Probabilistic evaluation of n traces with no putative source: a likelihood ratio based approach in an investigative framework

I. De March*, E. Sironi, F. Taroni

University of Lausanne, School of Criminal Justice, 1015 Lausanne-Dorigny, Switzerland

Abstract

Analysis of marks recovered from different crime scenes can be useful to detect a linkage between criminal cases, even though a putative source for the recovered traces is not available. This particular circumstance is often encountered in the early stage of investigations and thus, the evaluation of evidence association may provide useful information for the investigators. This association is evaluated here from a probabilistic point of view: a likelihood ratio based approach is suggested in order to quantify the strength of the evidence of trace association in the light of two mutually exclusive propositions, namely that the n traces come from a common source or from an unspecified number of sources. To deal with this kind of problem, probabilistic graphical models are used, in form of Bayesian networks and object-oriented Bayesian networks, allowing users to intuitively handle with uncertainty related to the inferential problem.

Keywords: Evidence interpretation; Likelihood ratio; Bayesian networks; Object-oriented Bayesian networks; Case linkage.

1   Introduction

In forensic practice, scientists are frequently asked to deal with cases in which several traces of the same type are collected. These items of evidence can be found on a unique crime scene or in different locations related to distinct criminal activities. If the scientific comparative analysis performed on the recovered traces shows that there are some relevant similarities between them, and if the traces are considered relevant to the criminal activity, it can be assumed that a linkage may exist between the different cases because a unique perpetrator could be the source of the traces [1, 2].

Moreover, if any putative source that could be at the origin of traces is known, it is crucial to evaluate the strength of the potential link, due to the fact that this kind of information could

---

*Corresponding author: `ilaria.demarch@unil.ch`

allow investigators to treat the cases as a series. Therefore, this work aims to investigate an inferential model to deal with such an investigative framework. A similar analysis was performed by Taroni et al. [3] in a scenario considering two items of evidence recovered in two locations. Here, the goal is to extend the reasoning to a scenario involving n traces recovered on n different crime scenes. The likelihood ratio (LR) will be used as a measure of the degree to which the evidence (in this case, the observed similarities between traces) is capable of discriminating between two competing propositions, i.e. the traces come from the same (unknown) source or from at least two different sources. Graphical models are developed in order to qualitatively and quantitatively describe the problem of interest. The advantages of the probabilistic graphical environment will be discussed, putting forward its flexibility and dynamism. This work's objective is to support a probabilistic use of investigative information in order to provide a structured approach to assess scientific elements issued from crime scene investigations for intelligence purposes.

The paper is structured as follows: Section 2 describes the scenario of interest and the questions related to the management of uncertainty in our investigative setting. In Section 3 graphical models, such as Bayesian networks and object-oriented Bayesian networks, are used to represent the proposed scenario and investigate the dependencies between the variables of interest. Section 4 presents the development of the likelihood ratio formula considering the two propositions previously defined. Finally Section 5 illustrates a practical example with a fixed number of traces; the adjustment of the formula and the graphical models will be described.

2 Scenario

A forensic scientist is asked to deal with n traces recovered at n separate crime scenes. All the traces are of the same kind, and each trace is characterized by a single discrete attribute A. For sake of illustration, consider n shoe marks found on n crime scenes, characterized by their general pattern only, or n blood stains defined by their genetic DNA profile. Assume the traces have the same degree of quality, and according to the scientific comparative analysis, they present a relevant amount of similarities, so that it cannot be excluded that they share a common origin. Finally, no suspect who could be at the origin of the traces is known. Because of the absence of a suspect, the problem is treated under an investigative point of view, and the main (source level) propositions of interest can be whether or not the recovered traces come from the same source.

3 Graphical models

The recovered items of evidence should be evaluated in order to support investigators in their inquiry. Should the investigators focus on a unique perpetrator (source) of the recovered material or are more than one source involved in the cases? In order to perform

the evaluation correctly, a graphical representation of the relevant variables for the problem of interest may prove to be particularly convenient. Bayesian networks and object-oriented Bayesian networks are probabilistic graphical models that have been used for decades to study forensic problems [4, 5], and they are therefore used to analyze the proposed scenario.

3.1   Bayesian networks

Bayesian networks (BN) are probabilistic graphical tools based on elements of graph and probability theory [6, 5] that allow one to deal with uncertainty in a coherent and rigorous manner. BNs are directed acyclic graphs (DAG) composed by a set of variables, represented by nodes, that model the entities of interest for the problem. The probabilistic relationships between two nodes are represented by directed edges (arrows). A conditional probability table is associated with each node, and shows the probabilities of the different states of the node conditioned by the states of its parent nodes, i.e. the nodes on which they directly depend. More formally, a BN illustrates the representation of the joint probability distribution for all the variables of the structured model. Moreover, because a BN is a complete model for all the variables of interest and their relationships, it can be used to answer probabilistic queries about them. For example, the network can be used to obtain updated knowledge of the states of a subset of variables when other variables are observed; this aspect is called probabilistic inference.

In the similar but simpler scenario treated in [3], the authors developed a Bayesian network for a case where $n = 2$ traces were involved; the qualitative part of the Bayesian network was described by node H, accounting for the propositions of interest, nodes S_1 and S_2, considering the type (A or $\bar{A}$) of the true source of the first and second trace respectively, and nodes E_1 and E_2, representing the features of the first and second trace.

