Bayesian inferential reasoning model for crime investigation

WANG, Jing <http://orcid.org/0000-0002-5418-0217> and XU, Zhijie

Bayesian Inferential Reasoning Model for Crime Investigation

Jing WANG and Zhijie XU[1]
Visualisation, Interaction and Vision (VIV) Research Group, School of Computing and Engineering, University of Huddersfield, UK

Abstract. Forensic inferential reasoning is a "fact-finding" journey for crime investigation and evidence presentation. In complex legal practices involving various forms of evidence, conventional decision making processes based on human intuition and piece-to-piece evidence explanation often fail to reconstruct meaningful and convincing legal hypothesis. It is necessary to develop logical system for evidence management and relationship evaluations. In this paper, a forensic application-oriented inferential reasoning model has been devised base on Bayesian Networks. It provides an effective approach to identify and evaluate possible relationships among different evidence. The model has been developed into an adaptive framework than can be further extended to support information visualisation and interaction. Based on the system experiments, the model has been successfully used in verifying the logical relationships between DNA testing results and confessions acquired from the suspect in a simulated criminal investigation, which provided a firm foundation for the future developments.

Keywords. Bayesian Networks, Inferential Reasoning, Digitised Forensic Evidence

1. Introduction

Gathering and testing forensic evidence is one of the most important tasks in a criminal investigation. It forms the basis for building up causality relationships between crime suspects and the victims, and reveals the hidden stories behind the chaotic crime scenes. In recent years, Digitised Forensic Evidence (DFE) has become increasingly popular in continuous advancing forensic science and technology domain with much evidence being directly generated from electronic equipment or computer network. Recent developments in forensic science research have shown that DFE should be treated as an integral part of the concept of "Big Data" [1]. Those incremental DFE datasets present serious challenges to the traditional forensic investigation approaches.

Psychological research has revealed that the accumulative style piece-by-piece DFE presentation often failed to generate linear aggregation of "knowledge" regarding a particular case, although the so-called "truth" kept stacking up. This difficulty has

[1] Corresponding Author

been highlighted in many complicated cases involving large quantity of DFE especially for reinvestigating and reviewing so-called "cold" cases.

The research introduced in this paper aims at developing a DFE inferential modelling framework for forensic information management through enabling DFE to be represented as "nodes", which can be edited and logically tested in customisable manners. The prototype system formulates the causality relationships between DEF nodes and verifies their statistical "likeliness" across the entire DFE chains.

The fundamental theory for the aforementioned analysis is the Bayesian Networks (BNs), which has been studied extensively in many forensic literature such as [2, 3]. In this research, the effort has been concentrated on building up an adaptable and flexible modelling formwork for practical forensic inferential reasoning. The model follows a hierarchical topological structure based on the distinctive nature of forensic investigations. It focuses on highlighting the logical relationships between crime suspects and their victims through evidence evaluation.

The rest of this paper is organised in the following order: a brief review of recent advancements in forensic research has been provided in Section 2. Section 3 focuses on designing a forensic application-oriented topological structure based on BNs. This structure has then been extended into the so-called forensic inferential reasoning framework in Section 4. Section 5 tests the model by using simulated crime data and legal hypothesis, which lead to discussions and future improvements illustrated in Section 6.

2. Literature Review

Forensic science includes many research disciplines across biology, psychology, chemistry, information science and computer theories. DEF drawn from fingerprint analysis [4], Internet and WWW [5], CCTV systems [6] and even dental identification [7] *et al*. have been widely used in many real world applications.

Although forensic experts can extract "accurate" information from a single piece of evidence such as DNA profile, establishing (or denying) the relationships between those pieces of evidence are still a challenging task. Since 2009, with the great improvements on mobile and network technologies, new crime evidence formats, such as website logs, text messages and photos shared across the Internet, have been introduced to court proceedings, which further increased the amount of DEF during crime investigation. The organising and interpreting of those DFE have become a daunting task even for the "trained" hands. Those problems have been summarised in the official reports such as [8].

