Digital scene of crime: technique of profiling users*

Clara Colombini†
External researcher at University of Milan
Milan, Italy
cmcolombini@email.it

Antonio Colella‡
Italian Army
Rome, Italy
colella@acm.org

Abstract

Nowadays the investigations are becoming more difficult than in the past due to the complexity of the scene of crime and the implication that the technology has in this new environment. This article describes a new technique called "Digital Profiling". The technique is an investigative method of computer forensics that offers a new prospective for analyzing digital memories of electronic devices. It works applying the traditional techniques of Criminal Profiling and Intelligence to the electronic devices in order to obtain information for reconstructing the users' identity. The process starts by researching and analyzing the information gathered from the "digital footprints of users" discovered on the device (for example a personal computer). Actually, although a computer is a machine, its user is an human being that customizes all the environment around her or him. Moreover, the user cannot avoid to leave on the device, even unconsciously, some evidence that can be detected, recognized and compared. The method applied by the Digital Profiling is based on a mathematical principle: the Set Theory, and can be addressed to any electronic device, such as personal computers, mobile phones, storage areas, and so forth. This particular type of analysis may be very helpful in various investigative fields, related to crimes that involve a device in which it is necessary to analyze a digital memory in order to identify a potential criminal. It is possible to reach positive results through this technique especially in operations against organized crime, anti-terrorism and intelligence operations.

Keywords: Hacking profiling, modus operandi, data mining, criminal behaviour, hackers signature.

1 Introduction

Due to development of technology, digital devices are changing from data containers to a sort of "digital diaries" where it is impossible not to leave evidence of our lifestyle, even if unintentionally [2]. Chat rooms, blogs, forums, social networks, etc..., are now the new keeper of our personal habits, with the feature of easily allowing the structuring and mining of data [3]. All these devices are part of scene of a crime, now virtual, which consists of most devices that can contain a digital memory and so evidence useful to reconstruct the crime generation. To meet these new requirements, in this work, we suggest a methodology based on the analysis of all those files that contain data about the users in order to provide investigators with a useful tool for identifying the author of a crime. The technique is particularly valid in all the cases in which it is difficult to match a digital device to a specific user, especially when we have a single device with several users. The analysis of Digital Profiling (DP) includes a cycle of six steps describing a process starting from research and analysis of all information that can be gathered from

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, volume: 3, number: 3, pp. 50-73

*This paper is an extended version of the work originally presented at the 6th International Conference on Availability, Reliability and Security (ARES'11), Vienna, Austria, August 2011[1].

†Digital Forensics Consultant for the Italian Prosecutor's Office, Italy and member of IISFA Italian Chapter, International Information System Forensics Association (http://www.iisfa.it).

‡Corresponding author: Criminologist and Computer Forensics Expert, Lecturer at Master of Art in Forensics Science, University of Rome La Sapienza and member of IISFA Italian Chapter, International Information System Forensics Association (http://www.iisfa.it), Via Boccapaduli 9 - 00137 ROMA- ITALY, Tel: 00390647357388, Email: colella@acm.org

"digital footprints" left on a Personal Computer (PC) and ends with comparison of behavioural models and with extrapolation of each user's profile. The goal is to extrapolate behavioural patterns and compare them in order to outline a user's profile. The work concludes with the definition of the profile obtained in this manner that it also outlines his or her modus operandi [4].

2 From Traditional to Digital Profiling

The application of investigative methods including profiling techniques in the cyberspace is not easy, especially in terms of appropriate investigation method. The main causes can be summarized as follow:

- inappropriate and incomplete documentation on this subject;
- difficulties to combine the human nature to computer science;
- manifested distrust towards traditional criminal profiling and in general psychological investigations.

To better clarify the differences between Digital and Traditional Profiling in Table 1 has been reported the main parallelisms with profiling model of Douglas, Ressler, Burgess, Hartman, and Digital Profiling. Before to show the use of computer science in support of investigation, we summarize in the Table 2 the possible fields in which Digital Profiling can be a valuable technique of forensics analysis. The table contains just an example of possible use. Behaviour of offender, as well, can be similar to everyday, but it can also be unique to the individual in question, and occur only sporadically. If there are repeated crime scenes (as with a serial or repeat offender), it is much more likely, with proper examination, that any unique behaviour, need, and pattern will be uncovered.

Three elements link crimes in a series:

- method of operation (modus operandi);
- ritual (signs of fantasy or psychological need);
- signature (unique combinations of behaviours).

With regard to Modus Operandi (MO), Douglas & Olshaker define it as "what an offender has to do to accomplish a crime". MO contains at minimum the following elements:

- ensure success of the crime;
- identity protection;
- escape effect.

According to Keppel, before the 1800s the expression "modus operandi" was considered a description of animal behaviour. Only after this period, when the term started to appearing in English utilitarian literature, "modus operandi" indicated a description of human behaviour [5].

In Criminology the following definition was put forward: "Modus operandi is the principle for which criminal is likely to use the same technique repeatedly, and any analysis or record of that technique used in every serious crime will provide a means of identification in a particular crime" [6]. This definition can be applied to cybercrime as well as it is possible to identify ritual and signature elements as for traditional crimes [7]. Ritual is a behaviour that exceeds the means necessary to commit the crime. By definition, it is a subtype of signature sometimes called "ritual signature". According to this definition and the Crime Classification Manual [8] ritual can be apply to cybercrime without more difficulties even

Profiling Phase	Traditional Profiling	Digital Profiling
Profiling Input	Data acquisition and information on the crime, snap-shots and testimony of subject involved.	Data related to structures and system architecture, Incident Response procedures and Computer Forensics acquisition data. Other data related to physical and infrastructure aspects.
Decision Process Model	Organization of information gained through pre-established classification schemes and questions relevant to the case.	Collection and data entry in the log file analysis software and databases, data processing, categorization, creating a data model appropriate to the characteristics of computer investigation.
Crime Assessment	Behavioural reconstruction of criminal and victim.	Assessment of the characteristics of computer systems involved, methodologies and tools used for the crime and subsequent impact. Analysis of possible connections and sociopolitical characteristics. Extracting behavioural data RPE (Reverse Engineering Profile).
Criminal Profiling	Development of an initial profile based on information from previous steps. Comparison of each hypothesis with data on phase two. Inductive analysis and processing through historical data.	Link Analysis, Data Mining, development of informatics and psychological connections.
Investigation	Elaboration of the profile assuming the counterparty investigation by comparison with the suspects. Any other data emerging from the phase of investigation will be used to update the profile.	Rationalization and skimming of the links obtained. Deepening connections and elements of the previous phase. Processing of behavioural data and use in the investigation stage. Possible report feedback.

Table 1: Comparison between traditional profiling and digital profiling

though more of hackers behaviour can be ritual. On the other hand, criminology signature concept better fits hackers world. In general, signature is a combination of behaviours. Douglas & Olshaker define it as “something the offender has to do to fulfill himself emotionally ... it is not needed to successfully accomplish a crime, but it may be the reason he undertakes the particular crime in the first place”. Signature, in a hacker behaviour, is a sort of “trademark” and reflects a compulsion on the part of criminals to go beyond just committing the crime to “express themselves”, reflecting in some way their personality. In a defacing attack, for instance, this aspect is more evident than in others because the acting of hack is visible to everyone. Anyway the motivations, actions, and modus operandi of traditional crimes respect to cybercrimes are different. For example, it appears that as of 2009, we have entered a new era where organized cybercriminals can operate identity theft resale operations, as well as engaged in cyber-

war. The approach to hacking as multi-stage process leads to individuate three main phases according to the bestseller Hacking Exposed (now in its 6th edition 2010). These stages are: casing, scanning, and enumeration. It is not the aim of this paper describe each of these stages but the same considerations can arise. For example, the time of action can change from 48/72 hours of constantly working during a network intrusion, to a longer period of time like in an action of pedophiles [9]. According to Richard Stiennon of IT Harvest Inc., hackers have found ways to streamline the efficiency of the classic methodology. In particular, the most recent development has been the use of viruses and trojans as part of the modus operandi. The Figure 1 shows the new anatomy of the hacking and as the modus operandi of hackers has change in the last few years [10].

