Forensic Science International 348 (2023) 111674

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

Ok Google, Start a Fire. IoT devices as witnesses and actors in fire investigations

Francesco Servida *, Manon Fischer, Olivier Delémont, Thomas R. Souvignet

University of Lausanne, School of Criminal Justice, Batochime, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 3 October 2022
Received in revised form 20 March 2023
Accepted 28 March 2023
Available online 1 April 2023

Keywords:
Fire reconstruction
Origin
Cause
Digital forensic
Internet of Things (IoT)
Smart devices
Smart buildings

ABSTRACT

Fire incidents are amongst the most destructive events an investigator might encounter, completely transforming a scene with most of the objects left in ashes or highly damaged. Until now, fire investigations relied heavily on burn patterns and electrical artifacts to find possible starting locations, as well as witness statements and more recently witness imagery.

As Internet of Things (IoT) devices, often seen as connected smart devices, become more common, the various sensors embedded within them provide a novel source of traces about the environment and events within. They collect and store information in different locations, often not touched by the event, such as remote servers (cloud) or companion smartphones, widening the investigation field for fire incidents. This work presents two controlled fire incidents in apartments that we furnished, equipped with IoT devices, and subsequently burnt. We studied the traces retrievable from the objects themselves after the incident, the companion smartphone apps, and the cloud and assessed the value of the information they conveyed. This research highlighted the pertinence to consider traces from IoT devices in the forensic process of fire investigation.

1. Introduction

Fire incidents constitute one of the most challenging investigation environments for a forensic examiner; indeed the scene is often damaged by the fire itself as well as by firefighters during fire suppression and overhaul. The loss of structures, goods, information, as well as the destruction of traces that could reveal the mechanism that triggered the fire are direct consequences of the fire. Fire investigations then rely on forensic examination principles with the aim to reconstruct the course of the event, and in particular 1) to locate the origin of the fire, 2) to determine the cause of the fire (at least to articulate the possible hypotheses of cause), and 3) to explain the development of the fire and its effluents [5].

Whilst in some scenes the origin and cause might be evident and not necessitate expert intervention, it is when a scene is damaged to an extreme extent that such an intervention might be needed [8].

Although the determination of the cause is most often considered the goal of the fire investigation, it is largely conditioned by the location of the origin of the fire. The latter usually relies on the examination of burn patterns. By taking into account the physical and chemical knowledge of fire and heat propagation, backward reasoning is applied to define, by hypothesis, the location where the fire started. The artifacts produced by the propagation of the fire on the electrical facilities (circuit breaker tripping sequence, arc mapping…) are also very often considered, as well as statements from witnesses and first responders [5]. More recently witness imagery has turned out to have an ever-increasing influence. Indeed fire incidents tend to attract crowds and with the widespread adoption of smartphone devices comes a large number of sources of images and videos; this comes in supplement to surveillance systems which previously were the only source of information during a fire incident. The first images were commonly taken only on arrival on scene of the first responders [10].

Cause determination then relies on the determination of the manner in which the elements of the fire triangle were combined to enable the ignition. Once a possible origin is determined, the investigator has to thoroughly examine all possible combinations of combustible materials and heat sources (and sometimes oxidizers as well), and consider the mechanisms that could explain such combinations [8].

1.1. Internet of Things (IoT)

IoT devices, often seen as connected smart devices, are a component of the recent evolution towards an interconnection of sensors and actuators called the Internet of Things [1]. IoT devices are increasingly common and are becoming an important source of

* Corresponding author.
E-mail addresses: francesco.servida@unil.ch (F. Servida), manon.fischer@unil.ch (M. Fischer), olivier.delemont@unil.ch (O. Delémont), thomas.souvignet@unil.ch (T.R. Souvignet).

https://doi.org/10.1016/j.forsciint.2023.111674

traces for investigations [7,13]; they are often equipped with sensors and can function as actuators for remote or automated control. Their sensors allow them to record environmental changes and events such as motion or speech, traces which in some cases could be used to support investigations or court decisions, paving the way to digital witnesses [4,12,14]. However, they can also malfunction or be misused, in which case they can become actors themselves in the incident or crime.

Traces from these devices can be stored either on the devices themselves, although in a limited fashion, on the companion devices used to control them, or remotely on the service provider servers. As such, some of these traces extend beyond the physical location of the device and widen the investigation field, sometimes even crossing jurisdictional boundaries and preserving them after the destruction of the devices.

Whilst to some extent IoT devices are already exploited in forensic investigations where their relevance is obvious, with the notable example of the highly connected Tesla vehicles [3], they are more often than not overlooked in traditional investigations. Such is the case with fire investigations. That is also the case for scientific research, as to our knowledge no contribution focused on the exploitation of IoT devices in fire cases is publicly available.

1.2. Research

This work presents two controlled large-scale experiments of arson in which IoT devices acted as witnesses as well as actors. Section 2 describes the devices analyzed and the study methodology; Section 3 details the raw results obtained from the digital forensic analysis of the components of those IoT ecosystems which are then discussed in light of their usefulness for the fire investigation in Section 4.

The contributions of this research are multiple. It highlights the potential of IoT traces for fire investigation and the extension of investigation fields they involve. But this study also provides a more general understanding of the interest of digital traces in crime scene investigations, as well as demonstrates the high potential of privacy lawful requests to conduct proactive cloud traces investigation.

2. Methodology

This study was conducted in collaboration with the fire investigation team of the School of Criminal Justice of the University of Lausanne, and the forensic service of the Police Neuchateloise. It involved two different arson scenarios, hereafter Scenario 1 and Scenario 2.

2.1. Scenario 1

The first scenario was performed in a multi-room apartment in an multi-storey building destined for demolition. It represented an arson case with accelerant spills. The objective of the fire investigation team was to study several strategies of ignitable liquid detection, and IoT devices were included as witnesses in order to study the chronology and spread of the fire in the apartment for the hypothetical smart home.

2.1.1. Devices and locations

For the scenario we decided the owner had equipped his home with multiple IoT devices presented in Table 1, ranging from temperature and motion sensors to smart cameras, smoke detectors, and a voice assistant. The choice of the devices was based mainly on the expected type of trace (temperature/humidity, presence, smoke, connection state) independently of the location (device/smartphone/cloud) where the traces were expected to be found, if any. The devices were installed mimicking a reasonable smart home setup, with temperature sensors in each room and more specific devices in the living room and the "kid's bedroom" (Bedroom 2). The schematic plan of the apartment with locations of the IoT devices is presented in Fig. 1. For this scenario the Trådfri light in the living room was linked to the Trådfri motion sensor with a timeout of 30 s and the light in Bedroom 2 was kept turned on.

Table 1
Devices installed for Scenario 1.

