

Modelling crime linkage with Bayesian Networks

Jacob de Zoete, Marjan Sjerps, David Lagnado, Norman Fenton

Keywords: crime linkage, serial crime, Bayesian networks, combining evidence, case linkage

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

38 not be very realistic but which contains the essence of the problem, showing the most important
39 arguments. In reality all sorts of detail will complicate the problem but the essence will remain the
40 same. Thus, although the model does not incorporate all the difficulties involved when dealing with
41 crime linkage in practice, it can highlight flaws in the reasoning and create a better understanding
42 of the main line of reasoning.

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

49 2. Literature and state of the art on crime linkage

50 Crime linkage is a broad topic that has been extensively reported (for example in [3, 4, 5]). Here
51 we focus only on two aspects of the literature that are relevant for our analysis, namely: (1) how to
52 identify linked cases, and (2) how to model crime linkage. We discuss a (non extensive) selection
53 of some key papers on these topics.

54 2.1. Literature on how to identify linked cases

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

58 The authors of [6, 7, 8] investigate the behavioural aspects of sexual crime offenders in solved
59 cases. These studies concentrate on the consistency of the behaviour of serial sexual assault offend-
60 ers. The authors conclude that certain aspects of the behaviour can be regarded as a signature of
61 the offender. These aspects can be used to identify possibly linked crimes.

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

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

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

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

80 *2.2. Literature on modelling crime linkage*

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

82 Taroni et al. [13] introduce Bayesian networks that focus on hypothesis pairs that distinguish
83 situations where two items of evidence obtained from different crime scenes do or do not have a
84 common source. They show how Bayesian networks can help in assigning a probability to the event
85 there is one offender responsible for both crimes. We concentrate on a different topic, namely the
86 offender configuration (who is the offender in which case) and on how evidence implies guilt¹ in one
87 case influences the probability that a suspect is guilty in another case. Taroni et al. also present a
88 Bayesian network for linking crimes with a utility and a decision node, which can help determine
89 the direction for further investigation. Their study concentrates on how evidence from different
90 cases influences the belief that there is a single offender responsible for both cases.

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

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

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

117 In what follows, we extend the work of Evett et al. by developing a generic Bayesian network.
118 While they presented the necessary probabilities and relatedness structure needed for a Bayesian
119 network they did not actually model a Bayesian network themselves. We further extend their work
120 to situations with more than two crimes and present a type of evidence that they did not recognize
121 in their paper, namely evidence supporting the claim that there is one offender responsible for

¹For 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.

122 multiple cases while simultaneously supporting the claim that the suspect is this offender. We will
123 use an example to introduce and explain how different situations can be modelled using a Bayesian
124 network. Most importantly, we show that it is not possible to ‘unlink’ crimes. When you have
125 evidence that crimes are linked, all cases should be presented in court even when the suspect is
126 charged for only a selection of them.

127 3. Using Bayesian networks when there are two linked crimes

128 In this section we introduce as a “thought experiment” the simplest example of two linked cases.
129 In order to focus on the essence of crime linkage, we ignore in this paper important issues like the
130 relevance of the trace, transfer-, persistence-, and recover probabilities, and background levels (see
131 [1] for more realistic models). Also, we ignore all details in assessing the degree of similarity of
132 observations, and simply say they ‘match’ or not, although we are aware that from a scientific
133 point of view this is a problematic concept. We emphasize that in practice, these issues cannot be
134 ignored.

135 3.1. The basic assumptions

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

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

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

157 The Bayesian network shows what inserting evidence does to the probability that the suspect
158 is the offender. By setting the state of the *evidence case i* nodes to ‘match’ we get the posterior
159 probability that suspect i is the offender, given the evidence. In this example, the posterior prob-
160 ability that suspect 1 is the offender in case 1 is 0.83 and the posterior probability that suspect 2

²The 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.

