# A Quantitative Approach to Triaging in Mobile Forensics

**CONFERENCE PAPER** · NOVEMBER 2011

DOI: 10.1109/TrustCom.2011.75



**4 AUTHORS**, INCLUDING:

Fabio Marturana
University of Rome Tor Vergata
**11** PUBLICATIONS **33** CITATIONS
SEE PROFILE

Rosamaria Bertè
University of Rome Tor Vergata
**7** PUBLICATIONS **28** CITATIONS
SEE PROFILE

Gianluigi Me
University of Rome Tor Vergata
**43** PUBLICATIONS **176** CITATIONS
SEE PROFILE

# A quantitative approach to Triaging in Mobile Forensics

Fabio Marturana
Department of Computer Science, Systems and Production
University of Rome Tor Vergata, Rome, Italy
marturana@libero.it

Gianluigi Me
Department of Computer Science, Systems and Production
University of Rome Tor Vergata, Rome, Italy
me@disp.uniroma2.it

Rosamaria Bertè
Department of Computer Science, Systems and Production
University of Rome Tor Vergata, Rome, Italy
rosamariaberte@libero.it

Simone Tacconi
Polizia di Stato e delle Comunicazioni
Rome, Italy
simone.tacconi@interno.it

*Abstract*— Forensic study of mobile devices is a relatively new field, dating from the early 2000s. The proliferation of phones (particularly smartphones) on the consumer market has caused a growing demand for forensic examination of the devices, which could not be met by existing Computer Forensics techniques. As a matter of fact, Law enforcement are much more likely to encounter a suspect with a mobile device in his possession than a PC or laptop and so the growth of demand for analysis of mobiles has increased exponentially in the last decade. Early investigations, moreover, consisted of live analysis of mobile devices by examining phone contents directly via the screen and photographing it with the risk of modifying the device content, as well as leaving many parts of the proprietary operating system inaccessible. The recent development of Mobile Forensics, a branch of Digital Forensics, is the answer to the demand of forensically sound examination procedures of gathering, retrieving, identifying, storing and documenting evidence of any digital device that has both internal memory and communication ability [1]. Over time commercial tools appeared which allowed analysts to recover phone content with minimal interference and examine it separately. By means of such toolkits, moreover, it is now possible to think of a new approach to Mobile Forensics which takes also advantage of "Data Mining" and "Machine Learning" theory. This paper is the result of study concerning cell phones classification in a real case of *pedophilia*. Based on Mobile Forensics "Triaging" concept and the adoption of self-knowledge algorithms for classifying mobile devices, we focused our attention on a viable way to predict phone usage's classifications. Based on a set of real sized phones, the research has been extensively discussed with Italian law enforcement cybercrime specialists in order to find a viable methodology to determine the likelihood that a mobile phone has been used to commit the specific crime of *pedophilia*, which could be very relevant during a forensic investigation.

***Keywords:*** **Triaging, Mobile Forensics, Data Mining, Knowledge Analysis, Machine Learning.**

## I. INTRODUCTION

Cell phone, PDA and new generation smartphone proliferation is on the increase all over the world. Worldwide sales of mobile devices to end users totaled 428.7 million units in the second quarter of 2011, a 16.5 percent increase from the second quarter of 2010, according to Gartner, Inc. (Fig.1) [2].

**Worldwide Mobile Device Sales to End Users by Vendor in 2Q11 (Thousands of Units)**

| Vendor | 2Q11 Units | 2Q11 Market Share (%) | 2Q10 Units | 2Q10 Market Share (%) |
|---|---|---|---|---|
| Nokia | 97,869.3 | 22.8 | 111,473.7 | 30.3 |
| Samsung | 69,827.6 | 16.3 | 65,328.2 | 17.8 |
| LG | 24,420.8 | 5.7 | 29,366.7 | 8.0 |
| Apple | 19,628.8 | 4.6 | 8,743.0 | 2.4 |
| ZTE | 13,070.2 | 3.0 | 6,730.6 | 1.8 |
| Research In Motion | 12,652.3 | 3.0 | 11,628.8 | 3.2 |
| HTC | 11,016.1 | 2.6 | 5,908.8 | 1.6 |
| Motorola | 10,221.4 | 2.4 | 9,109.4 | 2.5 |
| Huawei Device | 9,026.1 | 2.1 | 5,276.4 | 1.4 |
| Sony Ericsson | 7,266.5 | 1.7 | 11,008.5 | 3.0 |
| Others | 153,662.1 | 35.8 | 103,412.6 | 28.1 |
| **Total** | **428,661.2** | **100.0** | **367,986.7** | **100.0** |