Recent research have seen attention been paid to alleviate those problems by using intelligent decision making techniques such as machine learning and logical theories [9]. For example, Biedermann *et al*. [10] has introduced an inferential algorithm to test different DNA profiling possibilities through a BNs model called qualitative probabilistic networks (QPNs). Halliwell *et al*. [11] has investigated the parameter optimisation problems in forensic statistics. Recently, Han *et al*. [12] has introduced a high level contextual cue with the observed evidential information being applied into forensic reasoning in the form of subjective option function. A comprehensive report of intelligent forensic research has been composed by Aitken *et al*. [13], which focused on solving the statistic and probability problems of DFE within the context of criminal trials.

Those techniques and algorithms have shown that although the testing methods may vary in evidence types, it is possible to model their logical and causality relationships by using contextual information extracted from their testing results, which is the motivation of this research to enable the construction of the DEF's relationship models through a flexible and adaptive framework.

3. Bayesian Networks for DEF Inferential Reasoning

3.1. DEF nodes definition

Forensic investigation is a so-called truth-rebuilding task based on testing and evaluating different forms of evidence. Logical relationships across those evidence pieces are full of uncertainties and possibilities. Those criminal stories are usually built upon logical inferential reasoning. Bayesian networks (BNs) provide useful mathematical tools for handling those uncertainties.

BNs are mainly used for inferential reasoning and decision making, which is a research hot-spot for many artificial intelligence-based applications. BNs are a convergence of graph theory and probability theory composed by "nodes", "arcs", and corresponding sets of probabilities. The graph of BNs model is known as "directed acyclic graph" (DAG) with finite number of nodes and arcs. The nodes represent events of interest and directed arcs denote the probabilistic relationships, such as causality and spatial-temporal locations.

For forensic applications introduced in this paper, the nodes of the DAG have been defined as "claims" (denote as C) during crime investigation, which have multiple possibilities during the inference. Each claim contains a factual predicate. Such as C_1: "I was not at the crime scene"; C_2: "two blood stains are matched"; C_3: "Tom was guilty". Those claims contain uncertainties since people may tell lies (C_1); the testing may contain false positive results (C_2); or the judgement comes from one of many possible scenarios (C_3).

The uncertainty is modelled by the statistic possibility which come from knowledge and understanding based on personal experience, published surveys from social science, physiological studies and biological research.

3.2. DFEs Chain Formulation

In this paper, a crime example has been used for testing and discussing. It is described as follow (detailed case circumstances can be referred to [14]):

"*A balaclava was discarded by an offender at the scene of a robbery, which was quickly retrieved by a witness and handed to the police. A suspect was arrested six hours after the incident and combings were taken from his head hair. The suspect said the balaclava was his. He used to wear it regularly but hadn't seen it since last two months and he assumed it was taken by someone else.*"

In this example, the logical question is: "Is the suspect wore the balaclava at the scene of the robbery?" The forensic scientist requested an examination to link the suspect with the balaclava. By using BNs, some essential claims should be modelled:

C_1: The suspect is the man who wore the balaclava at the relevant time;

C_2: A scientist's report of a match between the suspect's profile and the profile of the sample by using DNA testing technology;

C_3: The suspect said he lost the balaclava two month ago;
C_4: The suspect is the offender.

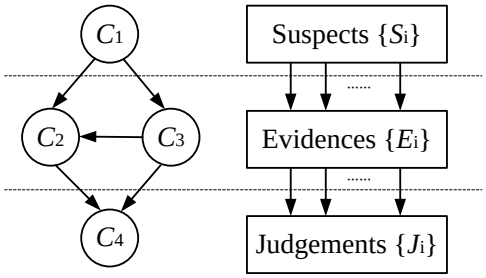

Figure 1. DAG of example case and its hierarchical framework

The DAG can be defined on the left hand side of Figure 1 based on those claims. The Conditional Probability Distribution (CPD) of each node has been listed from Table 1 to 4, which are based on the experiences from the related report. At the top of this graph, a claim about the relationship between a suspect and a balaclava is defined and should be tested. This claim has two possible answers, "true" or "false". It would be true if the observer matched the claim, otherwise it would be false, *i.e.* "the suspect is not the man who wore the balaclava at the relevant time".