Main Area of Applicability & Possible Contribution
Incident Response & understand the type of attack delimitation of investigation area batter exploitation of insiders threat data finding understand social engineering technique
Computer Forensics & finalize investigation investigate on data hiding understand use of anti forensics password guess

Table 2: Main area of applicability of Digital Profiling

The difference is that the "new" method uses a virus or trojan that is either custom made or standard and has the same effect as if anyone had been put into the target system to install a keylogger on the computer. The new method is considered easier than the old one. This last process best is suited to new content of World Wide Web (WWW), where the programming language has become more sophisticated and dynamic, especially in the new meeting places like the Networked Virtual Environments (NVEs).

3 Techniques of digital profiling analysis of a computer system

The process of digital profiling that has been developed in this research includes a cycle of six steps:

(1) identify the goal: what to look for in relation with the type of problem;

(2) collect and assess targeted data from mass memory that contain useful information;

(3) selection of relevant information and extraction of *indicators* for each analysed area;

(4) information matching of data *(indicators*) in order to identify deficiencies, discrepancies or similarities;

(5) collection of indicators shown by the comparison and build a *"digital profile*";

(6) analysis of the resulted profile in comparison to the initial goal.

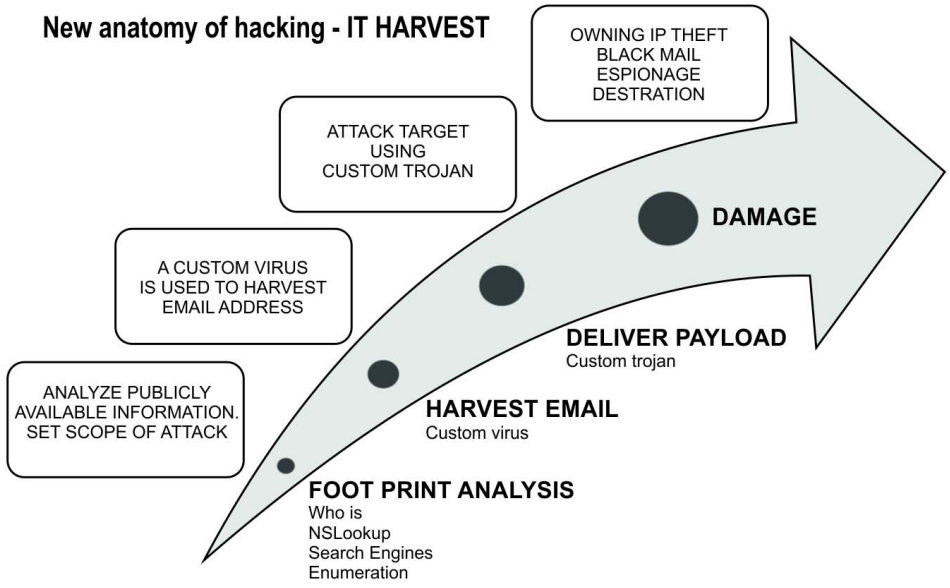


Figure 1: New anatomy of hacking –IT HARVEST

3.1 Research Areas inside a computer system. Where?

The following paragraph provides a list of the main areas from which it possible to extract useful indicators and portrait a draft digital profile. For simplicity, here it has been considered a case study in which the computer target is based on Microsoft Windows Operation System (OS).

3.1.1 User analysis

The Windows installation is recorded in a Registry file that contains date of the installation, user names/organization and serial numbers. From this file it is possible to obtain information about users who have access to the PC, such as date installation, the number of hits, the last access, last change of password etc...

Task: Extrapolation from the files “SAM” and “NTUSER” of all the information about who and how many people had access to the machine.

3.1.2 The analysis of the text file

Everyone has his own style of writing, using specific idioms, the same mistakes of syntax or[1] grammar, spelling, that are the so-called “signatures” that distinguish it.

Task: Research and analysis of all the files containing text written by the user: e-mail messages, chat conversations, notes, documents, etc... in order to extract the so-called “signature” characteristics, which can be compared and recognized in other documents, not just digital.

[1]Details of syntax and grammar errors will recognize whether the user is Italian or foreign.

3.1.3 The analysis of the personal file folders

User is used to making the same collections of music or photos, often placing them in a folder with the same name, often placed in the same position.

Task: Research and analysis of all the folders that contain: music, pictures, photographs, films, documents, publications, etc...

3.1.4 The analysis of the organization of folders

User is used to repeat the same pattern of organization of files and folders in order to recover them more quickly especially when he wants to move files from one machine to another.

Task: Research and extrapolation of the scheme of organization of folders and files.

3.1.5 The analysis of the nickname

The nickname is used to access to instant messaging, blogs, forums, social networks, in addition to be used for e-mail addresses. If repeated on different devices can be easily recognized.

Task: Research and extrapolation of the e-mail addresses, Instant Messenge(IM) screen names, blogs, social networks.

3.1.6 The analysis of log files and the history of connections

The usual browsing on specific Web sites, such as forums, web mail, ftp connections, require user account often made of the same logon user name and password. All this data are detectable and recognizable when repeated on different devices.

Task: Extrapolation of the files that contain the history of websites visited regularly (URL access webmail, forums, blogs, online access to current accounts, FTP, etc.).

3.1.7 The analysis of hardware facilities

The user that utilises more than a PC, often is used to connect to them the same devices: USB memory sticks, cell phones, cameras, MP3 players and so on. It is possible to find evidence examining log file of all these devices and compare the names, serial numbers, etc...

Task: Search and extrapolation from the log files of the information on which devices (printers, external drives, USB memory sticks, cameras , mobile phones etc...) have been installed on the machine in order to detect date of installation, brand, type of hardware, serial numbers).

3.1.8 The analysis of software installations

As for the hardware, the user who utilizes more than one PC, often gets the same software, which left more evidence (name, version, serial number) in the log files.

Task: Search and extrapolation from the log files what software applications are installed on the machine in order to detect installation date, version and serial number.

3.1.9 The analysis of the code listings

Every programmer has his own personal style: the choice of the functions that he best knows, and in what way he use them, without forgetting his comments into the lines of code, a sort of signature that distinguish each programmer from others.

Task: Research and extrapolation of the files containing code lists.

From the analysis of items 3.1.8 and 3.1.9 is also possible to obtain information on knowledge and skills of the user's machine (e.g. the presence of programming tools, multiple operating systems, virtual machines etc...).

3.1.10 The analysis of the timeline

Same log files provide useful information about date and time (even time gaps). These logs are either important to detect information about the time of machine on and off. They are, as well, useful to know about files/folders operations, internet access and e-mail, etc... [11]

Task: Analysis of the timeline of operations about files in memory, depending on the time frame in which the crime in question was made (determination of a possible alibi).

3.1.11 The analysis in virtual machine

"*Do not judge a person unless you first walked for 5 moons in his moccasins*"(called Apache).

Task: Loading the image memory in a virtual machine in order to browse the contents of the computer as well as the user sees: this allows to verify, among other things, the automatic execution of startup applications, automation of the updates, the provision of desktop, etc...

