

Integrating forensic information in a crime intelligence database

Quentin Rossy ^{a,*}, Sylvain Ioset ^{a,b}, Damien Dessimoz ^b and Olivier Ribaux ^a

^a Université de Lausanne, Institut de Police Scientifique, Batochime, 1015 Lausanne-Dorigny, Switzerland.

^b Police cantonale vaudoise, route de la Blécherette 101, 1014 Lausanne, Switzerland.

* Corresponding author. Tel.: +41 21 692 46 13; fax: +41 21 692 46 05. E-mail address: quentin.rossy@unil.ch (Q. Rossy).

Abstract

Since 2008, Intelligence units of six states of the western part of Switzerland have been sharing a common database for the analysis of high volume crimes. On a daily basis, events reported to the police are analysed, filtered and classified to detect crime repetitions and interpret the crime environment. Several forensic outcomes are integrated in the system such as matches of traces with persons, and links between scenes detected by the comparison of forensic case data. Systematic procedures have been settled to integrate links assumed mainly through DNA profiles, shoemarks patterns and images.

A statistical outlook on a retrospective dataset of series from 2009 to 2011 of the database informs for instance on the number of repetition detected or confirmed and increased by forensic case data. Time needed to obtain forensic intelligence in regard with the type of marks treated, is seen as a critical issue. Furthermore, the underlying integration process of forensic intelligence into the crime intelligence database raised several difficulties in regards of the acquisition of data and the models used in the forensic databases. Solutions found and adopted operational procedures are described and discussed. This process form the basis to many other researches aimed at developing forensic intelligence models.

Keywords

Forensic intelligence, forensic case data, crime analysis, crime intelligence, database

Introduction

Forensic processes are traditionally conceived to support and follow the investigative process from the crime scene to the trial on a case-by-case basis. Nevertheless forensic case data and the results of their analysis convey important but often underestimated information to support the detection of crime repetitions and understand the size, extent, and evolution of crime phenomena [1], [2], [3]. Measuring how forensic case data contribute to

the understanding of crime problems is a complex tasks [4]. Indeed, the performance of forensic science is more often than not evaluated in terms of crime detection rates or number of identifications [5]. But forensic outcomes can be assessed by considering many other aims like linking crimes, excluding suspects, impacting on the duration of investigations or even reducing the fear of crime (the perception of public safety) [6]. In this article, we focus on the impact of forensic case data to the sustained analysis of crime series mostly, but not exclusively, for high volume crime.

Within intelligence units, crime analysis proceeds through the daily interpretation of crime data coming from new reported cases. They mainly assume the repetitive activity of offenders or group of offenders on the basis of circumstantial data, modus operandi, vehicles, images and others useful situational information collated in separated and dedicated databases. The integration of accurate, timely and useful information produced by the analysis of forensic case data into this intelligence process, what we call forensic intelligence [7], is assessed and discussed in this article. This contribution of forensic outcomes to the detection and management of crime series is evaluated on the basis of a retrospective dataset extracted from a common database shared by six police forces covering the western part of Switzerland.

The next section addresses the integration of how forensic information has been integrated into this intelligence database. The management of links between crime events through a dedicated model, embedded in the structure of the database, is then explained. The dataset studied is described in the third section.

Results are presented and discussed through three specific questions: *How many series are detected by forensic case data ? How many forensic links detect, increase or confirm series? How long does it take to detect series and what is the impact of the term to integrate links?*

Integration of forensic information into the intelligence database

Swiss police forces are organized on three levels, represented by city, state and federal police. The federal police is responsible to conduct investigations mostly in fields related to organised crime, provide international single point of contacts and manages national databases (such as AFIS and DNA databases). 26 state police and 2 city police, of very different sizes, are in charge of the investigations of all the remaining types of crime. Each jurisdiction has a crime analysis team dedicated to the sustained monitoring and analysis of repetitive crimes, mostly high volume crime. Since 1994, four years after the creation of the first crime analysis unit in one state, a regional approach has been settled in the French and Italian speaking parts of Switzerland in order to coordinate intelligence efforts. Intelligence

units from the seven states have been grouped into a regional analysis centre, called CICOP (an acronym in French stating the coordination of intelligence effort for operational and preventative efforts) [8] Similar structures have been developed later across the other parts of the country.

Since 2008 a common interstate information platform has been implemented, collating automatically information from the manifold databases located in the six French-speaking states. It is now reachable by all the analysts across police forces. This development was a real challenge as these police are of very different sizes, and cover various types of territories. They are structured differently, and have developed their own computer infrastructure. Legal challenges had also to be overcome.

This shared intelligence database has been developed according to a common methodology that has been devised by crime analysts and continue to iteratively develop over time [8]. The implemented intelligence process collects information from criminal events, investigations and other sources of data and integrates them into a memory. This memory is specially designed and organized to support various analytical processes. The delivery of targeted products contributes to operational and strategic decisions.

The aim of the platform is to support the follow up of crime phenomena and to detect specific crime series. In this article we will use the term series for cases that are assumed to be perpetrated by the same offender, or by offenders belonging to a same group of criminals. In order to detect and assume the actuality of a series, several types of links between cases are integrated in the database. They are generated by the comparison of situational (e.g. MO, loot, spatiotemporal) and forensic information (e.g. DNA, shoemarks, images), as well as the analysis of stolen and recovered vehicles. In this context, links are considered as an aid to interpret globally the crime environment for intelligence purpose or for investigations, but not as proof of common source dedicated to a Court [1], [9]. Crime events data are integrated on a daily basis by each state unit and classified according to a harmonized doctrine including modus operandi, loot, spatiotemporal information and a dedicated classification system of events. The information flow is described in Figure 1.