Table 1. Conditional Probability Table (CPT) of C_1

C_1=false (F)	C_1=true (T)
0.5	0.5

Table 2. CPT of C_2

C_1	C_3	C_2=no-match (0)	C_2=single-match (1)	C_2=multi-match (2)
F	F	0.12	0.76	0.12
F	T	0.07	0.24	0.69
T	F	0.01	0.92	0.07
T	T	0.01	0.85	0.14

Table 3. CPT of C_3

C_1	C_3=F	C_3=T
F	0.5	0.5
T	0.99	0.01

Table 4. CPT of C_4

C_2	C_3	C_4=F	C_4=T
0	F	0.83	0.17
0	T	0.96	0.04
1	F	0.10	0.90
1	T	0.27	0.73
2	F	0.18	0.82
2	T	0.31	0.69

Based on those claimed assumptions, many evidence pieces can be collected and tested during the police investigation. Two evidence pieces are used as claims for this example. Those claims are also contains many possibilities which approve or disapprove their claims.

A judgement C_4 has been claimed based on above mentioned suspects and evidence, which should connect with all related evidence in the DAG.

Most claims, such as C_1, C_3, and C_4 only have true or false value. But based on the nature of the forensic application, the probability tables of those claims are not necessarily be binary. C_2, for example, which usually contains many DNA matching possibilities such as no matching, single profile match or multiple profile matches.

4. General Forensic Inferential Reasoning Framework

4.1. Operational Principles

In this research, a forensic application-originated inferential reasoning framework has been developed. As illustrated on the right hand side of Figure 1. The framework has a three-layer hierarchical structure that contains the Suspect (*S*), Evidence (*E*) and Judgement (*J*). The arrows between each layer represent the directions of their inferential reasoning segments. *i.e*. suspects should have causality relationships with evidence and those evidence pieces should support some suitable judgements.

The container *S*, *E*, *J* are sets $\{\bullet\}$ of its members S_i, E_i, J_i. By using this general framework, each claim C_i from specific crime case should be categorised into different containers based on the context of claims. As the example illustrated in the Figure 1, $S=\{S_1=C_1\}$, $E=\{E_1=C_2, E_2=C_3\}$, and $J=\{J_1=C_4\}$.

Suspects *S* is a collection of hypothesis claims. It can be criminal behaviours, crime scenarios, or related victims and suspects, which can trigger a series of police investigations built up on the hypothesis.

During the investigation, a group of evidence pieces can be gathered and tested. The claimed test results or assumptions belong to the Evidence container *E*. It can be defined by biological tests outputs, clips of CCTV video footages, mobile messages, suspect/witnesses/victim interviews, even some social network informations.

The evidence (*i.e.* the subset of *E*) is aimed to explain certain details of a crime scenario which is used for inferring the crime judgements *J*. The context of the claims in *J* should describe the possible relationship between Suspects *S* and the criminal behaviours. It is clear that evidence has been recognised as a bridge between crime suspect and juristic judgment.

4.2. Evidence Chain

Some crime evidence pieces are not existed independently. The causality, spatial and temporal relationships between two evidence items usually have significant impact on the judgement making during the inferential reasoning. For example, in the Figure 1, there is an arc point to C_2 from C_3. If C_3 was true, higher possibility of two different DNA profile would be matched. Since the possible testing outputs may contain single or multiple DNA profile matching results, the observed value of C_3 should have impact on the possibility distribution of C_2.

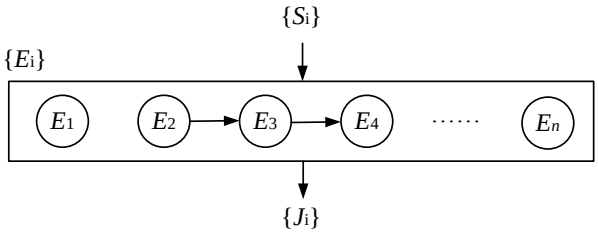

Figure 2. Evidence chain

In general, we can define the evidence chain inside the container. As illustrated in Figure 2, the chain can be recognised as extra arcs between related evidence, which represent the causality, spatial and temporal impacts from certain evidence to another. Under this framework, a logical relationship between evidence *i* and *j* with its suspect *S* can be recognised as the topological structure contains two parents (*S* and *Ei*) and one child E_j. For example, in Table 2, those parents' nodes were used for building up the CPD of E_j during inference, which means claims from *S* and E_i has logical impact on E_j.