3.1.12 The analysis of the "modus operandi"

To go back, for example, the identity of the perpetrators of a cybercrime (e.g. phishing - attacks on servers - etc...), it is useful to make the reconstruction and subsequent analysis of the modus operandi. To do this, the following information can be obtained from analysis of data:

- target of the attack (fraud, interruption of service, political attack etc..);
- tools and techniques used for intrusion (rootkits, shells, worms, social engineering, etc. .);
- technical skills used;
- possible correlation with measures of social engineering;
- time chosen for the attack (day / night, intra / end of week, etc..);
- duration of the attack and possible frequency (single or fragmented in predetermined time intervals, etc..);
- correlation of the moment chosen for the attack with external events;
- choices of victims (Italian or foreign government institution, bank, commercial organization etc.);
- typology of anti-forensic techniques used;
- achievement of goal.

3.2 The method

The Digital Profiling is based on a method which includes mining, comparison and recognition of digital profiles of a user's digital device. Identification is done through the comparison of a digital first profile taken from a PC certainty attached to a known subject and the profiles extracted from other digital devices with which those crimes were committed, but cannot be attributed with certainty to the subject. It should be noted that the principle upon which the method is based, is two-way, that is, it can also start from user's digital profile "anonymous" of the device, for comparison with profiles of other devices (also not involved in the offense) attributed with certainty to particular subjects [12]. It can also extract a digital profile of a "modus operandi" (e.g. cyber-attack) to compare with others in order to recognize and identify the author. The method comprises the following steps which describe a cycle that can be repeated whenever new information is added:

- extrapolation of the digital profile of the user (or users) of the device chosen as "standard profile";
- extrapolation of the profiles of users of digital devices in any other analysis;
- comparison of the profiles obtained in order to highlight convergences-divergences;
- quantitative and qualitative analysis of the convergences-divergences between the profiles for granting and then identification of the subject;

3.3 The model

The creation of the model starts from the study of information characterizing the detected files on a PC, chosen for the memory capacity, and for the high degree of customization allowed by the many applications available. [13]. It should be noted here that we construct a profile for each users found for each operating system installed on the device, including virtual machines.
The model describes:

- the elements;
- the profiles;
- the features and functions of the elements;
- the sequence of operations to create the digital profile;
- the comparison;
- the evaluation of the result.

3.4 The characteristics and functions of the elements

3.4.1 D - digital device

For digital device "D_i" is intended:

- **any digital device** provided with permanent memory capable of storing files. Example: PC, mobile phone, navigation system, telephone exchange, etc. . . ;
- **a storage device** capable of storing data. Example: Hard drive, flash card, memory card, smart card, USB pen, CD, DVD, DAT, etc...);

- **area of memory** where they are stored in remote data files created by users. Example: Stock in remote;
- **virtual machine** containing an operating system;
- **set of files** about the accesses. Example: log file.

3.4.2 f - feature

For feature "f_i" is intended the single basic hardware or software feature, not further decomposable into more elementary analysed in the context of the study, because it contains information that describes the "digital behaviour" of the device user. The feature is derived from the files stored inside the device and selected on the basis of objective investigation. It may consist of:

- file properties (metadata type);
- content of the file (type of information).

A file may contain one or more feature: they are considered basic features, depending on the purpose of the investigation:

- Filename. Example: the discovery of the same names of personal files (texts, photographs, music, movies, videos, etc..) in a variety of devices can be inferred that they have been stored by the same user;
- Path. Example: some files seem identical so that, this feature indicates that this file has the same location in the folder tree with respect to another (same folder name or set of folders);
- MD5 (or other hash algorithm). The feature provides the mathematical certainty of the coincidence of the contents of the same file found on various devices;
- Date of creation, modification, cancellation. These three features provide a history of saving, editing, deleting the same files found on other devices;
- Any type of information relevant to the target can be taken from its content.

3.4.3 A - Area of file

In order to better identify, within the memory of the device D_i, the files that may contain features, they were divided according to type in specific areas, defined as "$\mathbf{A}_i$" (D_i)

$$\bigcup_i A_i(D_i) \subset D_i$$

It then defines $A_i(D_i)$ as the homogeneous subset of D_i that contains, divided by type, all files that may contain features, relative to the device D_i.

3.4.4 Classification of areas of file A

Each device has its own specific classification of the areas containing features, according to its specific characteristics and installed applications. Here we find a generic classification of the basic areas related to Personal Computer. The number of the research areas of the feature is flexible, because it depends on the target of the research and the applications available on the device.

(1) $\mathbf{A}_1$ - Registry File: system users.

(2) $\mathbf{A}_2$ - Registry File: hardware installations.

(3) $\mathbf{A}_3$ - Registry File: software installations.

(4) The areas of personal files: here are considered "personal files" all those files stored on the user device outside of installed programs, and which may contain information characterizing the "digital behaviour" of the users. The feature can be obtained either from the metadata describing the files itself. The area of personal files has been divided by type of file in the following categories: A_4 - text file personal Text files written by the user's hand (DOC, DOCX, TXT, RTF, ODT, PDF, XLS, etc...). reveal the writing style. Their analysis can highlight several features. In addition to information that can provide through metadata analysis, other features can be detected:

- signature;
- nickname;
- proper name;
- password to access;
- idiom;
- misspelling;
- typing mistakes;
- reference to a specific event;
- reference to a particular person;
- reference to a given object;
- reference to a place;
- particular phrase;
- email address;
- Etc.

$\mathbf{A}_5$ - Personal Email messages (except for newsletters, advertising, etc.).

$\mathbf{A}_6$ – Chats.

$\mathbf{A}_7$ - Images ((BMP, JPG, TIF, etc.).

They are particular payable with regard to the photographs taken from cameras, cell phones, etc.

$\mathbf{A}_8$ - Graphic images (JPG, TIF, DWG, etc.).

For instance, collections of graphic images, such as DVD covers, CD, thematic collections of pictures, art, comics, etc...

$\mathbf{A}_9$ - Movies video (MPG, AVI., etc..)

The movies made from video cameras, cell phones, etc... are mainly important.

$\mathbf{A}_{10}$ - Audio files (WAV, MP3, etc.).

The collections of audio files stored by the user are particular payable.

$\mathbf{A}_{11}$ – URL

3.4.5 F- collection of features

From the analysis of the different areas, it is collected a set of basic features. As set of Feature F, however, it is intend a set of all the individual background characteristics analysed in a digital device.

$$F = \{f_1(A_i)(D_i), f_2(A_i)(D_i), \ldots f_n(A_i)\,(D_i)\}$$

3.4.6 m -minimum feature

Once the set of the maximum possible features removable from the device is fixed, it must be reduced to the features actually present on the device under analysis, according to the specific requirements of the investigation. The action is made on the initial selection of features, for a particular device, to restrict the number in order to form the minimum set of features.
For "$\mathbf{m}_i$" is intended, therefore, a consistent feature, which belongs to all the basic features, selected in relation with the specific investigation.

$$m_i(A_i)(D_i) \in F(D_i)$$

The name of this minimal feature is therefore:
$\mathbf{m}_i(\mathbf{A}_i)(\mathbf{D}_i)$ where:
$\mathbf{m}_i$ - identifies the minimum feature;
$\mathbf{A}_i$ - identifies the area belonging to the source file;
$\mathbf{D}_i$ - identifies the digital device from which it was extracted.

3.4.7 M - minimum set of features

It is defined as the minimum set of features a subset of $S(D_i)$ of size at least in relation to the individual case investigation.

$$M(D_i) \in F(D_i)$$

$$M(D_i) = \{m_1(A_i)(D_i), m_2(A_i)(D_i), \ldots m_n(A_i)(D_i)\}$$

The set of features is the minimum set of filters to be applied to the files for the extraction of characteristic information *(indicators)* that will make the digital profile.