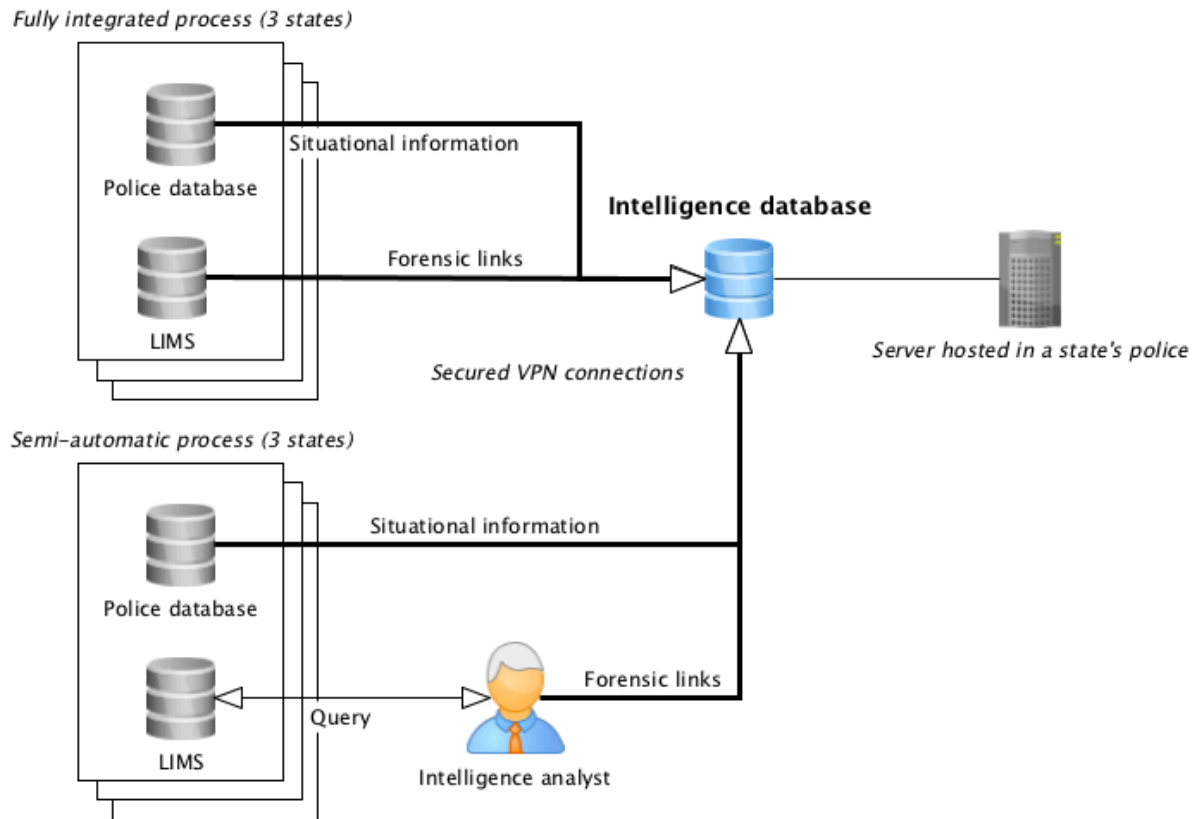


Figure 1 : technical setup of the system. Each state's police database automatically feeds the shared intelligence database through a secured VPN connection. Forensic links are extracted from LIMS, with an integrated interface for three states or communicated through other channels to intelligence analysts (emails, meetings, etc.).

DNA

Links between crime scenes are detected through the centralised Swiss DNA database. When a scene-to-scene DNA match is found, messages are sent to the forensic unit of each state concerned. Depending on a procedure designed in each state, links are then integrated into the shared intelligence database by the forensic unit itself or by the crime analysis unit upon notice from the forensic unit.

Time span between the detection of the mark and the integration of a link in the database is critical. It notably depends on when the decision to send the specimen for profile extraction is taken and on the response time from the laboratory. The later is regulated by an ordinance¹ and must be about three weeks, potentially less in cases of serious crime. The response time

¹ RS 363.11 - Ordonnance du DFJP sur les exigences de prestations et de qualité requises pour les laboratoires forensiques d'analyse d'ADN (Ordonnance du DFJP sur les laboratoires d'analyse d'ADN) 29.06.2005

for the analysis of samples taken from a suspect is shorter (less than 10 days). DNA laboratories suffer in many countries from much longer backlogs that can reduce their relevance for contributing to such intelligence processes.

Shoemarks

Scene linking is conjointly performed by the use of shoemark databases located in each state's forensic unit. Links are generated by the use of a classification system of shoe sole patterns. The comparison process must be rapid and simple in order to absorb the flow of information. It works roughly through a first selection of similar shoe sole patterns from the databases, and then by visual comparison, as no automatic matching algorithm is used [10]. Due to the fragmentary nature of shoe marks, the systematic comparisons search for same type of sources, rather than for the same source. When similar patterns are detected on different scenes, similarity of possible type(s) of source(s) is thus assumed, showing a proximity between the marks that helps to generate hypotheses about links [11].

Because the uncertainties prevailing during this process, the rarity of the pattern and known situational information about the cases are also taken into account to infer link's strength from this classification process. In particular, this evaluation is performed according to the structure of crime events during the time period of the crimes when the sole marks have been collected. Even if links based on shoemarks patterns are mostly weak, they are integrated for intelligence purpose to help detect series. Further comparisons and refinement are carried out later when crime series develop and necessitate some crystallisation or shoes of suspects are available and lead to an evaluation that may reach a Court.

Since forensic databases are handled at the state level, regional meetings are organized every two or three months by crime scene examiners to compare marks across jurisdictions. Between-states (regional) links are then identified and stored in a dedicated database. This system has been developed by the crime analysis unit in order to facilitate automatic importations in the shared intelligence database after each meeting. In parallel, several states integrate almost in real-time (see results below) the links identified within their own database (state level). The integration procedures vary between each state depending on their specific organisation: some forensic units have a direct access to the shared intelligence database, others send listings to intelligence units and, in some states, analysts have access to the forensic databases. For different reasons ranging from their degree of intelligence and forensic awareness to political priorities, only three of the six states introduce systematically these links in the database. Other states integrate them selectively depending on the significance of specific cases. This has been the best consensus that has been found

till now, but the methodological framework tends to aspire progressively every state in participating at the same level.

Images

Depending on the crime type, several kind of images are integrated in the database. When banks, ATMs or financial organizations are targeted, CCTV images are directly obtained in collaboration with security agencies and magistrates. In commercial locations (shops, petrol station, etc.) no such systematic procedure exists. The collection of images is done by the police officer in charge on a case-by-case basis. Images taken by victims or witnesses with camera or phone, from speed trap or from the Internet may also be integrated. Delay of integration may then greatly vary or images may not be sent to the intelligence unit or may even not be collected at all.

