



Contents lists available at ScienceDirect

Science and Justice

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

Modelling crime linkage with Bayesian networks

Jacob de Zoete^{a,*}, Marjan Sjerps^b, David Lagnado^c, Norman Fenton^d

^a University of Amsterdam, The Netherlands

^b Netherlands Forensic Institute, The Netherlands

^c University College London, United Kingdom

^d Queen Mary University of London, United Kingdom

ARTICLE INFO

Article history:

Received 23 May 2014

Accepted 20 November 2014

Available online xxxx

Keywords:

Crime linkage

Serial crime

Bayesian networks

Combining evidence

Case linkage

ABSTRACT

When two or more crimes show specific similarities, such as a very distinct *modus operandi*, the probability that they were committed by the same offender becomes of interest. This probability depends on the degree of similarity and distinctiveness. We show how Bayesian networks can be used to model different evidential structures that can occur when linking crimes, and how they assist in understanding the complex underlying dependencies. That is, how evidence that is obtained in one case can be used in another and vice versa. The flip side of this is that the intuitive decision to “unlink” a case in which exculpatory evidence is obtained leads to serious overestimation of the strength of the remaining cases.

© 2014 Forensic Science Society. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Suppose that two similar burglaries occur in a small village within a small time span. In the second one, a suspect is identified. The question whether this person is also responsible for the first crime arises. Clearly this depends on possible incriminating or exculpatory evidence in this first case, but also on the degree of similarity between the two burglaries. Several interesting questions arise in such common situations. For instance, can one “re-use” evidence incriminating the suspect in the second case as evidence in the first case? How does the evidence “transfer” between the two cases? How does the degree of similarity between the two cases affect this transfer? What happens when the evidence in the two cases partially overlaps, or shows dependencies? How can we make inferences for more than two cases?

In practice, it is generally assumed by the police, prosecution and legal fact finders that when there are two or more crimes with specific similarities between them there is an increase in the belief that the same offender (group) is responsible for all the crimes. The probability that there is only one offender (group) depends on the degree of similarity between the crimes. Even for a small number of crimes, the probabilistic reasoning rapidly becomes too difficult. In such situations it is recognized that a Bayesian Network (BN) model can help model the necessary probabilistic dependencies and perform the correct probabilistic inferences to evaluate the strength of the evidence [1]. We can

use BNs to examine how evidence found in one case influences the probability of hypotheses about who is the offender in another case.

In this paper, we will show how BNs can help in understanding the complex underlying dependencies in crime linkage. It turns out that these complex dependencies not only help us understand the impact of crime similarities, but also produce results with important practical consequences. For example, if it is discovered that in one of the similar crimes the suspect is not involved, then simply discarding that crime from the investigation could lead to overestimation of the strength of the remaining cases due to the dependency structure of the crime linkage problem. Hence, the common procedure in law enforcement to select from a series of similar crimes only those cases where there is evidence pointing to the suspect and disregard the other cases and evidence can be misleading. Our analysis thus extends the analysis of Evett et al. [2].

The notion of ‘crime linkage’ may be perceived and dealt with differently at different levels in the judicial process. During investigation (i.e., not at trial) considerations are typically very broad and connections among crimes may be made on other criteria than probability. In this paper, we focus on understanding the underlying logic regarding crime linkage. The examples we present serve as “thought experiments”. Such experiments are commonly used in mathematics to focus on the logic of the argumentation. In a thought experiment, a simple situation is considered that may not be very realistic but which contains the essence of the problem, showing the most important arguments. In reality all sorts of detail will complicate the problem but the essence will remain the same. Thus, although the model does not incorporate all the difficulties involved when dealing with crime linkage in

* Corresponding author.

E-mail addresses: j.c.dezoete@uva.nl (J. de Zoete), m.sjerps@nfi.minvenj.nl (M. Sjerps), d.lagnado@ucl.ac.uk (D. Lagnado), n.fenton@qmul.ac.uk (N. Fenton).

practice, it can highlight flaws in the reasoning and create a better understanding of the main line of reasoning.

The paper is structured as follows: In [Section 2](#) we present a selection of the relevant literature and state of the art on crime linkage. In [Section 3](#) we will model different situations in crime linkage using BNs, starting with the simplest example of two linked cases. We introduce and extend, step-by-step, to a network with three cases in [Section 4](#) where the evidence is directly dependent on each other (the extension to more cases is presented in the Appendix). In [Section 5](#) we discuss our conclusions and give some ideas for future research.

2. Literature and state of the art on crime linkage

Crime linkage is a broad topic that has been extensively reported (for example in [\[3–5\]](#)). Here we focus only on two aspects of the literature that are relevant for our analysis, namely: (1) how to identify linked cases, and (2) how to model crime linkage. We discuss a (non extensive) selection of some key papers on these topics.

2.1. Literature on how to identify linked cases

For identifying linked cases, it is necessary to assess how similar two crimes are, how strong the link between the cases is and how sure we are that the offender in one case is also the offender in another case.

The authors of [\[6–8\]](#) investigate the behavioural aspects of sexual crime offenders in solved cases. These studies concentrate on the consistency of the behaviour of serial sexual assault offenders. The authors conclude that certain aspects of the behaviour can be regarded as a signature of the offender. These aspects can be used to identify possibly linked crimes.

The notion of such a 'signature' is discussed by Petherick in the chapter *Offender Signature and Case Linkage* [\[9\]](#). It is noted that a signature in criminal profiling is a concept and not a 'true' signature. A signature may suggest that it is unique, whereas in criminal profiling it can only serve as an indication of whether or not two or more crimes are connected to each other.

Bennell and Canter [\[10\]](#) are interested in the probability (or indication) that two commercial burglaries are linked, given the modus operandi of these crimes. They use a database of solved commercial burglaries. Some of the burglaries studied had the same offender, which made it possible to identify behavioural features that reliably distinguish between linked and unlinked crime pairs. The authors present a model in which the distance between burglary locations and/or the method of entry can be used to determine the probability that the crimes are linked.

Tonkin et al. did a similar study [\[11\]](#). They concentrate on the distance between crime locations and the time between two crimes to distinguish linked and unlinked crimes. They conclude that the distance between crime locations found and/or the temporal proximity is able to achieve statistically significant levels of discrimination between linked and unlinked crimes.

The discussed papers show that, in practice, it is possible to select certain features of crimes (like the distance or temporal proximity) to assign a probability to the event 'the crimes are committed by the same person'. Taroni [\[12\]](#) discusses how such crime-related information may be used for the automatic detection of linked crimes.

