Decision Support Systems for Forensic Science in Crime Investigation

1,2,3Noor Maizura Mohamad Noor, 2Ahmad Faiz Ghazali, 3Md Yazid Mohd Saman
1,2,3School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu
E-mail: 1maizura@umt.edu.my, 2ahmad.faiz14@yahoo.com, 3yazid@umt.edu.my

Abstract

Multidisciplinary reviews presented in this paper aims to demonstrate the applicability of decision support systems (DSS) concept in relation to forensic science specialization field to assist crime investigations. There are numerous different including integrated methods applied in a DSS. This review specifically focus on current available commonly applied methods for forensic science in crime investigation. They are including but not limited to Bayesian, likelihood ratios, fuzzy, and many more. After the introduction, which is followed by the general concept of DSS, forensic science, and crime investigation, subsequent sections are arranged according to applied methods in order to help researchers. It is arranged in such a way to ease continuous study by researchers related to these fields.

Keywords: Decision Support System (DSS), Computer Science (CS), Forensic Science (FS)

1. Introduction

Decision support systems (DSS) together with the proposed frameworks and applied methods can contribute a lot towards the advancement of science and technology in crime investigation. In order to address decision-making issues in problem domains of forensic science during crime investigation, numerous studies have been conducted [1, 2] Research work involving concept of multi-disciplinary research fields, such as DSS for computer science and forensic science, is very significant. Scientists can combine their domain knowledge and data in one model, apply techniques into generating hypotheses, and solve problems of uncertain information systematically. In the same problem domain, there are also related studies that have used various computational methods, even though the authors do not relate their work for decision support [3, 4].

If additional decision support concepts are integrated into this type of research, then it may be able to seriously address the decision-makers point-of-view during problem solving. This is important because on multiple levels of forensic science in crime investigation, the managers or leaders that handle the cases need to make important decisions that affect the next level in the courts. Although there have been broad discussions on these issues, this review paper focused on currently available DSS for forensic science during crime investigation as well as the common computational methods used, including Bayesian, likelihood ratio, image processing, and fuzzy. Moreover, expert systems and knowledge management problems can also be solved together with DSS. Some studies have applied these methods together in a DSS, while others either applied these methods on their own or did not relate it specifically with DSS. The aim of this paper is not to present any new theories or methods but strive to serve as a compiled related reference mainly for researchers in DSS for forensic science.

2. Decision Support Systems (DSS)

Decision support systems (DSS) aim to assist decision-makers in making important decisions. They are not meant to replace humans, but rather to give suggestions and recommendations based on the available information from the provided options. The future of forensic science is discussed in a previous review [5]. Artificial Intelligence (AI) and Data Mining (DM) are also interrelated concepts with DSS. Probability and Bayesian methods, which originate from AI and DM, can be applied in DSS, underscoring the significance of employing DSS during a crime investigation.
DSS was originally defined as a system with the purpose of supporting management decision-makers in semi-structured decision conditions [6]. Though not specifically detailed, it was implied that the system would be computer-based, operate online, and generate graphical output. DSS can gather information from various sources, choose relevant and related knowledge intelligently, and structure the decision process efficiently. Artificial intelligent methods can also be applied to address heuristically problems into possible solutions. Applications of decision-making tools that are properly built can increase productivity, efficiency, and effectiveness in order to make optimal choices for technological processes. Elements of intelligence can be integrated into DSS, and an intelligence cycle is shown in Figure 1.

![Intelligence Cycle](image)

**Figure 1.** Intelligence cycle as demonstrated by previously [7].

A previous study discussed that intelligence is the result of a process that aims at transforming raw data into a suitable form for decision-making. It is a continuous process that is iterative and includes data collection, evaluation, collation, analysis, interpretation, dissemination, and re-evaluation; this is summarized in numerous publications [7].

Several other definitions can be compared and contrasted by examining multiple concepts used to define DSS. Thus, the many aspects of DSS contribute to the various interpretations of it in the literature. Developers need to consider the process of designing alternatives and provide a list of options from which to choose. Besides supporting choices, providing aid in modeling, and analyzing data, DSS must be designed in such a way that it can identify decision opportunities and structure solutions for problems. In order to improve the quality of decisions [8], it has been suggested that the decision problems need to be decomposed into simpler components that are well understood and defined appropriately. Thus, DSS can be applied in another field, such as in forensic science.