Source: Gartner (August 2011)

Fig 1. Worldwide Mobile Device Sales – 2Q11

The reasons for this commercial success can be found in the advances in semiconductor technologies, the increase of computing power of mobile phones, which also led to an increase of functionality while keeping the size of such devices small enough to fit in a pocket, moderate prices and ever-increasing storage capacity and supported features.
As a consequence, law enforcement have to deal with a proliferation of crimes related to mobile devices usage [3], thus cell phones, PDAs and smartphones are considered important items to analyze during investigations, providing useful intelligence about suspect habits, interests, social relations and technical skills [4].

978-0-7695-4600-1/11 $26.00 © 2011 IEEE
DOI 10.1109/TrustCom.2011.75

After a few interviews with Italian law enforcement cybercrime specialists, we noticed a growing complexity in today's forensic investigations due to the quantity of new mobile phones released every year, each with its own Operating Systems (i.e. Android, Symbian, Apple iOS, RIM, Windows Phone etc.) and a different File System and memory organization. In Fig.2 Worldwide Smartphone selling figures by Operating System in 2Q11:

**Worldwide Smartphone Sales to End Users by Operating System in 2Q11 (Thousands of Units)**

| Operating System | 2Q11 Units | 2Q11 Market Share (%) | 2Q10 Units | 2Q10 Market Share (%) |
|---|---|---|---|---|
| Android | 46,775.9 | 43.4 | 10,652.7 | 17.2 |
| Symbian | 23,853.2 | 22.1 | 25,386.8 | 40.9 |
| iOS | 19,628.8 | 18.2 | 8,743.0 | 14.1 |
| Research In Motion | 12,652.3 | 11.7 | 11,628.8 | 18.7 |
| Bada | 2,055.8 | 1.9 | 577.0 | 0.9 |
| Microsoft | 1,723.8 | 1.6 | 3,058.8 | 4.9 |
| Others | 1,050.6 | 1.0 | 2,010.9 | 3.2 |
| **Total** | **107,740.4** | **100.0** | **62,058.1** | **100.0** |

Source: Gartner (August 2011)

Fig 2. Worldwide Smartphone Sales by O.S.– 2Q11

In recent years a number of hardware and software tools have emerged to recover evidence from mobile devices and provide investigators with all the needed instruments to conduct a technical inquiry. Most tools consist of a hardware portion, with a number of cables to connect the phone to the acquisition machine, and some software, to extract the evidence and, occasionally, to analyze it [5][6]. It is important to notice, however, that different products extract different amounts of information from different devices. A critical drawback currently facing mobile OS and FS forensic development is the extremely short OS release cycles. Major OS release are normally delivered every twelve months or less with minor releases coming in between those major releases. This issue makes timely development, testing and release of forensic tools and updates that deal with the newer OS releases difficult to achieve; This leads to a very complex landscape when trying to overview the products [7].

In a former research, we realized that the 4-steps mobile forensic workflow, based on *device identification*, *acquisition*, *analysis and reporting*, was found to be inadequate to current investigations. We proposed thus to modify it introducing an intermediate step, called *triage*, located between *acquisition* and *analysis*, with the aim of limiting the area of interest and reduce the number of relevant devices to focus on [8]. As an example, in a complex case involving more persons, crimes and mobile equipment, by means of *triage*, it is possible to both split up relevant and less important aspects of the case and assign a priority to every seized device [9][10]. By means of some "Data Mining" and "Machine Learning" algorithms, indeed, we are able to discriminate the device importance between critical (to analyze as soon as possible) and less important (to analyze with lower priority), imitating the hospital "triage" protocol which assign disease seriousness priority codes [11].

In this paper we offer the reader a different point of view with regards to [8], by adopting *triage* methodology to predict the likelihood that a mobile phone has been used to commit a specific crime i.e. *pedophilia* and may contain evidence to be extracted and analyzed with more detail in Lab. To do so, we needed first to collect a set of seized devices regarding different cases and crimes and later on to use the extracted *data-instances* to train a "Machine Learning" algorithm in order to provide phone's classifications.