2.2. Literature on modelling crime linkage

The papers discussed here focus on how to model possibly related crimes.

Taroni et al. [\[13\]](#) introduce Bayesian networks that focus on hypothesis pairs that distinguish situations where two items of evidence obtained from different crime scenes do or do not have a common source. They show how Bayesian networks can help in assigning a probability to the event where there is one offender responsible for both

crimes. We concentrate on a different topic, namely the offender configuration (who is the offender in which case) and on how evidence implies guilt¹ in one case influences the probability that a suspect is guilty in another case. Taroni et al. also present a Bayesian network for linking crimes with a utility and a decision node, which can help determine the direction for further investigation. Their study concentrates on how evidence from different cases influences the belief that there is a single offender responsible for both cases.

In Evett et al. [\[2\]](#) the hypothesis of interest does concern the offender configuration. Two case examples of similar burglaries are considered. In the first case the evidence consists of a DNA profile with a very discriminative random match probability and in the second case the only evidence is the report of an eye witness. The influence of the evidence in the first case on the question of guilt in the second case is investigated. They vary the strength of the evidence that suggests that there is one offender responsible for both cases to see how this influences the event that a suspect is guilty in the individual cases. The most important observation from their work is that when there is evidence that there is one offender responsible for both cases, the evidence in the individual cases becomes relevant to the other cases as well. This can either increase or decrease the probability that the suspect is the offender in a particular case. Evett et al. classify evidence into two categories that concern: (1) a specific crime only and (2) evidence that relates to similarities between the two crimes. We will introduce a third type of evidence that concerns both specific crimes as well as the similarity between crimes.

The case examples discussed by Evett et al. are viewed from the decision perspective of a prosecutor. The model they present should help to decide whether the prosecutor should charge a suspect with none, one or both crimes. However, Evett et al. do not consider the interesting question of what evidence should be presented when the suspect is charged with only one crime. We will show that it is wrong to select a subset of cases from a group of possibly linked cases and present only the evidence obtained in these cases. This is because evidence that is relevant in an individual case becomes of interest for the other cases when there exists a link between them.

In practical casework, the degree of similarity between crimes is usually poorly defined and lacks a rigorous mathematical treatment. While not solving this problem, we believe that the Bayesian network framework which we develop in this paper is a step in the right direction. It shows how to draw rational inference given certain assumptions and judgements of similarity (but where these judgements come from, and how they should be assessed are still difficult questions, and the topics of the literature mentioned in [Section 2.1](#)).

In what follows, we extend the work of Evett et al. by developing a generic Bayesian network. While they presented the necessary probabilities and relatedness structure needed for a Bayesian network they did not actually model a Bayesian network themselves. We further extend their work to situations with more than two crimes and present a type of evidence that they did not recognize in their paper, namely evidence supporting the claim that there is one offender responsible for multiple cases while simultaneously supporting the claim that the suspect is this offender. We will use an example to introduce and explain how different situations can be modelled using a Bayesian network. Most importantly, we show that it is not possible to 'unlink' crimes. When you have evidence that crimes are linked, all cases should be presented in court even when the suspect is charged for only a selection of them.

3. Using Bayesian networks when there are two linked crimes

In this section we introduce as a "thought experiment" the simplest example of two linked cases. In order to focus on the essence of crime

¹ For simplicity, we shall assume that 'guilty' and 'being the offender' are equivalent even though in practice they are not. For instance, when a 4-year old kills someone, he may be the offender but he is not guilty of murder.

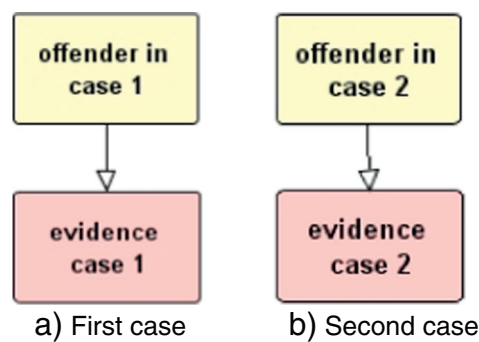


Fig. 1. Bayesian networks for the two cases. (a) First case. (b) Second case.

linkage, we ignore in this paper important issues like the relevance of the trace, transfer-, persistence-, and recover probabilities, and background levels (see [1] for more realistic models). Also, we ignore all details in assessing the degree of similarity of observations, and simply say they ‘match’ or not, although we are aware that from a scientific point of view this is a problematic concept. We emphasize that in practice, these issues cannot be ignored.

3.1. The basic assumptions

Suppose that two crimes – each involving a single (but not necessarily the same) offender – have occurred and are investigated separately. In each case a piece of trace evidence, assumed to have been left by the offender, is secured. Our notion of a ‘trace’ is very general (in the sense described in [14]). It includes biological specimens like blood, hair and semen (from which e.g. a full or partial DNA profile can be determined), marks made (such as fingermarks, footmarks) or physical features as seen by an eye witness (such as height, hair colour or tattoos). In each case, the police has a suspect that ‘matches’ the trace. We label them as suspect 1 and suspect 2 for the suspects in crimes 1 and 2 respectively.

The Bayesian networks for these cases are as in Fig. 1. The (yellow) *offender in case i* nodes ($i = 1, 2$) have two states, ‘suspect i ’ and ‘unknown’. The (pink) evidence nodes are conditionally dependent of the *offender in case i* nodes. They have two states, ‘match’ and ‘no match’. The probability tables for the *offender in case i* nodes are based on the possible offender population. Suppose that this possible offender population consists of 1000 men for each of the two crimes.²

Assuming that every person is equally likely to be the offender when no other evidence is available gives a prior probability of 0.001 for the suspect being the offender in each case. For the (pink) evidence case i nodes, the probability that a random person matches determines the probability tables (for example, random match probabilities when the evidence concerns DNA profiles). Suppose that the random match probability³ for the evidence in case 1 is 0.0002 and for case 2 is 0.0003. Here, we assume that no errors occurred in the analysis of the evidential pieces and that the offender matches with certainty. So, the probability tables for the *evidence case i* nodes are as in Table 1.

The Bayesian network shows what inserting evidence does to the probability that the suspect is the offender. By setting the state of the *evidence case i* nodes to ‘match’ we get the posterior probability that suspect i is the offender, given the evidence. In this example, the posterior probability that suspect 1 is the offender in case 1 is 0.83 and the posterior probability that suspect 2 is the offender in case 2 is 0.77. The difference in posterior probability occurs because the

² The number of men in the possible offender population only sets the prior on all the hypotheses of interest. Using another number of men will give another outcome but the conclusions we draw still hold.

