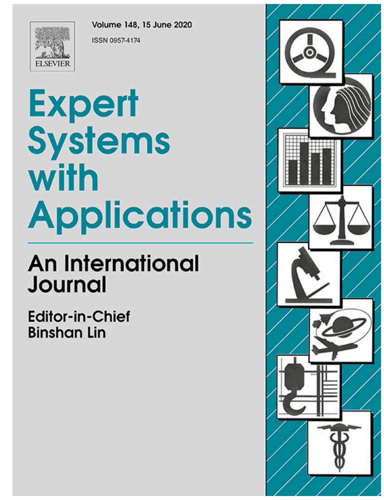
Journal Pre-proofs



A Knowledge-based Reasoning Model for Crime Reconstruction and Investigation

Litao Wang, Meisheng Jia, Yi Shi, Feiyu Chen, Shunjiang Ni, Shifei Shen

PII: S0957-4174(20)30435-8
DOI: https://doi.org/10.1016/j.eswa.2020.113611
Reference: ESWA 113611

To appear in: *Expert Systems with Applications*

Received Date: 11 February 2020
Revised Date: 14 May 2020
Accepted Date: 27 May 2020

Please cite this article as: Wang, L., Jia, M., Shi, Y., Chen, F., Ni, S., Shen, S., A Knowledge-based Reasoning Model for Crime Reconstruction and Investigation, *Expert Systems with Applications* (2020), doi: https://doi.org/10.1016/j.eswa.2020.113611

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Ltd.

Journal Pre-proofs

A Knowledge-based Reasoning Model for Crime Reconstruction and Investigation

Litao Wang[abc]: 327781466@qq.com

Meisheng Jia[ab]: meisheng_jia17@163.com

Yi Shi[d]: shiyi@cifs.gov.cn

Feiyu Chen[ab]: 13120165540@163.com

Shunjiang Ni[abc*]: sjni@tsinghua.edu.cn

Shifei Shen[abc*]: shensf@tsinghua.edu.cn

[a] *Institute of Safety Science and Technology, Department of Engineering Physics, Tsinghua University, Beijing 100084, China*

[b] *Institute of Public Safety Research, Tsinghua University, Beijing 100084, China*

[c] *Beijing Key Laboratory of City Integrated Emergency Response Science, Tsinghua University, Beijing 100084, China*

[d] *Institute of Forensic Science, Ministry of Public Security, Beijing 100038, China*

*Corresponding author.

Journal Pre-proofs

Journal Pre-proofs

A Knowledge-Based Reasoning Model for Crime Reconstruction and Investigation

Abstract

Artificial intelligence has been successfully applied in many areas including forensic sciences. Perhaps all forensic works can be regarded as helping reconstruct crimes, i.e. clarify and sequence the events that took place in the commission of a crime through evidence. However, there are few researches on the crime reconstruction using artificial intelligence methods. In this paper, we present a model based on Bayesian networks to help solve crimes. The model, which is termed 'case-type based model', is based on the knowledge of a type of crimes. We use Bayesian networks to represent the knowledge and conduct the uncertainty reasoning. We propose a growth algorithm of Bayesian networks to adapt the model to different cases. The model was tested through a real case, and the results indicate that the model can provide effective investigation suggestions and achieve the crime reconstruction.

Keywords: Artificial intelligence; Forensic science; Bayesian networks; Criminal investigation; Uncertainty reasoning; Evidence

1 Introduction

For a crime, criminal police cooperate with forensic scientists in reconstructing the events at a crime scene, namely who committed the crime, what was used to commit the crime, when/where did the crime occur, how did the crime occur, why did the crime occur and so on (Houck, Crispino, & Mcadam,2018;Rossy & Ribaux,2014). 'Crime reconstruction' is the term used to describe that process, which can be defined as analyzing the related physical evidence and information to clarify and sequence the events that took place in the commission of a crime (Chisum,2006; Chisum & Turvey,2011). Crime reconstruction is important for social justice and security, but it is actually a difficult task in the real world.

One of the most important things for crime reconstruction is evidential reasoning, namely reason the associated criminal actions through evidence, which can also be termed as 'Abduction' (Jackson, Aitken, & Roberts,2015; Schum,2002). The nature of evidential reasoning is in fact the uncertainty reasoning, and there exist different hypotheses concerning actions that can explain the same evidence. Nevertheless, we can express the uncertainty degree of different hypotheses through qualitative and quantitative methods. The Bayesian network (Geiger, Verma, & Pearl,1990; Pearl,1986) is a promising artificial intelligence method to conduct uncertainty reasoning (Huang, Cai, Yuan, & Chen,2019; Kabir & Papadopoulos,2019), where it uses probability to express the degree of uncertainty. Besides, the reasoning process can operate on Bayesian network by propagating information in any direction, hence it has an ability in both forward reasoning (prediction) and backward reasoning (abduction) (Pearl, 2011). Therefore, we can use the Bayesian network to model the forward crime-commission process and conduct backward reasoning, which is very helpful for crime reconstruction.

The Bayesian networks have gained increasing attention in the area of forensic science in recent years, especially in evaluating evidence. Gittelson, Biedermann, Bozza, and Taroni (2012) used Bayesian networks to study the value of the evidence for the forensic two-trace transfer problem. Wieten, De Zoete, Blankers, and Kokshoorn (2015) evaluated traces found on adhesive tapes using Bayesian networks. Taylor, Biedermann, Hicks, and Champod (2018) proposed a template of constructing Bayesian networks for evaluating forensic biology traces. In addition, Biedermann and Taroni (2012) gave a review of evaluating forensic DNA profiling evidence through Bayesian networks. For evaluating evidence, forensic scientists focus on the probability of evidence given different propositions regarding activities, while for crime reconstruction, it should be concentrated on the probability of events/scenarios given evidence.

There have been some studies on the crime reconstruction or criminal investigation using Bayesian networks. Wang and Xu (2014) proposed a model based on Bayesian networks for crime investigation, which was designed as three layers structure including suspects, evidence and judgments, and an example “Lost Balaclava” was tested under the structure of the developed framework. Verheij et al. (2016) studied three normative reasoning frameworks from the literature: arguments, scenarios and probabilities, discussed the hybrid model of them (which includes Bayesian networks), and utilized them to analyze a specific murder case. Noor, Asmara, Saman, and Hitam (2014) utilized Bayesian networks, where the nodes were separated into two kinds: ‘Causes’ and ‘Evidences’, to assist crime investigation, and analyzed a hypothetical crime case.

However, with the guidance of those studies, we need to establish specific Bayesian networks for each crime case, that is, we need to define the specific nodes according to the case, link those nodes under a reasoning framework, and determine the parameters of prior probability and conditional probability, only then can we conduct the analysis. It is clearly that the key and difficult point is building Bayesian networks, which is a challenge for criminal police. Criminal police have professional skills and rich experience, but they are not familiar with Bayesian networks, which implies that such technology may be seldom used by them for real cases.

Motivated by this, we aim to design a more general and applicable Bayesian model for crime reconstruction and investigation. More specifically, we aim to build a case-type based Bayesian model, namely a model that can be utilized to analyze any crime that belongs to a specific case type, like murder, rape, arson and so on. Such model is based on the generality (knowledge) of a kind of crimes, utilizes Bayesian networks to conduct uncertainty reasoning, and can automatically adjust the structure of Bayesian networks according to the condition of specific crime. When applying such model in real cases, criminal police or forensic scientists do not need to build any model or Bayesian networks, they just need to give the related input to the case-type based model, then the model can automatically adjust the structure of Bayesian networks and output the results.

