Taspinar, S., Mohanty, M., & Menon, N. (2020). Camera fingerprint extraction via spatial domain averaged frames. *IEEE Transactions on Information Forensics and Security*, *15*, 3270–3282.

Thangavelu, V., Divakaran, D. M., & Sairam, R. (2018). Deft: A distributed IoT fingerprinting technique. *IEEE Internet of Things Journal*, *6*(1), 940–952.

Thiyagarajan, K., Lu, R., El-Sankary, K., & Zhu, H. (2018). Energy-Aware Encryption for Securing Video Transmission in Internet of Multimedia Things. *IEEE Transactions on Circuits and Systems for Video Technology*, *29*, 610–624.

Vastel, A., Laperdrix, P., & Rudametkin, W. (2018). Fp-stalker: Tracking browser fingerprint evolution. In *The 2018 IEEE Symposium on Security and Privacy (SP).* IEEE.

Xie, P., Feng, J., & Cao, Z. (2017). Genewave: Fast authentication and key agreement on commodity mobile device. In *The 25th IEEE International Conference on Network Protocols (ICNP).* IEEE.

Xue, S., Zhang, L., & Li, A. (2018). Appdna: App behavior profiling via graph-based deep leaning. In *The 2018 IEEE INFOCOM Conference on Computer Communications*. IEEE.

Yang, Y., Zhou, W., Zhao, S., Liu, C., Zhang, Y., Wang, H., & Zhang, Y. (2021). Survey of IoT security research: Threats, detection and defense. *Journal of Communication*, *42*(8), 188–205.

Yin, X. M., Hu, Z. L., Chen, G. L., & Huang, H. Y. (2016). Research on IP Video Network Access Detection Based on Decision Tree Classification of Device Fingerprint. *Netinfo Security*, *12*, 68–73.

Zhang, Y., Qian, C., Lv, J., & Liu, Y. (2017). Agent and Cyber-Physical System Based Self-Organizing and Self-Adaptive Intelligent Shopfloor. *IEEE Transactions on Industrial Informatics*, *13*, 737–747.

Zheng, K., & Wang, X. (2018). Feature selection method with joint maximal information entropy between features and class. *Pattern Recognition*, *77*, 20–29.

Zhou, J., Cao, Z., Dong, X., & Vasilakos, A. V. (2017). Security and Privacy for Cloud-Based IoT: Challenges. *IEEE Communications Magazine*, *55*, 26–33.

Zou, H., Zhou, Y., Yang, J., Gu, W., Xie, L., & Spanos, C. (2018). WiFi-Based Human Identification via Convex Tensor Shapelet Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *32*(1).