In this work, the network is extended to cases involving $n > 2$ traces, and new nodes have been introduced in order to take the increased complexity of the scenario into account.

The developed network is shown in Figure 1, and Table 1 shows a summary of the different nodes and their definitions.

Here, node H represents again the two propositions of interest. Its states are H_1 and H_2, formally defined as:

- H_1: The n traces come from the same source;

- H_2: The n traces come from at least two different sources,

where the number of possible different sources is not specified.

Node K depends directly on node H. It represents the number of possible sources at the origin of the n traces, and its states are all the numbers from 1 to n. If proposition H_1 is true, only one source is at the origin of the traces, so $Pr(K = 1|H_1) = 1$ and $Pr(K \neq 1|H_1) = 0$. If H_2 is true, $Pr(K = 1|H_2) = 0$ and $Pr(K = k|H_2)$, with $k = 2 \ldots n$, represents the probability that exactly k sources are involved knowing that the traces come
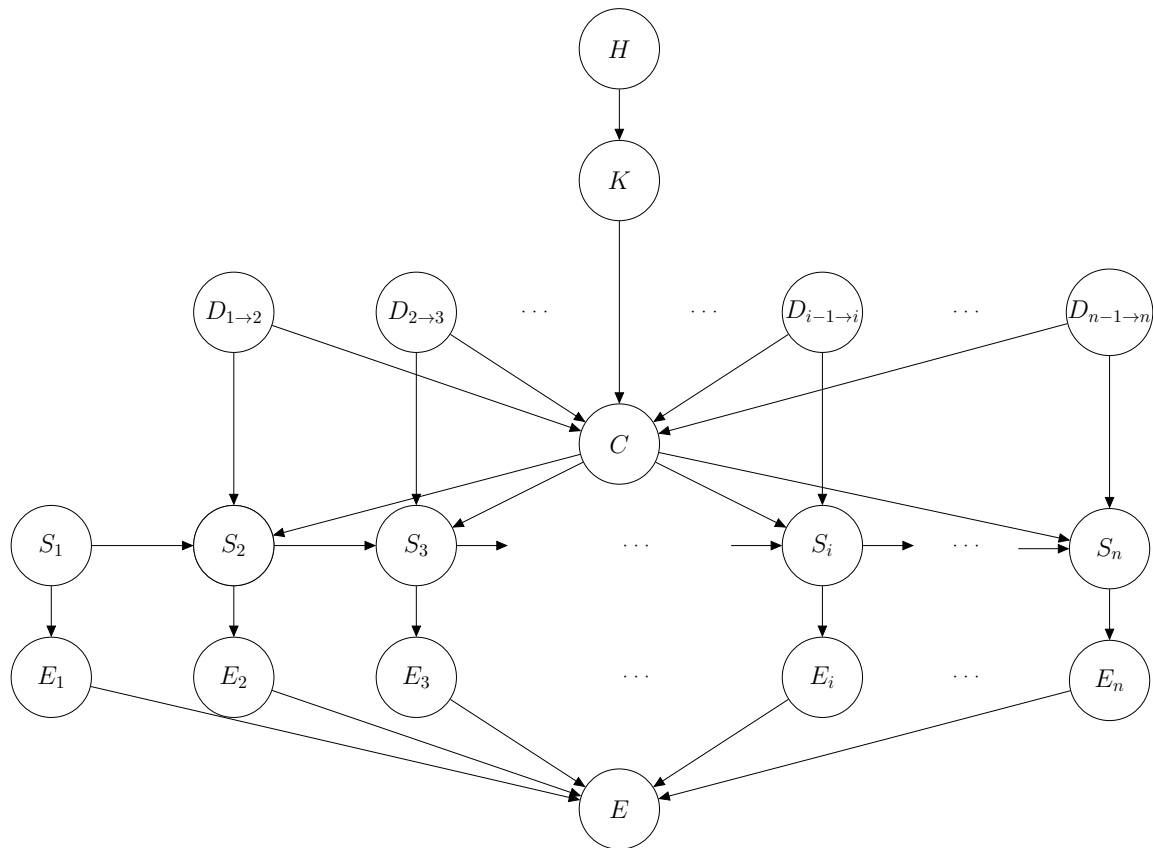
Figure 1: Bayesian network for a scenario in which n traces are found. New nodes have been introduced with respect to the network developed in [3], in particular the node K, accounting for the number of possible different sources, nodes $D_{i-1 \to i}$ that consider the dependency between the true sources of $(i-1)$th and ith traces, and the constraint node C that allows the network to consider only the logically possible configurations.

Table 1: Table representing the names of different nodes (1st column) of the Bayesian
network (Figure 1), their definition (2nd column) and their states (3rd column).