Vendor	Device	Function	Location
Google	Google Home Mini	Smart Assistant	Living Room
Google	Google Nest Protect v2	Smoke Detector	Living Room
Meross	Meross Gateway	Gateway	Living Room
Meross	Meross Thermostat	Radiator Valve	Bedroom 2
Xiaomi	Mi Control Hub	Gateway	Living Room
Xiaomi	Mi Motion Sensor	Motion Sensor	Bedroom 2
Xiaomi	Mi Motion Sensor 2	Motion Sensor	Living Room
Xiaomi	Mi Temperature and Humidity Sensor	Environment Sensor	Living Room
Xiaomi	Mi Temperature and Humidity Sensor 2	Environment Sensor	Bedroom 2
Xiaomi	Mi Temperature and Humidity Sensor 3	Environment Sensor	Bathroom
Xiaomi	Mi Temperature and Humidity Sensor 4	Environment Sensor	Bedroom 1
Xiaomi	Mi Temperature and Humidity Sensor 5	Environment Sensor	Kitchen
QBee/Askey	QBee Camera	Camera	Living Room
IKEA	Trådfri Bulb Multicolor	Smart Bulb	Bedroom 2
IKEA	Trådfri Bulb White	Smart Bulb	Living Room
IKEA	Trådfri Gateway	Gateway	Living Room
IKEA	Trådfri Motion Sensor	Motion Sensor	Living Room

2.1.2. Timeline

As explained previously scenario 1 represented an arson case with the use of ignitable liquid. One liter of diesel fuel was poured in Bedroom 2, as well as a liter of gasoline in the living room at 11:01. Music had been started previously on the Google Home through a

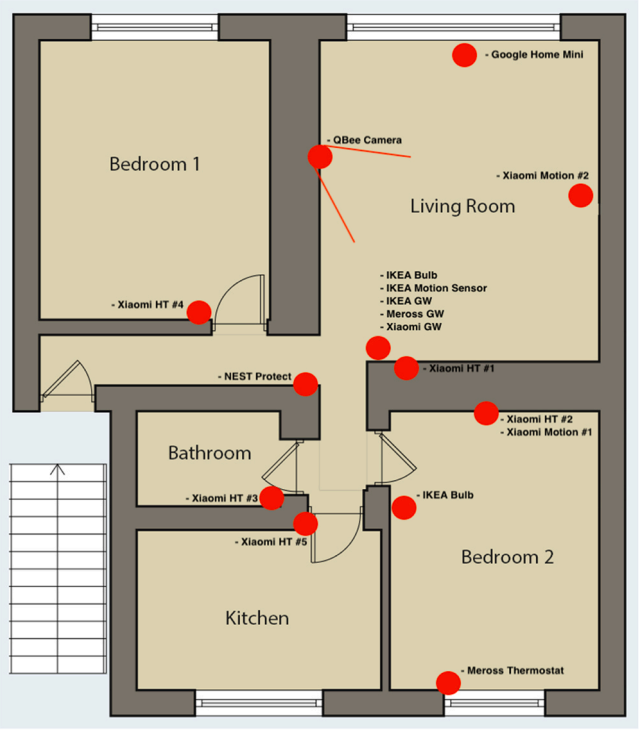

Fig. 1. Floorplan of Scenario 1 with device locations.

Fig. 2. Scene of Scenario 1. Before the fire (left); After the fire (right). *Credits: Service Presse, Police Neuchateloise.*

Table 2
Devices installed for Scenario 2.

Vendor	Device	Function	Location
Netatmo	Smart Home Weather Station Indoor module	Gateway	Outdoor
Netatmo	Additional Smart Indoor Module	Environment Sensor	Living Room Area
myStrom	myStrom WiFi Switch - Meter	Power Meter	Outdoor
myStrom	myStrom WiFi Switch - Kitchen	Smart Outlet / Power Meter / Temperature	Kitchen Area

YouTube playlist and was left running. At 11:02 the gasoline was ignited and flames rapidly engulfed the living room; in order to simulate a realistic case the fire was not controlled for around 20 min, after which firefighters entered on scene and started extinction operations. Even though the fire was left uncontrolled for a relatively long time it mainly damaged the living room which was heavily burned, but only created heat and smoke damage to the other rooms (Fig. 2).

2.2. Scenario 2

The second scenario was performed in a single-room apartment in a dedicated firefighters' training building. This scenario represented an arson case in which IoT devices played not only a witness role but were also an active part in the fire ignition. Indeed in this case the owner of the apartment decided to rig an immersion heater in a box with some Styrofoam as a rudimentary ignition device and connect it to a smart outlet in the kitchen. The smart outlet was used to turn the heater on at a later time remotely.

2.2.1. Devices and locations

For this scenario, a reduced set of IoT was chosen, with only two devices left inside the apartment: the rigged smart outlet and a Netatmo Indoor Weather Module. An additional smart outlet was positioned outside and used as a power meter whereas the Netatmo base station was left outside the house. The complete list of devices with their locations is presented in Table 2 and Fig. 3.

2.2.2. Timeline

In this scenario the owner prepared the ignition device in the morning and left the house, leaving all electronic devices in standby mode. At the beginning of the afternoon, at 14:06, being several kilometers away, he toggled remotely the power outlet with his smartphone, which turned on the immersion heater and led to the Styrofoam catching fire and subsequent propagation to the entire apartment. In the scenario timeline the fire was detected only after some minutes by some passers-by with firefighters entering the scene only after 20 min; this was sufficient for the complete destruction of the apartment with almost every non-metallic furniture item in ashes (Fig. 4).

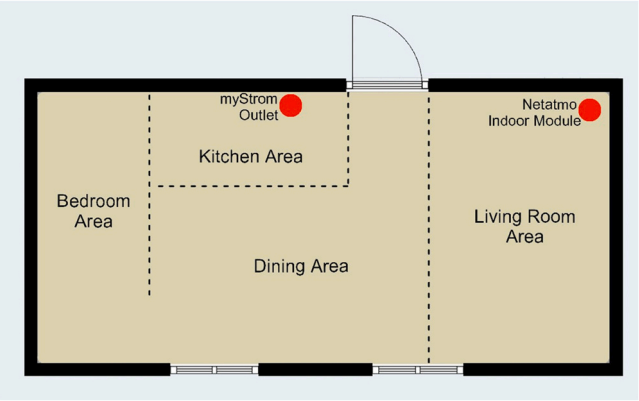

Fig. 3. Floorplan of Scenario 2 with device locations.

2.3. Environment setup

The objective of the study was to understand which data would be available in a real-world situation to a forensic specialist conducting an investigation of the two scenarios. Data produced by IoT devices can be found in multiple locations, notably on the devices/sensors themselves, gateways, and smartphones with companion apps linked to the devices as well as the cloud; some data might also be found on network devices [11].