## II. Related work

In the scientific field there is a growing interest about *triage* tools [9], techniques and methods that could ease Police Agencies tasks and help them to steer investigations quickly.

Moreover, the scientific literature has produced guidelines that formalize and standardize the live *triage* process to be performed on crime's scene as well as the basic requirements to automate *triage* tools [7].

Some implementations are based on data–driven approach that quickly parse mobile information without performing further analysis or on specialized software, such as DECODE system of Massachusetts University [10].

Other tools, developed for mobile *triaging*, ease police work through a more immediate and comprehensive view of the evidence [6]. It's important to notice that, however, the field is still at an embryonic stage of development and further experiments are underway, given the lack of reports and data concerning real forensic cases.

## III. Knowledge acquisition and analysis workflow

Unlike other approach to Mobile Forensics *triage*, where live analysis techniques regarding on-scene inspection and examination are emphasized, we decided here to change our point of view by focusing our attention on a "cold" exhibit analysis method.

This paragraph describes how the methodology proposed in [8] fits to a "real pedophilia case". The study is based on a collection of reports and data extracted from seized mobile devices, delivered by Servizio Polizia Postale e delle Comunicazioni (the Italian Cybercrime Low Enforcement) and regarding different crimes such as pedophilia and pedopornography, homicide, non-disclosure agreement violation, human trafficking and extortion.

The extracted data are first normalized, in order to remove the misalignments produced by the different extraction tools

and then processed in order to create the "dataset". Finally normalized data are elaborated by means of Knowledge Analysis algorithms in order to predict the likelihood that a mobile phone has been used to commit *pedophilia*.
The following figure shows the described process:

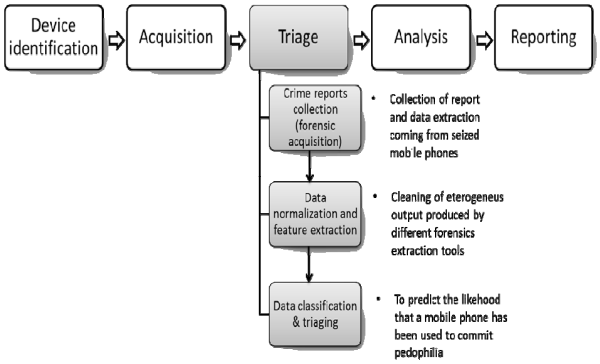


Fig 3. Process Workflow

The ***first stage*** of the Triage workflow is the data collection and it's called ***crime reports collection (forensic acquisition)***.
It is based on a classical forensic acquisition of a mobile phone memory with the appropriate extraction tool. It can be carried out both on crime scene or in a forensics lab.

The ***second stage*** of our workflow is called ***data normalization and feature extraction***; i.e. the extracted data, regarding common/serious and cybercrimes, are firstly normalized, in order to remove the different output produced by different extraction tools and then loaded into databases. This stage is fundamental because of the large, often incomplete, noisy due to outliers, inconsistent and redundant amount of data. According to the so called "*Knowledge Discovery Process*" [12], before submitting the *"dataset"* to the classification process, it is necessary to perform *data cleaning* (noise reduction), *relevance analysis* (redundant attributes elimination) and *data transformation* (normalization).
In particular, we focus our attention on the following parameters/data:

- Phone model (Smartphone, GSM);
- Number of phonebook contacts (stored both on SIM and phone);
- Number of dialed/received/missed calls;
- Percentage of dialed/received/missed calls (with regards to the specific time slot: Morning, Afternoon, Evening and if generated or received from phonebook contacts or not);
- Average duration of dialed/received calls (with regards to the specific time slot: Morning, Afternoon, Evening and if generated or received from phonebook contacts or not);
- Number of received/sent SMS/MMS;
- Percentage of received/sent SMS/MMS (with regards to the specific time slot: Morning, Afternoon, Evening and if they are sent or received from phonebook contacts or not);
- Number and percentage of visited URLs, with regards to the specific time slot: Morning, Afternoon, Evening and if they are bookmarked or not;
- Number and percentage of downloaded images/videos/audio files stored on the device or created by means of the device camera and microphone;
- Number of sent/received E-mail;
- Number of stored Notes.