³We ignore here all practical difficulties in estimating these frequencies

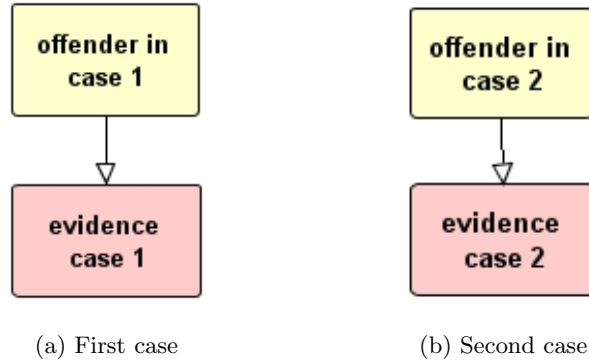


Figure 1: Bayesian networks for the two cases

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

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

161 is the offender in case 2 is 0.77. The difference in posterior probability occurs because the random
 162 match probability in case 1 is lower than the random match probability in case 2. The probative
 163 value of the evidence in case 1 is therefore stronger. The same result is easily obtained by using
 164 formulas, see [15].

165 3.2. Similarity evidence

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

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

179 prosecution believe that there might be one person responsible for both crimes. Therefore, they
 180 want to link the crimes.

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

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

- 185 1. The suspect is the offender in both cases.
- 186 2. The suspect is the offender in the first case; an unknown⁴ person is the offender in the second
 187 case.
- 188 3. An unknown person is the offender in the first case; the suspect is the offender in the second
 189 case.
- 190 4. An unknown person is the offender in both cases.
- 191 5. An unknown person is the offender in the first case; another unknown person is the offender
 192 in the second case.

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

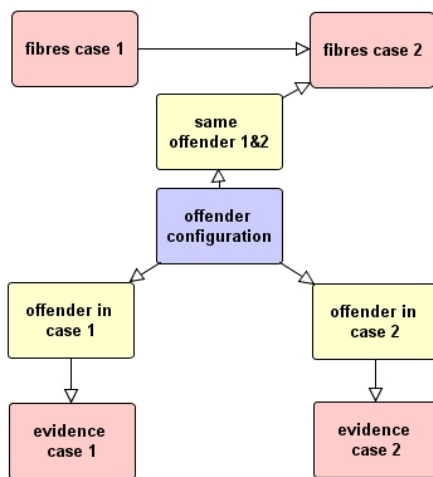


Figure 2: Bayesian network for linking two cases with similarity evidence

196 Again, the probability table for the offender configuration node is based on the assumption that
 197 the potential offender population consists of 1000 men. We added a (yellow) node, *same offender*
 198 *1&2*. This node summarizes the scenarios in which the offender in the first case is the same person
 199 as the offender in the second case. The conditional probability table of the node is given in Table
 200 2. Also, two (pink) evidence nodes are added, the *fibres case i* evidence nodes. These nodes have

⁴We 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.

201 two states ‘type A ’ (the type that is found on the crime scene and the balaclava) and ‘other’. To
 202 get the probability tables for these nodes, we need to determine how probable it is to observe fibres
 203 of type A . Suppose that the probability of observing this type of fibres in case 1 is 0.0001. We
 204 assume that if one person is responsible for both crimes, we will observe the same type of fibres in
 205 both cases.⁵ So, the probability tables are as in Table 3.

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 2: Probability table for the *same offender 1&2* node

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

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

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

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

⁵In 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].

220 Note that the likelihood ratio (LR), which is nowadays commonly reported by forensic experts,
221 depends on the assumptions made about the prior probabilities of scenarios 1-5, see [16],[17]. This
222 poses interesting reporting problems. However, this is not the main argument of this paper. In the
223 following, we will focus on the posterior probabilities.

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

231 3.3. “Dependent” evidence

232 In the last example, we assumed that the evidence obtained in the individual crimes is independ-
233 ent of each other, given the offender(s). However, if the pieces of evidence are of the same type
234 (DNA, footmarks, eyewitness descriptions), knowing that the offender in both crimes is the same
235 person makes them conditionally dependent. If one person is responsible for both burglaries, and
236 we know that his DNA profile matches with the DNA profile obtained from the crime stain in case
237 1, it is certain (ignoring all considerations of relevance and various types of errors) that his DNA
238 profile will also match the crime stain in case 2. For our example, we will concentrate on a situation
239 where the evidence in the individual cases consists of two pieces, one of a type that is also found in
240 the other case and one of a ‘case individual’ type.