In addition to the data collected from the various devices, the companion smartphone apps, and the service providers, we collected proactively all network traffic generated by the devices with the internet and between them, as well as additional measurements for the devices' state to use as a reference in interpreting the meaning of the traces.

The network was set up with an OpenWRT router with a 4G dongle, acting as mobile wifi access point, and a Raspberry-Pi with Home Assistant[1] for the collection of the reference data similarly to what is described by [9]; both were outside the fire perimeter so that data could be extracted and they could provide service as long as possible. In both scenarios precautions were taken so that an electrical fault in the system would not immediately stop the collection

[1] https://www.home-assistant.io/

Fig. 4. Scene of Scenario 2. Before the fire (left); After the fire (right). *Credits: Fire investigation team of the School of Criminal Justice.*

of data from battery-powered devices, mimicking a multi-phase setup with multiple circuit breakers.

2.3.1. Reference data collection

As discussed previously, reference data was collected locally on a Raspberry Pi running Home Assistant. Home Assistant is a home automation software allowing centralization of management of devices from different vendors, with integrations for a large number of devices. These integration poll either the device directly, or pull data from the cloud APIs provided by the vendor, thus allowing the storage of a local copy of any data produced by the devices. Data collected from the device memory, the smartphone, or the cloud might be unstructured, or not in sync between the different vendors. Having a local reference would allow us in the analysis step to correctly interpret the traces.

Network traffic generated by the devices, both local and outbound, was also collected as reference data. Using network traffic we could evaluate if network traces, the likes of which might be collected from an internet service provider (ISP), could be used to estimate which devices were operating in the network, their actions, and their connection state. Collection was done locally on the WiFi router, and for Scenario 1, mirrored to a laptop running Arkime.[2] This was to ensure preservation of data in case the fire also damaged the router. For Scenario 2 collection was done only locally on the router as there was a much lower risk of loss. However, analysis of the network capture only highlighted some DNS requests specific to the involved IoT ecosystems, which could be useful to enumerate the ecosystems present in a network; traffic was mainly encrypted and no additional insight could be obtained and as such will not be presented further. Additional analysis of network data, for example, classification by machine learning was not done but might be the object of further studies.

2.3.2. Companion smartphone

To control the devices installed we used a dedicated Android smartphone per scenario, a Samsung S9, and S8 respectively. This ensured a clean environment was available for the generation of traces during the scenario.

Both phones were afterward extracted using Cellebrite's UFED 4PC. Data available on the device after a forensic collection might be limited to the data loaded on the last application use; as such two extractions were performed in both cases: one extraction just after collection, with no network refresh, and a second one after having manually used and refreshed the applications. The second extraction as well as manually manipulating the applications would mimic data available if a suspect or victim is collaborating or the smartphone can be unlocked. It would also allow gaining further insight into all the data available, irrespective of security limitations on the devices. It thus might not be representative of data available on a smartphone in a real case investigation where the device cannot be unlocked and/or connected to the network.

2.4. Device analysis

A main issue with fire incident investigation is that the scene is often destroyed and changed from the original state. This is both due to the very nature of the incident but also the actions first responders take to preserve it, with firefighting actions liable to throw objects even meters away from their original location.

The first challenge is therefore the location and identification of IoT devices. In our case, this meant looking for remaining electronics, often by following electrical lines and collecting the devices. In some cases, it also meant collecting pieces of the floor itself as the devices were melted within.

The second step of device analysis consisted in scavenging the logic boards and memory chips from the melted or carbonized devices in order to use classical techniques for physical access to memory chips. Two techniques were used for this step:

- Dumping memory using programmer devices: this was employed in the case of intact logic boards with memory chips embedded in the processor itself (SoC); in this case, we connected over debugging interfaces to the processor and requested access to the memory. The main problem with this technique is that it relies on the memory not being read-protected, which unfortunately was the case for both devices where it was employed.
- Chip Off and Raw Memory Dump: this was used when a discrete memory chip was identified; in this case, we desoldered the memory chip and read the raw memory using specialized tools such as the MTK-II.

The last stage consisted in analyzing the raw dump with tools the likes of binwalk and string analysis as well as trying to reconstruct the file system from the memory dumps when a supported filesystem was identified. The main goal here was to retrieve information related to the configuration of the devices as well as possible cached data or log files of forensic interest.

2.5. Smartphone analysis

2.5.1. Manual interaction

After the initial extraction, the smartphone was unlocked and reconnected to the network; we manipulated every companion application to force a refresh from the cloud of possible historical data concerning the timeframe of the fire incident. This would allow pertinent information to be displayed and manually captured, as

[2] https://arkime.com/

well as potentially create additional interesting traces in the databases, caches, and log files of the applications which would be preserved by the second extraction and could be analyzed with forensic tools.

2.5.2. Extractions

Smartphone analysis was performed both automatically with UFED PA and manually using both X-Ways and Autopsy. As IoT companion apps are not generally supported by commercial tools, manual analysis of the extractions combined with custom plugins for Autopsy[3] was performed.

2.6. Cloud data analysis

In addition to data obtained from the devices themselves and from the companion apps on the smartphones, a goal of the research was to retrieve and analyze data the devices might have sent and stored remotely on the service provider servers.

Extraction of remote data comes with multiple challenges, both from a technical and legal perspective. For the purpose of the research we relied on two main ways of access: from the user interface of the account and data requests according to General Data Protection Regulation (GDPR), mimicking a lawful request from judiciary authorities.

2.6.1. Data on the user interface

Logging in to the account is required to access to personal data on the user interface. From it, it may be possible to visualize and/or export data. This data could also be recoverable through API requests. This type of access was not considered in this study.

2.6.2. Lawful requests

As researchers, we do not have a lawful right to access (even simulated) cases related data. However, to mimic law enforcement lawful requests, we relied on the "right of access" carried by data privacy laws. In Europe, GDPR[4] allows individuals to, among other rights, request a copy of personal data detained by a service.[5] The requests could be automatically processed, in the case of some larger services (eg. Google with Google Takeouts), or were completely manual with contact only via email or mail. For the two scenarios, we did automatic data request (for Google, Netatmo, Xiaomi), email request (for Ikea, Meross, QBee/Askey, Xiaomi, Netatmo, MyStrom), and also mail request (Xiaomi). Multiple requests were sent to Xiaomi through different means, as no answer was initially received in an acceptable delay (30 days).

3. Results

3.1. Device analysis

Device analysis required extensive work to access the logic board themselves as well as to extract data afterward. It was not always successful; in the case of the Google Home the device as well as it chips were recognizable however the memory chips were destroyed by the extreme temperatures and could not be read; the Meross thermostat and gateway as well as the Xiaomi sensors and the Ikea bulbs had no memory chips and therefore access to the internal memory of the SOC was required. This was tried for the Meross thermostat as well as Xiaomi sensors without success because of read protections whereas for the Meross gateway as well as the IKEA bulbs no connection could be made because we did not have access to the required debugging tools nor could we acquire them in a timely manner. Had we obtained the required tools, we expect the results would have been similar to that of Xiaomi sensors, with read protections preventing access to the device data.