We choose arson cases as an example to build such model in this paper, because arson cases are serious crimes and difficult to solve in the real world. With the same framework, one can build case-type based models for other crime types. The paper is organized as follows. In the next section, we propose the method of building a case-type

Journal Pre-proofs

based model. Then, a reasoning model for arson cases is established in section 3. A real arson case is used to test the model in section 4, and some conclusions are made in the last section.

2 Method

In this section, we instruct how to build a case-type based Bayesian model for crime reconstruction and investigation. We firstly introduce the knowledge for a type of crimes, which is the basis of reasoning model. Then, we propose the methods of building a case-type based reasoning model with the knowledge.

2.1 Knowledge

In a crime, events will cause the related evidence, which can be explained by Locard's exchange principle. The Locard's exchange principle is a core principle in the area of forensic science, which states that when two objects come into contact, a transfer of materials will occur (Bisbing,2006). The transferred fragments of materials lead to a connection between those two objects (Mistek, Fikiet, Khandasammy, & Lednev,2018), which can be collected as the evidence to infer what events took place. For a type of crime, there are commonalities in events along with their causing evidence. Those commonalities are termed as the knowledge, which is the basis of Bayesian inference model, because the evidential reasoning (abduction reasoning) is in fact based on it. Very briefly, assumed that for the knowledge that "event 'A' causes evidence 'i', event 'B' causes evidence 'k', and event 'C' causes evidence 'i' ", then for the abduction reasoning, if we find evidence 'i', we can infer that event 'A' or 'C' may have occurred.

To sort out the knowledge for a type of crime, we present a causal framework, which is illustrated in Figure 1. For forward evolution, a crime can be broken down into several scenarios, where a few transfers may occur, which will cause the related evidence. On the contrary, for backward reconstruction, separated scenarios will be inferred through the related evidence, then a crime can be reconstructed by sequencing scenarios.

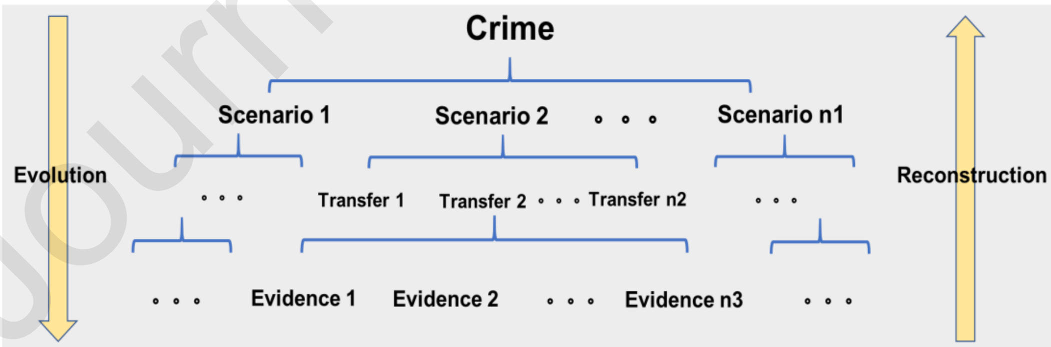


Figure 1. Causal framework for crimes.

With the causal framework, we can organize specific knowledge for a type of crimes. Firstly, we need to organize potential scenarios for a type of crimes, which highly depends on the type of crime. Then, we need to determine potential transfer conditions under each scenario. Finally, we should identify potential evidence under each transfer condition. The above three steps require experts' knowledge and experience, and a method called

Journal Pre-proofs

'scenario-entity' analysis can help to systematically sort out those knowledge (Wang, Jia, Peng, Ni, & Shen,2019).

Using the above method, we can obtain a list that associates scenarios with evidence, which is depicted in Table 1. For example, for arson cases, a potential scenario is that "a perpetrator used a container to splash accelerant", where the potential corresponding evidence is shown in table 2.

Scenarios	Evidence
Scenario 1	E 11, E 12, E 13, …, E $1n_1$
Scenario 2	E 21, E 22, E 23, …, E $2n_2$
Scenario 3	E 31, E 32, E 33, …, E $3n_3$
…	…
Scenario m	E m1, E m2, E m3, …, E mn_m

Table 1. Scenario-evidence list for a type of crimes.

Scenario	Evidence
A perpetrator used a container to splash accelerant	E1: Traces of the accelerant are on the perpetrator E2: Traces of the accelerant are on the container E3: The fingerprints of perpetrator are on the container E4: The DNA of perpetrator is on the container; …

Table 2. Potential evidence for a scenario

It is important to note that the scenario-evidence list is a logical knowledge, i.e. not a list instantiated by specific data. To help better understand this, we use some symbols to represent the mentioned scenario. Specifically, we can use X to denote a hypothesis set of the perpetrator, where each element represents a possible perpetrator. Similarly, Y denotes a hypothesis set of the container, and Z indicates a hypothesis set of the accelerant. For any person $x \in X$, object $y \in Y$ and $z \in Z$, a possible scenario is that x used y to splash z, and the corresponding evidence is shown in Table 3.

Scenario	Evidence
x used y to splash z	E1: Traces of z are on x E2: Traces of z are on y E3: Fingerprints of x are on y E4: The DNA of x is on y …

Table 3. Logical meaning of a scenario-evidence list

2.2 Logical Bayesian network

There exist uncertainty underlying the scenario-evidence list, that is, the potential scenario may or may not occurred, and the potential evidence of each scenario may or may not exist. As mentioned in the Introduction, we use Bayesian networks to model the knowledge and conduct backward uncertainty reasoning to achieve crime reconstruction. We show how to build a Bayesian network in this subsection, which is the core of case-type based inference model. Such network is termed as 'Logical Bayesian network' in this paper, which clearly models the logic of scenario-evidence list (but the logical model usually cannot be directly used to solve specific cases, see section 2.3).

There are two main principles of building logical Bayesian networks, which are shown below. Firstly, the building of Bayesian network is based on the causal logic, that is, cause nodes are the parent nodes of result nodes (Darwiche,2009; Fenton, Neil, & Lagnado,2013). More specifically, scenario nodes should be the parent nodes of corresponding evidence nodes. Then, with the discovered evidence at the crime scene, the Bayesian model can conduct backward reasoning to reconstruct scenarios. Secondly, the Bayesian network should consider both coarse-grained crime facts and fine-grained scenarios. The coarse-grained crime facts are in the light of a whole crime, such as the perpetrator, the crime tool and so on. In criminal investigation, separate scenarios are often inferred through evidence, and when we clarify and sequence the related scenarios, we can infer a coarse-grained crime fact. Based on the above principles, the established Bayesian network has three layers, which is shown in Figure 2.