The comparison of images to detect links is mainly visual and manual. It is above all based on facial and garments information. A dedicated interface is built in the intelligence database to facilitate the process. Furthermore, the classification system implemented in this database greatly facilitates the selection of relevant images for comparison.

Scene-to-scene matches are not only detected by analysts through the database, but also by other police officers since images are often disseminated across police forces. Also for image, only qualitative evaluation of the strength of the links is performed in this intelligence framework, as the quality of images and the flow of information do not allow it systematically. But if a series need crystallisation or a suspect is arrested, a full evaluative process occurs.

Fingermarks

No systematic process of linking cases is implemented with fingermarks through the national AFIS database. Only relations between marks and prints from a know suspect are integrated in the intelligence database. However, comparison of fingermarks may be performed on a case-by-case basis, if this type of comparison is judged to be relevant.

Earmarks, glovemarks and toolmarks

Each forensic unit at the state level manually performs earmarks, glovemarks and toolmarks comparisons. Except for particular cases, no systematic regional comparison is performed. In comparison with shoemarks, DNA and Images, these types of marks are less often detected on crime scenes. This could be due notably to the specific crime situations and modus operandi from which they result, but also because no systematic comparison process is implemented for them. As a consequence, very few links based on earmarks, glovemarks and toolmarks are detected (see description of dataset below). Furthermore no systematic procedures are implemented for their integration in the intelligence database. Nonetheless it

happens that during short periods of time, crime scene examiners alert about particular use of types of gloves or tools, or similarities between earmarks collected. This disparate links may drive the analysis units in their crime series detection process. The structure of the intelligence database is devised in order to accept them.

Structure of the memory and counting of links

Of importance is to consider that the database must be simple, input of data rapid, and it must conform to specific organisations that separate forensic from intelligence unit. The fluidity of the process is critical to its usability and performance. Many choices about the devise of the database are derived from these constraints.

Forensic case data, excepting images, are not directly integrated into the database. Indeed, links detected by separated forensic comparisons processes are stored.

Figure 2 shows the model designed to handle links between events (forensic and situational links). It is only a part of the whole database schema underlying the information platform.

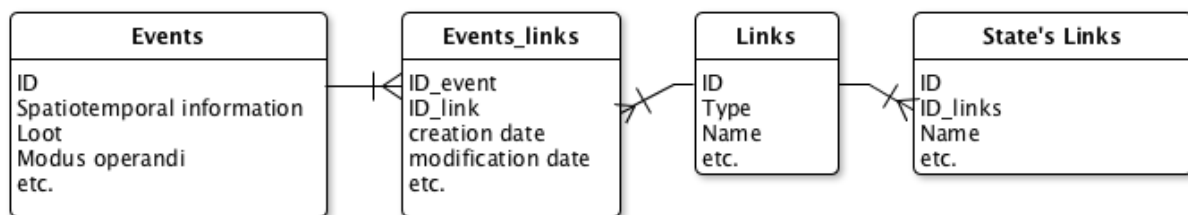


Figure 2 : database design applied to manage links between events.

The main table of the database stores data on the events with most of the information in the form of both codified (e.g. MO, location's type, etc.) and free text fields (e.g. for description, loot, etc.). Links between events are stored in a separate table called "Links". Each link has a specific type: MO, loot, spatiotemporal, series, operations, DNA, shoemarks, earmarks, toolmarks and images. For the purpose of this article we consider all links that are not based on forensic information as grouped into one single category called "situational links", since they are based on environmental information surrounding the cases and modus operandi. The "Events" and "Links" tables have a many to many relationship (Events_links table) since an event may have several links and links are obviously related to many events. A complementary table is used (State's Links) to store the specific names of links used by each state. Indeed, except for DNA links that are handle at the national level, all other forensic links are detected at the state level. When a regional (cross states) link is detected, links are merged and a new name is created, but state's names are stored in this dedicated table, as the process must respect state's specificity.

Series of events are reconstructed with all links between events (see Figure 3). A series is thus the overall set of events linked with all types of links registered. To describe each subset of events linked by one specific type of link, the term “group” (of events) is used. Obviously, these groups overlap, and the junction of them is what we call the series.

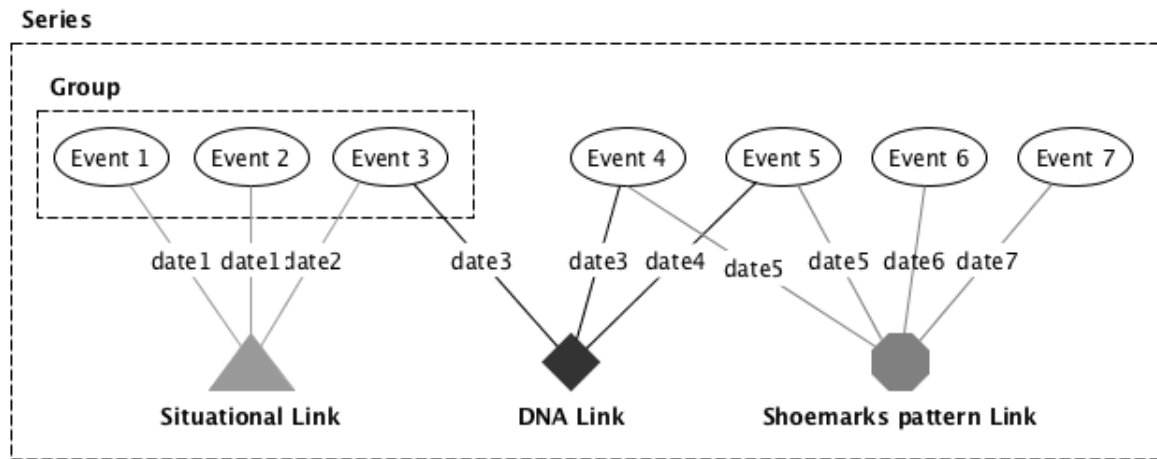


Figure 3 : reconstruction of series and counting of links

In this article, we count one link for each instance in the “Links” table regardless of the number of events connected by each type of link. Thus links are not counted based on mark-to-mark hits, but rather the number of forensic links expresses the number of DNA profiles, shoemarks patterns or same suspect on images. For instance, the series in Figure 3 is composed of 7 events and 3 links. The “Events_links” table is used to count the number of events linked with a specific link. A new link can have one of three functions:

1. detect: events that were considered isolated previously are connected for the first time; a new series is detected;
2. increase: new events are integrated in an already known series;
3. confirm: the events belonging to the same series are linked again by another type of link.