Node	Definition	States
H	Propositions	H_1, H_2
K	Number of potential sources	$k = 1 \ldots n$
C	Allowed configurations	*on*, *off*
$D_{i-1 \to i}$	Dependence	$\mid$, $\circ$
S_1	Type of Source 1	A, $\bar{A}$
S_i	Type of Source i	A, $\bar{A}$
E_i	Type of Trace i	A, $\bar{A}$
E	Combination of nodes E_i	A, $\bar{A}$

from at least two different sources. This probability depends on the different configurations
in which many sources can be at the origin of the n traces. According to the considered
scenario, all the potential sources, being unknown, are indistinguishable from one another.
On the contrary, no assumption has been made on the crime scenes, which means that
non-scientific information could exist, allowing one to model a probability distribution
which considers the prior belief "two or more crime scenes are more related than others",
which is the initial belief that the same source is at the origin of traces collected on two
or more particular crime scenes. According to this consideration, all the crime scenes
can be considered distinguishable and, for the sake of simplicity, they will be treated as
equiprobable in this work. The number of ways in which the n traces can come from exactly
k sources corresponds to the Stirling number of the second kind $\left\{ {n \atop k} \right\}$ [7, 8] which represents
the number of ways in which a n-elements group can be partitioned in exactly k subsets.
The total number of ways in which n traces can come from 2 to n sources is $B_n - 1$, where
B_n represents the nth Bell number [1] [7, 8]. $Pr(K = k|H_2)$ can therefore be written as
$Pr(K = k|H_2) = \left\{ {n \atop k} \right\}/B_n - 1$. The conditional probability table for node K is shown in Table
2.
Following the nodes' definitions presented in [3], the type of each trace recovered and of
its true source are represented by nodes E_i and S_i ($i = 1 \ldots n$) respectively. The states of
these nodes are A and $\bar{A}$, with A being the considered discrete attribute (for example a
given DNA profile). It is important to stress that the numbering of the traces (1 to n) and
of their actual sources (1 to n) is not related to information issued from investigations (for
example, the order of recovery). Numbers 1 to n are only useful to label the traces, not to
order them.

---

[1]The Stirling number of the second kind $\left\{ {n \atop k} \right\}$ indicates the number of partitions of an n-elements set into
k non-empty subsets. It can be calculated using its explicit formula $\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n$. The Bell
number counts the number of partitions of an n-elements set, regardless the number of subsets. Bell number
can be obtained as $\sum_{k=0}^{n} \left\{ {n \atop k} \right\}$.

Table 2: Conditional probability table for node K.

	H:	H_1	H_2
K:	*1*	1	0
	2	0	$\{{}^n_2\}/B_n-1$
	...	...	...
	k	0	$\{{}^n_k\}/B_n-1$
	...	...	...
	n	0	$\{{}^n_n\}/B_n-1$

Each node E_i ($i = 1 \ldots n$) is dependent by the respective node S_i, and, since it is sensible to accept the assumption that if a source is of type A then the trace recovered on the crime scene will be of the same type and assuming that no laboratory or manipulation errors occurred, $Pr(E_i = A|S_i = A) = 1$ and $Pr(E_i = \bar{A}|S_i = A) = 0$. Given the fact that the presence of at least one source at the origin of the traces recovered is certain, the first node of type S (S_1) does not depend upon any other node. $Pr(S_1 = A)$ corresponds to the probability that a source of the relevant population is of type A, which can be written as γ_A. The conditional probability table for node S_1 is shown in Table 3.

Table 3: Conditional probability table for node S_1. γ_A represents the probability that a source from the relevant population is of type A.

S_1:	A	γ_A
	$\bar{A}$	$1 - \gamma_A$

For a generic node S_i with $i \neq 1$, the construction of the conditional probability table must consider if the ith source is dependent (i.e. it is the same source) from one of the previous sources, represented by nodes $S_{i-1} \ldots S_1$. As stressed before, the fact of numbering the traces does not have any theoretical meaning, so it can be considered that all nodes S_i representing the same source can be placed one after another. In this way, a generic node S_i depends only on node S_{i-1} and no more on nodes $S_{i-2} \ldots S_1$. $n-1$ nodes $D_{i-1 \to i}$ consider the a priori probability that the $(i-1)$th and the ith trace share a common source. The states of these nodes are defined by the symbols $|$ and $\circ$, the former meaning that the sources of the $(i-1)$th and ith traces are the same, and the latter that the two sources are different. Note that investigative information such as modus operandi or eyewitness evidence could play an important role in such probability assignments. Nevertheless, in the present work, the two states are considered equiprobable.

Information from nodes $D_{i-1 \to i}$ and from node K, representing the number of potential

sources, are combined into node C. Node C is a constraint node [9] [2] having two states, *on* and *off*, that has been introduced in order to allow the network to separate the n nodes representing the true sources into k groups, k clearly being one of the states of node K. Node C is conditioned by node K and by all nodes $D_{i-1\rightarrow i}$, and for a given number of traces n it can be built as follows: given k possible sources, a probability value of 1 to the state called *on* must be assigned to one of the sequences of $|$ and $\circ$ (states of nodes $D_{1\rightarrow 2}\ldots D_{n-1\rightarrow n}$) including exactly $k-1$ $\circ$. Note that any of the sequences respecting this rule can be valid: in fact, they all represent a situation involving exactly k sources. The difference between the sequences lies in the size of the groups, i.e. in the number of traces coming from each source. Being this parameter not considered in this work, all the sequences can be considered equivalent[3].