³ We ignore here all practical difficulties in estimating these frequencies.

Table 1
Probability tables for the *evidence case 1* and *evidence case 2* nodes.

	Offender in case 1	
	Unknown	Suspect 1
No match	0.9998	0
Match	0.0002	1
	Offender in case 2	
	Unknown	Suspect 2
No match	0.9997	0
Match	0.0003	1

random match probability in case 1 is lower than the random match probability in case 2. The probative value of the evidence in case 1 is therefore stronger. The same result is easily obtained by using formulas, see [15].

3.2. Similarity evidence

Now, suppose that suspect 1 and suspect 2 are the same person. Since this suspect matches with the evidence obtained in both cases, it appears that the cases are linked by a common offender, the suspect. In what follows, we will construct a Bayesian network that models these (possibly) linked cases.

Next suppose that, in addition to evidence of similarity of offender, there is other evidence of similarity of the crimes. In contrast to evidence of ‘similarity of offender’ (which is human trace evidence in the sense explained in Section 3.1), evidence of ‘similarity of crime’ is not necessarily a human biological trace. This evidence could be, for example: a similar modus operandi, the time span between the two crimes, and the distance between the two crime scenes. In this example, we will use the evidence that fibres were recovered from the crime scenes that “matched” each other. In the second case, a balaclava is found at the crime scene. In the first case, fibres that match with the fibres from this balaclava are found. Since it is more likely to observe these matching fibres when the same person committed both burglaries than when two different persons did, the prosecution believe that there might be one person responsible for both crimes. Therefore, they want to link the crimes.

The network follows the description of the probability tables given in Evett et al. [2]. They discuss a crime linkage problem with two cases and use matching fibres as similarity evidence. However, they do use different individual crime evidence.

With two crimes, there are five possible scenarios regarding the offender configuration, namely:

1. The suspect is the offender in both cases.
2. The suspect is the offender in the first case; an unknown⁴ person is the offender in the second case.
3. An unknown person is the offender in the first case; the suspect is the offender in the second case.
4. An unknown person is the offender in both cases.
5. An unknown person is the offender in the first case; another unknown person is the offender in the second case.

The new Bayesian network, where we include the matching fibres evidence, is given in Fig. 2. Note that this BN implies that the evidence from the individual cases is conditionally independent given the offender(s). We will examine the influence of this assumption in Section 3.3.

Again, the probability table for the offender configuration node is based on the assumption that the potential offender population consists

⁴ We do not distinguish between related and unrelated ‘unknowns’. Obviously, that does affect the random match probabilities but we are ignoring that for simplicity. Also, it does not affect the main argument of this paper.

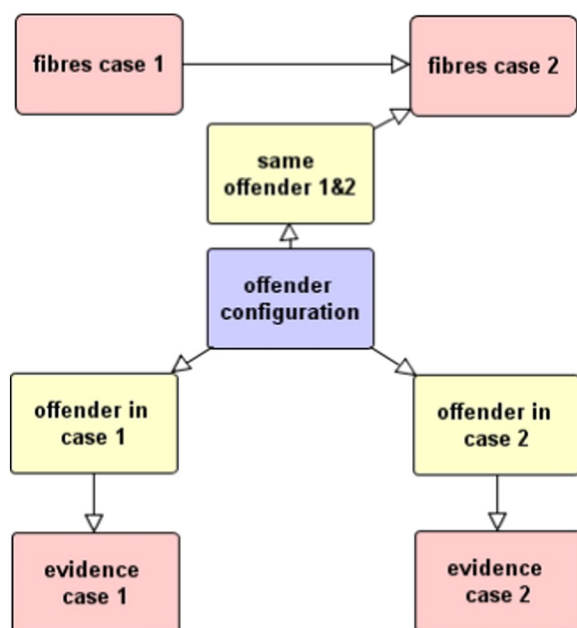


Fig. 2. Bayesian network for linking two cases with similarity evidence.

of 1000 men. We added a (yellow) node, *same offender 1&2*. This node summarises the scenarios in which the offender in the first case is the same person as the offender in the second case. The conditional probability table of the node is given in Table 2. Also, two (pink) evidence nodes are added, the *fibres case i* evidence nodes. These nodes have two states 'type A' (the type that is found on the crime scene and the balaclava) and 'other'. To get the probability tables for these nodes, we need to determine how probable it is to observe fibres of type A. Suppose that the probability of observing this type of fibres in case 1 is 0.0001. We assume that if one person is responsible for both crimes, we will observe the same type of fibres in both cases.⁵ So, the probability tables are as in Table 3.

Now, by inserting the matching fibres evidence and the matching evidence from the individual cases 1 and 2, we can compute the posterior probabilities for the suspect being the offender. The probability that the suspect is the offender in case 1, given the evidence, has increased to 0.9999. In case 2, this posterior probability also increased to 0.9999. The probability that the suspect is the offender in both case 1 and case 2 follows from the offender configuration node. This posterior probability is 0.99988.

The example shows that, by including evidence that increases the belief that the offenders in case 1 and 2 are the same, this also increases the belief that the suspect is the offender in the individual cases. The similarity evidence makes it possible that the value of evidence obtained in one case is 'transferred' to another case. The simple line of reasoning is as follows. There is evidence that the two crimes are committed by the same person. There is evidence that crime 1 is committed by the suspect. The combination of these two pieces of evidence increases our belief that the suspect is the offender in crime 2, even without including the evidence found in crime 2. This also works the other way around, from crime 2 to crime 1.

*Note that the likelihood ratio (LR), which is nowadays commonly reported by forensic experts, depends on the assumptions made about the prior probabilities of scenarios 1–5, see [16,17]. This poses interesting

Table 2
Probability table for the *same offender 1&2* node.

Same offender 1&2					
Configuration	Both suspect	Suspect first	Suspect second	Same unknown	Different unknowns
Yes	1	0	0	1	0
No	0	1	1	0	1

Table 3
Probability tables for the *fibres case i* nodes.

Fibres case 1				
Other	0.9999			
Type A	0.0001			
Fibres case 2				
Same offender 1&2	No		Yes	
Fibres case 1	Other	Type A	Other	Type A
Other	0.9999	0.9999	1	0
Type A	0.0001	0.0001	0	1

reporting problems. However, this is not the main argument of this paper. In the following, we will focus on the posterior probabilities.