For the other devices, notably the Nest Protect, the Xiaomi Gateway, the IKEA Gateway, and the QBee camera, it was possible to obtain the full binary dump of the memory chips. The binary dumps were analyzed for existing file systems. Only the dump from the Xiaomi Gateway presented the file system of type UBIFS, however, corruption in the memory dump rendered the file system unmountable; as such the only traces recovered from these and the other three dumps were obtained by carving and string parsing.

As detailed in Table 3 most of the data recovered from the physical devices in this way consisted in various device IDs such as MAC addresses or serial numbers, user IDs, wireless credentials, as well as some information about connected sensors or room configuration; these identifiers could be exploited to target devices when requesting data via judiciary requests. On no device dump, it was possible to retrieve logs about actions, events, or environmental conditions.

3.2. Smartphone

Analysis of data present on the linked smartphones was performed manually in both scenarios, indeed the forensic tool available (UFED PA) did not support any of the companion apps installed. Artifacts retrieved from the application data folders are presented in Tables 4 and 5. It includes expected configuration data for all of the configured applications, at a minimum account information and access tokens, which might be useful to retrieve information from the cloud either programmatically or via judiciary requests (see Sections 2.6 and 3.3); however, it is also possible to find event logs and environmental conditions for some of the applications.

3.2.1. Scenario 1

For the first scenario, the most interesting findings related to the Xiaomi Mi app: a history of sensor values both for temperature and humidity as well as the motion detection is present, stored in an XML file for each configured sensor[6]; due to the application architecture this log is updated from the cloud during usage of the application. Of interest is the difference in the handling of motion sensor logs, which are refreshed with the historical values from the cloud, even if the sensor is not connected anymore, and the temperature/humidity sensors, which are only refreshed from the cloud if the sensor is still connected. Additionally refreshing the app, even if no sensor or gateway is connected will still pull from the cloud the last sensor state and store it in the "home_env_info.xml" file in the Xiaomi Mi app preference directory. The last motion events recorded in the historical logs were at 11:02:30 (Living Room) and 11:03:21 (Bedroom 2) whereas the last sensors state indicated the living room experienced the highest temperature of 130°C and lowest humidity of 9%RH at 11:02:57 (cf. Fig. 6).

Through manual interaction with the QBee Camera application, we were able to download video recorded during the event; interestingly as videos were triggered by motion detection, the camera recorded not only the accelerant spill at 11:01, but was also triggered by the flames and smoke at 11:02:30. Whilst one could argue it is unrealistic to leave a camera on during an arson scenario, it is

[3] https://github.com/fservida/autopsy_plugins/tree/main/mi_home

[4] EU General Data Protection Regulation: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

[5] By the right of access, art.15 GPDR; services operating in Europe have to provide this data (if it has any) upon request within 30 days.

[6] /data/com.xiaomi.smarthome/files/plugin/install/rn/1001810/data/ < device_id > /data/config.xml

Table 3
Device Analysis Results for Scenario 1.

Meross	Google Home	Nest Protect	Xiaomi Home	Ikea Trådfri	QBee Camera
Thermostat & Gateway • No extraction possible	Damage to memory chips too extensive	**Detector** • Account ID • Structure ID (Home ID) • SSID and WiFi PSK	**Gateway** • Serial Number • Model • MAC Address • SSID and WiFi PSK • References to some sensors • BUT no evident logs Unable to reconstruct filesystem (UBIFS) due to memory corruption **Sensors** • No extraction possible	**Gateway** • Serial Number • Room Configuration • Device Configuration or Logs • No evident timestamped information **Lights** • No extraction possible	**Camera** • Serial Number • Model • SSID and WiFi PSK

Table 4
Smartphone Analysis Results for Scenario 1.

Meross	Google Home	Nest Protect	Xiaomi Home	Ikea Trådfri	QBee Camera
Model/SN/MAC/IP…	Account Information	User ID	Last home sensors state	Total Devices Installed, Accessory IDs and groups	Username
DeviceSettings LastActive Time *	Tokens	Tokens Geofence Transition Events	Rooms and device setup Device specific logs (sensor readings)	Google/Amazon API tokens	Login Cookie
Email/ID DeviceInformation		Objects Cache Events Cache	Account settings		

Table 5
Smartphone Analysis Results for Scenario 2.

Netatmo	myStrom
Home ID Station MAC Address	Auth Tokens Device Details Type MAC Address/ID Name Configured Schedules Energy Reports Cached HTTP requests (raw) reponses (JSON)

nonetheless interesting to see that fire and smoke triggered a motion sensor, which could still give interesting chronological information in other circumstances.

3.2.2. Scenario 2

Concerning the second scenario, traces of particular interest were found in the data directory for the smart outlet's "myStrom" application; indeed in addition to the configuration details, the app keeps a cache of HTTP requests performed. Looking at those and the corresponding response it was possible to identify the time and the request responsible for the activation of a smart socket at 13:06:23, which could be cross-referenced with the information in the configuration file to identify the socket as the one in the kitchen (Fig. 5). This not only highlights the intentional enabling of the socket but also allows to attribute the action to the device that performed it, which is often an issue with cloud-enabled systems where multiple devices could have access to the same user account.

3.3. Cloud

From the lawful requests, different situations arose. Most of the service providers answered and provided us with the data that was requested. However, we did not receive any answer from QBee Camera service provider, and thus we could not obtain any data. Finally, IKEA claimed that no data are collected and stored on their side.

All the answers did not necessarily contain traces of interest in terms of event reconstruction or chronology. This was the case for Meross. Indeed, only identifying information about the account or the device was provided (see Table 6).

Fig. 5. myStrom smartphone analysis. Cached HTTP request turning on the smart socket (left); Extract of myStrom app configuration database (right).

Table 6
Cloud Analysis Results for Scenario 1.

Meross	Google Home	Nest Protect	Xiaomi Home	Ikea Trådfri	QBee Camera
UUID and model of the linked smartphone	Google Account ID	Events history	Measurements of Temperature and Humidity	No data	No answer
UUID and MAC address of the gateway	Registered email	Tokens			
Login and Google tokens (expired)	Android app usage	Sensors measurements history			
Registered email	Voice commands history IKEA bulbs state information YouTube history				

From Google, Xiaomi, myStrom and Netatmo, interesting data were obtained and detailed in the following subsections.

Main results for the cloud data are summarized in Tables 6 and 7.