Given k sources, there will be only $k-1$ nodes S_i that will be independent from node S_{i-1}. This corresponds to a separation of the potential sources at the origin of traces in exactly k groups.

Node C has to be instantiated to its state *on* when using the Bayesian network. This has to be done in order to avoid the separation of nodes S_i into a number of groups not coherent with the number of true sources fixed by node K. An example of the conditional probability table for node C for a case of $n=4$ traces is presented in Table 4, and the conditional probability table for nodes S_i is shown in Table 5.

This BN can be seen as a composition of two blocks: the first block consists of nodes K, D and C and focuses on *how* different sources can be at the origin of the traces; it is specific to every different n. The second block is formed by all couples of nodes S_i and E_i and considers the type of the recovered traces. According to this, the use of object-oriented Bayesian networks is very reasonable in our case.

3.2 Object-oriented Bayesian networks

An object-oriented Bayesian network (OOBN) [10] can be seen as a hierarchically ordered BN. OOBNs are composed of a master network that includes instance nodes along with nodes as defined in the previous section. Instance nodes represent themselves a network fragment on a lower hierarchical level, and can be connected by using so-called input and output nodes, which allow the network to propagate the belief through different hierarchical levels. Instance nodes can be reused multiple times in the same master network, so OOBNs are particularly useful when dealing with problems that present a modular structure. With

---

[2] As stated in [9], it is possible to face a situation where a group of variables not directly dependent on each other cannot assume simultaneously a certain configuration of their states. This constraint can be modeled by using a constraint variable, which states reflect the legal and illegal configurations respectively. The constraint is enforced by instantiating the variable to the state that represents the allowed configurations.

[3] For example, in the case of $n=4$ traces and $k=2$ sources the allowed sequences are $||\circ$, $|\circ|$ and $\circ||$. In the first and third case, three traces come from a same source and one from the other, whilst in the second case two traces come from one source and two traces from another. The size of the group of traces sharing a common source changes, but always the same number of sources (in this case 2) is involved.

Table 4: Example of the conditional probability table for node C for the case $n = 4$ traces. In this case, given $n = 4$ a probability of 1 to the state *on* is assigned to sequences ||| for $k = 1$, ||○ for $k = 2$, |○○ for $k = 3$ and ○○○ for $k = 4$.

	K:	1								2							
	$D_{1\to2}$:	|				○				|				○			
	$D_{2\to3}$:	|		○		|		○		|		○		|		○	
	$D_{3\to4}$:	|	○	|	○	|	○	|	○	|	○	|	○	|	○	|	○
C:	*on*	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	off	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
	K:	3								4							
	$D_{1\to2}$:	|				○				|				○			
	$D_{2\to3}$:	|		○		|		○		|		○		|		○	
	$D_{3\to4}$:	|	○	|	○	|	○	|	○	|	○	|	○	|	○	|	○
C:	*on*	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
	off	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0

Table 5: Conditional probability table for nodes S_i. Given that node C must be instantiated to its state *on*, only this part of the table is represented here.

	C:		*on*		
	$D_{i-1\to i}$:	|		○	
	S_{i-1}:	A	$\bar{A}$	A	$\bar{A}$
S_i:	A	1	0	γ_A	γ_A
	$\bar{A}$	0	1	$1 - \gamma_A$	$1 - \gamma_A$

respect to the BN, the OOBN has many structural and practical advantages, in particular it
facilitates hierarchical construction, utilizing small modular networks (network fragments)
as building blocks. The use of object-oriented Bayesian networks is largely attested in
forensic science [5], in particular in problems dealing with DNA evidence [e.g., 11, 12, 13].
In the present scenario, it is possible to define a first instance node ($n\ traces$), representing
a network's fragment composed by nodes K, $D_{i-1 \to i}$ and C (see Figure 2a), which is specific
to the number of traces n. Another instance node ($Trace\ i$) represents the part of the
network composed by couples of nodes S_i and E_i (see Figure 2b), and a third instance node
($Trace\ 1$) which, similarly to the previous one, is composed by nodes S_1 and E_1 (see Figure
2c). The latter is necessary because probabilities on states of node S_1 do not depend upon
any other variable.

It is possible to reconstruct the original Bayesian network by using the composition of these
three instance nodes (Figure 3).

(a) Network's fragment represented by instance node $n\ traces$.

(b) Network's fragment represented by instance node $Trace\ i$.