5. Evaluation and Discussion

Establishing a BNs-based forensic inferential reasoning framework provide a theoretical solution for helping people organise the DEFs and to understand their relationships. Based on the nature of police investigation, legal practices often demond the abilities to test DFEs by changing their parameters interactively. For example, investigators wish to see if it is necessary to add more evidence by increasing the possibility of a lie told by the suspect.

A series of experiments has been carried out for testing the availability of the developed inferential model. The experiments have been designed to maintain the parameters interactively, *i.e*. the CPD from evidence container *E*, to review the possibility changes in other containers and the impact to the entire logical framework.

As introduced in Section 3.2, the experiment was based on the "Lost Balaclava" crime case. In the experiment, it is believed the DNA tests results and its statistical parameters listed in Table 2 are accurate. People interest in the impact from E_2 to J_1 that is the possibility of recognising the suspect was the offender if the he told a lie. In the experiment, the parameter lists in the Table 3 was maintained.

The possibility distribution $P(E_2=F|S_1=F)$ was set from 0 to 1 by step 0.01 which means the experiment reduced the trust of the suspect's defence 1% present step by step based on the belief that the suspect didn't wear the balaclava. The distribution of $P(J_1=T|S_1=F,E_1=1)$ was illustrated in Figure 4. The lines are all monotonically increasing, which is easy to understand as, in this case, if the suspect told lies, he/she had more chance to be found guilty.

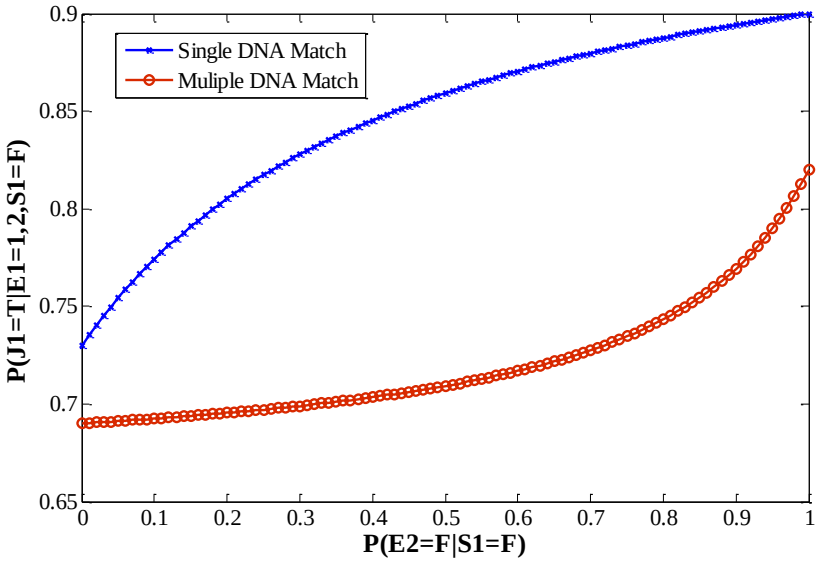

Figure 3. Judgment distributions based on the trust of suspect's interviews

Defendant lawyer could argue the test results if there are multiple DNA matching in the balaclava, the distribution, $P(J_1=T|S_1=F,E_1=2)$ has also been listed in Figure 3 for comparison. It is clear that under the same possibility of the suspect told a lie, the single DNA matching make people believe the suspect was guilty than the multiple matched one. In fact, scientific test results, such as DNA matching, fibres matching, body liquid analysis results *et al.*, are trusted by the prosecutors and also have stronger impact on their decision making.

It is worth noting that after the 40% threshold being exceeded of $P(E_2=F|S_1=F)$, the single matched DNA result rises slower, which means the impact from the suspect confession was reduced. The line never reaches 100% since some false positive test possibilities have to be considered. In the Table 4, $P(J_1=T|E_1=1,E_2=F)$=90%.