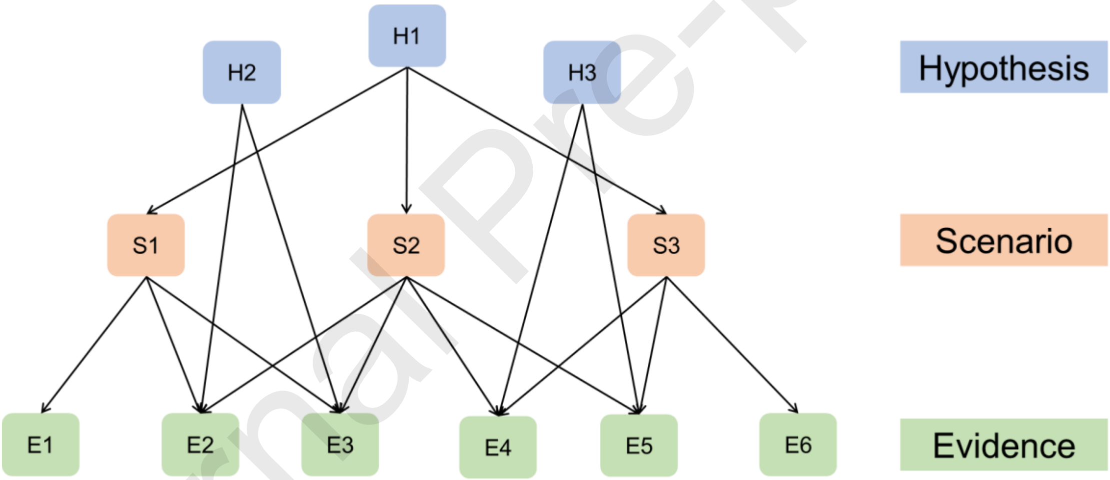


Figure 2. The framework of logical Bayesian network.

From bottom to top, the first level is 'Evidence'. Each evidence is turned into an evidence node (noted that we should delete duplicate nodes because different scenarios may cause the same evidence), and it has two exclusive states, including existence and inexistence. For example, for E3 listed in Table 3, 'Fingerprints of possible perpetrator x are on the possible container y ' is turned into an evidence node: 'Fingerprints on the possible container y have the same characteristics as the possible perpetrator x's fingerprints', and if we find such potential evidence, the state of this evidence node is 'existence', otherwise the state is 'inexistence'.

The second layer is 'Scenario', where each potential scenario is turned into a scenario node, which has two exclusive states, namely 'the concerned scenario occurred' and 'the concerned scenario did not occur'. Noted that unlike scenario nodes presented in other works, some important information is extracted to form the nodes in 'Hypothesis' level, that is, to reason the coarse-grained crime facts. Modeling the scenario node in this way can also help to avoid the ambiguity of the assignment of conditional probabilities. For example, the scenario listed in Table 3 may be directly turned into a scenario node in other work, i.e. "A possible perpetrator x used a possible container y to splash possible accelerant z", but the state of the scenario not occurring has different explanations: that possible perpetrator x did not do that, that possible container y was not used, and z was not the accelerant. That uncertainty can cause ambiguity and affect the assignment of conditional probabilities. Therefore, in our model, the possible crime tools y and z should be extracted to form two hypothesis nodes, i.e. "y was the container" and "z was the accelerant", then the scenario node is changed to 'x used a container to splash accelerant' (or abbreviated as 'x splashed accelerant'), thus any ambiguity is avoided.

Through the above process, it is clearly that the hypothesis node is used to infer whether a hypothesis is the crime fact or not, and each node has two exclusive states, that is, 'yes' and 'no'. Besides, to infer the probability that a possible perpetrator x is the perpetrator, a hypothesis node "x was the perpetrator" should be modelled, which can output a single probability value to help crime investigation.

According to framework of Bayesian networks, we can build Bayesian networks nodes of each layer. Next, we need to connect those nodes, and the methods are shown below. Firstly, evidence nodes are child nodes of the related scenario nodes, because they are the results of related scenarios (Table 1). Secondly, remember that hypothesis node is used to infer whether an element is the crime fact or not, hence, hypothesis node should be the parent node of those evidence and scenario nodes which have the element of that hypothesis node, because whether the element is the crime fact or not affects the occurrence of scenarios and the existence of evidence which involve that element. For example, in Figure 2 assumed that scenario node 'S1' denotes the mentioned scenario 'x splashed accelerant', hypothesis node 'H2' represents 'y is the container', and evidence node 'E2' means 'The fingerprints on y have the same characteristics as x's fingerprints', then both scenario node 'S1' and hypothesis node 'H2' are parent nodes of evidence node 'E2'.

To complete the Bayesian networks, we need to further determine the parameters of networks, i.e. prior probability and conditional probability (Darwiche,2009). For a case-type based model, the prior probability of hypothesis nodes can be set to 50% for each state, or it can be set by criminal police for specific cases. The conditional probability is mainly used to denote the probability of evidence given the scenario, which can be obtained by statistical methods or expert consultation (Koller & Friedman,2009; Taylor, Kokshoorn, & Biedermann,2018). After we set the parameters, the logical Bayesian network has then been established.

2.3 Growth algorithm

The logical Bayesian network clearly models the logic of knowledge, which is the basis of case-type based reasoning model. Nevertheless, the logical Bayesian network usually cannot be directly used to solve crimes, which will be explained in detail below. To adapt the case-type based model to different cases, we propose an algorithm that can automatically adjust the structure of logical Bayesian networks according to the condition of a specific crime.

Firstly, we explain why the logical Bayesian network cannot be directly used to solve crimes through the mentioned scenario-evidence list. In section 2.1, X, Y, Z denote hypothesis sets of the perpetrator, the container and the accelerant, respectively. For example, assumed for an arson case people x_1 and x_2 are possible perpetrators, y_1 and y_2 are possible containers, and z_1, z_2 and z_3 are possible accelerant, then x_1 and x_2 make up a hypothesis set of the perpetrator, y_1 and y_2 constitute a hypothesis set of the container, and z_1, z_2 and z_3 form a hypothesis set of the accelerant.

According to the method in section 2.2, the established logical Bayesian network only has three hypothesis nodes to infer the perpetrator, the container and the accelerant, respectively, i.e. nodes of 'x is the perpetrator', 'y is the container', and 'z is the accelerant'. However, for this assumed specific case, we need two hypothesis nodes to infer the perpetrator, namely nodes of 'x_1 is the perpetrator' and 'x_2 is the perpetrator', two hypothesis nodes to infer the container, i.e. nodes of 'y_1 is the container' and 'y_2 is the container', and three hypothesis nodes to infer the accelerant, including nodes of 'z_1 is the accelerant', 'z_2 is the accelerant' and 'z_3 is the accelerant'. Besides, for the scenario node 'x splashed accelerant' in the logical Bayesian network, we need two scenario nodes for this assumed case, namely scenario nodes of 'x_1 splashed accelerant' and 'x_2 splashed accelerant'. Moreover, evidence nodes should also be added. Such as for the evidence node 'The fingerprints on y have the same characteristics as x's fingerprints' in logical Bayesian network, we need $2 \times 2 = 4$ nodes for the assumed case (the number of different combinations between x_1, x_2 and y_1, y_2).

Through the above analysis, it is clearly that the logical Bayesian network is just a logic model and can only directly solve cases where the number of elements in each hypothesis set is one, thus we need to adjust the structure of logical Bayesian network according to the condition of a specific crime for crime reconstruction and investigation. Therefore, to make the case-type based model more powerful and practical (avoid adjusting Bayesian network manually), we propose an algorithm to automatically adjust the structure of logical Bayesian network.

The thinking of the algorithm is easy to understand, which is that elements in the same hypothesis set have the same characteristics, hence they should have the same structure

Journal Pre-proofs

in the Bayesian network after growth. To achieve that, the algorithm has three main steps: Find the related nodes; Copy nodes and connections; Complete the connection between nodes. The detail of each step is described below, and the flow of algorithm is depicted in Figure 3.