241 In case 1, the evidence consists of a fingerprint and a footmark of size 12. In case 2, the evidence
242 consists of a partial DNA profile and a footmark of size 12. The suspect’s DNA profile matches
243 with the partial DNA profile, his shoe size is 12 and his fingerprint matches the fingerprint from
244 case 1. Clearly, the footmarks from the individual cases are conditionally dependent given whether
245 or not the offender in both cases is the same person. We assume that shoes with size 12 have
246 a population frequency of 0.01. The random match probability of the fingerprint from case 1 is
247 0.02. The random match probabilities of the partial DNA profile from case 2 is 0.03. Using these
248 numbers, the combined evidential value of the evidential pieces in an individual case (which we
249 assume to be conditionally independent) is the same as in the situations of Figure 1 and 2. The
250 Bayesian network describing this situation is given in Figure 3.

251 The probability tables for the *partial DNA profile* and the *fingerprint* evidence nodes are similar
252 to the ones given in Table 1 (with different random match probabilities). Only the probability table
253 for the (pink) evidence node *footmark size 12, case 2* is different. This node has three parents. It
254 depends on who is the offender in case 2, but it also depends on whether there is one person who is
255 the offender in both cases and the state of *footmark size 12, case 1*. The probability table for this
256 node is given in Table 4.⁶

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

⁶The 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.

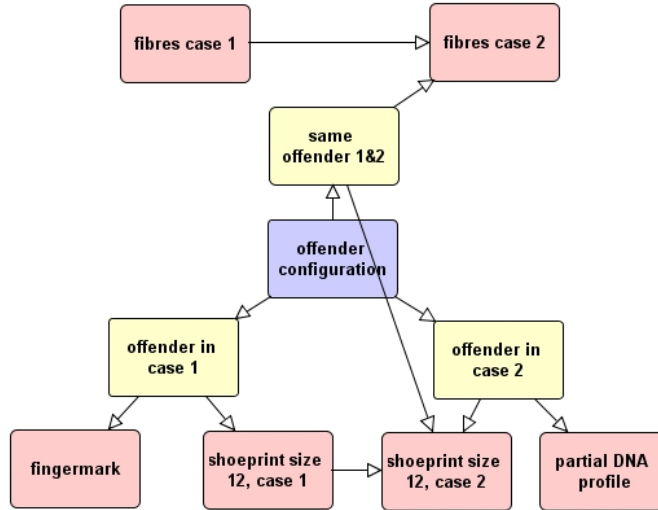


Figure 3: Bayesian network for linking two cases with dependent DNA evidence

footmark size 12, case 2							
offender in case 2	unknown				suspect		
same offender 1&2	no		yes		no		yes
footmark size 12, case 1	no match	match	no match	match	no match	match	match
no match	0.99	0.99	1	0	0	0	0
match	0.01	0.01	0	1	1	1	1

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

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

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

268 4. Three linked cases

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

- 272 1. The suspect is the offender in all three crimes.

- 273 **2.** The suspect is the offender in the first two crimes. An unknown person is the offender in the
274 third crime.
- 275 **3.** The suspect is the offender in the first and the third crime. An unknown person is the offender
276 in the second crime.
- 277 \vdots \vdots
- 278 **14.** An unknown person is the offender in the first crime. Another unknown person is the offender
279 in the second and the third crime.
- 280 **15.** Three different unknown persons are the offenders in the three crimes.

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

283 4.1. Assumptions about the evidence

284 Again, the evidence in this case consists of footprint size 12. The same fibres as in case 1 and 2
285 are found at the crime scene. The new *same offender* node summarises which cases have a common
286 offender and has 5 states; (1) *one offender for all cases*, (2) *one offender for the first two cases,*
287 *another for the third*, (3) *one offender for the first and third case, another for the second*, (4) *one*
288 *offender for the first case, another for the second and third* and (5) *three different offenders*. The
289 probability tables of the nodes *fibre evidence case i* are similar to those in Table 3 and are also
290 based on the assumption that the fibre type occurs with probability 0.0001.