It is also worth noting that after the 80% threshold been passed of $P(E_2=F|S_1=F)$, the line of multiple matched DNA rise significantly, which means although multiple DNA profiles was found, if people believe a lie was told, the suspect still have more possibility to be recognised as an offender. Figure 3 has proved that when people making decisions, the impact from confessions should be considered if the scientific test results were not convinced. In addition, many useful logic theories, such as argument theory [15], can support the evaluation of the truth of a confessions based on a set of assumptions concerning mutually acceptable conclusions.

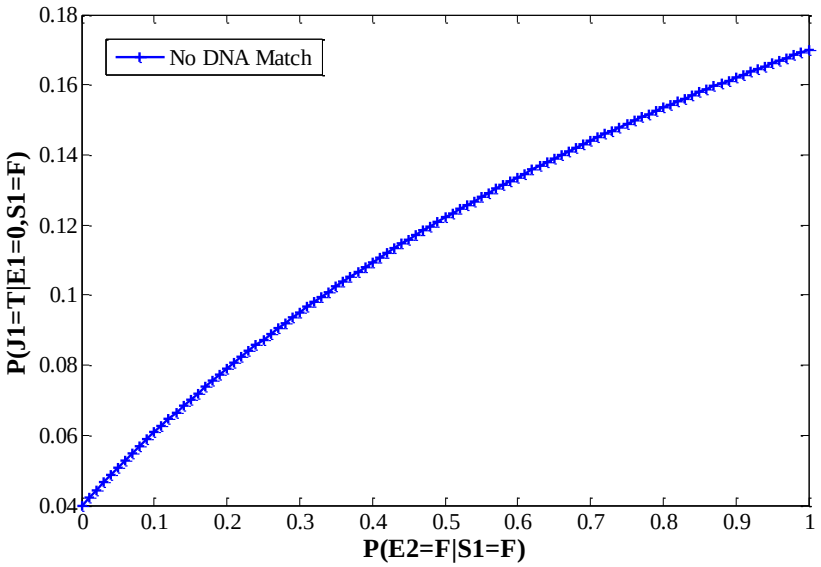

Figure 4. Judgment distributions on "No DNA profile was found"

Figure 4 listed the results on Judgment distributions on "No DNA profile was found". It means even people recognise what the suspect's saying is a 100% lie, just because there was no DNA matching, there were only 17% possibilities to charge the suspect. More police investigations and evidence will need to be provided.

The experiment result shows that, the developed forensic inferential reasoning framework is a valid model to describe the relationships between the pieces of evidence and judgements. The simulation output can describe the evidence relationship and their impact on the decision making. In real applications, prosecutors and defendant lawyers usually pay different attention on evidence then make different conclusions. This method can quantify the relationships between suspects and judgements based on the logical impacts coming from evidence pieces, which minimised the ambiguous decision making during the police investigation and legal debate.

6. Conclusion and future work

A forensic application-orientated inferential reasoning method has been introduced in this paper. The model is based on BNs and has been designed as three layers hierarchical structure including suspects, evidence pieces and judgments. The model highlights the impact from forensic evidence which bridge the gap between assumptions and conclusions. Inferential reasoning progress of a pedagogical example "Lost Balaclava" has been tested under the structure of the developed framework. The system has been proved as a valid approach to understanding the forensic inferential reasoning progress.

As mentioned in the Section 1, the complicated inter- and intra-relationships of detailed DEFs are one of the main challenges of modern forensic science. In some complicated crime cases, a huge number of detailed evidence pieces are collected for

analysis, which means in the Figure 2, large quantity of evidence nodes and some complicated evidence chains are existed. When interacting with these parameters, even professional BNs experts cannot easily tell the meaningful changes behind the BNs model, never mentioning those lay users involved in the process. Therefore, it is necessary to build up a dynamic BNs model for editing the evidence nodes to handle partially -unknown topological structures with partial observability.

In addition, when the claims contain multiple possibility distributions, the interaction of those multi-dimensional features can also be a huge challenge for the real-world applications. For tackling those problems, a user friendly and multi-dimensional information visualisation and interaction system need to developed.

References

[1] A. Singh. Big Data Scalability. *The International Journal of Big Data*, vol. 1, no. 3, 2014.

[2] M. Buscema and W. J. Tastle. *Intelligent Data Mining in Law Enforcement Analytics: New Neural Networks Applied to Real Problems*: Springer, 2013.