We are able to create the following input data structure, called *input matrix*:

| Attribute name | Nokia N73 | HTC Magic |
|---|---|---|
| Phone model | GSM | Smartphone |
| Number_phonebook_contacts | 19 | 451 |
| Number_received_calls | 18 | 166 |
| Number_dialled_calls | 42 | 307 |
| Number_missed_calls | 16 | 27 |
| Percentage_received_calls | Low | Medium |
| Percentage_dialled_calls | Medium | Medium |
| ---------- | ---------- | ---------- |
| ---------- | ---------- | ---------- |
| ---------- | ---------- | ---------- |
| Mean_duration_received_calls | 71 | 168 |
| Mean_duration_dialled_calls | 31 | 192 |
| ---------- | ---------- | ---------- |
| ---------- | ---------- | ---------- |
| Number_read_sms | 261 | 90 |
| Number_sent_sms | 270 | 22 |
| Percentage_read_sms | Medium | High |
| ---------- | ---------- | ---------- |
| Number_URL_visited | 16 | 15 |
| ---------- | ---------- | ---------- |
| user_class | Pedo | Non-Pedo |

Fig 4 – Input matrix

Attribute's name is indicated in the left column (i.e. the parameter called Number_phonebook_contacts etc.) while, in the two right columns two sample instances (i.e. the collection of extracted and normalized data) of mobile phones (i.e. a Nokia N73 and an HTC Magic) are shown, where each attribute has its own correspondent value that can be either nominal or numeric.

The ***last stage*** of our workflow is called ***data classification & triaging*** and it is based on techniques enabling quick

evidence classification in order to predict the likelihood that a mobile phone has been used to commit *pedophilia*.
We analyzed a collection of 21 phones and, after the normalization process, we created an input matrix with 21 columns, called *instances*, one for every phone and 114 rows, called *Attributes*, concerning the parameters of every instance; the cited matrix represents the so called *Relation* in our *Attribute-Relation model*, analyzed by means of WEKA Data Mining algorithms [13].

## IV. CLASSIFICATION ANALYSIS

In order to choose the best performing classifier, we analyzed the WEKA suite [13], making a detailed comparative analysis among some of the supported algorithms and evaluating their performance on our dataset by means of key performance indicators such as *Precision*, *Recall* and *F-measure*; in particular:

a. ***Precision = True Positive / (True Positive + False Positive)*** is the ratio of true positive and the sum of those with false positive (note: if the number of false positive is low the *Precision* is close to 1);

b. ***Recall = True Positive / (True Positive + False Negative)*** is the ratio of true positive and the sum of those with false negative (note: if the number of false negative is low the *Recall* is close to 1);

c. ***F-measure = 2 * Recall * Precision / (Recall + Precision)*** is the harmonic mean of Recall and Precision.

As a result, we compared the following three classification algorithms: **Bayesian Networks, Decision Tree and LWL-Locally Weighted Learning**.
All the aforementioned algorithms train themselves by inspecting all the instances included in the *training-set* and create different statistical models.
With **Bayesian Networks**, what is being estimated is the conditional probability distribution of the values of the class attribute (i.e. the *usage class*) given the values of the other attributes. Ideally, the classification model represents this conditional distribution in a concise and easily recognizable way. Bayesian networks are drawn as a network of nodes, one for each attribute, connected by directed edges in a directed acyclic graph [14].
Adopting a **Decision Tree** algorithm, we apply a "divide-and-conquer" approach to the problem of learning from a set of independent instances; the corresponding binary tree representation consists of nodes implying a test on one or more attributes and leafs giving a classification to all the instances that reach it. For example, to classify an unknown instance it is routed down the tree according to the values of the attributes tested in successive nodes, and when a leaf is reached the instance is classified according to the class assigned to that leaf [16].
**Locally Weighted Learning** (aka LWL) is a general algorithm and can be applied with any learning technique that can handle weighted instances. In particular, it can be used for classification. It assigns weights using an instance-based method and builds a classifier from the weighted instances; *LWL* only assumes independence within a neighborhood, not globally in the whole instance space [15].
At the end of the "training" process, WEKA creates a statistical model; each classifier is able to assign an *usage class* to all the instances within the *trainingset*, comparing the predicted class with the original one.