It is important to note that the use of matching fibres as similarity evidence is provided just for convenience. As mentioned in Section 2.1, the distance between crime scenes, the time between crimes, the modus operandi or certain behaviour of the offender can provide very strong evidence that two crimes are committed by the same person. We could include any combination of these as similarity of crime evidence, but one can imagine that it is harder to come up with the probability for observing a certain modus operandi in a case. We could also include the work of Taroni [13] that concentrates on the question of the strength of the link between the cases.

3.3. "Dependent" evidence

In the last example, we assumed that the evidence obtained in the individual crimes is independent of each other, given the offender(s). However, if the pieces of evidence are of the same type (DNA, footmarks, eyewitness descriptions), knowing that the offender in both crimes is the same person makes them conditionally dependent. If one person is responsible for both burglaries, and we know that his DNA profile matches with the DNA profile obtained from the crime stain in case 1, it is certain (ignoring all considerations of relevance and various types of errors) that his DNA profile will also match the crime stain in case 2. For our example, we will concentrate on a situation where the evidence in the individual cases consists of two pieces, one of a type that is also found in the other case and one of a 'case individual' type.

In case 1, the evidence consists of a fingerprint and a footprint of size 12. In case 2, the evidence consists of a partial DNA profile and a footprint of size 12. The suspect's DNA profile matches with the partial DNA profile, his shoe size is 12 and his fingerprint matches the fingerprint from case 1. Clearly, the footprints from the individual cases are conditionally dependent given whether or not the offender in both cases is the same person. We assume that the shoes with size 12 have a population frequency of 0.01. The random match probability of the fingerprint from case 1 is 0.02. The random match probabilities of the partial DNA profile from case 2 is 0.03. Using these numbers, the combined evidential value of the evidential pieces in an individual case (which we assume to be conditionally independent) is the same as in the situations of Figs. 1 and 2. The Bayesian network describing this situation is given in Fig. 3.

The probability tables for the *partial DNA profile* and the *fingerprint* evidence nodes are similar to the ones given in Table 1 (with different random match probabilities). Only the probability table for the (pink)

⁵ In a more realistic setting, these numbers could be obtained by using a database of fibres. Also, other probabilities are involved, like the probability that the same balaclava was used by two different offenders. However, for this example the actual numbers are not that important, and we have chosen to follow the approach of Evett et al. [2].

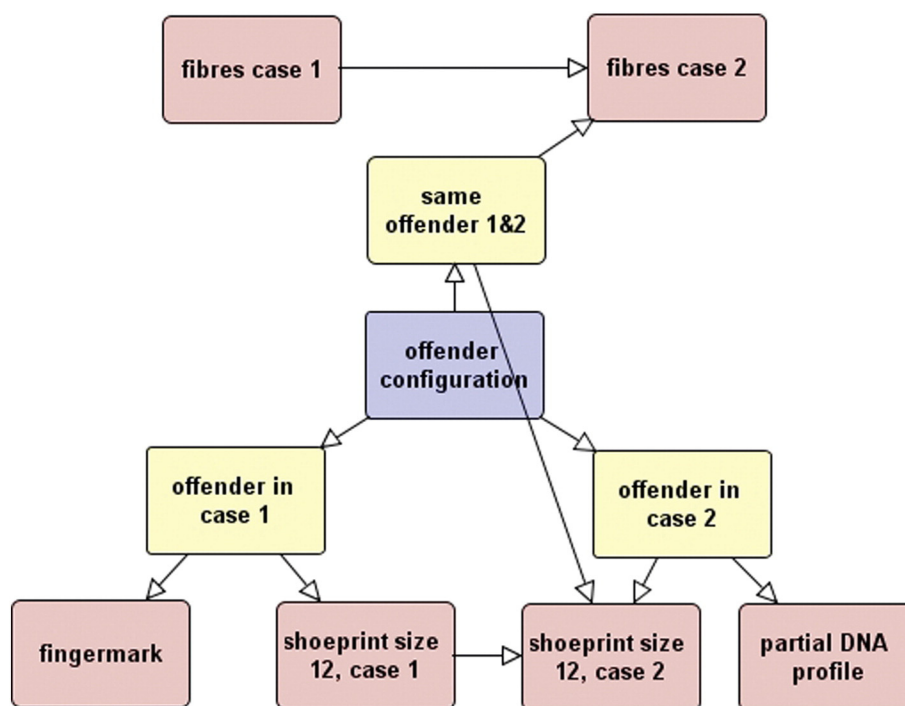


Fig. 3. Bayesian network for linking two cases with dependent DNA evidence.

evidence node *footmark size 12, case 2* is different. This node has three parents. It depends on who is the offender in case 2, but it also depends on whether there is one person who is the offender in both cases and the state of *footmark size 12, case 1*. The probability table for this node is given in Table 4.⁶

The important difference in Table 4, when we compare them to the probability tables given in Table 1, is that when we know that an unknown person is the offender in both cases, the *footmark size 12, case 1* evidence must match with the *footmark size 12, case 2* evidence.

If we insert all the evidence (DNA, fingerprint, footmark and fibres evidence), the posterior probability that the suspect is the offender in the first case is 0.99402. The posterior probability that the suspect is the offender in the second case is 0.99401, and the posterior probability that the suspect is the offender in both cases is 0.99399. Compared to the previous situation, we are slightly less confident that the suspect is the offender in both cases. This happens because the two items of footmark evidence are conditionally dependent on each other.

Although the posterior probabilities are slightly lower, it is still very likely that, given the evidence, the suspect is the offender in both cases.

4. Three linked cases

Now suppose that a third burglary comes up which is similar to the first two. Naturally, the prosecution would like to add this case in the link. With three crimes the number of possible scenarios for the offender configurations grows from 5 to 15, namely:

1. The suspect is the offender in all three crimes.
2. The suspect is the offender in the first two crimes. An unknown person is the offender in the third crime.

3. The suspect is the offender in the first and third crimes. An unknown person is the offender in the second crime.

:

:

14. An unknown person is the offender in the first crime. Another unknown person is the offender in the second and third crimes.
15. Three different unknown persons are the offenders in the three crimes.

In Appendix A we discuss the number of scenarios, given an arbitrary number, n , of cases.