In order to identify the functions of the links, the sequence of their integration is critical. The creation date of the link in the table “Events_links” is used to determine the temporal order in which the groups have been built. If a specific link was the first created in the series of events, we consider it has detected the series (date 1 of the situational link in Figure 3). If the link is the first for the event or the only one, we consider it increases the series (for instance date 2 inform that event 3 increase the series based on the situational link). If the link was posterior to another link, we consider it confirms the integration of the case into the series (for instance the shoemarks pattern link confirms the link between event 4 et event 5).

Description of the dataset

In order to evaluate the role of forensic information into the intelligence process, we have selected a subset of series from the database. All the links between crime events introduced between 2009 and 2011 have been taken into account to build the dataset.

Link types	2009	2010	2011	All
Situational Information	299 (57.1%)	432 (53.6%)	580 (53.4%)	1273 (54.6%)
Shoemark pattern	134 (25.6%)	236 (29.3%)	316 (29.1%)	676 (29%)
DNA	77 (14.7%)	99 (12.3%)	102 (9.4%)	244 (10.5%)
Image	12 (2.3%)	37 (4.6%)	81 (7.5%)	127 (5.4%)
Earmarks	1 (0.2%)		8 (0.7%)	9 (0.4%)
Glovemarks	1 (0.2%)			1 (0%)
Toolmarks		2 (0.2%)		2 (0.1%)
Total	524	806	1087	2332

Table 1 : links in the dataset (N = 2343 links). Percentages are calculated for each year.

No dedicated analysis has been done for earmarks, glovemarks and toolmarks since very few of them are integrated. However they were used to build the series and participate to the description of them. The dataset contains mainly cases from 2009 and 2011, but they have brought previous cases, when links transcend this period. In turn other 55 older links have been selected. They have been eliminated from the calculations of number of links, but, as for anecdotic types of links, they have been used to build the series.

Table 2 shows the number of series according to their size (in number of events) and the total number of events for each size group (including cases prior to 2009).

Sizes of series	Number of series	Ratio of series	Total number of events	Ratio of events
2	779	49.7%	1'558	17.0%
3 - 10	645	41.2%	2'881	31.5%
11 - 20	71	4.5%	989	10.8%
21 - 30	24	1.5%	585	6.4%
31 - 40	18	1.2%	645	7.1%
41 - 50	10	0.6%	447	4.9%
51 - 60	2	0.1%	113	1.2%
61 - 70	2	0.1%	127	1.4%
71 - 80	3	0.2%	227	2.5%
81 - 90	3	0.2%	254	2.8%
91 - 100	2	0.1%	194	2.1%
> 100	8	0.5%	1'130	12.4%
Total	1567	100.0%	9'150	100.0%

Table 2 : series in the dataset (N = 1567 series). Overall 20% of the biggest series contains 65 % of all linked cases.

Table 3 below presents the number of crime events belonging to series in the dataset. The total number of events for the period is also shown. The dataset of assumed crime series includes 6.9% of all crime events in the database for the period considered.

Event type	2009-2011		Ratio of events	before 2009
	Total	In series		In series
Burglary	60301	5400	9.0%	679
Theft from a vehicle	13044	415	3.2%	85
Larceny	10771	270	2.5%	15
Pickpocketing	8310	217	2.6%	0
Distraction theft	4535	581	12.8%	40
Shoplifting	3307	51	1.5%	0
Racketeering	2513	43	1.7%	4
Vehicle theft	2432	207	8.5%	20
Mugging	2264	32	1.4%	0
Aggression - Battery - Brawl	2027	5	0.2%	0
Fraud	1870	151	8.1%	0
Vehicle plates stolen	1690	94	5.6%	3
Sexual offense	1508	29	1.9%	0
Paying machine breaking	1030	259	25.1%	3
Card Fraud	993	162	16.3%	2
Arson	960	64	6.7%	2
Counterfeit banknotes	886	96	10.8%	0
Robbery	572	128	22.4%	8
Waiter purse theft	447	12	2.7%	1
Cloakroom / locker breaking	391	23	5.9%	0
Property damage	281	40	14.2%	4
Chiselling	138	2	1.4%	0
Murder	74	3	4.1%	0
Total	120344	8284	6.9%	866

Table 3 : events in the dataset, only « In series » events are used (N = 9150 events).

Of importance is to note that the total amount of events presented isn't the total amount of crime events for the time period. Indeed, crime analysts filter some events that are irrelevant to serial crimes analysis. For instance, some sexual and violence crimes where the perpetrator is known as a member of the family of the victim may not be included. The decision to integrate some crime types may also vary between states.

Results

How many series are detected by forensic case data?

Detection	Number of series			
	2009	2010	2011	Total
Situational information	230	298	446	974 (62.2%)
Shoemarks pattern	76	127	165	368 (23.5%)
DNA	39	42	47	128 (8.2%)
Image	9	28	56	93 (5.9%)
Earmaks	1		3	4 (0.2%)
Total	355	495	717	1'567 (100%)

Table 4 : amount of series detected for each link types.

Results presented in Table 4 seem to be directly proportional to the number of links integrated into the database. Indeed, the database contains 29% of shoemarks patterns links, 10.5% of DNA links and 5.4% of links based on images (see Table 1).

Globally, 37.8% of all series registered are initially detected with forensic information. A direct comparison of the detection potential of each type of marks may not be relevant because events linked with shoemarks, DNA or Images are not all from the same type. The number of series detected for each crime type is presented in Table 5. If a link is detected between different crime types, the link is counted for each crime type.