(1) Find the related nodes: Given a hypothesis node in the logical Bayesian network, and assumed that the hypothesis node is utilized to infer the crime fact of element p, find the nodes involving element p in the growing Bayesian network, and make those nodes into a list of growing nodes. Noted that the growing Bayesian network is the network directly manipulated by the algorithm, which will grow from the logical Bayesian network to the final Bayesian network that can be used to deal with a specific crime.

(2) Copy nodes and connections: Given a list of growing nodes, copy those nodes and the connections between those nodes, and add them to the growing Bayesian network.

(3) Complete the connection between nodes: Given a list of growing nodes, after they have been copied and added to the growing Bayesian network, complete the connections between those new nodes and the other nodes in the growing Bayesian network, where the connections are the same as those between growing nodes and the other nodes in the network.

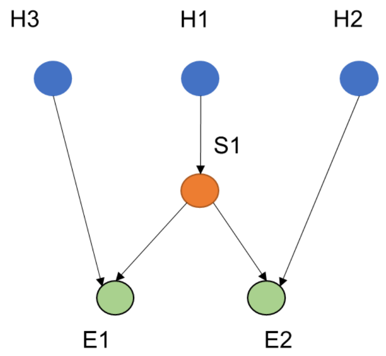
```
Growing Bayesian network= Logical Bayesian network;
for ( i=1; i ≤ n_Hypothesis; ++i) {
    Find the related nodes;
    for ( j=2; j ≤ n(i); ++j ) {
        Copy nodes and connections;
        Complete the connection between nodes;
    }
}
```

Figure 3. The flow of algorithm. The symbol 'i' denotes the ith hypothesis node in the logical Bayesian network, 'n_Hypothesis' denotes the number of hypothesis nodes in the logical Bayesian network, and the symbol 'n(i)' represents the number of elements in the ith hypothesis set. It can be found that if the number of elements in each hypothesis set is one, the Bayesian network does not grow, which agrees with the above analysis.

To better understand the algorithm, we give a simple example here. Given a logical Bayesian network shown in Figure 4, and assumed that for a crime there is one possible perpetrator, one possible accelerant and two possible containers, thus only the nodes regarding container y needs to be copied and connected by the algorithm. In the step of 'find the related nodes', nodes 'H2' and 'E2' make up a list of growing nodes, and those nodes along with the connection between them are copied and added to the network in the step of 'copy nodes and connections', after which the copied nodes are connected to the other nodes in the network in the step of 'complete the connection between nodes'. The

Journal Pre-proofs

above is the growth process of network, and it is illustrated in Figure 5. After the growth process, it can be found that those two possible containers have the same structure in the network. Specifically, the final Bayesian network has two hypothesis nodes to infer containers, i.e. nodes of 'y_1 is the container' and 'y_2 is the container', and two evidence nodes to describe the potential evidence regarding y_1 and y_2, namely nodes of 'The fingerprints on y_1 have the same characteristics as x's fingerprints' and "The fingerprints on y_2 have the same characteristics as x's fingerprints', which are consistent with the crime's condition.



Nodes
H1 : x is the perpetrator
H2 : y is the container
H3 : z is the accelerant
S1 : x splashed the accelerant
E1 : Some traces on x have the same characteristics as z
E2 : The fingerprints on y have the same characteristics as x's fingerprints

Figure 4. An example of logical Bayesian network.

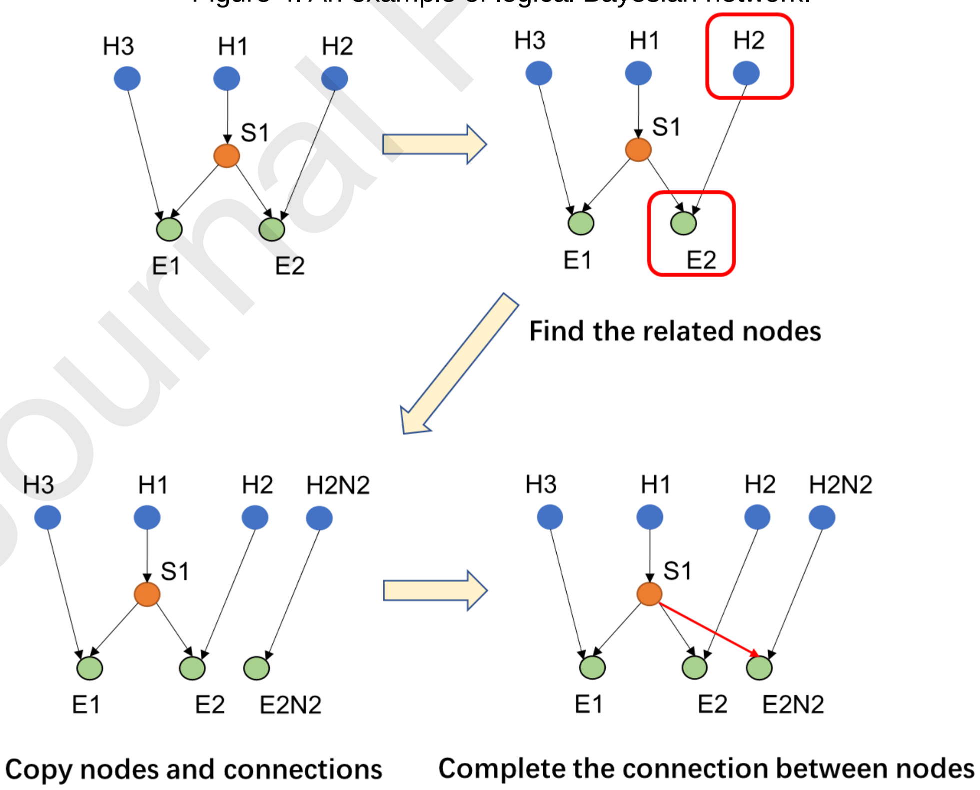


Figure 5. The growth process of Bayesian network.

3 Model for Arson Cases

We choose the crime type of arson cases as an example to build such model in this paper, with the same method one can build case-type based models for other crime types.

3.1 The establishment of model

Based on the proposed method, a case-type based model for arson cases has been established in this section.

We firstly sort out the scenario-evidence list (knowledge) for arson cases with the aid of domain experts, and 12 types of scenarios (shown in Table 4) and 48 types of evidence are identified.

Phases	Scenarios
Preparation	S1P1: Perpetrators scouted the scene
	S2P1: Perpetrators prepared accelerant
	S3P1: Perpetrators prepared ignition tools
	S4P1: Perpetrators destroyed entrances and exits
	S5P1: Perpetrators destroyed surveillance systems.
Implementation	S1P2: Perpetrators splashed accelerant
	S2P2: Perpetrators camouflaged the scene
	S3P2: Perpetrators implemented ignition
Escape	S1P3: Perpetrators self-rescued
	S2P3: Perpetrators discarded tools
	S3P3: Perpetrators fled the scene
	S4P3: Perpetrators destroyed clothing

Table 4. Possible scenarios for arson cases.

With the scenario-evidence list for arson cases, the logical Bayesian network has been built based on the method shown in section 2.2 using the Netica software (www.norsys.com/netica.html). As a result, there are 5 hypothesis nodes, 12 scenario nodes and 48 evidence nodes in the logical Bayesian network, which are presented in Table 5 and Figure 6. From the meaning of nodes shown in Table 5, we can also clearly find that the logical Bayesian network is just a logical model, and it can only directly solve cases where the number of elements in each hypothesis set is one, thus the growth algorithm of Bayesian networks is needed.