4.1. Assumptions about the evidence

Again, the evidence in this case consists of footmark size 12. The same fibres as in cases 1 and 2 are found at the crime scene. The new *same offender* node summarises which cases have a common offender and has 5 states; (1) *one offender for all cases*, (2) *one offender for the first two cases, another for the third*, (3) *one offender for the first and third case, another for the second*, (4) *one offender for the first case, another for the second and third* and (5) *three different offenders*. The probability tables of the nodes *fibre evidence case i* are similar to those in Table 3 and

Table 4
Probability table for the *footmark size 12, case 2* node.

Footmark size 12, case 2							
Offender in case 2	Unknown				Suspect		
	No		Yes		No	Match	Yes
Same offender 1&2	No	Match	No	Match	No	Match	Match
Footmark size 12, case 1	No match	0.99	0.99	1	0	0	0
	Match	0.01	0.01	0	1	1	1

⁶ The situation *offender in case 2 = suspect, same offender 1&2 = yes, footmark size 12, case 1 = no* cannot occur. If the suspect is the offender in the second case and the offender in both cases is the same person, we know that he is also the offender in the first case. Hence, the *footmark size 12* evidence in case 1 will be a match (assuming no mistakes). This is not a problem because the Bayesian network will never use these numbers in the computation.

are also based on the assumption that the fibre type occurs with probability 0.0001.

The Bayesian network for this situation is given in Fig. 4. The prior probabilities for the offender configuration have changed. Again we assume that the potential offender population consists of 1000 men. Under the assumption that each of these men is equally likely to be the offender in the individual cases, independently from each other we can compute the prior probabilities for all the scenarios. These are given in the fourth column of Table 5.

If we insert the evidence, matching DNA profile, matching fingerprint, footmarks of size 12 and matching fibres between the cases, we get the posterior probabilities for the offender configuration, given the evidence. These are given in the last column of Table 4. The distribution of posterior probabilities shows that it is very likely that the suspect is the offender in all cases (with probability 0.99294). For the individual cases, the posterior probabilities that the suspect is the offender are 0.99399, 0.99397 and 0.99299 respectively.

4.2. Evidence proving innocence in the third case

A piece of exculpatory evidence is found in the third case. In our example, an eyewitness description of the offender states that the offender has a permanent tattoo on his left arm. If the suspect does not have a tattoo on his left arm, and we assume that the eyewitness description is correct, i.e. the actual offender has a permanent tattoo, it is certain that the suspect is not the offender in the third case. Now, the prosecution can do two things, (1) drop the third case and go to court with the first two cases, where they have strong evidence of the suspect's guilt, or (2) go to court with all three cases. The prosecution could argue that both options amount to the same outcome. In the first one, they drop the third case and use the evidence of the first two cases. In the second

Table 5

Prior and posterior probabilities for the offender configuration node, given that the offender population consists of 1000 men. The posterior probabilities are obtained by inserting the evidence. X represents the suspect, 1, 2 and 3 are other unknown men. The configuration 1, 2, X stands for: an unknown man is the offender in the first case, another unknown man is the offender in the second case and the suspect is the offender in the third case.

Offender configuration				
Offender 1	Offender 2	Offender 3	Prior probability	Posterior probability
X	X	X	$1.00 \cdot 10^{-9}$	0.99
X	X	1	$9.99 \cdot 10^{-7}$	$9.92 \cdot 10^{-4}$
X	1	X	$9.99 \cdot 10^{-7}$	$2.98 \cdot 10^{-5}$
1	X	X	$9.99 \cdot 10^{-7}$	$1.98 \cdot 10^{-5}$
X	1	1	$9.99 \cdot 10^{-7}$	$2.98 \cdot 10^{-5}$
X	1	2	$9.99 \cdot 10^{-4}$	$2.97 \cdot 10^{-6}$
1	X	1	$9.99 \cdot 10^{-7}$	$1.98 \cdot 10^{-5}$
1	X	2	$9.97 \cdot 10^{-4}$	$1.98 \cdot 10^{-6}$
1	1	X	$9.99 \cdot 10^{-7}$	$5.95 \cdot 10^{-7}$
1	2	X	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-8}$
1	1	1	$9.99 \cdot 10^{-7}$	$5.95 \cdot 10^{-3}$
1	1	2	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-6}$
1	2	1	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-4}$
2	1	1	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-6}$
1	2	3	0.99	$5.92 \cdot 10^{-7}$

option, the prosecution uses all three but, since they have evidence that the suspect is innocent in the third crime, they are only interested in whether the suspect is guilty in the first two cases. We will show that the first option is wrong since it withholds exculpatory evidence from the court for the first two cases.

When linking crimes, one needs to be aware that the sword cuts both ways. As we saw, if there is evidence in a case suggesting that there is one person responsible for both cases, evidence in one case is of interest for the question whether or not a suspect is guilty in another case. This means that if there is evidence in the first case that increases your belief that the suspect is the offender in the first case, it will also increase your belief that the suspect is the offender in the second case. This also works the other way around and is just as relevant: if there is evidence that a suspect is innocent in one case, this should also increase your belief that the suspect is innocent in the second case. This is illustrated by the example.

Suppose that it is known that the proportion of men with a tattoo on their left arm is 1/25.⁷ The Bayesian network representing the situation is given in Fig. 5. Remember that if we do not include the third case, we are in the situation of Fig. 3.

To compare the outcome in terms of the posterior probabilities when one drops or includes the third case, we compare the posterior probabilities of the offender configuration node of the models from Figs. 3 and 5. This is done in Table 6. The posterior probabilities for the suspect being the offender in the individual cases under the situation where the third case is dropped and the situation where the third case is included are given in Table 7.

The tables show that excluding or including the third case has serious consequences for the posterior probabilities, and thus, the outcome of a possible trial. When the third case is dropped, one can confidently state that it is very likely that the suspect is the offender in the first two cases. If we use the following two hypotheses,

Hp. The suspect is the offender in cases 1 and 2.

Hd. The suspect is not the offender in case 1 nor in case 2.