3.4.8 i - Indicator

The indicator represents the single featuring information collected and analysed in the context of study for the purpose of profiling. It is obtained from the files selected by the application of minimum features as filters during the operation of creating the digital profile. It is defined as $\mathbf{i}_i(\mathbf{l}_i)(\mathbf{A}_i)(\mathbf{D}_i)$, which represents the information detected by applying, as a filter to collections of files, the feature minimum, $\mathbf{m}_i$ on a specific area $(\mathbf{A}_i)$ of a specific device $(\mathbf{D}_i)$. $(\mathbf{l}_i)$ identifies the file from which the indicator has been extracted.
The indicator is, in effect, a *digital evidence*, and as such can be:

- detected;
- compared;
- recognized.

3.4.9 I - set of Indicators

It is defined as the set of indicators **I(D$_i$)**:

$$I(D_i) = \{i_1(l_i)(A_i)(D_i), i_2(l_i)(A_i)(D_i) \ldots i_n(l_i)(A_i)(D_i)\}$$

The set of indicators that characterize all the information is collected from the files.
It describes the "*digital behaviour*" of the user of the device under analysis.

3.4.10 k - file that contains Indicators

k$_i$(A$_i$)(D$_i$) uniquely identifies every file that contains one or more indicators, when:

- **(A$_i$)** identifies the area where you found the file;
- **(D$_i$)** identifies the device.

The file that contains one or more indicators is the *"source of digital evidence*" confirming the source of the indicator, that is the information of interest extracted from it (see Table 3).

3.4.11 K - set of files containing the Indicators

K(D$_i$) defines the set of files that contain information related to a specific device (D$_i$).

$$K(D_i) = \{k_1(A_i)(D_i), k_2(A_i)(D_i) \ldots k_n(A_i)(D_i)\}$$

3.5 The sequence of operations useful for the creation of digital profile

The sequence of operations includes the following extrapolation of five profiles from a PC:

- the profile obtained from the log files;
- the profile obtained from the files in user folder;
- the profile obtained from the files in the remaining areas of memory.

From them the followings are derived:

- the user profile, formed by their union;
- the sample profile, which coincides with the user profile, but refers to a device selected as the "*sample*" for the comparison with others.

From the sample profile the followings are drawn:

- the indicators, or the information characterizing to be used for comparison with other profiles for user identification;
- the files containing them (test wells).

Bearing in mind that a PC can detect the presence of multiple users, the following explanation of the method, presents an example of the digital profile of a personal computer attached to a single user in relation to a Microsoft Windows Operating System.

3.5.1 Ps - System Profile

Starting point are the log files (A_1 area), providing all the information *(indicators)* about the user machine configuration.
They will form the system profile $\mathbf{Ps_i(D_i)}$, where $(\mathbf{D_i})$ identifies a specific device.

$$Ps_i(D_i) = I(Ps_i)(D_i) \cup K(Ps_i)(D_i)$$

where:

- the set of indicators collected from the log files is called $\mathbf{I(Ps_i)(D_i)}$ where:
  - **I** - identifies all the indicators measured;
  - $\mathbf{Ps}_i$ - identifies the specific system profile;
  - $\mathbf{D}_i$ - identifies the specific device.
- the set of files that contains them is called $\mathbf{K(Ps_i)(D_i)}$ where:
  - **K** - identifies the set of files;
  - $\mathbf{Ps}_i$ - identifies the specific system profile;
  - $\mathbf{D}_i$ - identifies the specific device.

in which:

- each indicator consists of a single piece of information is not further decomposable;
- each indicator refers to one or more files;
- each file can contain one or more indicators.

3.5.2 PC - User Folder Profile

The second step is the analysis of files stored in folders created for any user by the operating system. In fact, they contain the most "personalized" files made by the user. This action creates a profile called $\mathbf{Pc(D_i)}$ (user folder profile), based on the analysis of files in the folder created for the user on the device (D_i) by the Operating System. There is a folder for each user found in the PC memory. (e.g., PC OS Windows XP: c:\ Documents and Settings\ SuperPippo\ ...) If there are multiple operating systems (including OS contained in virtual machines), each of them should be treated as a separate device.
The user folder profile $\mathbf{Pc_i(D_i)}$ is defined as:

$$Pc_i(D_i) = I(Pc_i)(D_i) \cup K(Pc_i)(D_i)$$

where:

- $\mathbf{I(Pc_i)(D_i)}$ is the set of indicators collected by the files in your user folder, where:
  - $\mathbf{I}_i$ - identifies the set of collected indicators;
  - $\mathbf{Pc}_i$ - identifies the user folder profile;
  - $\mathbf{D}_i$ - identifies the device.
- $\mathbf{K(Pc_i)(D_i)}$ is the set of files that contains the indicators, in which:

- **K** - identifies the set of files;
- **Pc**$_i$ - identifies the user folder profile;
- **D**$_i$ - identifies the device.

in which:

- each indicator consists of a single piece of information not further decomposable;
- each indicator refers to one or more files;
- each file can contain one or more indicators.

3.5.3 Pd - Device Profile

The creation of the user folder profile is not sufficient to delineate the entire profile of the user's machine, since other features can be detected from files stored in areas not included in the generic users folders. The Device Profile includes those files contained, for example, into directories on other partitions, additional hard disks, including also deallocated files, etc... A second round is done, so that the use of the Minimum set of features aims to highlight all those feature contained in the files stored outside the user folders.
The Device Profile **Pd**$_i$**(D**$_i$**)** is defined as:

$$\mathrm{Pd}_i(\mathrm{D}_i) = \mathbf{I(Pd}_i\mathbf{)(D}_i\mathbf{)} \cup \mathbf{K(Pd}_i\mathbf{)(D}_i\mathbf{)}$$

where:

- the set of indicators drawn from the files contained in the user folder, is called **I**$_i$**(Pd**$_i$**)(D**$_i$**)**, where:
  - **I** - identifies all the detected indicators;
  - **Pd**$_i$**-**identifies the device profile;
  - **D**$_i$ - identifies the device.
- all the file that contains them are called **K**$_i$**(Pd**$_i$**)(D**$_i$**)**, where:
  - **K** - identifies the set of files;
  - **Pd**$_i$ - identifies the device profile;
  - **D**$_i$ - identifies the device.

in which:

- **e**ach indicator consists of a single piece of information not further decomposable;
- **e**ach indicator refers to one or more files;
- **e**ach file can contain one or more indicators.