### A. *First scenario: Complete trainingset and 10 folds cross-validation*

In detail, in the first scenario we used the 21 instances *dataset* to train Bayesian Networks, WEKA Decision Tree (aka J48) and Locally Weighted Learning by means of an iterative and predictive method called *10 folds cross-validation*. This method splits up the training-set into ten approximately equal partitions, each in turn used for training and the remainder for testing. That is, use nine-tenth of the data for training and one-tenth for testing, and repeat the procedure ten times so that in the end, every instance has been used exactly once for testing [13]. Applying *10 folds cross-validation* to the each classifier we calculated the following performance indicators

Tab. 1: first scenario comparative results

| Model Classifier | **Precision** | **Recall** | **F-Measure** |
|---|---|---|---|
| **BayesNet** | 0.553 | 0.579 | 0.56 |
| **J48** | 0.68 | 0.684 | 0.644 |
| **LWL** | 0.644 | 0.632 | 0.636 |

In this case, WEKA Decision Tree (aka J48) performs better than the other classification schemes, nevertheless to reduce further the average classifier's error rate, we decided to adopt different strategies.

### B. *Second scenario: Complete trainingset and* two "*real*" phones/instances *testset*

In the second scenario, for instance, instead of 10 folds cross-validation we used 19 phones/instances out 21, associated with different crime profiles, to build a *training-set* and train the classifiers and the remaining 2, both with a *non-pedophile* profile, to create a *test-set*. After the initial training, WEKA evaluated and classified the 2 testing instances with resulting performance indicators summarized in the table below:

Tab. 2: second scenario comparative results

| Model Classifier | Precision | Recall | F-Measure |
|---|---|---|---|
| **BayesNet & J48** | 0 | 0 | 0 |
| **LWL** | 1 | 1 | 1 |

In this case the Locally Weighted Learning scheme is able to correctly classify both the testing instances while Bayesian Networks and WEKA Decision Tree (aka J48) performs worse than the first case, each with two wrong classifications out of two. As a result, in this scenario we noticed that at least one classifier performs better than in the first one.

To give the reader an idea of how classifiers' performance could be improved, we noticed, for instance, that reducing the number of attributes of the input matrix M (rif. fig.2), a rectangular m × n where m = 114 and n = 23, could have a positive influence on the measured error rate. Thus, in the next two scenarios (third and fourth), we applied some linear algebra concepts to justify the reduction of redundant attributes (i.e. rows in the input matrix), with special regards to those with more missing values (i.e. called outliers) in the training instances.

We created, thus, two reduced datasets, respectively with numeric attributes only (third scenario - *numeric training-set*) and with both numeric and nominal attributes (fourth scenario - *nominal training-set*). Firstly, we recall the rank or maximum number of linearly independent rows and columns of a generic m × n matrix cannot be greater than m nor n; thus rank (M) <= min (m,n). In our case rank (M) is equal or less than 23 thus there are no more than 23 linearly independent rows or attributes in the problem. Therefore, in the first simplified dataset, we reduce manually the number of rows first eliminating the nominal attributes, obtaining a m'× n matrix with m' = 63 and n = 23. We apply then the Gauss Elimination algorithm to reduce the m'× n input matrix to a row-echelon form, finding out that the value of rank (M) is exactly 23. In the second simplified scenario it is not possible to apply the Gauss Elimination algorithm since we have both numeric and nominal attributes that cannot be elaborated thus we assume that the result is the same and select only 23 attributes.

In the following two scenarios, we applied the same approach adopted in the second scenario.

### C. *Third scenario: Reduced Trainingset (Numeric)*

In the third scenario, we use a reduced *trainingset*, (based on the 23 numeric attributes detailed in the table below), to train the model in order to classify a *testset* made of 2 phones, each with a *non-pedophile* profile:

Tab. 3: Trainingset with numeric attributes

| Attribute name | Attribute type |
|---|---|
| Phone model | {GSM,Smartphone} |
| Number_phonebook_contacts | numeric |
| Number_received_calls | numeric |
| Number_dialled_calls | numeric |
| Number_missed_calls | numeric |
| Mean_duration_received_calls | numeric |
| Mean_duration_dialled_calls | numeric |
| Number_read_sms | numeric |
| Number_sent_sms | numeric |
| Number_read_mms | numeric |
| Number_sent_mms | numeric |
| Number_downloaded_picture_files | numeric |
| Number_downloaded_video_files | numeric |
| Number_downloaded_audio_files | numeric |
| Number_produced_picture_files | numeric |
| Number_produced_video_files | numeric |
| Number_produced_audio_files | numeric |
| Number_URL_visited | numeric |
| Number_URL_bookmarks | numeric |
| Number_sent_email | numeric |
| Number_received_email | numeric |
| Number_notes_memo | numeric |
| user_class | {Pedo, Non-Pedo} |

Calculated performance indicators are summarized in the table below:

Tab. 4: third scenario comparative results

| Model Classifier | Precision | Recall | F-Measure |
|---|---|---|---|
| **BayesNet, J48 & LWL** | 1 | 1 | 1 |

As we can notice, the three classification algorithms are able to correctly classify all the testing instances with the implication that reducing the attribute space to a set of non-redundant numeric attributes means that classifiers are less stressed and better performing.

### D. *Fourth scenario: Reduced Trainingset (Nominal)*

In the fourth scenario, we use a reduced *trainingset*, (based on the 22 numeric attributes detailed in the table below), to train the model in order to classify a *testset* made of 2 phones, each with a *non-pedophile* profile:

Tab. 5: Trainingset with nominal attributes

| Attribute name | Attribute type |
|---|---|
| Telephone type | {GSM,Smartphone} |
| Number_phonebook_contacts | numeric |
| Percentage_received_calls | {Low,Medium,High} |
| Percentage_dialled_calls | {Low,Medium,High} |
| Percentage_missed_calls | {Low,Medium,High} |
| Mean_duration_received_calls | numeric |
| Mean_duration_dialled_calls | numeric |
| Percentage_read_sms | {Low,Medium,High} |
| Percentage_sent_sms | {Low,Medium,High} |
| Percentage_read_mms | {Low,Medium,High} |
| Percentage_sent_mms | {Low,Medium,High} |
| Percentage_downloaded_picture_files | {Low,Medium,High} |
| Percentage_downloaded_video_files | {Low,Medium,High} |
| Percentage_downloaded_audio_files | {Low,Medium,High} |
| Percentage_produced_picture_files | {Low,Medium,High} |
| Percentage_produced_video_files | {Low,Medium,High} |
| Percentage_produced_audio_files | {Low,Medium,High} |
| Number_URL_visited | numeric |
| Number_URL_bookmarks | numeric |
| Number_sent_email | numeric |
| Number_received_email | numeric |
| user_class | {Pedo,Non-Pedo} |

Calculated performance indicators are summarized in the table below:

Tab. 6: fourth scenario comparative results

| Model Classifier | Precision | Recall | F-Measure |
|---|---|---|---|
| **BayesNet & J48** | 1 | 1 | 1 |
| **LWL** | 1 | 0.5 | 0.667 |

As we can notice, in this case two classification algorithms out of three are able to correctly classify all the testing instances with the implication that reducing the attribute space to a set of non-redundant nominal attributes means that classifiers are, on average, less stressed and better performing.

## V. FUTURE WORK

The result of our research can be considered a first step towards new dead forensic "triage" scenarios where analysts could, for instance, examine the digital evidence under different perspectives in order to find, as an example, associations between the extracted data and different criminal behaviors. We foresee a growing interest towards this topic since, the relatively new approach based on the adoption of Machine Learning techniques to mobile forensics, will allow to isolate potential evidence and provide useful intelligence about suspects; If this investigative sector will evolve, as we foresee, researchers could be interested in developing integrated tools able to collect the extracted data in different format (HTML, XML, Plain text, DB etc.) normalize and elaborate them in order to find relations and associations among persons, facts, evidence and crimes. Some manufacturers and software houses in the marketplace would probably be interested in acknowledging the research field and collaborating with researcher and law enforcement in order to commercialize such highly specialized tools.

## VI. CONCLUSION

This paper summarize the application of Knowledge Management classification algorithms to Mobile Forensics, the branch of Digital Forensics that deals with forensically sound methods to isolate potential evidence from mobile phones or any other digital device with both internal memory and communication ability.