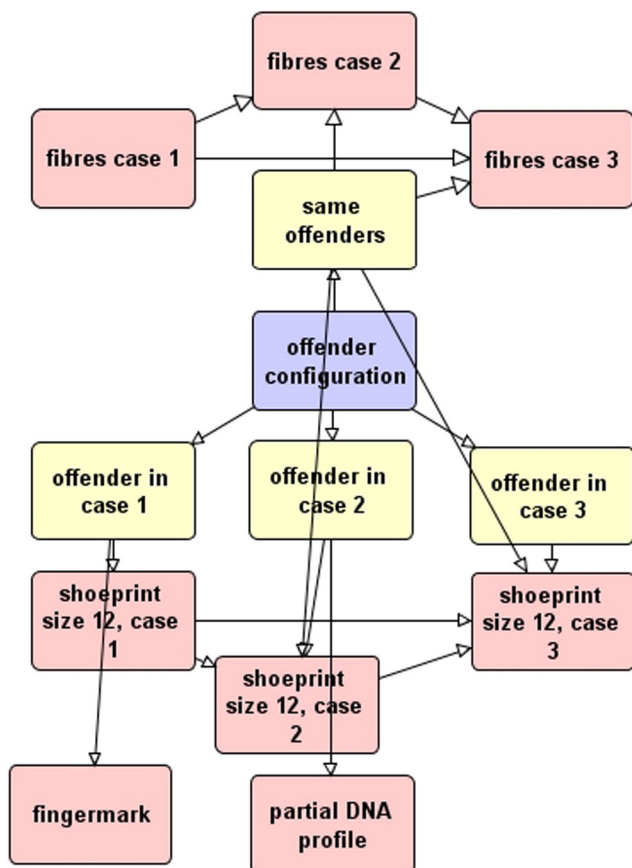


Fig. 4. Bayesian network for linking three cases with “dependent” evidence.

⁷ In this case, where we insert as evidence that there is no match with the suspect, the probability is irrelevant since it is impossible to observe no match when the suspect is the donor (assuming no errors were made). In a situation where the evidence does not directly show that the suspect is innocent but where it only increases one's belief that he is innocent, the random match probability is relevant.

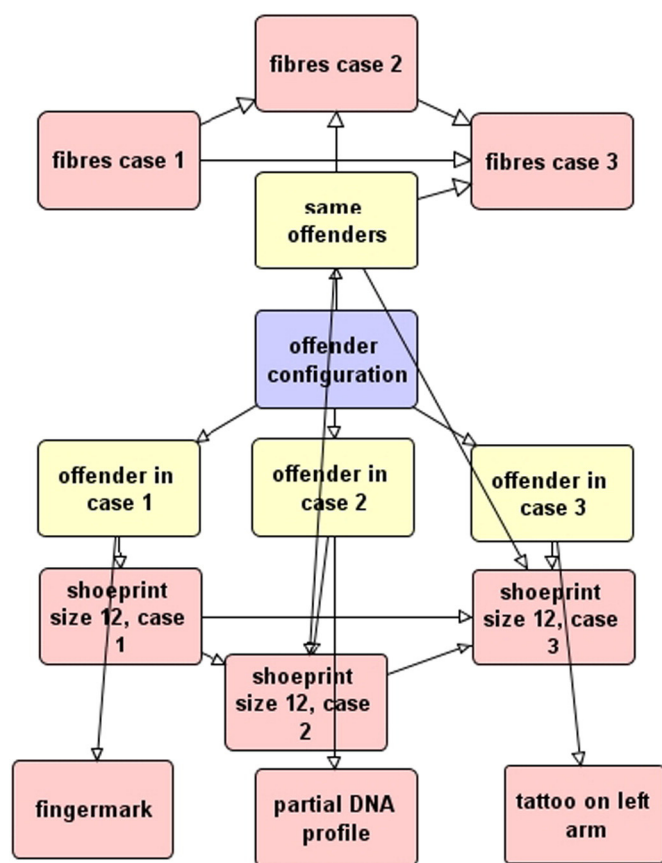


Fig. 5. Bayesian network for linking three cases with exculpatory evidence in the third case.

The posterior odds can be computed as

$$\frac{\mathbb{P}(H_p|E)}{\mathbb{P}(H_d|E)} = \frac{0.99399}{5.958 \cdot 10^{-3} + 5.946 \cdot 10^{-6}} = 167.$$

The conclusion would be: *Based on the observed evidence and the prior assumptions, it is 167 times more likely that the suspect is the offender in cases 1 and 2 than that he is not the offender in case 1 nor in case 2.*

Table 6

Posterior probabilities for the offender configuration node, for a situation where the third case is dropped and a situation where the third case is included.

Offender configuration				
Offender 1	Offender 2	Offender 3	Consider 2 cases	Consider 3 cases
X	X	X	0.99399	0.0
X	X	1		0.14148
X	1	X	$2.979 \cdot 10^{-5}$	0.0
X	1	1		0.0042445
X	1	2		$4.236 \cdot 10^{-6}$
1	X	X	$1.986 \cdot 10^{-5}$	0.0
1	X	1		0.0028296
1	X	2		$2.824 \cdot 10^{-6}$
1	1	X	0.005958	0.0
1	1	1		0.84889
1	1	2		$8.472 \cdot 10^{-4}$
1	2	X	$5.946 \cdot 10^{-6}$	0.0
1	2	1		$8.472 \cdot 10^{-4}$
2	1	1		$8.472 \cdot 10^{-4}$
1	2	3		$8.447 \cdot 10^{-7}$

Table 7

Posterior probabilities for the suspect being the offender in cases 1 and 2 for a situation where the third case is dropped and a situation where the third case is included.

	Consider 2 cases	Consider 3 cases
Case 1	0.992	0.146
Case 2	0.991	0.144

When we include the third case and use the same hypothesis pair, the posterior odds become,

$$\frac{\mathbb{P}(H_p|E)}{\mathbb{P}(H_d|E)} = \frac{0.14148}{0.84889 + 8.472 \cdot 10^{-4} + 8.472 \cdot 10^{-4} + 8.472 \cdot 10^{-4} + 8.447 \cdot 10^{-7}} = 0.17.$$

Now, the conclusion would be: *Based on the observed evidence and the prior assumptions, it is 6 times more likely that the suspect is not the offender in case 1 nor in case 2 than that he is the offender in cases 1 and 2.*

It is important to understand that the probability tables and assumptions under both models are the same. The only thing we changed is including the third case. The underlying reasoning is as follows. There is evidence that there is a common offender in the three cases. Both the matching fibres and the footmark evidence account for this. Also, there is evidence that the suspect is the offender in the first and in the second case. This is the partial DNA profile, the footmark and the fingerprint evidence. Due to the similarity evidence, this not only increases our belief that the suspect is the offender in the cases where this evidence was found, it also increases the belief that the suspect is the offender in the other cases. However, the evidence in case 3 shows that the suspect is innocent in the third case. So, there is another unknown person responsible for the third crime. Due to the similarity evidence, this also increases our belief that this unknown person is responsible for the first and second crimes. This is the double-edged sword when linking crimes.