Reference files (sources)	**FEATURE (filter applied)**	**Indicator**
$k_1(A_1)(D_1)$ - SAM	$m_8(A_1)$ – computer name	$i_1(k_1)$ $(A_1)(D_1)$ - PC_SuperPippo
	$m_9(A_1)$ – user system name	$i_2(k_2)$ $(A_1)(D_1)$ - SuperPippo
$k_2(A_1)(D_1)$ - SYSTEM.DAT	$m_{10}(A_1)$ - name of installed hardware	$i_3(k_3)$ $(A_1)(D_1)$ - USBpen Trust
	$m_{14}(A_1)$ - hardware installed-serial	$i_4(k_4)$ $(A_1)(D_1)$ - A01234567
$k_3(A_1)(D_1)$ - SOFTWARE.DAT	$m_{13}(A_1)$ - software installed: nome	$i_5(k_5)$ $(A_1)(D_1)$ - AVAST v1.34
	$m_{14}(A_1)$ - software installed: serial	$i_6(k_6)$ $(A_1)(D_1)$ - AD1234DC1234
$k_4(A_2)(D_1)$ - XXX.DOC	$m_1(A_2)$ - Nome file	i_7 $(k_4)(A_2)(D_1)$ - xxx.doc
	$m_6(A_2)$ - Path	i_8 $(k_4)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_{16}(A_2)$ - nickname	i_9 $(k_4)(A_2)(D_1)$ - ilgiaguaro
	$m_7(A_2)$ - MD5	I_{10} $(k_5)(A_2)(D_1)$ - B1E5CBE1E019E12E5B73EB4AFB619B5A
$k_5(A_2)(D_1)$ - NOTAMIA.TXT	$m_1(A_2)$ - Nome file	i_{11} $(k_5)(A_2)(D_1)$ - Notamia.txt
	$m_{16}(A_2)$ - nickname	i_{12} $(k_5)(A_2)(D_1)$ - superpippo
	$m_6(A_2)$ - Path	i_{13} $(k_5)(A_2)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\
	$m_7(A_2)$ - MD5	i_{14} $(k_5)(A_2)(D_1)$ - C1E5CBE1E019E12E5B73EB4AFB619B5A
k_6 $(A_3)(D_1)$ - message01.eml	$m_{28}(A_3)$ – email address	i_{15} $(k_6)(A_3)(D_1)$ - superpippo@myemail.com
	$m_{28}(A_3)$ - email address	i_{16} $(k_6)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
k_7 $(A_3)(D_1)$ - message02.eml	$m_{28}(A_3)$ - email address	i_{17} $(k_7)(A_3)(D_1)$ - superpippo@myemail.com
	$m_{28}(A_3)$ - email address	i_{18} $(k_7)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
k_8 $(A_3)(D_1)$ - message03.eml	$m_{28}(A_3)$ - email address	i_{19} $(k_8)(A_3)(D_1)$ - superpippo@myemail.com
	$m_{28}(A_3)$ - email address	I_{20} $(k_8)(A_3)(D_1)$ - ilgiaguaro@jahoo.com
k_9 $(A_4)(D_1)$ - 0261f112b3f57021.dat	$m_{19}(A_4)$ - idiomatic expression	i_{21} $(k_9)(A_4)(D_1)$ - ola hombre
	$m_{16}(A_4)$ - Nickname	i_{22} $(k_9)(A_4)(D_1)$ - ilgiaguaro
	$m_{16}(A_4)$ - Nickname	i_{23} $(k_9)(A_4)(D_1)$ - superpippo
	$m_{27}(A_4)$ - particular phrase	i_{24} $(k_9)(A_4)(D_1)$ - you left me no cigarettes in the same place yesterday
	$m_{24}(A_4)$ – reference to an object	i_{25} $(k_9)(A_4)(D_1)$ - cigarettes
	$m_{25}(A_4)$ - reference to a place	i_{26} $(k_9)(A_4)(D_1)$ - usual place
	m_{26} (A_4) - reference to a data	i_{27} $(k_9)(A_4)(D_1)$ - 24/12/2009
	$m_{22}(A_4)$ - reference to an event	i_{28} $(k_9)(A_4)(D_1)$ - no delivery
	$m_{23}(A_4)$ - reference to a person	i_{29} $(k_9)(A_4)(D_1)$ - giaguaro

k_{10} $(A_5)(D_1)$ - DSC_0001,jpg	$m_1(A_5)$ - file name	i_{30} $(k_{10})(A_5)(D_1)$ - DSC_0001.jpg
	$m_6(A_5)$ - path	I_{31} $(k_{10})(A_5)(D_1)$ - c:\Documents and Settings\SuperPippo\101ND040\
	$m_{32}(A_5)$ – image of a specific object	i_{32} $(k_{10})(A_5)(D_1)$ - yellow car with plate nr. MI01234567
	$m_{24}(A_5)$ - reference to an object	i_{33} $(k_{10})(A_5)(D_1)$ - yellow car
	$m_{24}(A_5)$ - reference to an object	i_{34} $(k_{10})(A_5)(D_1)$ - plate MI01234567
	$m_7(A_5)$ - MD5	i_{35} $(k_{10})(A_5)(D_1)$ - D1E5CBE1E019E12E5B73EB4AFB619B5A
k_{11} $(A_6)(D_1)$ - Dvd01.tif	$m_1(A_6)$ - nome file	i_{36} $(k_{11})(A_6)(D_1)$ - Dvd01.tif
	$m_6(A_6)$ - path	i_{37} $(k_{11})(A_6)(D_1)$ - Dvd01.tif c:\Documents and Settings\SuperPippo\Desktop\XXX\copertine dvd\
	$m_7(A_6)$ - MD5	i_{38} $(k_{11})(A_6)(D_1)$ - A2E5CBE1E019E12E5B73EB4AFB619B5A
k_{12} $(A_6)(D_1)$ - Dvd02.tif	$m_1(A_6)$ - nome file	i_{39} $(k_{12})(A_6)(D_1)$ - Dvd02.tif
	$m_6(A_6)$ - path	I_{40} $(k_{12})(A_6)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\dvd covers\
	$m_7(A_6)$ - MD5	I_{41} $(k_{12})(A_6)(D_1)$ - A3E5CBE1E019E12E5B73EB4AFB619B5A
k_{13} $(A_6)(D_1)$ - Dvd03.tif	$m_1(A_6)$ - nome file	i_{42} $(k_{13})(A_6)(D_1)$ - Dvd03.tif
	$m_6(A_6)$ - path	i_{43} $(k_{13})(A_6)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\dvd covers\
	$m_7(A_6)$ - MD5	i_{44} $(k_{13})(A_6)(D_1)$ - B6E5CBE1E019E12E5B73EB4AFB619B5A
k_{14} $(A_8)(D_1)$ - La cumparsita.mp3	$m_1(A_8)$ - nome file	i_{45} $(k_{14})(A_8)(D_1)$ - La cumparsita.mp3
	$m_6(A_8)$ - path	i_{46} $(k_{14})(A_8)(D_1)$ - c:\Documents and Settings\SuperPippo\Desktop\XXX\mymp3\
	$m_7(A_8)$ - MD5	i_{47} $(k_{14})(A_8)(D_1)$ - C3E5CBE1E019E12E5B73EB4AFB619B5A
k_{15} $(A_8)(D_1)$ - El dindondero.mp3	$m_1(A_8)$ - nome file	i_{48} $(k_{14})(A_8)(D_1)$ - El dindondero.mp3
	$m_6(A_8)$ - path	c:\Documents and Settings\SuperPippo\Desktop\XXX\mymp3\
	$m_7(A_8)$ - MD5	i_{49} $(k_{14})(A_8)(D_1)$ - E6E5CBE1E019E12E5B73EB4AFB619B5A
k_{16} $(A_9)(D_1)$ - History.dat	$m_{40}(A_9)$ - URL	I_{50} $(k_{16})(A_9)(D_1)$ - http://www.facebook.com/superpippo2d945.php
	$m_{40}(A_9)$ - URL	I_{51} $(k_{16})(A_9)(D_1)$ - http://www.mysite.com/ superpippo2goap43.php
	$m_{37}(A_9)$ - URL	I_{52} $(k_{16})(A_9)(D_1)$ - http://www.lamiaposta.com/superpippo3df567.php

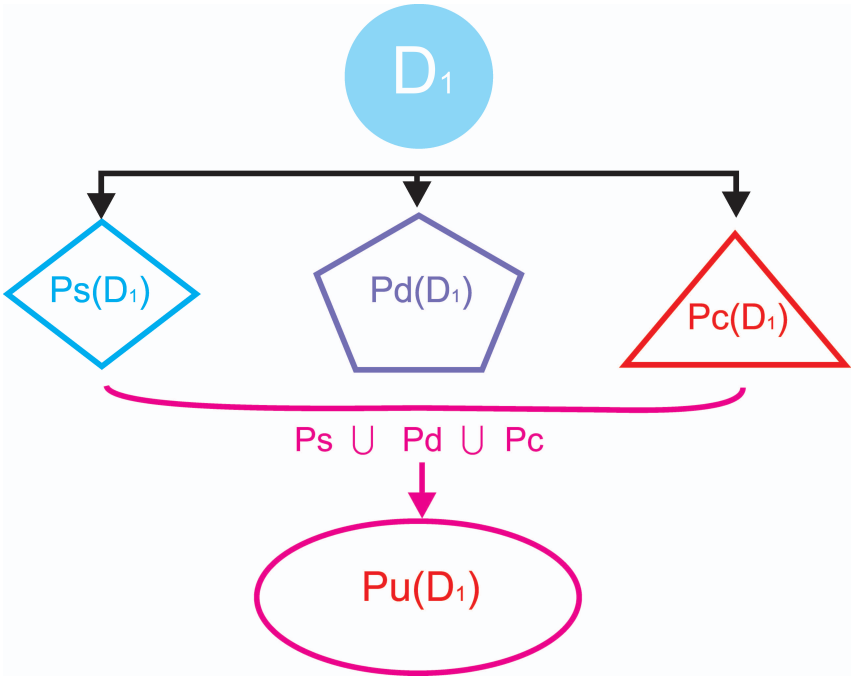


Figure 2: The user profile Pu.