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

296 If we insert the evidence, matching DNA profile, matching fingerprint, footmarks of size 12 and
297 matching fibres between the cases, we get the posterior probabilities for the offender configuration,
298 given the evidence. These are given in the last column of Table 4. The distribution of posterior
299 probabilities shows that it is very likely that the suspect is the offender in all cases (with probability
300 0.99294). For the individual cases, the posterior probabilities that the suspect is the offender are
301 0.99399, 0.99397 and 0.99299 respectively.

302 4.2. Evidence proving innocence in the third case

303 A piece of exculpatory evidence is found in the third case. In our example, an eyewitness
304 description of the offender states that the offender has a permanent tattoo on his left arm. If the
305 suspect does not have a tattoo on his left arm, and we assume that the eyewitness description is
306 correct, i.e. the actual offender has a permanent tattoo, it is certain that the suspect is not the
307 offender in the third case. Now, the prosecution can do two things, (1) drop the third case and go
308 to court with the first two cases, where they have strong evidence of the suspect's guilt, or (2) go
309 to court with all three cases. The prosecution could argue that both options amount to the same
310 outcome. In the first one, they drop the third case and use the evidence of the first two cases. In
311 the second option, the prosecution uses all three but, since they have evidence that the suspect is
312 innocent in the third crime, they are only interested in whether the suspect is guilty in the first two
313 cases. We will show that the first option is wrong since it withholds exculpatory evidence from the
314 court for the first two cases.

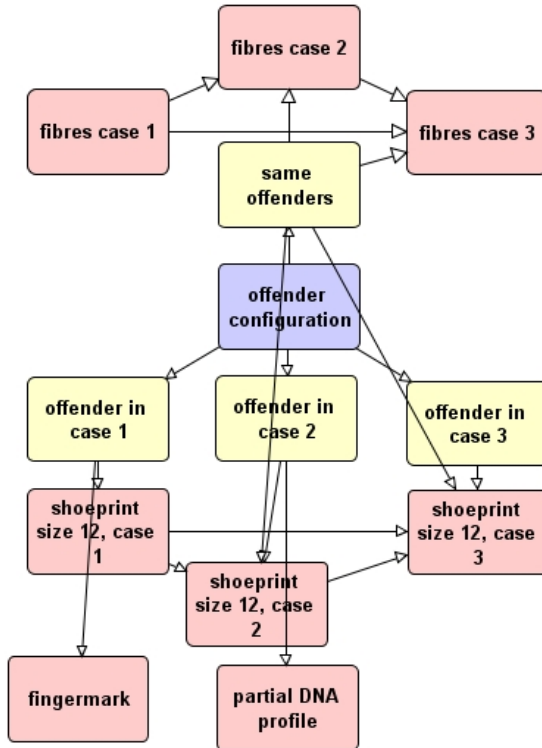


Figure 4: Bayesian network for linking three cases with “dependent” evidence

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

323 Suppose that it is known that the proportion of men with a tattoo on their left arm is $1/25$.⁷
 324 The Bayesian network representing the situation is given in Figure 5. Remember that if we do not
 325 include the third case, we are in the situation of Figure 3.

326 To compare the outcome in terms of the posterior probabilities when one drops or includes the
 327 third case, we compare the posterior probabilities of the offender configuration node of the models

⁷In 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.

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	$9.97 \cdot 10^{-4}$	$2.97 \cdot 10^{-6}$
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	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-6}$
2	1	1	$9.97 \cdot 10^{-4}$	$5.94 \cdot 10^{-6}$
1	2	3	0.99	$5.92 \cdot 10^{-7}$

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.

328 from Figure 3 and 5. This is done in Table 6. The posterior probabilities for the suspect being
329 the offender in the individual cases under the situation where the third case is dropped and the
330 situation where the third case is included are given in Table 7.