[3] N.-H. Chiu, C.-E. Pu and M.-C. Hsieh. An Intelligent System for Reconstructing the Ripped-up Paper-Moneys. *ICEIS 2013*, pp. 461, 2013.

[4] S. Kiltz, M. Hildebrandt, J. Dittmann and C. Vielhauer. Challenges in contact-less latent fingerprint processing in crime scenes: Review of sensors and image processing investigations. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, IEEE, pp. 1504-1508, 2012.

[5] S. Mukkamala and A. H. Sung. Identifying significant features for network forensic analysis using artificial intelligent techniques. *International Journal of digital evidence*, vol. 1, no. 4, pp. 1-17, 2003.

[6] S. Han, A. Hutter and W. Stechele. Toward contextual forensic retrieval for visual surveillance: Challenges and an architectural approach. In *Image Analysis for Multimedia Interactive Services, 2009. WIAMIS'09. 10th Workshop on*, IEEE, pp. 201-204, 2009.

[7] T. Chomdej, W. Pankaow and S. Choychumroon. Intelligent dental identification system (IDIS) in forensic medicine. *Forensic science international*, vol. 158, no. 1, pp. 27-38, 2006.

[8] The judgment of the Court of Appeal in RvT. EWCA Crim 2439, 2010.

[9] V. Civie and R. Civie. Future technologies from trends in computer forensic science. In *Information Technology Conference, 1998. IEEE*, IEEE, pp. 105-108, 1998.

[10] A. Biedermann and F. Taroni. Bayesian networks and probabilistic reasoning about scientific evidence when there is a lack of data. *Forensic science international*, vol. 157, no. 2, pp. 163-167, 2006.

[11] J. Halliwell, J. Keppens and Q. Shen. Linguistic bayesian networks for reasoning with subjective probabilities in forensic statistics. In *Proceedings of the 9th international conference on Artificial intelligence and law*, ACM, pp. 42-50, 2003.

[12] S. Han, B. Koo, A. Hutter and W. Stechele. Forensic reasoning upon pre-obtained surveillance metadata using uncertain spatio-temporal rules and subjective logic. In *Image Analysis for Multimedia Interactive Services (WIAMIS), 2010 11th International Workshop on*, IEEE, pp. 1-4, 2010.

[13] C. Aitken et al. Expressing evaluative opinions: a position statement. *Science & justice*, vol. 51, no. 1, pp. 1-2, 2011.

[14] G. Jackson, C. Aitken and P. Roberts. Case Assessment and Interpretation of Expert Evidence Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses, Royal Statistical Society's Working Group on Statistics and the Law, 2013.

[15] T. J. Bench-Capon and P. E. Dunne. Argumentation in artificial intelligence. *Artificial intelligence*, vol. 171, no. 10, pp. 619-641, 2007.

Revision Specification:

Based on the requirements from the reviewers, following content of this paper has been changed:

1. To address the suggestion from Reviewer #1 from Section 3.1, Paragraph 1: "*…Before an accused people or victims can "tell" a real criminal story, all the scenarios and their logical relationships are full of uncertainties and possibilities..."* Has been changed to:
   "Forensic investigation is a so-called truth-rebuilding task based on testing and evaluating different forms of evidence. Logical relationships across those evidence pieces are full of uncertainties and possibilities. Those criminal stories are usually built upon logical inferential reasoning. Bayesian networks (BNs) provide useful mathematical tools for handling those uncertainties."

2. To address the suggestion from Reviewer #2 from Section 6, the First paragraph has been rewritten:
   "A forensic application-orientated inferential reasoning method has been introduced in this paper. The model is based on BNs and has been designed as three layers hierarchical structure including suspects, evidence pieces and judgments. The model highlights the impact from forensic evidence which bridge the gap between assumptions and conclusions. Inferential reasoning progress of a pedagogical example "Lost Balaclava" has been tested under the structure of the developed framework. The system has been proved as a valid approach to understanding the forensic inferential reasoning progress."

Proof-reading of the English Presentation

To address the suggestion from Reviewer #1 and #2, the English presentation, grammar and spelling have been double checked through the paper to guarantee the correctness.