Crime types	Number of series				Total
	Situational Information	Shoemark pattern	DNA	Image	
Burglary	588 (55.5%)	363 (34.3%)	107 (10.1%)	2 (0.2%)	1060
Theft from a vehicle	71 (83.5%)	3 (3.5%)	9 (10.6%)	2 (2.4%)	85
Robbery	32 (80%)	3 (7.5%)	2 (5%)	3 (7.5%)	40
Property damage	2 (22.2%)	1 (11.1%)	6 (66.7%)		9
Vehicle theft	65 (97%)	1 (1.5%)	1 (1.5%)		67
Aggression - Battery - Brawl	3 (75%)	1 (25%)			4
Waiter's purse theft	4 (57.1%)		3 (42.9%)		7
Arson	10 (76.9%)		3 (23.1%)		13
Sexual offense	9 (81.8%)		2 (18.2%)		11
Paying machine breaking	17 (89.5%)		1 (5.3%)	1 (5.3%)	19
Distraction theft	72 (67.3%)			35 (32.7%)	107
Pickpocketing	33 (56.9%)			25 (43.1%)	58
Larceny	62 (76.5%)			19 (23.5%)	81
Card Fraud	16 (66.7%)			8 (33.3%)	24
Fraud	31 (83.8%)			6 (16.2%)	37
Shoplifting	15 (75%)			5 (25%)	20
Cloakroom / locker breaking	3 (60%)			2 (40%)	5
Counterfeit banknotes	13 (92.9%)			1 (7.1%)	14
Vehicle plates stolen	33 (100%)				33

Racketeering	17 (100%)	17
Mugging	4 (100%)	4
Chiselling	1 (100%)	1
Murder	1 (100%)	1

Table 5 : Number of series detected according to types of crime and links

These results show that the combination of all forensic links seems complementary. For instance, events linked with images are mainly not the same as events linked with shoemarks or DNA. Furthermore, very few events are linked with several forensic outcomes (see Table 6).

Links types	Number of events			
	Situational Information	Shoemarks pattern	DNA	Image
Situational Information	6758	425 (22.5%)	311 (32.4%)	84 (21.9%)
Shoemarks pattern		1892	73	1
DNA			959	3
Image				383

Table 6 : number of events linked with different types of link.

The 73 events linked with both DNA and shoemarks are burglaries, except one vehicle theft. The event linked with a shoemark and an image is also a burglary. Two of the three events linked with DNA and images are robberies and the last is a card fraud.

The ratio of events linked with both forensic case data and situational information is nearly similar for shoemarks and images, but higher for DNA. But how many of these links have been integrated previously?

How many forensic links detect, increase or confirm series?

Regardless of the fact that the amount of forensic links introduced in the database has increased from 2009 to 2011 (see Table 1), the percentage of links that detect or increase series remains almost constant across the years. The results for the whole period are summarized in Tables 7 and 8.

Link types	Number of links			
	Detect	Detect and/or increase	Only confirm	Total
Situational Information	949 (74.6%)	1236 (97.1%)	37 (2.9%)	1273 (100%)
Shoemarks pattern	366 (54.1%)	626 (92.6%)	50 (7.4%)	676 (100%)
DNA	119 (48.8%)	216 (88.5%)	28 (11.5%)	244 (100%)
Image	92 (72.4%)	123 (96.9%)	4 (3.1%)	127 (100%)

Table 7 : potential of forensic information to detect, increase or confirm series. Percentages are calculated for each link type.

Nearly the same proportion of DNA and shoemarks patterns links allows the detection of new series (around 50%). A bigger proportion of image links detect series (72.4%), which is similar to the ratio of situational information links. DNA links have the biggest proportion of links that only confirm already detected links (11.5%).

Since one forensic link may link two or many events, calculations have also been done according to the total amount of events linked (see Table 8).

Link types	Number of linked events			
	Detect	Increase	Confirm	Total
Situational Information	1873 (20.5%)	4440 (48.7%)	920 (10.1%)	6758 (74.1%)
Shoemarks pattern	584 (6.4%)	1051 (11.5%)	420 (4.6%)	1892 (20.8%)
DNA	186 (2.0%)	546 (6.0%)	251 (2.8%)	959 (10.5%)
Image	162 (1.8%)	181 (2.0%)	59 (0.7%)	383 (4.2%)

Table 8 : potential of forensic information to detect, increase or confirm series. Percentages are calculated over all events from 2009 to 2011.

Globally, 29.7% of all events are linked (detect or increase series) with forensic information and only 8.1% of linked events are confirmed with forensic information. This result shows the great impact that forensic case data have on the detection and follow-up of series of events and also confirm the complementarity of the different types of marks. Nevertheless it is not possible to infer that all of these events wouldn't have been linked later based on other information.

How long does it take to detect series and what is the impact of the term to integrate links?

To test the hypothesis that forensic information help significantly to link cases, which would be hardly linked only on the basis of situational information, time periods between the occurrence of cases and when links are registered is calculated. This analysis is also used to test another hypothesis: the longer the delay to link case, the lesser the chance to detect series.

To test these hypothesis, the time spans, expressed in weeks between the date of each crime events firstly linked in each series and the date of creation of the link where computed and compared for each link types (see Figure 4). The exact date of the event is not always known. In order to simplify, we have chosen the latest date, representing most often when the victim discovered the crime event.