3.3.1. Scenario 1

Google. From Google requests, we obtain traces generated by the vocal assistant and the Nest[7] smoke detector. For these two devices, there is data about the account and the devices, but also data related to events. Data is organized in different folders.

Account and devices information. In the "Account Google" folder, there is a history of login IP addresses in addition of the general information about the owner of the account (such as Google ID, email address or phone number).

Data related to events. A lot of data is stored in the "My Activity" folder, composed of several subfolders. Most of it is generated by a generic Android use and should not be considered as part of the fire scenario. In the "Assistant" subfolder, there are traces generated by the voice assistant when using it. The file "My Activity.html" contains information about the voice commands. This is either information about the use of the assistant or voice commands transcribed, along with a timestamp. From this file, it is observed that the last voice commands happened at 09:58, so one hour before the event. Someone asked the vocal assistant to play a piece of music. No vocal record was available for this command, only the transcriptions. From the "Youtube" subfolder, we obtain the YouTube history. Fifteen music videos were played from 09:58 until 11:03.

In the "Nest" folder, there are the traces generated by the smoke detector. It consists of subfolders, but the more pertinent traces are in the "protect" subfolder. There are two files of interest: *events.csv* and *sensors.csv*. On one hand, the *events* file, stores logs about the proper functioning of the device, such as a speaker or buzzer tests, network connection or battery state. From this file, it is observed that two emergency alerts were triggered at 11:02. The first one, titled "HeadsUp Smoke2″, is a warning about an increase in the level of smoke.[8] The second one, titled "Emerg Smoke", due to the fact that the emergency smoke values have been reached.[8]

On the other hand, the *sensors* file stores logs about the measurements made by the sensors integrated into the smoke detector, including temperature, and humidity. They are taken every 15 min, and a maximum value is assigned to each measure. From this file, it is observed the last entry was at 14:00, which reported unusual measurements; high amounts of carbon monoxide and smoke were detected by the smoke detector. At 11:00, the temperature was 20.1°C, and the humidity was 68.4%RH.

Table 7
Cloud Analysis Results for Scenario 2.

Netatmo	myStrom
Measurements of - temperature - humidity - CO_2	Measurements of - consumption - power - energy - temperature

In the "App Home" folder, there are traces generated by the IoT devices set up in the "smart house" and managed by the *Google Home* application.[9] The folder contains several files. In the *Home-App.json* file, there is information about the IKEA Trådfri system and the state of the bulb. At 10:58, the smart white bulb in the living room had the following state: *on*, *online,* and a brightness of *100*.

Xiaomi. From Xiaomi requests, we obtain traces generated by Xiaomi sensors. Three files were sent, but only one (*user_device_data*) contains relevant information useful for event reconstruction. From this file, we gather measurements made by the motion or temperature and humidity sensors. First, motion sensors were triggered at 10:55, 11:01, 11:02 and 11:03. Then, we obtain measurements from temperature and humidity sensors only for Bedroom 1, the bathroom, and the kitchen. Before the event (from 10:30 to 10:58), the temperature in Bedroom 1, the kitchen, and the bathroom was around 18°C. After that, the temperature increases in Bedroom 1 and the bathroom.

3.3.2. Scenario 2

Netatmo. Data from Netatmo were acquired in two ways: by downloading data directly from the account and also by sending a email request to Netatmo's data protection office.

From the first one, we obtain CSV and XLS files containing the measurements (temperature, humidity, CO_2) made by the smart weather station. The temperature increased from 32.9 °C to 88.7 °C in less than 5 min while the humidity decreased from 45%RH to 12%RH at the same time. The CO_2 concentration was multiplied by 6 (from 474 ppm to 2′737 ppm).

From the second one, we received data organized in 5 files (*device*, *events*, *home*, *measurements* and *user*). Unfortunately, all the records provided were related to events that happened before the fire. However, if we had been able to obtain the information for the period of interest, it would probably have been similar to the data obtained through the first way.

myStrom. Data from myStrom was acquired from the account by two ways: by downloading data associated with the graph displayed on the account and by sending a data request from the account.

[7] Nest Protect is a Google product

[8] https://support.google.com/googlenest/answer/9249081?hl=en

[9] information from Google Takeout

From the data linked to the graph, we obtain measurements of consumption; there are made each 30 s. The consumption was multiplied by 4 (from 507 W to around 2000 W) in 30 s

From the automatic request, we obtain information about power (Watt), energy (Ws), estimation of the cost and also temperature. The energy increased from around 7′000 Ws to 29′000 Ws (or J) in 24 s whereas the power decreases from around 900 W to 300 W. The temperature stays stable; this is probably due to the lack of synchronization of information between the device and the cloud.

4. Summary of research findings

In the previous section we presented the result of the analysis of the devices for each type of source. Combination of those traces provided valuable information for the reconstruction of the fire, notably for fire origin location, cause determination and fire development.

4.1. Scenarios overview

4.1.1. Scenario 1

In the first scenario, IoT devices did not play an active role in the ignition of the fire, however they were present in the environment and collected multiple traces becoming, therefore, digital witnesses. With the traces recovered from their ecosystems, we were able to answer two of the main questions in fire investigation: determination of the origin and fire development.

Fig. 6 presents data recovered from two evidentiary sources, notably the smartphone apps and the cloud requests, and compares them to the reference data we collected locally. It is evident from the comparison that data from the sources are complementary: on the smartphone we recover single datapoints for the last sensor states, but only for four sensors out of five in the timeframe of interest; data from the cloud on the other hand is extremely detailed, with an even higher sampling rate than our reference data, but is only available for two sensors out of the five, missing the living room and one bedroom. The evidentiary sources show motion detected in Bedroom 2 at 11:02:22 just at the same time as a sustained increase in temperature for both Bedroom 1 and the bathroom. This is followed shortly after by motion detection in the living room, detection of flames by the camera and emergency alarm by the Nest Protect. While the temperature in Bedroom 1 and the bathroom grow slowly up to a maximum of 69°C and 49°C respectively at 11:03:40, the only evidentiary datapoint for the living room indicates the temperature reached 130°C well before, at 11:02:57; the slightly higher temperature in Bedroom 1 might indicate an increased combustion nearby. Based on these elements it is possible to formulate the hypothesis that the fire started, or at least changed proportion, around 11:02:22; developing rapidly in the living room, whilst only slightly affecting the bedrooms or the bathroom. Looking backward the motion sensor in the living room was triggered 50 s before the moment the temperature started rising in Bedroom 1. The increase could be due to human presence or the beginning of smoke or flames; however, it is not accompanied by rising temperatures, nor smoke alarms, and as such might weigh in favor of human presence.