Hypothesis: H1: {Possible perpetrators} (i) is the perpetrator H2: {Possible victims} (j) is the victim H3: {Possible containers} (k) is the container H4: {Possible accelerant} (p) is the accelerant H5: {Possible ignition tools} (q) is the ignition tool
Scenario: S1P1: {Possible perpetrators} (i) scouted the scene S2P1: {Possible perpetrators} (i) prepared accelerant

S3P1: {Possible perpetrators} (i) prepared ignition tools
S4P1: {Possible perpetrators} (i) destroyed entrances and exits
S5P1: {Possible perpetrators} (i) destroyed surveillance systems.
S1P2: {Possible perpetrators} (i) splashed accelerant
S2P2: {Possible perpetrators} (i) camouflaged the scene
S3P2: {Possible perpetrators} (i) implemented ignition
S1P3: {Possible perpetrators} (i) self-rescued
S2P3: {Possible perpetrators} (i) discarded tools
S3P3: {Possible perpetrators} (i) fled the scene
S4P3: {Possible perpetrators} (i) destroyed clothing

Evidence:
E1: Some fingerprints at the peripheral scene have the same characteristics as i 's fingerprints
E2: The DNA at the peripheral scene has the same characteristics as i 's DNA
E3: Some footprints at the peripheral scene have the same characteristics as i 's footprints
E4: Some i 's objects or their traces are at the peripheral scene
E5: The DNA at the central scene has the same characteristics as i 's DNA
E6: Some footprints at the central scene have the same characteristics as i 's footprints
E7: Some i 's objects or their traces are at the central scene
E8: Some traces on i have the same characteristics as p
E9: Some traces on i have the same characteristics as objects at the related scene
E10: Some fingerprints at the related scene have the same characteristics as i 's fingerprints
E11: The DNA at the related scene has the same characteristics as i 's DNA
E12: Some footprints at the related scene have the same characteristics as i 's footprints
E13: Some i 's objects or their traces are at the related scene
E14: Some traces at the related scene have the same characteristics as p
E15: Some fingerprints on k have the same characteristics as i 's fingerprints
E16: The DNA on k has the same characteristics as i 's DNA
E17: Some i 's objects or their traces are on k
E18: Some traces on k have the same characteristics as p
E19: Some traces on i have the same characteristics as q
E20: Some traces at the related scene have the same characteristics as q
E21: Some fingerprints on q have the same characteristics as i 's fingerprints
E22: The DNA on q has the same characteristics as i 's DNA
E23: Some i 's objects or their traces are on q
E24: Some scars on i have the same characteristics as the entrances or exits
E25: Some traces on i have the same characteristics as objects at the peripheral scene
E26: Some bloodstains at the peripheral scene have the same characteristics as i 's blood

Journal Pre-proofs

E27: Some scars on i have the same characteristics as the surveillance systems
E28: Some traces on j have the same characteristics as p
E29: Some traces at the central scene have the same characteristics as k
E30: Some j 's objects or their traces are on i
E31: Some traces on i have the same characteristics as objects at the central scene
E32: Some fingerprints on j have the same characteristics as i 's fingerprints
E33: The DNA on j has the same characteristics as i 's DNA
E34: Some bloodstains on j have the same characteristics as i 's blood
E35: Some bloodstains at the central scene have the same characteristics as i 's blood
E36: The burning mark is on i
E37: The smoke mark is on i
E38: The burning mark is on j
E39: The smoke mark is on j
E40: Some bloodstains on q have the same characteristics as i 's blood
E41: The burning mark is on q
E42: The smoke mark is on q
E43: Some self-rescue objects are on i
E44: Some bloodstains at the related scene have the same characteristics as i 's blood
E45: Some bloodstains at the related scene have the same characteristics as j 's blood
E46: Some j 's objects or their traces are at the related scene
E47: Some traces at the central scene have the same characteristics as p
E48: Some traces at the central scene have the same characteristics as q

Table 5. Nodes in the logical Bayesian network for arson cases. The symbol {…} represents a hypothesis set, and (i) means the i th element in a set. For example, {Possible perpetrators} represents the hypothesis set of the perpetrator, and {Possible perpetrators} (i) represents the i th person in the hypothesis set of the perpetrator.

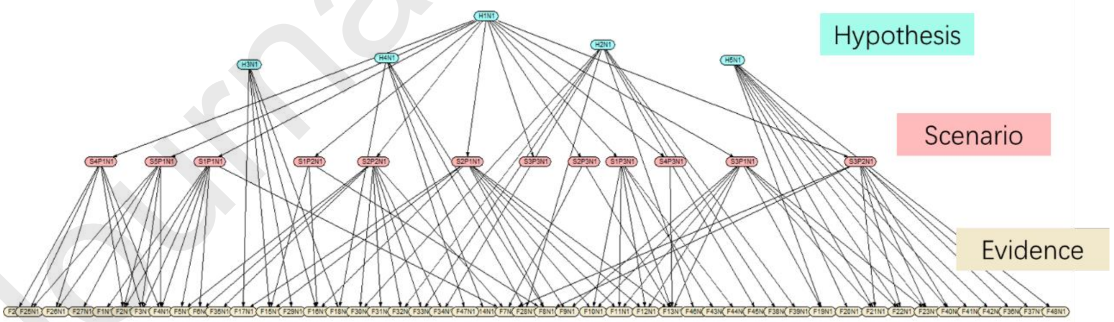


Figure 6. The illustration of logical Bayesian network.

For this example model, the probability parameters of the logical Bayesian network are determined by experts. The triangular fuzzy-number was adopted to represent the probability, which is denoted by (a, m, b) , and the membership function is shown in equation 1 (Liang, Liu, Pedrycz, & Hu,2013). Besides, the semantic value was also

Journal Pre-proofs

introduced to help get the experts' opinion friendly, and the relation between semantic values and triangular fuzzy-numbers is depicted in Table 6 (Ma et al.,2012).

$$\mu(x)=\begin{cases}0, x<a\\(x-a)/(m-a), a\le x\le m\\(b-x)/(b-m), m\le x\le b\\0, x>b\end{cases} \quad (1)$$

Serial number	Semantic values	Triangular fuzzy-numbers
1	Very high	(0.9,1.0,1.0)
2	High	(0.7,0.9,1.0)
3	Slightly high	(0.5,0.7,0.9)
4	Medium	(0.3,0.5,0.7)
5	Slightly low	(0.1,0.3,0.5)
6	Low	(0,0.1,0.3)
7	Very low	(0,0,0.1)

Table 6. The relation between semantic values and triangular fuzzy-numbers

The process to calculate the probability is shown as follows. We use the symbol n to represent the number of experts, and for a concerned probability, the semantic value given by the ith expert was transformed into the semantic value $Q_i=(a_i,m_i,b_i)$ according to table 6. Next, the mean value of all experts' opinions is calculated using equation 2.

$$\begin{aligned}\overline{Q}&=\frac{Q_1\oplus Q_2\oplus\ldots\oplus Q_n}{n}\\&=\frac{(a_1+a_2+\ldots+a_n, m_1+m_2+\ldots+m_n, b_1+b_2+\ldots+b_n)}{n}\\&=(\overline{a},\overline{m},\overline{b})\end{aligned} \quad (2)$$

Then, the mean area method was used to calculate the probability (Lan & Fan,2010), and the equation is shown below, where P represents the probability.

$$P=\frac{\overline{a}+2\overline{m}+\overline{b}}{4} \quad (3)$$

We consulted 10 domain experts and used the above method to determine the parameters of Bayesian network. Besides, instructed by the method in section 2.3, we have developed the growth algorithm for the logical Bayesian network of arson cases using the Netica-C API (www.norsys.com/netica.html). We have developed and encapsulated the whole case-type based model using C language, and the input and output are presented as follows.