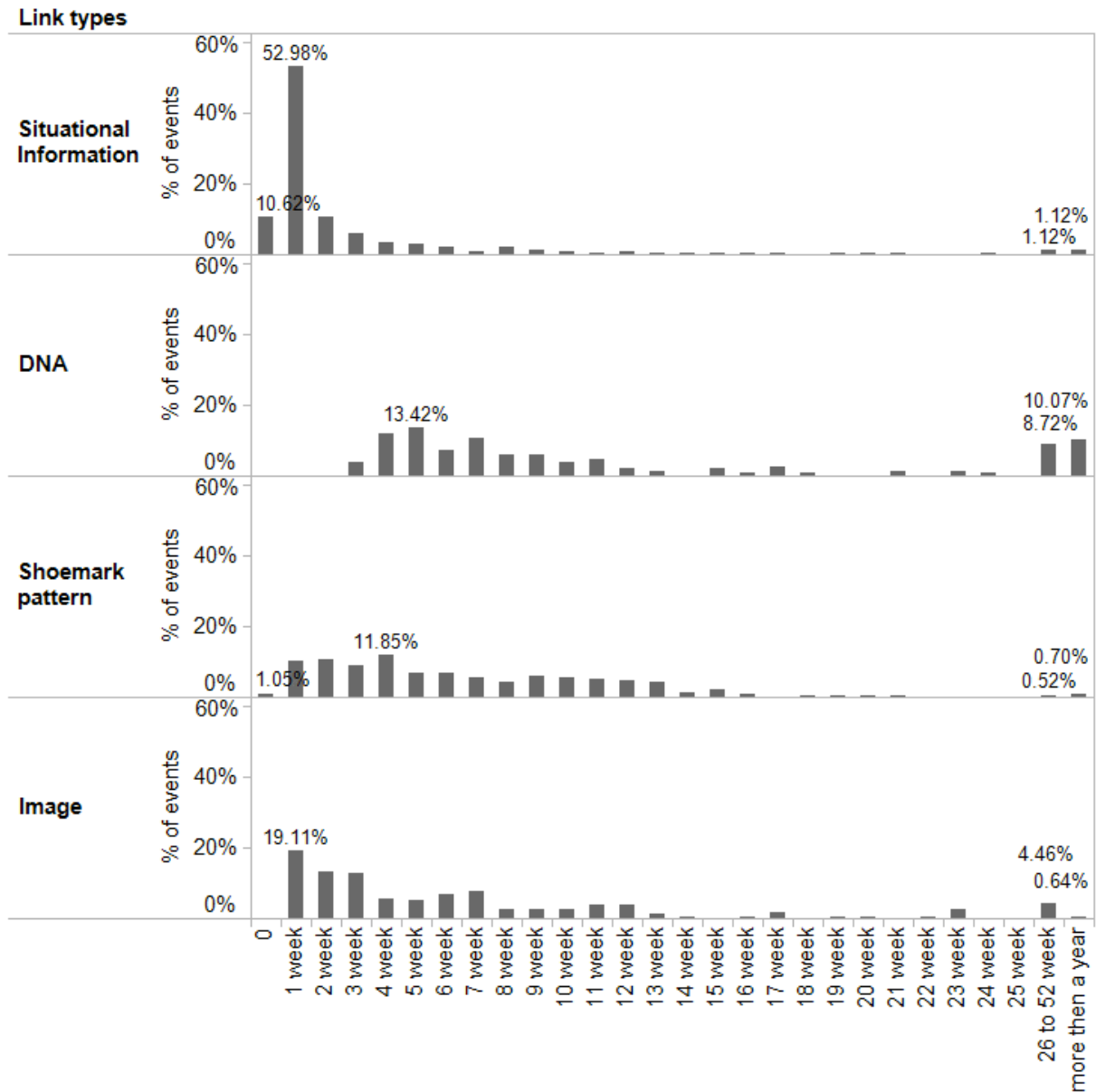


Figure 4 : Time (in weeks) between end date of events that detect the series and the date of creation of the link in the database.

More than 63% of events linked with situational information are registered during the first week after the occurrence of cases and 80% during the first three weeks. In comparison 96% of events are linked with DNA, 69% with shoemarks patterns, and 55% with images more than three weeks after the occurrence of cases. After three month, 95% of all events linked with situational information have been registered. But remaining 29.5% of DNA related events, 4% of shoemarks pattern and 14% of images have still not be detected at this point. This result tends to show that a substantial proportion of links detected with DNA wouldn't be detected with situational information. This is also true for an important part of events linked with shoemarks patterns and images.

Furthermore, only 2% of linked events with situational information are detected after 6 month, whereas this proportion increases to 19% with DNA and 5.1% with images. This result shows the particular potential of DNA to detect series long after events occurs and in a lesser extent with images. As already noted, time needed to detect links between images may be explained by the time to obtain the images.

Finally, since shoemarks patterns links are integrated by several processes in the database (see section “integration of forensic information into the intelligence database”), a comparison is made between links introduced at the state level and links introduced at the regional level (i.e. after forensic scientists of each state have exchanged shoemarks to detect regional links).