Comparing to reference data and the ground truth it is possible to see that the evidentiary elements are coherent. Indeed gasoline was spilled at 11:01 in the corridor and especially in the living room, and was ignited at 11:02:05 and reference data shows an extremely rapid rise in temperature in the living room. The final temperature of Bedroom 1, higher than in the Bedroom 2 and the bathroom can be explained by heat release rate of the combustion of gasoline spill in the corridor, occurring nearby the sensor.

While in this case origin determination is limited to a room, which was also possible by the study of burn patterns, the information is nonetheless useful as it could help establish it even if the apartment burned to a higher extent or even completely. Additionally, traces from the IoT ecosystems convey temporal information that could be extremely useful to the investigator; it allows to pinpoint the start of the fire almost to the exact moment, with incertitude in the order of seconds.

4.1.2. Scenario 2

In the second scenario, IoT devices played both an active and witness role, and the combination of the traces collected helps answer two of the main questions: fire cause and development.

In Fig. 7 it is possible to observe the temperature information recorded by the Netatmo Weather Station and the myStrom Smart Outlet, as well as a combination of the specialized measures of each, respectively CO_2 concentration and power consumption. From this evidentiary data it is clear that the fire must have started at or sometime before 14:12:19, the first datapoint indicating a higher temperature in the room. The sharp increase in CO_2 concentration, passing from an average 500 ppm until 14:08:04 to 2′737 ppm at 14:13:11 suggests the fire did not experience a slow buildup but instead grew rapidly after ignition; this could support the exclusion of the hypothetical scenario of a glowing combustion evolving in flame combustion, as would occur with a discarded cigarette left on a sofa. The information collected from the myStrom ecosystem moreover shows a sharp increase in electricity usage at 14:06:30, which almost as quickly drops and stops recording. This suggests an electrical appliance turned on at that time resulting in the increased power draw of the socket. Given the proximity to the temperature increase in the room, the electrical appliance could have been the cause of a fire, either by being defective, or by being rigged to a ignition mechanism. Analysis of the smartphone of the owner shows traces of an HTTP request at 14:06:23 turning on a smart socket categorized as “Kitchen Socket”. This points to a deliberate remote activation of the socket which can be used to explain the alibi of the owner saying he was not at home when the incident happened.

Fire investigators that trained on the scene as part of this project recognized the origin zone and the rests of the immersion heater and the smart socket as being “Some kind of device” but could not explain whether it was a timer, just a socket or anything else. Analysis of the smartphone therefore helps both with understanding the traces observed on scene and the mechanics of the fire, and weights greatly in favor of arson instead of device malfunction.

4.2. Forensic considerations

These controlled case studies provide several outcomes that are of interest to both crime scene, fire investigation and IoT forensic fields of research.

4.2.1. Crime scene investigation

IoT devices are already known to be possible witnesses in criminal cases [12,14], but this research study tends to extend this new paradigm to current smart homes, where numerous IoT devices are present in the same place. The combination of traces from several devices provides a refined understanding of the crime scene. Time stamped digital traces also permit activity dating which is quite challenging with most of the traditional traces.

4.2.2. Fire investigation

The potential of IoT related traces is even more heightened when it comes to fire scenes. As fire produced perturbations in the

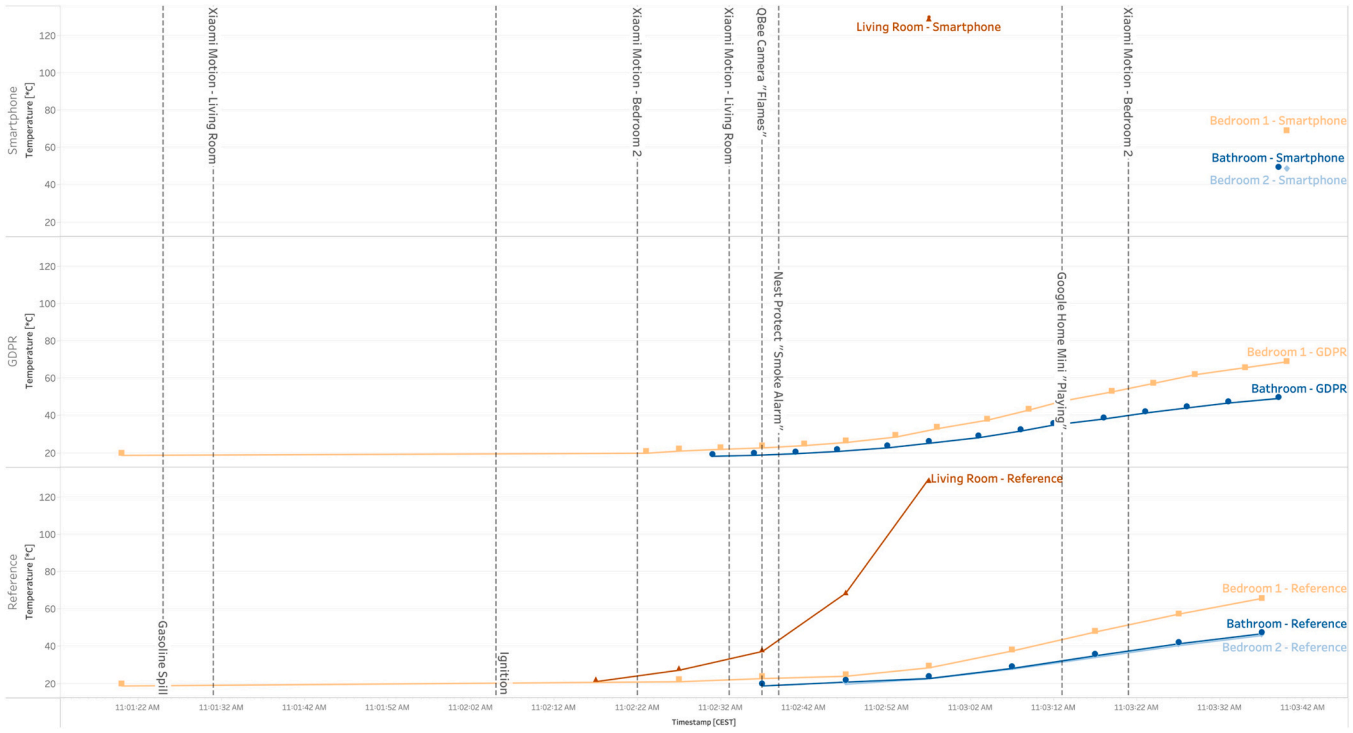

Fig. 6. Combination of traces for Scenario 1.

environment it takes place (temperature, atmosphere…) and alterations of electrical appliances, it inevitably produces sensor signals that are traces of the event. Unlike most traditional traces, digital traces appear to survive fires, due to the hardware shield case or, most of the time, due to their remote storage. As already underlined, even if the hardware is fully destroyed, IoT related traces can provide valuable information to support fire investigations, in the location of the fire origin, the clarification of the cause, or the explanation of the fire development.