3.5.4 Pu- User Profile

The profiles extrapolated so far (see figure 2) consist of all the elements necessary for creating the user profile called **Pu(D_i)**. It is the digital behavioural model that describes the user interaction with the digital device under analysis. It is therefore composed of:

- all the characterizing information(*indicators*) that are recognized on the entire machine during the analysis.
- all files that contain them.

The user profile Pu(D_i) is then defined by:

$$Pu(D_i) = I(Pu)(D_i) \cup K(Pu)(D_i)$$

where:

- **I(Pu)(D_i)** - born from the union of the three sets of indicators reported:

$$I(Ps)(D_i) \cup I(Pc)(D_i) \cup I(Pd)(D_i)$$

- **K(Pu)(D_i)** - born from the union of the three sets of files:

$$K(Ps)(D_i) \cup K(Pc)(D_i) \cup K(Pd)(D_i)$$

in which

- each indicator is no further information from a single piece;
- each indicator refers to one or more files;
- each file can contain one or more indicators.

K_{17} $(A_2)(D_1)$ - carved[123456789].doc	$m_{16}(A_2)$ - nickname	i_{53} $(k_{17})(A_2)(D_1)$ - superpippo
	$m_{18}(A_2)$ - password	i_{54} $(k_{17})(A_2)(D_1)$ - fasterthenlight
	$m_{28}(A_2)$ - indirizzo email	i_{55} $(k_{17})(A_2)(D_1)$ - superpippo@myemail.com
	$m_{26}(A_2)$ - riferimento a dato	i_{56} $(k_{17})(A_2)(D_1)$ - 339123456
	$m_7(A_2)$ - MD5	i_{57} $(k_{17})(A_2)(D_1)$ - D1E9ABE1E009E12E5B23EB4DFB689B5E
K_{18} $(A_5)(D_1)$ - carved[123456749].jpg	$m_{32}(A_5)$ – image of an object	i_{58} $(k_{17})(A_5)(D_1)$ - credit card Bankamericard
	$m_{26}(A_2)$ - reference to an object	i_{59} $(k_{17})(A_5)(D_1)$ - Bankamericard
	$m_{26}(A_2)$ - reference to a data	I_{60} $(k_{17})(A_5)(D_1)$ - 4935 1500 4556 5784
	$m_7(A_2)$ - MD5	I_{61} $(k_{17})(A_5)(D_1)$ - A1E5CBE1E019E12E5B73EB4AFB619B5A
K_{18} $(A_9)(D_1)$ - carved[123451049].3gp	$m_{30}(A_2)$ – person image	I_{62} $(k_{17})(A_9)(D_1)$ - Rossi Mario
	$m_{23}(A_2)$ - reference to a person	i_{63} $(k_{17})(A_2)(D_1)$ - Rossi Mario
	$m_7(A_2)$ - MD5	i_{64} $(k_{17})(A_2)(D_1)$ - B1E5CBE1E019E13E5B73EB4AFB619B5D
ORGANIZATION FILES Organization of personal files and folders you "Superpippo" in D_1. c:\Documents and Settings\SuperPippo\Desktop\XXX\ c:\Documents and Settings\SuperPippo\Desktop\XXX\dvd covers\ c:\Documents and Settings\SuperPippo\Desktop\XXX\mymp3\ c:\Documents and Settings\SuperPippo\101ND040\		

Table 3: Sample summary of coincident indicators detected by the comparison of profiles taken from one PC.

3.5.5 Puc – Sample User Profile

The sample user profile **Puc(D_i)** coincides with the user profile Pu(D_i), which differs only by definition because it is set as a benchmark for comparison with other devices.
In fact, the indicators collected will be used as filters to search for information within the overlapping memories of other devices.

4 The comparison

Once the sample profile Puc(D_1) is obtained from one device, the collected indicators are used as filters for the detection of the same on other devices, to detect connections and/or differences.

Final step, if necessary, is the comparison between the dates of creation / modification / deletion of files extracted by the two devices in order to reconstruct the history of user actions on devices over time. The example of Table 4 illustrated how the search for indicators, extrapolated from the device D_1, the files stored in the device D_2, has 30 information characterizing the user share (75% of filters applied). They show that both devices were used by the same user. However, this type of comparison is one way: the search of characteristic information is performed based on the indicators found in a single device, called "sample", leaving out the analysis and therefore the search for possible indicators on other devices. To work around this problem, it is be possible to take an additional step of refining the profiles through the cross referencing, which is based on the contents of memory of all the devices involved in

Nr.	Feature	Indicator	D_1	D_2
1	organization folders	... \SuperPippo \Desktop \XXX \	X	X
2	organisation folders	... \SuperPippo \Desktop \XXX \dvd covers \	X	X
3	organization folders	... \SuperPippo \Desktop \XXX \mymp3 \	X	X
4	path file	... \SuperPippo \Desktop \dvd covers \Dvd01.tif	X	X
5	path file	... \SuperPippo \Desktop \dvd covers \Dvd02.tif	X	X
6	path file	... \SuperPippo \Desktop \XXX \mymp3 \The cumparsita.mp3	X	X
7	path file	... \SuperPippo \Desktop \XXX \mymp3 \dindondero.mp3	X	X
8	personal file	Dvd01.tif	X	X
9	personal file	Dvd02.tif	X	X
10	personal file	The cumparsita.mp3	X	X
11	file personal	dindondero.mp3	X	X
12	sender email	superpippo@myemail.com	X	X
13	email recipient	ilgiaguaro@jahoo.com	X	X
14	nickname sender skype	SuperPippo	X	X
15	skypenickname	recipient's friend jaguar	X	X
16	skype password	fasterthenlight	X	X
17	idiomatic expression	ola hombre	X	X
18	nickname	ilgiaguaro	X	X
19	particular sentence	you left me no cigarettes in the same place yesterday	X	X
20	in reference cigarettes	cigarettes	X	X
21	to risereference	usual place	X	X
22	referenceat the date	24/07 / 2010	X	X
23	object reference	Bankamericardns 4935 1500 4556 5784	X	X
24	reference no phone	339123456	X	X
25	refers to vehicles	with yellow car number plate MI01234567	X	X
26	url	http://www.facebook.com/superpippo2345cdk0945.php	X	X
27	url	http://www.mysite.com/superpippo234sdfgoap43.php	X	X
28	url	http://www.myemail.com/superpippo3456asdf567.php	X	X
29	hardware	USBpen Trust s/n A01234567	X	X
30	software	v1.34 sn AVAST AD1234DC1234	X	X

Table 4: Sample of comparison

the investigation.