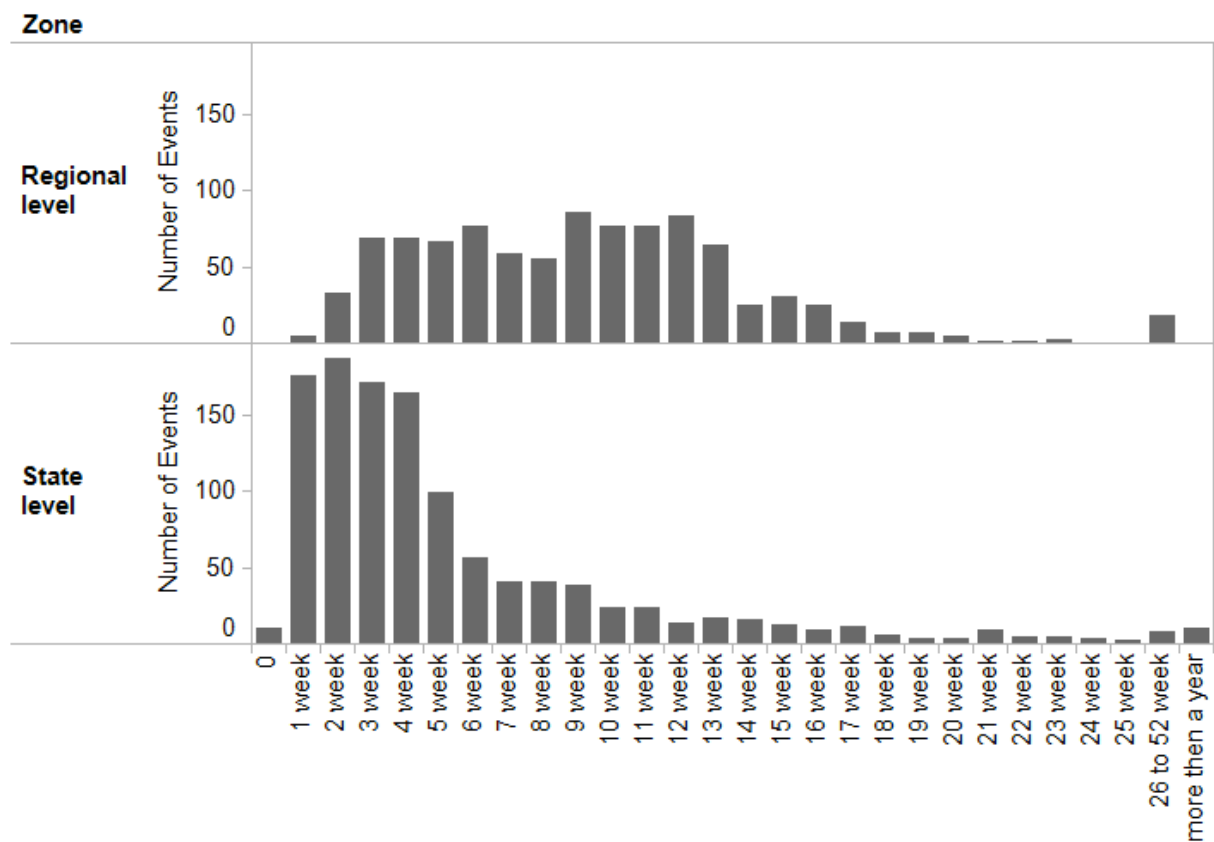


Figure 5 : Time (in weeks) between end date of events and the date of insertion of the link.

Almost 70% of links are introduced during the five weeks after events occurs at the state level, but only 25% of links detected at the regional level (see Figure 5). Nevertheless, the percentage of links that detect series is almost the same (near 54%) either at state or regional level. Furthermore, the percentage of links that only confirm series is lower at regional level (see Table 9).

Spatial extend	Number of shoemarks pattern links			Total
	Detect	Detect and/or increase	Only confirm	
State level	154 (53.8%)	249 (87.1%)	37 (12.9%)	286 (100%)
Regional level	212 (54.4%)	377 (96.7%)	13 (3.3%)	390 (100%)

Table 9 : difference between shoemarks pattern links introduced at state level and regional level.

These results show that the celerity of introduction of forensic links in the database is not the only important factor. Despite the fact that the detection of shoemarks pattern links at the regional level takes more time than at the state level, they have the same potential to detect series and have even a better potential to increase already detected series.

Since, the spatial dimension seems to be a critical issue to detect links, Table 10 presents a comparison of links detected between events of the same state and events from several states (regional links).

Link types	Number of links		Total
	Regional	State	
Situational Information	195 (15.3%)	1078 (84.7%)	1273 (100%)
Shoemark pattern	390 (57.7%)	286 (42.3%)	676 (100%)
DNA	137 (56.1%)	107 (43.9%)	244 (100%)
Image	75 (59.1%)	52 (40.9%)	127 (100%)

Table 10 : number of links detected between states.

Overall, almost 60% of all links detected with forensic case data are at the regional level, while only 15% of situational links are detected between states. An explanation can be that linking cases with situational information is more difficult across states. For instance, spatiotemporal relationships seem to be harder to detect. Indeed, the database contains 3 spatiotemporal links at regional level and 213 at state level. It is important to note here, that cases are introduced in the database by each state's crime intelligence unit. Linking cases with situational information is then much more easier at state level since analysts read and codify events from their states.

Discussion

Even if the presented results show the potential of forensic case data to detect and increase series, it is important to note that all kinds of relationships have been compared regardless of the underlying uncertainties of each type of links. Indeed, we have considered a series as a group of events assumed to have been perpetrated by the same offender or group of offenders. Series have been reconstructed based on all links introduced in the database. But the type of marks and their nature have an important impact on the degree of confidence of the detected link. For instances, shoemarks pattern inform of a possible same type of shoes

and not directly a possible common offender, images used to identify links may be blurred, links based on situational information may be assumed on the basis of very specific MO or more common ones, etc. In all cases, DNA offers the most certain relationships than other material evidences, than situational information (with the exception of cases committed the same day in the same place) and than behavioural aspects [11], [12].

Disparities observed between types of links may also depend on the nature of each mark. Shoemarks are visible at the crime scene, allowing crime scene officer to exclude shoes of victims. If attendant knew that a particular shoemarks pattern has already been collected on previous cases, he may recognise it and favour its collection [4]. In such cases, detection of links may be immediate and if a real-time procedure exists between forensic scientists and crime analysts the link may be quickly integrated in the intelligence database. Things are much more complicated with DNA since in many cases traces are latent (contact traces) and the detection of a common profile requires a longer analysis. Images show the biggest relative amount of links that detect series. They are complementary of DNA and shoemarks since they link other crime types.

During the three years covered by the dataset, the amount of links detected with images has increased more than links detected with shoemarks and DNA. This can be explained by the fact that shoemarks and DNA have been used to link cases since the early 2000s and systematics are well implemented. Each state's crime intelligence unit has increasingly performed the comparison of images over the last few years, due to the development of processes and the resulting implementation of functionalities in the platform to manage and compare images.

Globally, the potential of forensic outcomes to detect or increase series is certainly underestimated in this study. We can indeed assume that there are many more links between cases than those who are detected. Several arguments support this hypothesis. First of all, do scene of crime officers and forensic scientists fully search and exploit forensic case data for intelligence purpose? Since the potential of DNA and shoemarks to link crimes is well known in Switzerland for many years, many efforts are done to use them for this purpose. Nevertheless, other marks like glovemarks, earmarks and toolmarks are much less systematically used to detect links. The difficulty of the comparison processes is probably the main explanation. However, similar approaches as the one used for shoemarks comparisons could be undertook to detect links between types of source if links at the source level are too hard to detect. New developments in automatic comparison systems may also facilitate the detection of links between cases [13]. We should note here that there is even an underestimation of the potential of shoemarks patterns to link cases in the results since not

every state integrates them systematically. This situation is the second reason why we assume that much more links should be detected. The overall process from the detection of forensic case data at the crime scene to the integration of a detected link in the database is complex and shall be improved.