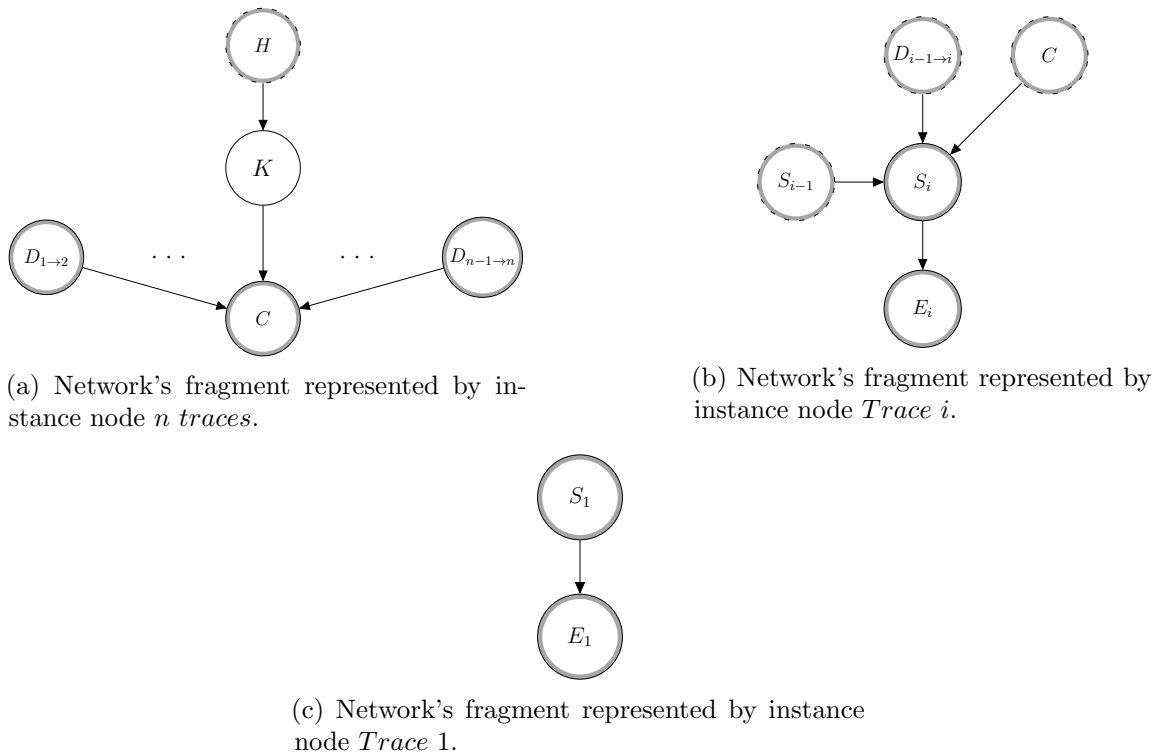(c) Network's fragment represented by instance node $Trace\ 1$.

Figure 2: Bayesian network's fragments corresponding to the three instance nodes used to
construct the object-oriented Bayesian network for the model. Input nodes are represented
by gray dotted contoured circles, whilst output nodes are represented by gray contoured
circles.

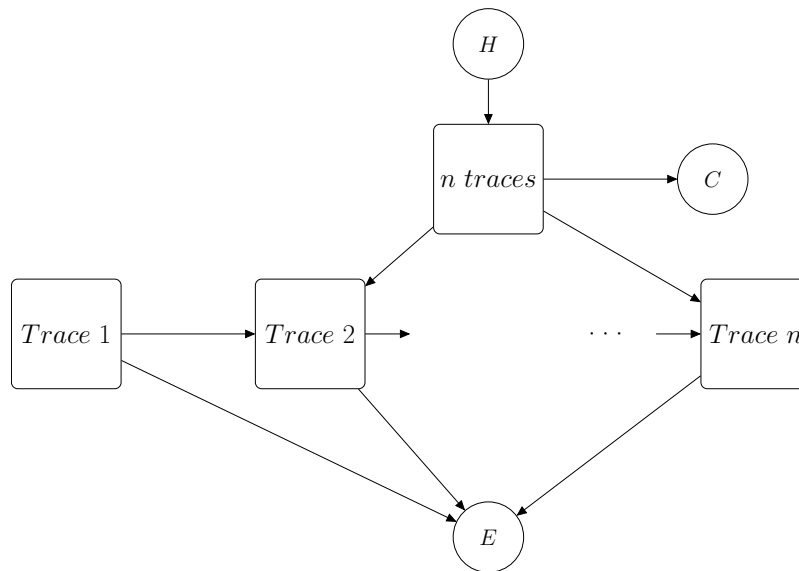Irrespectively to the number of traces, the proper OOBN can be constructed combining the

Figure 3: Structure of the object-oriented Bayesian network for n traces. Instance nodes are represented by squares. Node C, defined inside the instance node n *traces*, appears also in the master network. In this way, all nodes that must be instantiated when running the network are present on the master network.

three instance nodes previously defined.

4 Likelihood ratio formula

In order to obtain a verification of the results obtained through graphical models, it is possible to develop an analytical formula for the likelihood ratio. It is important to stress that both the graphical models and the formula are based on the same assumptions. For this reason, the formula can be considered as a validation of the numerical results issued from the graphical models.
Remind that the hypothesis of interest are:

- H_1: The n traces come from the same source;

- H_2: The n traces come from at least two different sources,

where the number of possible different sources is not specified.
The likelihood ratio can therefore be written as:

$$LR = \frac{Pr(E_1 = A, E_2 = A, \ldots, E_n = A|H_1)}{Pr(E_1 = A, E_2 = A, \ldots, E_n = A|H_2)} = \frac{Pr(E|H_1)}{Pr(E|H_2)},$$

where E represents the event that all traces E_i, with $i = 1 \ldots n$, share the same discrete characteristic A. In the following sections, the numerator and the denominator of the likelihood ratio will be studied separately.