3.2 The input of model

The evidence constitutes the main input of the model, based on which the model can

Journal Pre-proofs

infer both coarse-grained crime facts and fine-grained scenarios. Besides, the model can also consider the following three kinds of inputs, with which the model can achieve more accurate, more reasonable and faster inference.

(1) Definite crime facts and possible hypotheses. For some cases, criminal police or forensic scientists can give definite crime facts and possible hypotheses according to video surveillance information, expert experience and so on. Such as 'Mr M is a possible perpetrator', 'Gasoline is the accelerant' and the like. The model can take this kind of input into account, which helps to realize faster and more accurate inference.

(2) Determined scenarios. For real cases, criminal police or forensic scientists can also give the determined scenarios (including the determined occurrence scenarios and the determined non-occurrence scenarios) according to some investigation information, like the video surveillance information. The model can take this kind of input into consideration in the reasoning to give more accurate and reasonable scenario results.

(3) The evidence not used for reasoning. In real cases, some evidence should not be used for reasoning. For example, if the evidence is obtained illegally, then it should not be used for reasoning. The model can consider this kind of input, which helps to achieve more flexible and accurate reasoning.

3.3 The output of model

Based on the logical Bayesian network and the growth algorithm for Bayesian network, the model can give the following three kinds of outputs.

(1) Coarse-grained crime facts. The model can output the definite or possible hypotheses (crime facts) along with the evidence or scenarios that support them. For each hypothesis, the model can output the probability that the hypothesis is true. Besides, the model can give the probability optimal solution for each hypothesis set, which is based on the joint probability of state combination.

(2) Fine-grained scenarios. The model can output the determined or possible occurrence scenarios along with the evidence that supports them. For each possible scenario, the model can output the probability that the scenario occurred. Besides, the model can output the probability optimal crime-commission process, i.e. the optimal scenarios together with their time sequences. The optimal scenarios are obtained through the joint probability for each type of scenario given the result of optimal coarse-grained crime facts. For the time sequence of scenarios, the model has a built-in time sequence (shown in table 4), in addition, model users can assign the time sequence manually in real cases.

(3) Investigation suggestions. For a possible hypothesis (see results of crime facts), if the probability that the hypothesis is true is below a given hypothesis probability threshold, then the model will output the related investigation suggestions. More specifically, the model will suggest investigating the undetermined potential scenarios and undiscovered potential evidence that related to this hypothesis.

4 Results

In this section, we use a real arson case to test the established case-type based model.

Journal Pre-proofs

Journal Pre-proofs

The case, known as 'Xiamen Bus Fire', was a very serious criminal case in China. On 7 June 2013, a bus operating for the Xiamen BRT caught fire near the Jinshan stop, which resulted in 47 deaths and 34 injuries. The investigation was conducted by experts of the Ministry of Public Security and the provincial Public Security Department. Readers can refer to Wikipedia (https://en.wikipedia.org/wiki/Xiamen_bus_fire) for more information about the case and to China Daily (https://www.chinadaily.com.cn/china/2013-06/11/content_16605793.htm) for an overview of some evidence released by police.

To better verify the model's performance, we will follow the real police's case-solving process to test the model. In that process, the discovered evidence was enriched over time, thus the model's input should also be gradually enriched. We divide the real case-solving process into three stages, and we will compare the model's outputs with the real investigation results for each stage.

4.1 Stage I

In stage 1, only some basic information and physical evidence at the scene were discovered, which are listed below:

(a) Several evidence at the bus, including:
  (1) Dozens of dead bodies;
  (2) A handcart;
  (3) Fragments of a woven bag;
  (4) Burning traces of flowing fire;
  (5) The fuel tank of bus was not destroyed.

(b) The information of witnesses:
  (1) A man who sat opposite the back door of bus pushed down a handcart containing a woven bag, and ignited the woven bag, which caused the bus to burn violently.

According to the above evidence, we can form a hypothesis that it is probably an arson case and the woven bag might be the container of accelerant. We can use the model to conduct the further inference, and the inputs of model are shown in Table 7. The input value is divided into three classes, which is explained as follows. For hypothesis nodes, '1' represents definite crime facts and '0' denotes possible hypotheses; For scenario nodes, we preset all scenarios could occur, and we use '1' to indicate determined occurrence scenarios and '-1' to indicate determined non-occurrence scenarios; For evidence nodes, we preset all potential findings are not discovered, and we use '1' to denote the discovered findings and '0' to denote the findings not used for reasoning. These three types of values correspond to the three kinds of model's inputs, which are presented in section 3.2.

Layer	Nodes	Value	Meaning
Hypothesis	H3	0	The woven bag was the container
Evidence	E29	1	Fragments of the woven bag were at the central scene

Table 7. The inputs of model in stage 1.

The outputs of model for hypotheses are listed in Table 8. Noted that there are no inputs about suspects in this stage, thus the model does not output any scenarios. The model also outputs some investigation suggestions, which are shown in Table 9.

Possible hypotheses	Probability	Supporting evidence
The woven bag was the container	76.8%	E29

Table 8. The outputs of model for hypotheses in stage 1. The prior probability that the hypothesis is true was set at 50% throughout the example.

Number	Investigation suggestions
1	Investigate whether there are perpetrator's fingerprints, DNA and goods on the woven bag
2	Investigate whether there is accelerant on the woven bag.

Table 9. Investigation suggestions in stage 1. When the probability of hypothesis is below a given hypothesis probability threshold, the model will output investigation suggestions. We choose the threshold value of 90%, because it is acknowledged that the probability of criminal hypothesis is difficult if not impossible to reach 100% (Dhami,2008), and 90% is a reasonable threshold value to determine whether the criminal hypothesis is convincing or not (Magnussen, Eilertsen, Teigen, & Wessel,2014). It also should be noted that the threshold value is just used to determine whether the model needs to output investigation suggestions or not, which has no influence on the results of coarse-grained criminal hypothesis or fine-grained scenarios.

4.2 Stage II

In stage 2, some detailed investigation information and identification results were obtained, and we list important ones in Table 10. Comparing the suggestions provided by the model in last stage with the real discovered evidence in this stage, it is clearly that the model can give effective suggestions for crime investigation. More specifically, the second investigation suggestion listed in Table 9 corresponds to the second discovered evidence listed in Table 10, and the first suggestion is consistent with the first discovered evidence.

Number	Evidence
1	Video surveillance information: A man in white pulled a handcart containing a woven bag to get on the bus.
2	Gasoline was identified at the woven bag, and the bus is a diesel car.
3	A windproof cap of lighter was at the bus

Table 10. Investigation information and identification results in stage 2.

According to the above evidence, we can form the following four hypotheses. (a) The man in white was the perpetrator; (b) The woven bag was the container of accelerant; (c) The gasoline was the accelerant; (d) The lighter was the ignition tool. Then, we use the model to conduct the reasoning, and the inputs of model are listed in Table 11.