Conclusion

The results suggest that forensic outcomes have a great potential to detect crime series. DNA and shoemarks mainly detect burglaries, while images are better at detecting series of distraction thefts, pickpocketing and larcenies. It is then worth relying on a diversified set of forensic case data to gain better insight on the different types of crimes series. The vast majority of events are linked through only one forensic link type (99.2%), further demonstrating the necessity to use all types of marks for a better detection of crime repetitions.

In the current functioning of the database, results show that 37.8% of all series are initially detected with forensic outcomes (mainly DNA profiles, shoemarks patterns and images). This amount seems to be directly proportional to the ratio of forensic links among all links integrated. DNA and, in a lesser extent, images offer a great potential to detect series long after events occurred. More often than not, links detected with forensic case data would not be detected with situational information (e.g. MO, loot, spatiotemporal data, etc.). These results show the great impact of the integration of forensic outcomes to detect and follow up crime series.

The integrated processes of shoemarks patterns at state level and regional level have been compared. It shows that the detection of shoemarks patterns links at the regional level takes more time than at the state level. Nonetheless they have the same potential to detect series. The regional level links have even a better potential to increase already detected series.

The celerity of introduction of forensic outcomes in the database is not the only important factor to strengthen the contribution of forensic case data for intelligence purpose. Although it requires more time, efforts to exchange and compare marks across jurisdictions seems particularly effective. Indeed, multiple states are represented in about 60% of links detected with forensic case data, while this ratio drops down to 15% for situational links.

The use of forensic case data for intelligence purpose should impact at a higher degree the design and architecture of systems dedicated to the sustained monitoring and analysis of repetitive crimes. To achieve this, it is necessary that the forensic analyses in remote laboratories and the routing of information are either accelerated or, at least, better

integrated in a workflow tailored for intelligence production. This has still not been fully reached in the current state of the system, indicating that the potential of forensic case data is even better than what the statistical data tend to show.

A more coherent articulation between forensic tasks and the functions of policing should be promoted, in particular for the management of crime scene and the selection of specimens to analyse. Moreover, the integration of forensic case data can supplement information on the structure of criminality and give more insight either for operational or strategic decisions.

Despite the fact that we have limited our results to the contribution of forensic case data to detect, increase, and confirm series, many others analyses could have been carried out, that will be the subject of further researches. For instance, analyses of the results presented in this article could be performed considering the total number of marks collected at crime scenes and the number of marks actually analysed. It may contribute to the assessment of forensic science efficiency in complement of other indicators like forensic identification rate.

Furthermore, a deeper analysis of the links based on situational information may inform on criteria used by crime analysts to decide to link events. The study of spatiotemporal links can answer questions about the distances in time and space considered close enough, or about the kinds of MO considered specific enough to link events. A global spatiotemporal analysis of the groups of cases linked with forensic case data can provide insight about the mobility of suspects. Cases linked with forensic case data can also be compared to test hypothesis on how offender's activities transcend types of crimes (polymorphism).

We assume that, beyond detecting series and impacting operational actions, forensic outcomes can give interesting strategic insight about the crime environment and should be integrated with criminological researches.

Acknowledgements

The authors would like to thank all the crime analysts and forensic scientists who work hard daily to detect series and follow the evolution of the crime environment. Without their work this study couldn't have been done. We hope that the results presented here will encourage them to get further involved in the use of forensic outcomes for intelligence purpose.

References

- [1] O. Ribaux and P. Margot, "Inference structures for crime analysis and intelligence: the example of burglary using forensic science data," *Forensic Science International*, vol. 100, no. 3, pp. 193–210, 1999.
- [2] S. J. Walsh, C. Roux, A. Ross, O. Ribaux, and J. S. Buckleton, "Forensic DNA Profiling - Beyond Identification," *Law Enforcement Executive Forum*, vol. 2, no. 3, pp. 13–21, 2002.
- [3] P. Margot, "Forensic science on trial - What is the law of the land?," *Australian Journal of Forensic Sciences*, vol. 43, no. 2–3, pp. 89–103, 2011.
- [4] O. Ribaux, A. Baylon, C. Roux, O. Delémont, E. Lock, C. Zingg, and P. Margot, "Intelligence-led crime scene processing. Part I: Forensic intelligence," *Forensic Science International*, vol. 195, no. 1–3, pp. 10–16, 2010.
- [5] J. Burrows and R. Tarling, "Measuring the impact of forensic science in detecting burglary and autocrime offences.," *Science & justice*, vol. 44, no. 4, pp. 217–222, 2004.
- [6] F. R. Bieber, "Turning Base Hits into Earned Runs: Improving the Effectiveness of Forensic DNA Data Bank Programs," *The Journal of Law, Medicine & Ethics*, vol. 34, no. 2, pp. 222–233, 2006.
- [7] O. Ribaux, A. Girod, S. J. Walsh, P. Margot, S. Mizrahi, and V. Clivaz, "Forensic intelligence and crime analysis," *Law, Probability and Risk*, vol. 2, no. 1, pp. 47–60, 2003.
- [8] O. Ribaux and S. Birrer, "Iterative development of co-operation within an increasingly complex environment. Example of a Swiss regional analysis centre," in *International Police Cooperation. Emerging issues, theory and practice*, F. Lemieux, Ed. Portland: Willan Publishing, 2010, pp. 81–100.
- [9] P. D. Martin, "National DNA databases—practice and practicability. A forum for discussion," *International Congress Series*, vol. 1261, pp. 1–8, Apr. 2004.
- [10] A. Girod, C. Champod, and O. Ribaux, *Traces de souliers*. Lausanne, Switzerland: Presses polytechniques et universitaires romandes, 2008.
- [11] O. Ribaux, S. J. Walsh, and P. Margot, "The contribution of forensic science to crime analysis and investigation: forensic intelligence," *Forensic Science International*, vol. 156, no. 2–3, pp. 171–181, 2006.
- [12] M. Lammers, W. Bernasco, and H. Elffers, "How Long Do Offenders Escape Arrest? Using DNA Traces to Analyse when Serial Offenders Are Caught," *Journal of Investigative Psychology and Offender Profiling*, vol. 9, no. 1, pp. 13–29, Jan. 2012.
- [13] S. Junod, J. Pasquier, and C. Champod, "The development of an automatic recognition system for earmark and earprint comparisons," *Forensic Science International*. 2012.