4.1 Numerator

The likelihood ratio's numerator, $Pr(E|H_1)$ can be developed further by introducing a new variable S, that considers the type of the sources at the origin of the n traces. The event $S = A$ corresponds to all the sources involved being of type A, and $S = \bar{A}$ is the event that at least one of the sources involved is not of type A. Obviously, given H_1 only one source is involved.
Therefore, the numerator can be extended as follows:

$$Pr(E|H_1) = \underbrace{Pr(E|S = A, H_1)}_{1}\underbrace{Pr(S = A|H_1)}_{\gamma_A} + \underbrace{Pr(E|S = \bar{A}, H_1)}_{0} Pr(S = \bar{A}|H_1) = \gamma_A.$$

In fact, under the proposition that the n traces come from the same source, the probability of observing this kind of evidence knowing that the true source is of type A is 1, and the probability to observe the same traces knowing that their true source is of type $\bar{A}$ is 0. The factor $Pr(S = A|H_1)$ is equal to γ_A, which is again the probability that a source from the relevant population is of type A [3].

4.2 Denominator

Similarly, the denominator can be written as

$$Pr(E|H_2) = \underbrace{Pr(E|S = A, H_2)}_{1} Pr(S = A|H_2) + \underbrace{Pr(E|S = \bar{A}, H_2)}_{0} Pr(S = \bar{A}|H_2)$$
$$= Pr(S = A|H_2)$$

The probability that the sources at the origin of the traces are all of type A, knowing that the traces come from different sources, i.e. $Pr(S = A|H_2)$, depends on the number of sources involved. The number of sources potentially at the origin of traces can vary between 1 and n, therefore

$$Pr(S = A|H_2) = \sum_{k=1}^{n} Pr(S = A|K = k, H_2)Pr(K = k|H_2)$$

$$= Pr(S = A|K = 1, H_2)\underbrace{Pr(K = 1|H_2)}_{0} \tag{1}$$

$$+ \sum_{k=2}^{n} Pr(S = A|K = k, H_2)Pr(K = k|H_2).$$

$Pr(S = A|K = k, H_2)$ denotes the probability that all the sources are of the same type A knowing that exactly k sources are involved. If the k sources are assumed to be independent to each other, this term is equal to $\gamma_A{}^k$.

$Pr(K = k|H_2)$ represents the probability that exactly k sources are involved knowing that the traces come from different sources.

As explained in Section 3, this probability depends on the number of different configurations in which k sources can be at the origin of n traces. $Pr(K = k|H_2)$ can therefore be written as $Pr(K = k|H_2) = \{{}^n_k\}/_{B_n-1}$. Equation 1 becomes then:

$$\sum_{k=2}^{n} Pr(S = A|K = k, H_2)Pr(K = k|H_2) = \sum_{k=2}^{n} \gamma_A{}^k \frac{\{{}^n_k\}}{B_n - 1},$$

and the likelihood ratio's formula for the scenario considered is finally:

$$LR = \frac{\gamma_A}{\frac{1}{B_n-1} \sum_{k=2}^{n} \{{}^n_k\}\gamma_A{}^k}. \tag{2}$$

5 Model application

In order to illustrate how the n traces model works, a case involving $n = 4$ traces will be described.

Assuming a probability $\gamma_A = 0.2$, the proper BN and OOBN are shown in Figure 4.

The same result can be obtained by using equation (2), that can be written as:

$$LR = \frac{0.2}{\frac{1}{B_4-1} \sum_{k=2}^{4} \{{}^4_k\}0.2^k}$$

$$= \frac{0.2}{\frac{1}{14}\left[7 \cdot 0.2^2 + 6 \cdot 0.2^3 + 1 \cdot 0.2^4\right]}$$
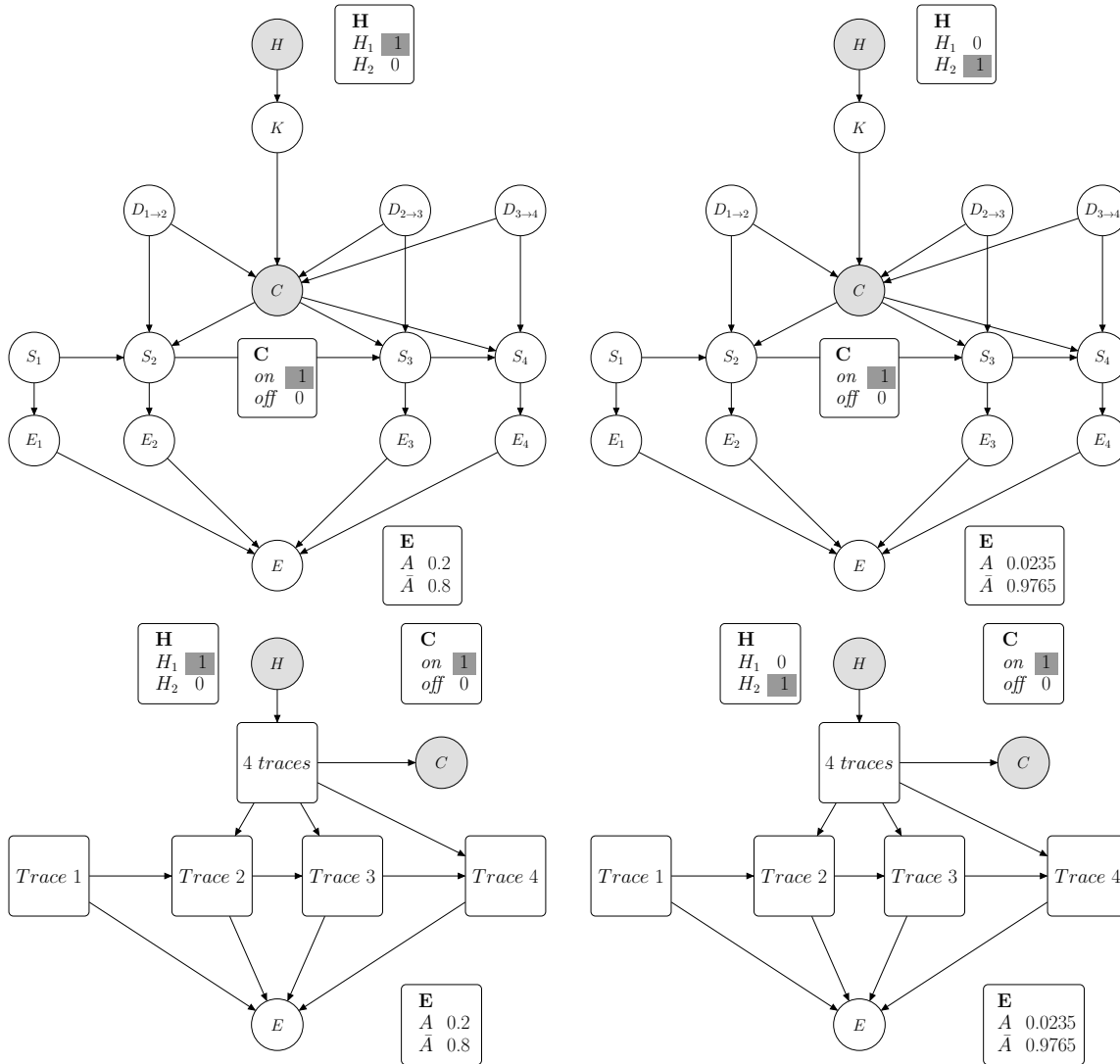
$$= \frac{0.2}{0.0235}$$

$$= 8.5$$

Figure 4: Bayesian networks (top) and object-oriented Bayesian networks (bottom) for $n = 4$ traces. On the left, node H is instantiated to its state H_1, and probability associated to state A of node E corresponds to the likelihood ratio's numerator. On the right, node H is instantiated to its state H_2, and so variable E's state A corresponds to likelihood ratio's denominator.

Simulations conducted using a number of traces varying between two and ten with a fixed value of γ_A show that results obtained with the formula 2 coincide with those obtained when using Bayesian networks and object-oriented Bayesian networks. Furthermore, results obtained in [3] coincide with those obtained when using this model in the particular case of $n = 2$ traces.

In Figure 5, the LR as a function of probability γ_A is represented in correspondence of an increasing number of traces.
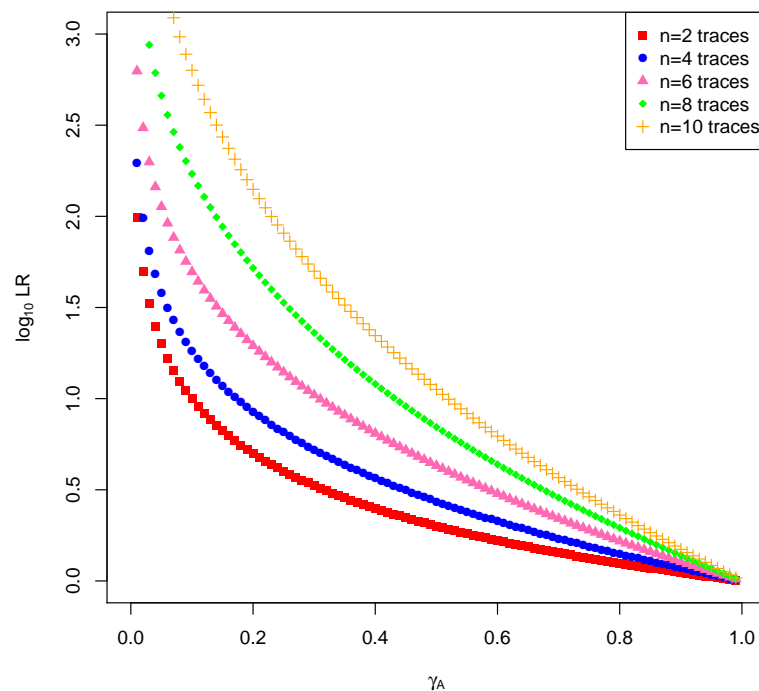

Figure 5: Logarithm of likelihood ratio as a function of probability γ_A for $n = 2, 4, 6, 8$ and 10 traces

As expected, the LR decreases as values of γ_A increase regardless of the number of traces recovered, but it always supports the proposition H_1.

It can also be observed that the number of traces plays a minor role when probability γ_A is high (i.e. when the considered characteristic A is common in the relevant population), whilst, for lower values of γ_A, the number of traces has a strong impact on the LR's magnitude.