Layer	Nodes	Value	Meaning

Journal Pre-proofs

Hypothesis	H1	0	The man in white was the perpetrator
	H3	0	The woven bag was the container
	H4	0	The gasoline was the accelerant
	H5	0	The lighter was the ignition tool
Evidence	E29	1	Fragments of the woven bag were at the central scene
	E18	1	The gasoline was on the woven bag
	E47	1	The gasoline was at the central scene
	E7	1	The goods of the man were at the central scene
	E17	1	The goods of the man were on the woven bag
	E41	1	The burning mark was on the lighter
	E42	1	The smoke mark was on the lighter
	E48	1	The lighter was at the central scene

Table 11. The inputs of model in stage 2.

The outputs of model for coarse-grained crime facts and fine-grained scenarios are presented in Table 12 and 13, respectively. Besides, the model also outputs investigation suggestions that are shown in Table 14.

Possible hypotheses	Probability	Supporting evidence
The man in white was the perpetrator	79.8%	E7, E17
The woven bag was the container	92.8%	E17, E18, E29
The gasoline was the accelerant	87.2%	E18, E47
The lighter was the ignition tool	96.7%	E41, E42, E48

Table 12. The outputs of model for hypotheses in stage 2.

Possible scenarios	Time sequence	Probability
The man scouted the scene	1	70.6%
The man prepared accelerant	2	76.8%
The man camouflaged the scene	3	62.9%
The man implemented ignition	4	76.1%
The man discarded tools	5	60.8%

Table 13. The outputs of model for scenarios in stage 2.

Number	Investigation suggestions
1	Investigate whether the man in white committed scenarios listed in Table 4.
2	Investigate whether there is gasoline on the man in white.
3	Investigate whether there is gasoline at the related scene.
4	Investigate whether there is gasoline on the victims.

Table 14. Investigation suggestions in stage 2.

4.3 stage III

In this stage, more evidence was collected in the real crime-solving process, which

Journal Pre-proofs

are presented in Table 15. Comparing the investigation suggestions given by the model in stage 2 with the discovered evidence in the real case, it can be easily found that the model can provide effective investigation suggestions to help solve crimes. Specifically, the first investigation suggestion listed in Table 14 agrees with the fourth discovered evidence listed in Table 15, and the third investigation suggestion consists with the second evidence.

Number	Evidence
1	The man in white was dead at the bus, and he is Mr. M through DNA testing
2	Gasoline was discovered at Mr. M's house
3	Mr. M's suicide note was found, which said that he wanted to revenge on society
4	Video surveillance: Mr. M bought some gasoline two days before the case

Table 15. Investigation information and identification results in stage 3.

With those new evidence, we can use the model to further help crime reconstruction, and the model's inputs are depicted in Table 16. Noted that for the input of evidence, we just list new evidence collected in this stage, while the model also takes the input of evidence listed in Table 11 into consideration while calculating.

Layer	Nodes	Value	Meaning
Hypothesis	H1	0	Mr. M was the perpetrator
	H3	0	The woven bag was the container
	H4	0	The gasoline was the accelerant
	H5	0	The lighter was the ignition tool
Scenario	S1P1	-1	Mr. M did not case the joint
	S2P1	1	Mr. M prepared accelerant
	S4P1	-1	Mr. M did not destroy entrances and exits
	S5P1	-1	Mr. M did not destroy surveillance systems
	S1P2	1	Mr. M splashed accelerant
	S2P2	-1	Mr. M did not camouflage the scene
	S3P2	1	Mr. M implemented ignition
	S3P3	-1	Mr. M did not flee the scene
Evidence	E14	1	The gasoline was at the related scene

Table 16. The inputs of model in stage 3.

The reconstruction results of model are shown in Table 17, which include both coarse-grained crime facts and fine-grained scenarios. The real conclusions of the case are that 'Because the individual demand is not met, Mr. M bought some gasoline, took the gasoline with the woven bag to get on the bus, pushed down the woven bag, and implemented ignition while the bus was driving, which caused dozens of deaths.' Comparing the reconstruction results of model with the real case conclusions, it can be found that the model can achieve crime reconstruction.

Crime facts:		
Possible hypotheses	Probability	Supporting items
Mr. M was the perpetrator	99.8%	S2P1, S1P2, S3P2, E7, E17
The woven bag was the container	94.8%	E17, E18, E29
The gasoline was the accelerant	94.2%	E14, E18, E47
The lighter was the ignition tool	96.7%	E41, E42, E48
Crime-commission process:		
Mr. M: prepared accelerant → splashed accelerant → implemented ignition → discarded tools		

Table 17. Reconstruction results of model

5 Discussion

To help police solve crimes, a case-type based reasoning model was proposed in this work. A real arson case was used to test the model, and the results indicated that the model can provide effective investigation suggestions and give reasonable crime reconstruction results.

Like other works that studied on Bayesian networks to help solve crimes, the presented model is also based on Bayesian networks to conduct uncertainty reasoning. Differently, with the guidance of those studies, we need to establish specific Bayesian networks for each crime case, namely we need to define the specific nodes according to the case, link those nodes under a reasoning framework, and determine the probability parameters, only then can we conduct the analysis. In the real world, criminal police have professional skills and rich experience, but they are not familiar with Bayesian networks, which implies that such technology may be seldom used by them for real cases. On the contrary, when using our model, criminal police or forensic scientists do not need to build any model or Bayesian networks, they just need to give the related input, and the model can automatically adjust the structure of Bayesian networks and give results. This advantage is due to the knowledge for a type of crime and the growth algorithm which can adapt the model to different cases.

The presented work has the following three aspects of implications:

(i) Practical implications

Our model can directly help to solve crimes in practice. For criminal investigation, it is essential for police to collect evidence, and our model can provide them with effective investigation suggestions. Besides, police and forensic scientists need to figure out what events took place given the collected evidence, and our model can output reasonable crime events together with the evidence that supports them.

(ii) Theoretical implications

Solving crimes can be regarded as a task with the following characteristics: (1) results need to be explained: we need to figure out crime events together with the evidence that supports them, and also evidence investigation suggestions; (2) the data is insufficient or hard to obtain: the criminal investigation data is hard to obtain due to some security reasons. Our presented work has theoretical implications in both areas of artificial intelligence and forensic science. In the area of artificial intelligence, our work indicates that in the face of tasks with the mentioned characteristics, a knowledge-based model can help a lot in

Journal Pre-proofs

specific fields, and our work presents a systematic method of using expert knowledge to build such an intelligent model, which is enlightening. In the area of forensic science, our work proposes a new way (i.e. a knowledge-based model) to help solve crimes, which is very innovative.

(iii) Social implications

On one hand, the presented model can directly help police solve crimes, which has significance for social justice and security. On the other hand, our work increases the capacity of criminal investigation, which will help to deter criminals, hence plays a role in crime prevention.

In this work, we provide a 'general' model that can analyze any crime that belongs to a specific case type, which means that to meet the demand of 'general', the model therefore reduces some precision. For criminal investigation, we do not need exact probability in practice. However, in a court trial, the result needs to be very accurate, which suggests that each criminal case may need a 'case-specific' model but not a 'general' model.

6 Conclusions

In this paper, we propose a case-type based model for crime reconstruction and investigation, which is based on the knowledge of a type of crimes.

The method of building the case-type based model is proposed. A logical Bayesian network is established to represent the knowledge and conduct the uncertainty reasoning. The logical Bayesian network has three layers: hypothesis, scenario and evidence, which can infer both coarse-grained crime facts and fine-grained scenarios. To adapt the model to different cases, a growth algorithm of Bayesian networks is proposed, which can automatically adjust the structure of Bayesian networks according to the condition of a specific crime.