4.1 Cross comparison

The step consists of crossing the analysis of all the information gathered for each device. Its implementation involves the following steps:

(1) extrapolation of sample user profiles **Puc** of all devices in the analysis, each of which will consist of union of the three sets of indicators reported:

$I(Ps)(D_i) \cup I(Pc)(D_i) \cup I(Pd)(D_i)$;

and union of three sets of files: $K(Ps)(D_i) \cup K(Pc)(D_i) \cup K(Pd)(D_i)$

(2) extraction and application of:

- set of the indicators $\mathbf{I(Pu)(D_i)}$ and its files $\mathbf{K(Pu)(D_i)}$ from each profile;

- each set of filters drawn from the indicators **I(Pu)(D$_i$)** to each of the devices.

(3) update individual profiles to new indicators identified.

The procedure, though having the disadvantage of lengthening lead times, may prove useful in cases where the information obtained from the analysis of a single device are not very significant because it allows to:

(1) analyse the data in all devices;

(2) increase the number of indicators obtained;

(3) make the user profiles more consistent.

It also allows detecting any additional user.

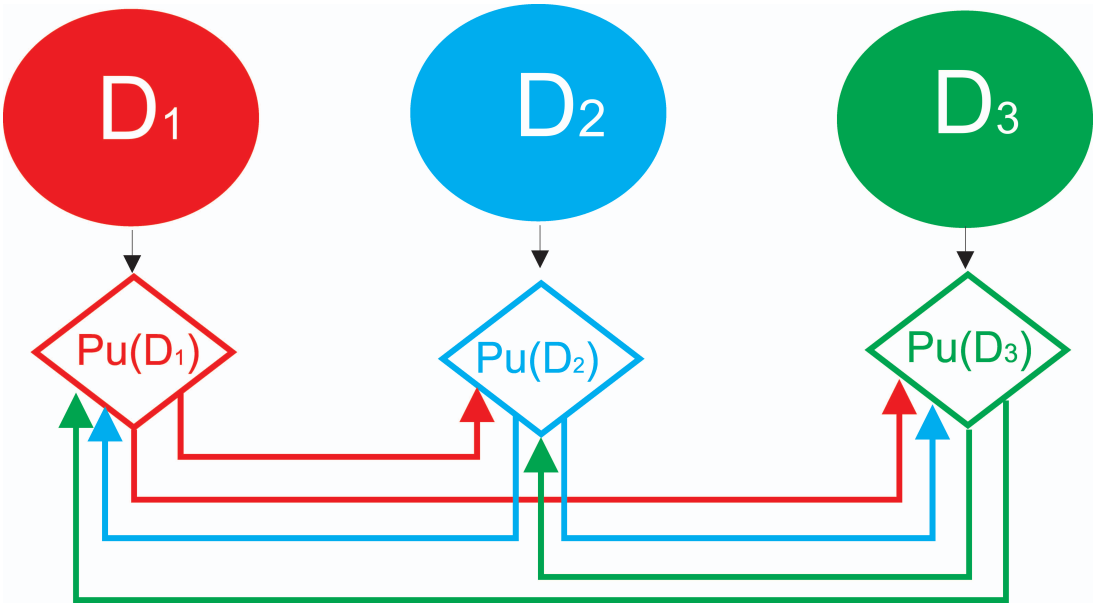


Figure 3: Comparison of cross: each of the three measured profiles are compared with the other two

4.2 Multi-User Devices

A more complex case can occur if the same device D$_i$ is used by more than one person (e.g. personal computer).
A profile for each user has to be extrapolated, with the following rules:

(1) build a user profile Pc for each user (i.e., Pc$_1$, Pc$_2$, etc..) on each of the user folders on the machine;

(2) build a system profile Ps (e.g., Ps$_1$, Ps$_2$, etc..) for each user;

(3) build a unique device profile Pd;

(4) cross compare each Pc and Pd profiles in order to produces one Pu profile for each detected user;

(5) each user profile Pu$_i$(D$_i$)will be defined as:

$$\mathbf{Pu}_i(\mathbf{D}_i) = \mathbf{Pc}_i(\mathbf{D}_i) \cup \mathbf{Pd}(\mathbf{Pc}_1)(\mathbf{D}_i) \cup \mathbf{Ps}(\mathbf{Pc}_1)(\mathbf{D}_i)$$

The compare between the different user folder profiles Pc, and the device profile Pd, are designed to:

(1) identify their own indicators in the areas of memory included in the device profile;

(2) extract the files containing them and add them to the relative **Pdu,** where Pdu means the user device profile, a subset of Pd, formed by:

(i) the set of indicators in common with the Pc;

(ii) all files that contain them.

(3) create many profiles $\mathbf{Pu}_i(\mathbf{D}_i)$ how many are the user folders (not empty) comprising:

$$\mathrm{Pc}_i(\mathrm{D}_i) \cup \mathrm{Pdu}_i(\mathrm{Pc}_i)(\mathrm{D}_i) \cup \mathrm{Ps}_i(\mathrm{D}_i)$$

(4) decrease the size of the Pd profile that will ultimately be composed of these indicators (and related files) not included in different user profiles.

The end result is:

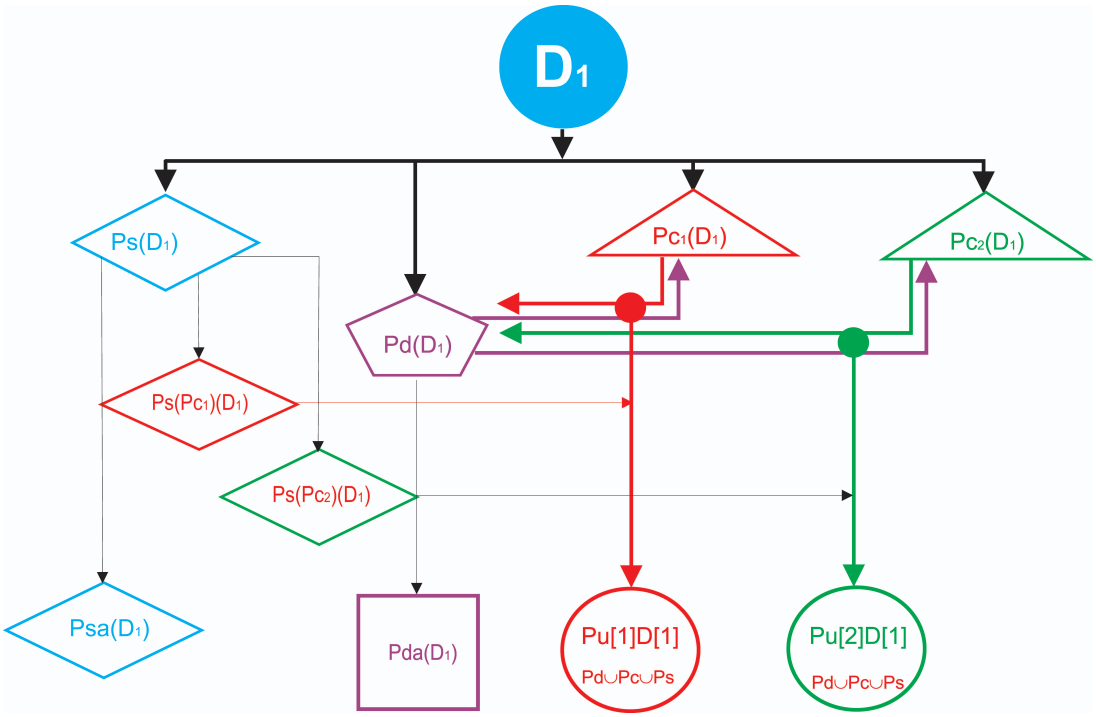


Figure 4: Process of detection of two profiles in a multi-user device

(1) *n* user profiles - the set of characteristic information that describes the behaviour of digital users found on the machine;

(2) nr. 1 anonymous Pda Device Profile (if any) - that is, a set of information characterizing not related to those users, which will also include that information on configuring the system which does not give users found.