4.2.3. IoT

By introducing our method to mimic Law Enforcement Agencies (LEA) lawful requests by using data privacy ones, this study also contributes to improving IoT-related research.

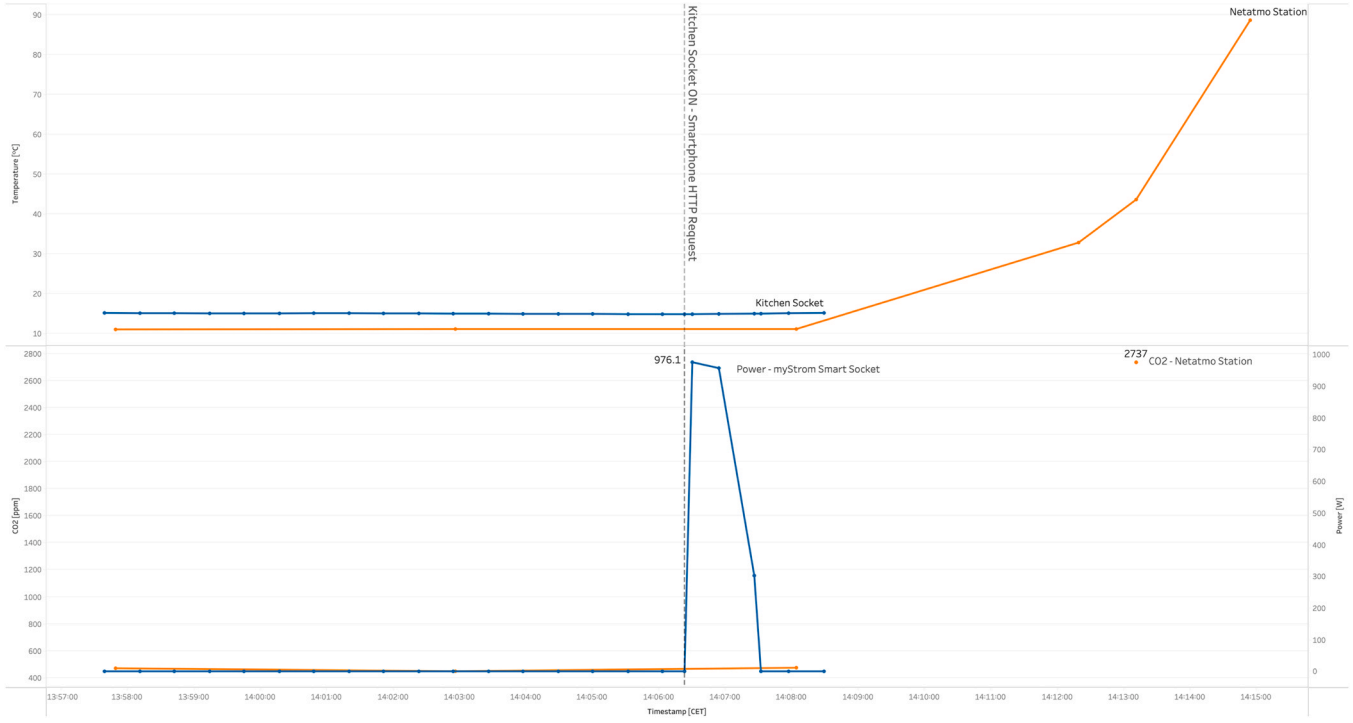

Fig. 7. Combination of traces for Scenario 2.

Actually, an important challenge in IoT forensics research consists in having access to remote data in the same way that a judiciary body would have. To do so, currently researchers can only rely on real case data provided to them by LEAs. This reactive contribution is due to the fact that judiciary lawful requests are, by their own nature, strictly limited to real case processing. To overcome this limit, we are currently conducting research to measure if data provided through privacy lawful requests (right of access) are and/or should be different from judiciary lawful requests. Using this mechanism, IoT researchers can move from a reactive position to a proactive one.

4.3. Limits

While IoT-related traces analysis looks extremely promising for crime scene investigation in general, and fire investigation in particular, they suffer from several drawbacks that remain to be considered.

4.3.1. Availability

First of all, considering private homes (as opposed as public buildings), only a few real scenes may present a large number of IoT devices, that will generate the expected traces. But these studies also show that just a few IoT devices (ie. scenario 2) can already provide valuable information to answer the investigation questions (what, when, how). Even when devices are present on scene, access to the data is not evident; analysis of physical devices often yields only limited data, with traces from the smartphone application or the cloud being much more exploitable. However even information from the cloud proved to be incomplete or difficult to access, with sometimes different results depending on the request method; as seen in Scenario 1, data for at least the last temperature of each device exists in the cloud, as it could be downloaded using the application. However, this same data was not included in the copy received through privacy requests. Responsiveness to the requests was also a limiting factor, as in multiple instances we had to repeat or use multiple contact channels before obtaining a response. It also meant we had no explanation as to the reason for the missing data from Xiaomi for example as no response was ever received to our manual requests.

The study of the difference between data stored on the cloud, and what is returned following either LEA requests or data privacy requests, is currently the subject of ongoing research.

4.3.2. Heterogeneity

Another limit related to IoT on crime scenes is the heterogeneity of the devices. The fragmentation of the market causes first responders and CSI specialists to have to recognize unusual devices they may never have met before. On fire investigations, this task is further hampered by the degraded state the objects might be found in.

Asking the property owners whether they had any IoT devices is extremely important, as they might not even be aware of all the data available. More than ever it might be pertinent to perform forensic analysis of digital devices, especially smartphones, to check whether IoT devices were installed in the home and might hold relevant data.

The same issue occurs when it comes to process the hardware or get the data from the cloud; specialists often have to start from zero as the IoT system was never before encountered, nor described in the literature.

4.3.3. Validation

Data from IoT devices is sometimes difficult to interpret or might be erroneous. Extensive tests might be needed to understand the traces collected, especially if extracted in raw form from the physical devices. However even information of easier interpretation, such as structured data obtained from the cloud must be evaluated with care; indeed IoT devices might not behave the expected way, especially during situations such as fire incidents where they might be operating extremely outside “working temperatures” and/or experiencing sudden disconnection from the cloud. In scenario 1 it was possible to observe that the Nest Protect did not record any information about CO and temperature after 11:00, but a record exist at 14:00 when the device was clearly destroyed and disconnected. In scenario 2 the temperature recorded by the smart socket in the kitchen remains stable even as the fire already started, and might suggest the socket was further away from the fire than it really was.