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

335 H_p : The suspect is the offender in case 1 and 2.

336 H_d : The suspect is not the offender in case 1 nor in case 2.

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 case 1 and 2 than that he is not the offender in case 1 nor in case 2.* When we include the third case and use the same hypothesis pair, the posterior odds become,

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

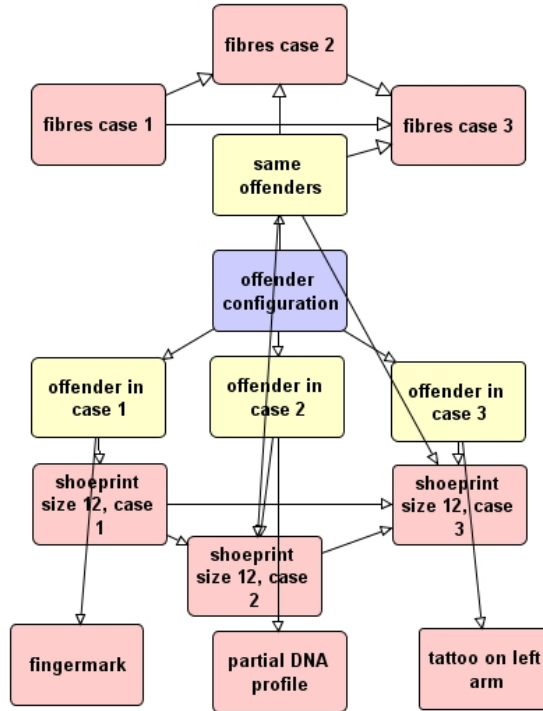


Figure 5: Bayesian network for linking three cases with exculpatory evidence in the third case

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

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

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

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 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.

	consider 2 cases	consider 3 cases
case 1	0.992	0.146
case 2	0.991	0.144

Table 7: Posterior probabilities for the suspect being the offender in case 1 and 2 for a situation where the third case is dropped and a situation where the third case is included.

356 5. Discussion

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

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

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

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

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

382 The number of possible offender configurations grows exponentially when the number of linked
383 cases increases. Although we could use a computer to build a Bayesian network linking, e.g., twenty
384 crimes and to fill the probability tables, the computation time will also increase according to the
385 increase in offender configurations. The number of offender configurations with twenty crimes is
386 474 869 816 156 751 [18]. When the number of linked crimes is not that high (say less than 10), the
387 method to present and understand the relations between evidence described should help provide
388 insight into the problem. For large numbers of linked crimes, further research needs to be done.
389 More on this can be found in Appendix Appendix A and Appendix B.

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

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

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

408 [1] F. Taroni, C. G. G. Aitken, P. Garbolino, and A. Biedermann. *Bayesian Networks and Prob-*
409 *abilistic Inference in Forensic Science*. John Wiley & Sons, Ltd, Chichester, UK, February
410 2006.

411 [2] I. W. Evett, G. Jackson, D. V. Lindley, and D. Meuwly. Logical evaluation of evidence when a
412 person is suspected of committing two separate offences. *Science & Justice*, 46(1):25–31, 2006.