With the collaboration of Italian Cybercrime law enforcement specialists, we had the opportunity, indeed, to collect and examine a considerable amount of evidence regarding "real" seized phones and concerning crimes such as pedophilia and pedopornography, homicide, non-disclosure agreement violation, human trafficking and extortion. By means of the evidence at our disposal, we created 21 *data-instances* and trained 3 WEKA algorithms (i.e. Bayesian Networks, Decision Trees, and Locally Weighted Learning) to provide phone's classifications with regards to a specific crime. The goal was to find a viable method to easily predict the likelihood that a sized mobile phone has been used to commit *pedophilia* and needs to be analyzed with more detail in Lab. We analyzed 4 different scenarios: 2 of them using the complete 114 attributes' set to create the *trainingset* and the others with a simplified attributes' set called, respectively, *numeric* and *nominal.* In each scenario we calculated 3 performance indicators (*Precision*, *Recall* and *F-measure*) to evaluate the effectiveness of each classification algorithm. We applied some linear algebra concepts to justify the reduction of redundant attributes in the training instances, with special regards to the so-called outliers and we demonstrated that, reducing the number of attributes, classifiers have generally better performance. It is important to notice that our results suffer the *collision-of-output* issue, as it happens in other classification problems. Different input indeed may produce the same output and could be misinterpreted by the classifier, even though they represent false positive instances. We address this potential issue by saying that the outcome of our classification actually represents a viable way to predict phone usage's classifications with a

measurable likelihood and not the absolute certainty that a phone has been used to commit a crime. Of course, the problem becomes negligible with the increase of training instances. Finally, it is important to highlight that the followed approach, focusing on "cold" data, is basically different from the live forensic triage described in [9] since the last is performed on the crime scene by mobile phone's technicians with basic skills.

## REFERENCES


[1] Wayne Jansen and Rick Ayers, "Guidelines on Cell Phone Forensics", NIST, May 2007.

[2] Gartner – 11 August 2011. "Gartner Says Sales of Mobile Devices in Second Quarter of 2011 Grew 16.5 Percent Year-on-Year; Smartphone Sales Grew 74 Percent".

[3] Polizia Postale e delle Comunicazione (2011). Data on complaints investigation and action in 2010. pp. 4 – 5.

[4] "Cybercrime: Italia quarta in Europa. Assintel: manca una cultura della sicurezza e un approccio legislativo equilibrato alla web society". (2011, March).
From: http://www.tecnomagazine.it/tech/1758. Last view july 2011.

[5] Access Data, "AD Triage 1.0.0". Relased in April 12, 2011.

[6] Dell Mobile Forensics Unit.
From: http://gcn.com/Articles/2011/04/07/Dell-Spektor-Digital-Mobile-Forensics, last view july 2011.

[7] NIST. (March 2007). "Cell Phone Forensic Tools: An Overview and Analysis Update".

[8] F. Marturana, R. Bertè, G. Me, S. Tacconi. Mobile Forensics "triaging": new directions for methodology. Being published in itAIS 2011 VIII Conference of the Italian Chapter of AIS.

[9] Richard P. Mislan, Eoghan Caseyb and Gary C. Kesslerc "The growing need for on-scene triage of mobile devices". Digital Investigation, Vol. 6, Issues 3-4, May 2010, pp. 112-124.

[10] Robert J. Walls Erik Learned-Miller Brian Neil Levine. "Forensic Triage for Mobile Phones with DEC0DE".

[11] "The American Heritage Dictionary of the English Language" - 4th Edition. (2000). Triage. Boston: Houghton Mifflin.

[12] Cios, K.J., Pedrycz, W., Swiniarski, R.W., Kurgan, L.A. "Data Mining. A Knowledge Discovery Approach". Springer, 2007, XV, 606 p.

[13] Ian H. Witten, Eibe Frank, Mark A. Hall. "Data Mining Practical Machine Learning Tools and Techniques" 3rd Edition- Elsevier.

[14] Remco R. Bouckaert. "Bayesian Network Classifiers in Weka for Version 3-5-7". May 12, 2008.

[15] Eibe Frank, Mark Hall, Bernhard Pfahringer: Locally Weighted Naive Bayes. In: 19th Conference in Uncertainty in Artificial Intelligence, 249-256, 2003.

[16] Ross Quinlan (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.