6   Discussion and conclusions

The evaluation of a comparative analysis of multiple pieces of evidence can be of high interest from a forensic point of view, even though a putative source of the traces is not available, notably in investigative settings. For example, this kind of approach can allow the forensic scientist to provide some case linkage, which can be extremely helpful in both an investigative and an evaluative framework [14]. Taroni et al. [3] already discussed the problem in the case of two items of evidence, for both discrete and continuous trace characteristics. Through this paper, the inferential model was extended in order to consider a number of items of evidence possessing discrete attributes greater than two ($n > 2$). Developments were performed based on graphical models, namely Bayesian networks and object-oriented Bayesian networks. Additionally, an analytical formula for the LR was developed, in order to verify the results issued from the graphical models.

Though the LR is usually applied to evaluate the weight of an item of evidence given two competing propositions involving a putative source, its use in investigative frameworks is widely attested [see, e.g. 15]. The current work further supports the use of LR as a powerful statistical tool that can provide relevant information also in an early stage of the investigations process.

The study shows that, besides the mathematical equations, the use of Bayesian networks allows one to structure the inferential problem of interest in a rigorous manner by expressing dependencies between the variables involved in the problem. Therefore, this probabilistic tool facilitates the handling of uncertainty related to the problem in a logical and coherent way, which does not request any particular expertise in statistics or mathematics. This is a great advantage for the forensic practitioners and experts, since they could focus on the evaluation procedure rather than on mathematical computation. The flexible nature of Bayesian networks is highlighted when dealing with multiple items of evidence. The model is extended taking advantage of repetitive sub-structures; in this respect, OOBN is adapted in a particular way.

These results have to undoubtedly be interpreted in the light of the assumptions only. In particular, one should not consider circumstantial and investigative information in the model, even though it could provide substantial elements to interpret the case more accurately.

A further development of the model investigating the influence of non scientific elements such as the modus operandi or location of the traces on the crime scenes can therefore be envisaged.

References

References

[1] J. Birkett, Scientific scene linking, Journal of the Forensic Science Society 29 (1989) 271–284. doi:10.1016/S0015-7368(89)73262-X.

[2] O. Ribaux, S. J. Walsh, P. Margot, The contribution of forensic science to crime analysis and investigation: forensic intelligence., Forensic science international 156 (2006) 171–81. doi:10.1016/j.forsciint.2004.12.028.

[3] F. Taroni, S. Bozza, A. Biedermann, Two items of evidence, no putative source: An inference problem in forensic intelligence, Journal of Forensic Sciences 51 (2006) 1350–1361. doi:10.1111/j.1556-4029.2006.00272.x.

[4] C. Aitken, F. Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists, 2nd Edition, John Wiley & Sons, Chichester, 2004.

[5] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino, C. Aitken, Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science, 2nd Edition, John Wiley & Sons, Chichester, 2014.

[6] F. Taroni, A. Biedermann, Bayesian Networks, in: J. Siegel, P. Saukko (Eds.), Encyclopedia of Forensic Sciences, 2nd Edition, Vol. 1, Academic Press, Waltham, 2013, pp. 351–356.

[7] R. A. Brualdi, Introductory Combinatorics, 4th Edition, Pearson Prentice Hall, Upper Saddle River, 2004.

[8] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics: A Foundation for Computer Science, 2nd Edition, Addison-Wesley, Boston, 1994.

[9] U. Kjærulff, A. Madsen, Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis, Springer Science, New York, 2008.

[10] D. Koller, A. Pfeffer, Object-Oriented Bayesian Networks, in: D. Geiger, P. Shenoy (Eds.), Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97), Morgan Kaufmann Publishers Inc., San Francisco, 1997, pp. 302–313.

[11] G. Cereda, A. Biedermann, D. Hall, F. Taroni, Object-oriented Bayesian networks for evaluating DIP-STR profiling results from unbalanced DNA mixtures, Forensic Science International: Genetics 8 (2014) 159–169. doi:10.1016/j.fsigen.2013.09.001.

[12] R. G. Cowell, S. L. Lauritzen, J. Mortera, Object-oriented Bayesian networks for DNA mixture analyses, Available: `http://citeseerx.ist.psu.edu/viewdoc/download;`

`jsessionid=2BAB58A71CBD3BBF7F9ED8600ADBD7A2?doi=10.1.1.102.1834&rep=`
`rep1&type=pdf` accessed: 11.05.16 (2006).

[13] A. P. Dawid, J. Mortera, P. Vicard, Object-oriented Bayesian networks for complex forensic DNA profiling problems, Forensic Science International 169 (2007) 195–205. doi:10.1016/j.forsciint.2006.08.028.

[14] G. Jackson, S. Jones, G. Booth, C. Champod, I. W. Evett, The nature of forensic science opinion - A possible framework to guide thinking and practice in investigations and in court proceedings, Science and Justice 46 (2006) 33–44. doi:10.1016/S1355-0306(06)71565-9.

[15] F. Taroni, R. Marquis, M. Schmittbuhl, A. Biedermann, A. Thiéry, S. Bozza, The use of the likelihood ratio for evaluative and investigative purposes in comparative forensic handwriting examination, Forensic Science International 214 (2012) 189–194. doi:10.1016/j.forsciint.2011.08.007.