- 413 [3] O. Ribaux, A. Girod, S.J. Walsh, P. Margot, S. Mizrahi, and V. Clivaz. Forensic intelligence
414 and crime analysis. *Law and Psychology Review*, pages 47–60, 2003.
- 415 [4] J. Woodhams, R. Bull, and C.R Hollin. Case linkage: identifying crimes committed by the same
416 offender. In *Criminal Profiling: International Theory, Research and Practice*, pages 117–133.
417 Humana Press, 2007.
- 418 [5] Crime linkage international network. Resource centre. [http://www.birmingham.ac.uk/
419 generic/c-link/activities/resources/index.aspx](http://www.birmingham.ac.uk/generic/c-link/activities/resources/index.aspx), 2014.
- 420 [6] A. Davies. Rapists’ behaviour: a three aspect model as a basis for analysis and the identification
421 of serial crime. *Forensic Science International*, 55(2):173–94, August 1992.
- 422 [7] D. Grubin, P. Kelly, and C. Brunson. Linking Serious Sexual Assaults Through Behaviour.
423 Technical report, Home Office - Research, Development and Statistics Directorate, 2001.
- 424 [8] R. R. Hazelwood and J. I. Warren. Linkage analysis: Modus operandi, ritual, and signature
425 in serial sexual crime. *Aggression and Violent Behavior*, 8(6):587–598, November 2003.
- 426 [9] W. Petherick. *Profiling and Serial Crime - Theoretical and practical issues*. Anderson, 3rd
427 edition, 2014.
- 428 [10] C. Bennell and D. V. Canter. Linking commercial burglaries by modus operandi: tests using
429 regression and ROC analysis. *Science & Justice*, 42(3):153–64, 2002.
- 430 [11] M. Tonkin, J. Woodhams, R. Bull, and J. W. Bond. Behavioural case linkage with solved and
431 unsolved crimes. *Forensic Science International*, 222(1-3):146–53, October 2012.
- 432 [12] F. Taroni. Serial crime : a consideration of investigative problems. *Forensic Science Interna-
433 tional*, 65:33–45, 1994.
- 434 [13] F. Taroni, S. Bozza, and A. Biedermann. Two items of evidence, no putative source: an infer-
435 ence problem in forensic intelligence. *Journal of Forensic Sciences*, 51(6):1350–61, November
436 2006.
- 437 [14] N. Fenton, M. Neil, and A. Hsu. Calculating and understanding the value of any type of match
438 evidence when there are potential testing errors. *Artificial Intelligence and Law*, 22(1):1–28,
439 October 2014.
- 440 [15] I. W. Evett, P. D. Gill, G. Jackson, J. Whitaker, and C. Champod. Interpreting small quantities
441 of DNA: the hierarchy of propositions and the use of Bayesian networks. *Journal of Forensic
442 Sciences*, 47(3):520–30, May 2002.
- 443 [16] J. S. Buckleton, C. M. Triggs, and C. Champod. An extended likelihood ratio framework for
444 interpreting evidence. *Science & Justice*, 46(2):69–78, 2006.
- 445 [17] F. Taroni and A. Biedermann. Inadequacies of posterior probabilities for the assessment of
446 scientific evidence. *Law, Probability and Risk*, 4(1-2):89–114, March 2005.
- 447 [18] OEIS Foundation Inc. The on-line encyclopedia of integer sequences. [http://oeis.org/
448 A000110](http://oeis.org/A000110), 2014.

449 **Funding**

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

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

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

cases	1	2	3	4	5	6	7	8	9	10
configurations	2	5	15	52	203	877	4140	21147	115975	678570

Table A.8: The number of possible offender configurations for 1 to 10 cases

457
 458 The n th Bell number represents the number of partitions of a set with n members, or equiv-
 459 alently, the number of equivalence relations on it. Bell numbers satisfy the recursion formula in
 460 A.1.

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

461 The n th Bell number corresponds with the number of offender configurations for $n - 1$ cases.
 462 We see that the number of possible offender configurations grows rapidly when we increase the
 463 number of cases. Therefore, drawing conclusions becomes more difficult when the number of cases
 464 increases. Not regarding all possible offender configurations to decrease the number of scenarios
 465 might not be the solution to overcome this problem. Every offender configuration represents an
 466 interesting different situation. Especially when the number of cases is large (say 50), using less
 467 offender configurations will most likely mean that the new number of scenarios is too limited or
 468 still too large. For example when we would only include *the suspect is the offender in all cases*
 469 or *another unknown person is the offender in all cases*, we are disregarding too many situations.
 470 A solution might be limiting the number of different offenders we allow. In a situation where
 471 there are 50 similar crimes, it may not be necessary to allow for the possibility that all crimes are
 472 committed by different men. Either way, the modelling of very large numbers of possibly linked
 473 cases is interesting for further research.