4.3.4. Time synchronization

This work underscores the usefulness of IoT devices in providing an additional type of traces to fire investigations, notably timestamp information. However, for time information to be useful in constructing timelines the information must be uniform and synchronized. In traditional investigations one can check the time of a device for possible time drift at the acquisition step and account for it during analysis [2,6]. With IoT devices, however, this is more complex as no easy way is present to check for the current time on either the device or the remote server. In addition to ensuring timezone uniformity, an investigator must therefore take into account the possibility that the timestamps reported by a device are skewed in relation to UTC, but also potentially skewed between each source.

Generating traces with known timestamps for each ecosystem/device under analysis might help with the interpretation of timestamp-related traces, but involves extensive work for the investigator and might be reserved only for cases where the exact order of events is extremely important.

In our study, we standardized all timestamps to local time (CEST/CET depending on the scenario) and assumed that time between all traces collected from the cloud/smartphone was synchronized. This was not however verified and may have lead to some events being reported in a different order. However, it should be safe to assume that for a given source (eg. Xiaomi Ecosystem) all related traces are synchronized between them.

4.3.5. Knowledge

Above all, no references (articles, nor even blog posts) were found on how to conduct forensic analyses, for most of the IoT devices placed in the scenarios. The IoT forensics literature is growing but many articles are related to process and security models. The few articles related to IoT forensic technical analysis have to face the market fragmentation syndrome exposed above and provide trace exploitation techniques related to single items, often outdated. Moreover, only rare articles are considering cloud traces as potential authors have to face remote data collection issues, which we overcame with our privacy lawful request mechanism.

Then, we have to conduct our own forensic examinations, often using reference devices, but we were not proactive in knowing the trace potential of each IoT device before processing them. For that reason, we have potentially lost some traces due to retention delays that are often present with remote data and of which we were not aware.

These observations confirm the need for knowledge sharing in IoT trace processing in order to proactively cover the fragmented IoT market. The current IoT forensics state is comparable to the mobile forensics one in early 2000′s where proprietary operating/file systems were used by numerous phone manufacturers and no major acquisition/analysis tool was in place. Nowadays challenges are even more complex as specialists also have to face encryption, dispersed traces, retention delays, and legal restrictions to access remote data.

5. Conclusion and future work

Through two controlled large-scale experiments of arson, this work highlighted the usefulness of IoT devices as potential digital witnesses and novel sources of traces in fire investigations. The traces obtained from the IoT ecosystems, provided useful information in answering the main investigative questions: origin, cause, and development of the fire. Those traces were preserved even when the devices were completely destroyed and no data could be directly acquired; outlining how IoT devices are contributing to an expansion of the field of investigation for fire incidents, from the physical scene to remote servers. Additionally, we demonstrated that data can be recovered even from heavily damaged devices, although due to the nature of IoT devices was limited to identifying information.

This research also outlines how IoT devices might be misused by criminals, and the importance, even in fire investigations, of increased awareness of their existence and capabilities.

This work is part of more general research that studies how digital traces, and IoT devices in particular, may change the investigative established processes, starting with crime scene investigation ones. Current outcomes include the presented mechanism to simulate a judiciary lawful request by submitting a privacy request directly to a cloud service provider. Through this study, we want to draw the attention of crime scene and fire forensic scientists and stimulate their awareness of the already increasing role of digital traces to support their investigations. We also want to highlight the paramount importance of recognizing these traces on the crime scene, of suitably recording and collecting them, and of evaluating their potential. There is a need to guide them from collection to interpretation and visualization.

CRediT authorship contribution statement

Francesco Servida: Conceptualization (equal), Investigation (equal), Visualization (lead), Writing – original draft preparation (lead), Writing – review & editing (lead). **Manon Fischer**: Conceptualization (equal), Investigation (equal), Visualization (supporting), Writing – original draft preparation (equal), Writing – review & editing (supporting). **Olivier Delémont**: Conceptualization (equal), Resources (equal), Writing – original draft preparation (supporting). **Thomas Souvignet**: Conceptualization (equal), Resources (equal), Supervision, Writing – original draft preparation (equal), Writing – review & editing (supporting).

Acknowledgments

Authors would like to thank the fire investigation team of the School of Criminal Justice of the University of Lausanne and the forensic service of the Police Neuchateloise for organizing the experiments, the opportunity to participate, as well as the useful insights and discussions without which this paper would not have been possible. We'd also like to acknowledge the firefighting teams from SDIS Valtra and SDIS EPFL who worked on the scenes and contributed to the realization of this research.

References

[1] K. Ashton, That 'internet of things' thing, RFID J. 22 (7) (2009) 97–114 (Publisher: Hauppauge, New York).
[2] C. Boyd, P. Forster, Time and date issues in forensic computing—a case study, Digit. Investig. 1 (1) (2004) 18–23.
[3] Carey, N. and Shepardson, D. (2021). Tesla's driving data decoded by Dutch forensic lab.
[4] Casey, E., Spichiger, H., Ryser, E., Servida, F., and Jaquet-Chiffelle, D.-O. (2020). IoT forensic science: principles, processes, and activities. In: Advances in Information Security, Privacy, and Ethics, 1–37. IGI Global.
[5] J.D. DeHaan, D.J. Icove, Kirk's Fire Investigation, seventh ed, Pearson, 2011.
[6] ISO/IEC (2012). 27037:2012 - Information technology — Security techniques — Guidelines for identification, collection, acquisition, and preservation of digital evidence. Technical Report 1, ISO/IEC.
[7] Kebande, V.R., Karie, N.M., Michael, A., Malapane, S.M., and Venter, H. (2017). How an IoT-enabled "smart refrigerator" can play a clandestine role in perpetuating cyber-crime. In: 2017 IST-Africa Week Conference (IST-Africa), 1–10. IEEE.
[8] J.-C. Martin, Incendies et explosions d'atmosphère, first ed, EPFL Press, 2008.
[9] Meffert, C., Clark, D., Baggili, I., and Breitinger, F. (2017). Forensic State Acquisition from Internet of Things (FSAIoT): a general framework and practical approach for IoT forensics through IoT device state acquisition. In: Proceedings of the 12th International Conference on Availability, Reliability and Security (Ares 2017), NEW YORK, NY, USA. Assoc Computing Machinery.
[10] Q. Milliet, O. Delémont, P. Margot, A forensic science perspective on the role of images in crime investigation and reconstruction, Sci. Justice 54 (6) (2014) 470–480 (Publisher: Elsevier).
[11] Oriwoh, E., Jazani, D., Epiphaniou, G., and Sant, P. (2013). Internet of Things forensics: challenges and approaches. In: Proceedings of the 9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. ICST.
[12] B. Oswald, Regensburger Gericht lädt Alexa als Zeugin, Bayer. Rundfunk (2020).
[13] F. Servida, E. Casey, IoT forensic challenges and opportunities for digital traces, Digit. Investig. 28 (2019) S22–S29.
[14] J. Völkerling, Killer-Stalker muss zehn Jahre in den Knast, Bild (2020).