The example shows that one cannot ‘unlink’ a case because of evidence suggesting that the suspect is not the offender. Although this is clear from the example shown, it seems likely that in practice people might simply drop a case without being aware of the consequences it has on the validity of their conclusions. One might even think that it benefits the suspect to drop the case, since the number of cases in which he can be found guilty decreases.

5. Discussion

In reality, crime linkage is very complex. We would like to emphasize again that issues like relevance of trace material and many other uncertainties are essential to consider. The presented crime linkage model simplifies the reality and does not capture all the problems that play a role when linking crimes. Hence, we do not recommend the use of the model presented here in actual casework (although more detailed BN models can be used to incorporate many of the other relevant issues also). We do think it is useful for uncovering the interesting aspects of the reasoning underlying crime linkage, and as such assists in understanding and dealing with it in practice.

We have shown, using a simple example, how a Bayesian network can help us understand and interpret evidence in cases where crimes are linked. Also, we have shown how to model cases with “dependent” evidence. It is possible to categorize evidence in a crime linkage problem into three categories.

1. Evidence relevant for the question of who the offender is in a specific case.
2. Evidence relevant for the question of whether the offender in two cases is the same person.
3. A combination of 1 and 2: Evidence relevant for both questions.

The first two categories are mentioned in Evett et al. [2]. The third category is a combination of the first two. For example if the similarity evidence is a match between fibres found at the different crime scenes, it falls in the second category. If, in addition, a sweater is found at the house of the suspect wherein fibres match with the fibres found at the crime scene, it belongs in the third category.

When linking more than two crimes, the combined effect of different pieces of evidence, i.e. how one observation influences another, rapidly becomes more complex. The use of Bayesian networks helps us understand the relations between observations. In our example we have shown a model where three crimes are linked. Using a Bayesian network the problem breaks down to filling the entries of some very straightforward probability tables. We have only shown Bayesian networks for two and three crimes. In Appendix B Bayesian networks for situations with four and five crimes are presented.

The number of possible offender configurations grows exponentially when the number of linked cases increases. Although we could use a computer to build a Bayesian network linking, e.g., twenty crimes and to fill the probability tables, the computation time will also increase according to the increase in offender configurations. The number of offender configurations with twenty crimes is 474 869 816 156 751 [18]. When the number of linked crimes is not that high (say less than 10), the method to present and understand the relations between evidence described should help provide insight into the problem. For large numbers of linked crimes, further research needs to be done. More on this can be found in Appendices A and B.

Most importantly, we have shown that one cannot ‘unlink’ cases. When there exists a link between cases, so there is evidence that there is one offender responsible for both cases, the cases should be treated simultaneously. In (forensic) practice, a similar thing occurs when multiple traces are secured of which the location of the traces suggest that they belong to one person, i.e. in a situation where fingerprints are recovered from an object. If these traces lie close to each other and form a grip pattern, it is likely that they belong to the same hand. Now, if only some of the fingerprints are similar to finger prints obtained from a suspect while the others are not, it is wrong to focus on the similarity evidence only.

It would be interesting to see how judges and police investigators deal in practice with cases that appear to be linked, where evidence in one case points towards a suspect whereas in the other case the evidence suggests that the suspect is innocent. Our own limited experience is that the relevance of exculpatory evidence found in one case for other similar cases is underestimated. This hypothesis can be tested in properly designed experiments.

Besides modelling linking of a large number of crimes, future research could study more complex situations of linked crimes. The research presented here could be expanded to situations where there is a group of criminals that e.g. rob houses together in various group compositions. In these cases it is possible to have a very similar and distinctive modus operandi, while the evidence in the different cases could point to different suspects.

Funding

This work was supported by the Netherlands Organisation for Scientific Research (NWO) as part of the project “Combining Evidence” in the Forensic Science programme [727.011.007].

Appendix A. The number of scenarios given the number of cases

The number of possible offender configurations increases very rapidly when the number of cases increases. For one case the number of configurations is 2 (suspect is the offender, unknown is the offender). For two cases it is 5 and for three cases it is 15. The number of scenarios given the number of cases is an example of the so-called Bell numbers [18]. For 1 to 10 cases, the number of configurations is given in Table A.8.

Table A.8

The number of possible offender configurations for 1 to 10 cases.

Cases	1	2	3	4	5	6	7	8	9	10
Configurations	2	5	15	52	203	877	4140	21,147	115,975	678,570

The n th Bell number represents the number of partitions of a set with n members, or equivalently, the number of equivalence relations on it. Bell numbers satisfy the recursion formula in Eq. (A.1).

$$B_{n+1} = \sum_{k=0}^n \binom{n}{k} B_k \quad (\text{A.1})$$

The n th Bell number corresponds with the number of offender configurations for $n - 1$ cases.

We see that the number of possible offender configurations grows rapidly when we increase the number of cases. Therefore, drawing conclusions becomes more difficult when the number of cases increases. Not regarding all possible offender configurations to decrease the number of scenarios might not be the solution to overcome this problem. Every offender configuration represents an interesting different situation. Especially when the number of cases is large (say 50), using less offender configurations will most likely mean that the new number of scenarios is too limited or still too large. For example when we would only include *the suspect is the offender in all cases* or *another unknown person is the offender in all cases*, we are disregarding too many situations. A solution might be limiting the number of different offenders that we allow. In a situation where there are 50 similar crimes, it may not be necessary to allow for the possibility that all crimes are committed by different men. Either way, the modelling of very large numbers of possibly linked cases is interesting for further research.

Appendix B. Bayesian networks for more cases

In this section we present networks for situations with four and five cases. As can be seen from Table A.8, the number of possible offender configurations is 52 for four cases and 203 for five cases. The probability tables are straightforward, and could therefore be made by the

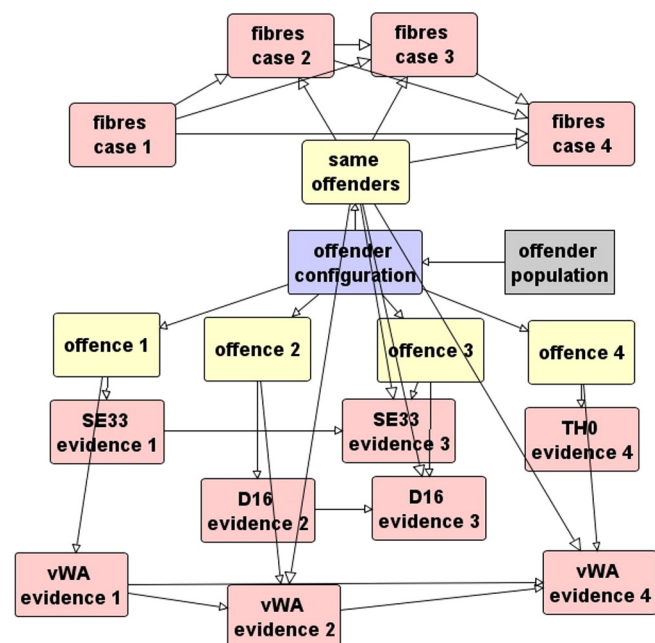


Fig. B.6. A Bayesian network linking four cases where the evidence consists of partial, overlapping, DNAProfiles.