474 **Appendix B. Bayesian networks for more cases**

475 In this section we present Networks for situations with four and five cases. As can be seen
 476 from Table A.8, the number of possible offender configurations is 52 for four cases and 203 for five
 477 cases. The probability tables are straightforward, and could therefore be made by the computer.
 478 We included an extra node. One that represents the number of men in the potential offender
 479 population. This node gives us the possibility to examine the influence of the prior on the posterior
 480 probabilities. The Bayesian networks in Figure B.6 and B.7 are similar to the networks we saw
 481 before and show that it is possible to use the same methods to model situations with more cases.

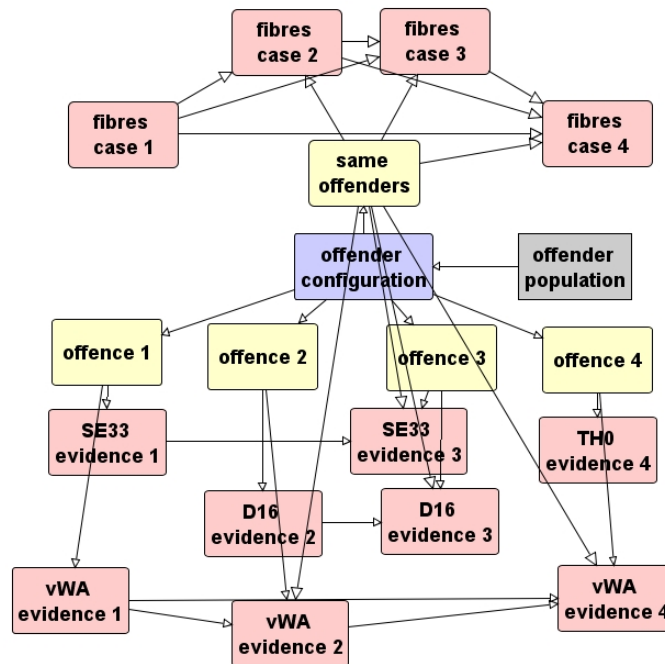


Figure B.6: A Bayesian network linking four cases where the evidence consists of partial, overlapping, DNA profiles

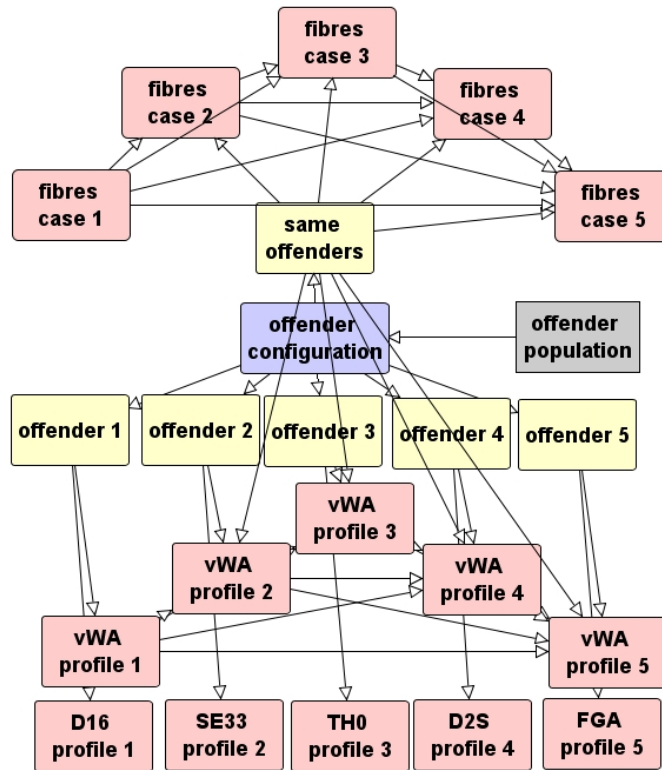


Figure B.7: A Bayesian network linking five cases where the evidence in the individual cases consists of partial DNA profiles on two loci. In every case we have information on the **vWA** locus and on one other locus.