This last profile is not deleted, but is listed as anonymous profile because it contains information that, just because so far attributed to anyone, might be useful for the identification of other entities, by comparison with other devices in subsequent analysis.

4.3 Evaluation of results

The evaluation of the result (operation of close investigators relevance) is carried out in a quantitative sense (i.e., considering the number of coincident indicators measured), and in a qualitative sense (i.e., the veracity of information), as even a only single information can be found as the solution of the problem posed by the analysis goal.

4.3.1 The quantitative assessment

It is carried out in statistical way by calculating the percentage of coincident indicators found by comparing the total of those used as a filter.

EXAMPLE:

Quantitative assessment of the results obtained by simple comparison (on the case presented in Table 1)

Task 1 - Create User Profile sample Pcu (D_1)

N. filters applied (taken from the minimum set of features) 44

RESULT: indicators extracted 40

Task 2 - Research using filters of the indicators on the device D_2 :

N. filters applied (taken from the set of indicators) 40

RESULT: coincident indicators found 30

On total of nr. 40 indicators/filter applied by simple comparison, has been detected Nr 30 coincident indicators, which is 75%.

4.3.2 The qualitative assessment

This analysis gives to the information obtained (indicators) a value of "relevance" based on the individual indicators in relation to their degree of usefulness for the target. In consideration of the digital nature of the analysis, the sources are not assessed: if properly extracted and verified by hashing algorithms, are considered *"completely reliable"*. With regard to the *information* obtained in case presented here, it shows no qualitative assessment (however the only investigator's responsibility, in the survey), as the specific research described by the example given here was aimed solely to collecting coincident (i.e. in possession of only two values: *match / no match*), which could bring with certainty the identity of the same subject in question.

5 Conclusions

In this paper, we propose a new investigative method, the Digital Profiling, based on the set theory that can extract useful information from a digital device to assist in the identification of criminal subjects. It analyses data and metadata memorized into a digital device by applying specific techniques taken from intelligence and traditional profiling in order to obtain information, with which it is possible to reconstruct a user's profile and her or his modus operandi. As we mentioned, the process starts from the research and analysis of all information that can be gathered from "digital footprints" left on a digital device by its user. This is possible because the computer user is a human being tending to customize all the environments with which she or he interacts. Thus, she or he cannot avoid to leave, even unconsciously,

digital evidence that can be detected, recognized and compared. The analysis described in this article is suitable to all the devices, starting from those in which multiple applications are available, such as personal computers, iPads, tablets, mobile phones and smartphones. However, embedded devices should not be excluded from this type of analysis: a GPS car navigation system device, even though at first glance it may seem not containing data useful to find a solution of a crime, can even provide valuable information on the movements of a subject, such as places where has gone, travel times, the usual routes that, when compared with the position of his home, may help to delineate the aim of his activities, his geographical position in a definite time range, for example to verify an alibi. Digital profiling technique can also be applied to the content of storage areas provided in remote areas (websites, social networks, cloud computing etc.) and data streams, selected, for example, in a certain time during a network interception. This particular type of analysis may be very helpful in various investigative fields, related to those crimes which involve a digital device (today, who among us does not have at least a cell phone or a credit card?), in which it is necessary to analyze a digital memory in order to identify a criminal. To conclude, this technique is particularly useful in the investigation about cybercrimes, as computer frauds, phishing, cyber stalking, child pornography, and hacker attacks, especially where anti-forensic techniques are applied to hide or delete the crime evidence. In this way, Digital Profiling is very helpful in operations against organized crime, anti-terrorism and intelligence operations, where it can be interfaced with the statistical study in the prediction and prevention of criminal events.

References

[1] Clara Colombini and Antonio Colella, "Digital profiling: A computer forensics approach," in *Proc. of the 6th International Conference on Availability, Reliability and Security (ARES'11), Vienna, Austria, LNCS*, vol. 6908. Springer-Verlag, August 2011, pp. 330–343.

[2] Anconelli M., "Introduzione al digital profiling," *www.cybercrimes.it*, 2010.

[3] Strano M, "Nuove frontiere delle tecniche di criminal profiling," *ANFP Forze Civili*, 2005.

[4] Colombini C.M., "Digital profiling, un nuovo strumento di indagine informatica," *IISFA Memberbook 2011*, 2011.

[5] Keppel, R, "Serial offenders: Linking cases by modus operandi and signature." *Forensic science Boca Raton: CRC Press*, pp. 605–614, 2005.

[6] Rogers, M.K., *Computer forensics: Steps toward defining a common body of knowledge. Paper presented at the Information Protection Association of Manitoba conference*. Winnipeg Manitoba, 2002.

[7] Corneli A., "Intelligence diffusa e cultura dell'intelligente," *Per aspera ad veritatem*, 2007.

[8] Douglas J.E., Burgess A., Ressler R., *Crime Classification Manual*. Jossey-Bass, Publishers, 1992.

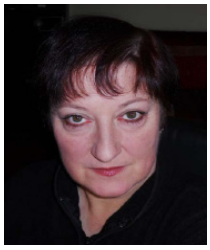
[9] DiPaolo Anna M., *Elementi di Intelligence e tecniche di analisi investigativa*. Ruffolo, 2009.

[10] Douglas J.E., Olshaker M., *Obsession: the FBI's legendary profiler probes the psyches of killers, rapists, and stalkers and their victims and tells how to fight back*. Scribner, 1998.

[11] Kruse, W.J. Heiser, J.G., *Computer forensics incident response essentials*. New York: Addison Wesley, 2002.

[12] Schultz E., Shumway R., *Incident Response: A Strategic Guide to Handling System and Network Security Breaches.* Sams., 2002.

[13] Turvey B, "Deductive criminal profiling: Comparing applied methodologies between inductive and deductive criminal profiling tecniques," *Knowledge Solutions Library*, 1998.



Clara Maria Colombini received the Degree in Computer Science, the Postgraduate Qualification in Criminology and Forensic Sciences, the Postgraduate Qualification in Computer Forensics and the Postgraduate Qualification in Digital Investigation from Universitá degli Studi di Milano, Italy, in 2007, 2008 and 2009 respectively. She received also the Certification in Digital Forensic from Universita degli Studi di Pisa-Ordine degli Ingegneri of Brescia in 2010. Now she is Forensic Consultant for the Italian Prosecutor's Office for the Forensic Analysis of Digital Devices and external research Collaborator at Professorship of Criminology and Forensic Sciences, Universitá degli Studi di Milano. Moreover, as speaker, she participated to several courses, seminars and international conference on Computer Forensics subject and Criminology. In particular, she participated in the International Conference on Availability, Reliability and Security (Vienna, Austria 2011). Her research interests include the application of new theories to the digital investigation.

Antonio Colella received the Degree in Strategic Science, Political Science and a Postgraduate Qualification in Criminology and Forensic Sciences, Information Security Management and Cybercrime IT Criminal Law. He also got the Certification as ISO/EAC 27001 Lead Auditor and International Information System Investigator (IISFA). Dr. A. Colella teaches Computer Forensics at Master of Arts in Forensics Science at Universitá La Sapienza in Rome and Cybersecurity at Libera Universita' Maria Ss. Assunta in Rome for the Course of Cybercrime & IT Criminal Law. Moreover, as speaker, he participated to several courses, seminars and international conference on Computer Forensics subject and Criminology. In particular, he has been part of reviewing team for the Conference on Cyber Conflict (Tallinn - Estonia 2010) and of Workshop Program Committee for the International Conference on Availability, Reliability and Security (Vienna, Austria 2011). He belongs to Italian Army with rank of Lieutenant Colonel.