Instructed by the method, we choose the crime type of arson cases as an example to build such model in this paper. With the same framework, one can build case-type based models for other crime types. The model can consider various kinds of evidence as the inputs, and output both investigation suggestions and reconstruction results. A real arson case was used to test the model, and the results indicate that the model can provide effective investigation suggestions and give reasonable crime reconstruction results.

As discussed, our model has great practical, theoretical, and social implications, but also has limitations in precision due to its 'general' nature. For future work, with more real investigation data collected, our model will be improved in precision with the aid of more accurate probability parameters. Besides, our presented model is designed to assist in criminal investigation, but for the stage of court trial, a 'case-specific' model may be required. Therefore, a knowledge-based reasoning model for the court trial is worth studying. Moreover, a share common modeling method for these two stages needs to be researched, which will make a good use of reasoning techniques for solving crimes.

Acknowledgements

This work was supported by the National Key R&D Program of China [grant number

Journal Pre-proofs

2017YFC0803800] and China Scholarship Council [number 201906210166].

References

Biedermann, A., & Taroni, F. (2012). Bayesian networks for evaluating forensic DNA profiling evidence: a review and guide to literature. Forensic Science International: Genetics, 6(2), 147-157.

Bisbing, R. E. (2006). Trace evidence in the real crime laboratory. In the Forensic Laboratory Handbook (pp. 265-290). Humana Press.

Chisum, W. J. (2006). Crime reconstruction. In the Forensic Laboratory Handbook (pp. 63-77). Humana Press.

Chisum, W. J., & Turvey, B. E. (2011). Crime Reconstruction, 2nd Edition. Academic Press, 1-674.

Darwiche, A. (2009). Modeling and reasoning with Bayesian networks. Cambridge university press.

Dhami, M. K. (2008). On Measuring Quantitative Interpretations of Reasonable Doubt. Journal of Experimental Psychology: Applied, 14(4), 353-363.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A General Structure for Legal Arguments About Evidence Using Bayesian Networks. Cognitive Science, 37(1), 61-102.

Geiger, D., Verma, T., & Pearl, J. (1990). Identifying independence in Bayesian networks. Networks, 20(5), 507-534.

Gittelson, S., Biedermann, A., Bozza, S., & Taroni, F. (2012). Bayesian networks and the value of the evidence for the forensic two-trace transfer problem. Journal of forensic sciences, 57(5), 1199-1216.

Houck, M. M., Crispino, F., McAdam, T., Houck, M. M., Crispino, F., & McAdam, T. (2018). Crime Scene Reconstruction. The Science of Crime Scenes; Academic Press: Waltham, MA, USA, 341-344.

Huang, L., Cai, G., Yuan, H., & Chen, J. (2019). A hybrid approach for identifying the structure of a Bayesian network model. Expert Systems with Applications, 131, 308-320.

Jackson, G., Aitken, C., & Roberts, P. (2015). Case assessment and interpretation of expert evidence. Guidance for judges, lawyers, forensic scientists and expert witnesses. Practitioner guide, (4).

Kabir, S., & Papadopoulos, Y. (2019). Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: A review. Safety science, 115, 154-175.

Koller, D., & Friedman, N. (2009). Probabilistic graphical models: principles and techniques. MIT press.

Lan, R., & Fan, J. (2010). Complete metric on triangular fuzzy numbers and its application to decision-making. Journal of Systems Engineering, 25(3), 313-319.

Liang, D., Liu, D., Pedrycz, W., & Hu, P. (2013). Triangular fuzzy decision-theoretic rough sets. International Journal of Approximate Reasoning, 54(8), 1087-1106

Ma, D., Zhou, Z., Yu, X., Fan, S., Xing, W., Guo, Z. (2012). Reliability analysis of multi-state Bayesian networks based on fuzzy probability. Systems Engineering and Electronics, 34(12), 2607-2611.

Magnussen, S., Eilertsen, D. E., Teigen, K. H., & Wessel, E. (2014). The Probability of Guilt in Criminal Cases: Are People Aware of Being 'Beyond Reasonable Doubt'? Applied Cognitive Psychology, 28(2), 196-203.

Mistek, E., Fikiet, M. A., Khandasammy, S. R., & Lednev, I. K. (2018). Toward Locard's Exchange Principle: Recent Developments in Forensic Trace Evidence Analysis. Analytical Chemistry, 91(1), 637-654.

Noor, N. M. M., Asmara, S. M., Saman, M. Y. M., & Hitam, M. S. (2014). Probabilistic knowledge base system for forensic evidence analysis. Journal of Theoretical and Applied Information Technology,

Journal Pre-proofs

Journal Pre-proofs

59(3), 708-717.

Pearl, J. (1986). Fusion, propagation, and structuring in belief networks. Artificial intelligence, 29(3), 241-288.

Pearl, J. (2011). Bayesian networks. UCLA: Department of Statistics, UCLA. Retrieved from https://escholarship.org/uc/item/53n4f34m

Rossy, Q., & Ribaux, O. (2014). A collaborative approach for incorporating forensic case data into crime investigation using criminal intelligence analysis and visualisation. Science & Justice, 54(2), 146-153.

Schum, D. A. (2002). Species of abductive reasoning in fact investigation in law. In the Dynamics of Judicial Proof (pp. 307-336). Physica, Heidelberg.

Taylor, D., Biedermann, A., Hicks, T., & Champod, C. (2018). A template for constructing Bayesian networks in forensic biology cases when considering activity level propositions. Forensic Science International: Genetics, 33, 136-146.

Taylor, D., Kokshoorn, B., & Biedermann, A. (2018). Evaluation of forensic genetics findings given activity level propositions: A review. Forensic Science International: Genetics, 36, 34-49.

Verheij, B., Bex, F., Timmer, S. T., Vlek, C. S., Meyer, J. J. C., Renooij, S., & Prakken, H. (2016). Arguments, scenarios and probabilities: connections between three normative frameworks for evidential reasoning. Law, Probability and Risk, 15(1), 35-70.

Wang, J., & Xu, Z. (2014). Bayesian inferential reasoning model for crime investigation. Frontiers in Artificial Intelligence and Applications, 262, 59-67.

Wang, L., Jia, M., Peng, C., Ni, S., & Shen, S. (2019). Scenario-Entity Analysis based on an entity-relationship model: Revisiting crime reconstruction. Forensic science international, 302, 109923.

Wieten, R., De Zoete, J., Blankers, B., & Kokshoorn, B. (2015). The interpretation of traces found on adhesive tapes. Law, probability and risk, 14(4), 305-322.

Journal Pre-proofs

- A more general uncertainty reasoning model is proposed to help solve crimes
- The knowledge for crime reconstruction is represented by Bayesian networks
- A growth algorithm of Bayesian networks is proposed to adapt to different cases
- The model can provide useful investigation suggestions and reconstruct crimes

Journal Pre-proofs

Litao Wang: Conceptualization, Methodology, Software, Writing - Original Draft

Meisheng Jia: Methodology, Investigation

Yi Shi: Investigation, Resources

Feiyu Chen: Investigation, Validation

Shunjiang Ni: Conceptualization, Methodology, Writing – Review,

Shifei Shen: Conceptualization, Writing – Review, Supervision

Journal Pre-proofs