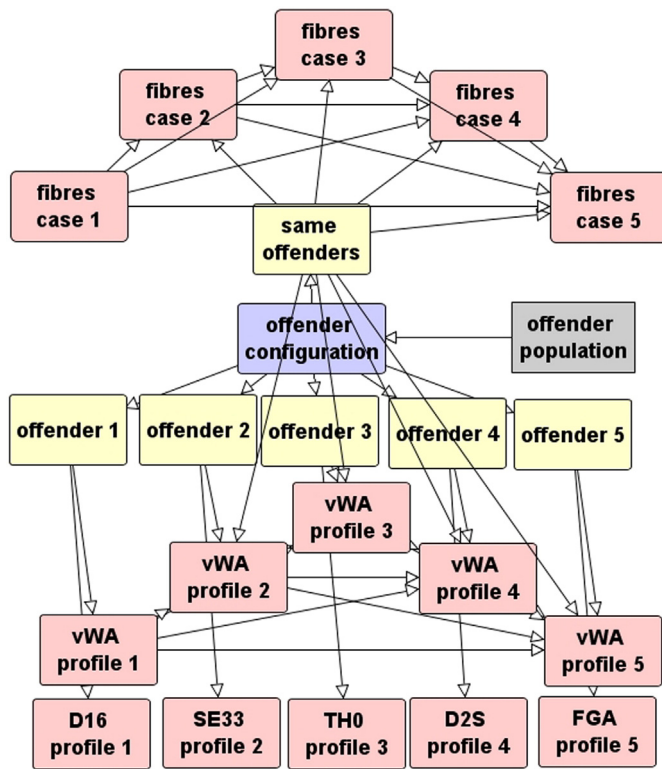


Fig. B.7. A Bayesian network linking five cases where the evidence in the individual cases consists of partial DNA profiles on two loci. In every case we have information on the vWA locus and on one other locus.

computer. We included an extra node. One that represents the number of men in the potential offender population. This node gives us the possibility to examine the influence of the prior on the posterior probabilities. The Bayesian networks in Figs. B.6 and B.7 are similar to the networks we saw before and show that it is possible to use the same methods to model situations with more cases.

References

- [1] F. Taroni, C.G.G. Aitken, P. Garbolino, A. Biedermann, Bayesian Networks and Probabilistic Inference in Forensic Science, John Wiley & Sons, Ltd, Chichester, UK, February 2006.
- [2] I.W. Evett, G. Jackson, D.V. Lindley, D. Meuwly, Logical evaluation of evidence when a person is suspected of committing two separate offences, *Sci. Justice* 46 (1) (2006) 25–31.
- [3] O. Ribaux, A. Girod, S.J. Walsh, P. Margot, S. Mizrahi, V. Clivaz, Forensic intelligence and crime analysis, *Law Psychol. Rev.* (2003) 47–60.
- [4] J. Woodhams, R. Bull, C.R. Hollin, Case linkage: identifying crimes committed by the same offender, *Criminal Profiling: International Theory, Research and Practice*, Humana Press, 2007, pp. 117–133.
- [5] Crime Linkage International Network. Resource centre <http://www.birmingham.ac.uk/generic/c-link/activities/resources/index.aspx>, 2014.
- [6] A. Davies, Rapists' behaviour: a three aspect model as a basis for analysis and the identification of serial crime, *Forensic Sci. Int.* 55 (2) (August 1992) 173–194.
- [7] D. Grubin, P. Kelly, C. Brunsdon, Linking serious sexual assaults through behaviour, Technical report, Home Office – Research, Development and Statistics Directorate, 2001.
- [8] R.R. Hazelwood, J.I. Warren, Linkage analysis: modus operandi, ritual, and signature in serial sexual crime, *Aggress. Violent Behav.* 8 (6) (November 2003) 587–598.
- [9] W. Petherick, Profiling and Serial Crime – Theoretical and Practical Issues, 3rd edition Anderson, 2014.
- [10] C. Bennell, D.V. Canter, Linking commercial burglaries by modus operandi: tests using regression and ROC analysis, *Sci. Justice* 42 (3) (2002) 153–164.
- [11] M. Tonkin, J. Woodhams, R. Bull, J.W. Bond, Behavioural case linkage with solved and unsolved crimes, *Forensic Sci. Int.* 222 (1–3) (October 2012) 146–153.
- [12] F. Taroni, Serial crime: a consideration of investigative problems, *Forensic Sci. Int.* 65 (1994) 33–45.
- [13] F. Taroni, S. Bozza, A. Biedermann, Two items of evidence, no putative source: an inference problem in forensic intelligence, *J. Forensic Sci.* 51 (6) (November 2006) 1350–1361.
- [14] N. Fenton, M. Neil, A. Hsu, Calculating and understanding the value of any type of match evidence when there are potential testing errors, *Artif. Intell. Law* 22 (1) (October 2014) 1–28.
- [15] I.W. Evett, P.D. Gill, G. Jackson, J. Whitaker, C. Champod, Interpreting small quantities of DNA: the hierarchy of propositions and the use of Bayesian networks, *J. Forensic Sci.* 47 (3) (May 2002) 520–530.
- [16] J.S. Buckleton, C.M. Triggs, C. Champod, An extended likelihood ratio framework for interpreting evidence, *Sci. Justice* 46 (2) (2006) 69–78.
- [17] F. Taroni, A. Biedermann, Inadequacies of posterior probabilities for the assessment of scientific evidence, *Law Probab. Risk* 4 (1–2) (March 2005) 89–114.
- [18] OEIS Foundation Inc., The On-Line Encyclopedia of Integer Sequences, 2014.