### 3. DSS for Forensic Science (FS)

Forensic science is a field of science that is important for solving crime cases and disaster recovery. There are a variety of strategies that are unique [9] and different from each other, but which can also be generalized [10]. One forensic science field is forensic DNA analysis, which involves familial searches in DNA databases [11]. The fundamental theory of forensic DNA analysis is related to forensic mathematics [12]. There are multiple forensic aspects to be considered in each case; examples involving cases such as sexual violence have been previously discussed by [13]. In other cases, many kinds of evidence can be retrieved from crime scenes, including earprints [14] and fingerprints. These research areas can be categorized under a multi-disciplinary field that includes computer science and
Forensic science, but not necessarily focused on a DSS. However, they can be improved by integrating them with the DSS concept in order to assist decision-makers in the forensic science field.

Forensic scientists have to cope with uncertainty in making decisions once receiving the evidence. They have to take these issues under consideration, as outlined by [15];

i) There is no complete theory regarding the problem domain in forensic examination and crime investigation; hence, there is always theoretical uncertainty.

ii) The space for including all relevant factors is limited, and it might require too many resources for a complete forensic examination and crime investigation.

iii) There is uncertainty about a particular individual (practical uncertainty) in the domain of interest.

iv) The decisions to be made are uncertain most of the time (decision-making under uncertainty), because it is necessary to make rational decisions even when there is not enough information to prove that the actions to be taken will work. This concept is the fundamental understanding related to decision theory.

Forensic scientists face the challenges of dealing with increasing casework and dozens of samples, including a long list of data from evidence that is waiting to be analyzed [16]. Scientists working in the forensic field will be able to provide DNA data for analysis. A previous study [17] defined forensics as the collection and study of traces mainly left at crime scene by suspects, as these materials can provide material or physical information from a crime scene. Knowledge that can be interpreted from forensics will be able to help scientists make decisions or provide a new alternative for solutions.

4. DSS for FS in Crime Investigation

The concept of decision-making, such as in a DSS, is also important for crime investigations [18-21]. COPLINK [20, 21] uses knowledge management to support decisions in crime analysis, which is a similar concept used on other previous work as well [18, 19]. The difference is in the applied case studies; some author have used knowledge management in police investigations [18, 19], while other authors use knowledge management in crime analysis but do not specifically mention police use [20, 21]. These studies show that knowledge management can have a significant contribution on decision support in crime investigation [18-22], as well as databases that are also important for search purposes and rational decision making [23]. A data-driven software tool has been proposed that enables the police department to share information [24]. Table 1 summarizes some of the related work regarding DSS in crime investigation as discussed in this section.

Table 1. Related work of DSS in crime investigation

<table>
<thead>
<tr>
<th>Topic (Author)</th>
<th>Main points</th>
<th>Applied method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision-theoretic analysis of forensic sampling criteria using Bayesian decision networks, 2012 [1].</td>
<td>This research addressed the methodology from decision theory that may help to cope with the wide range of sampling issues in forensic science applications.</td>
<td>Bayesian decision networks</td>
</tr>
<tr>
<td>Evidential calibration process of multi-agent based system: An application to forensic entomology, 2012 [2].</td>
<td>A decision support system (DSS) that was developed based on the belief function theory to validate and calibrate agent simulations. Results of this architecture are presented within the framework of forensics.</td>
<td>Decision support based on belief function theory</td>
</tr>
<tr>
<td>Topic (Author)</td>
<td>Main points</td>
<td>Applied method</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>Decision Support Systems for Police: Lessons From the Application of Data Mining Techniques, 2006 [25].</td>
<td>Authors discuss data mining and decision support technologies that have the potential to address issues and challenges faced by police with respect to detection and prevention of the volume crime of burglary.</td>
<td>Decision support and data mining</td>
</tr>
<tr>
<td>COPLINK Connect: information and knowledge management for law enforcement, 2003 [21].</td>
<td>Information technology professionals in law enforcement face information and knowledge management issues in a knowledge-intensive and time-critical environment.</td>
<td>Decision support and knowledge management</td>
</tr>
</tbody>
</table>

During a crime investigation, there are several sources of error in evidence collection and handling. Human factors such as corrupt scientists and officers, can alter evidence at many different points in the investigation process. Evidence might even be planted before or during the crime investigation itself. Lawyers and judges that are appointed to manage the case may compromise the evidence found by their own justification as well, even though it is obvious that DNA evidence directly links a suspect to a crime and cannot exonerate suspects from the list of criminal and convicted offenders. Database is one of the main components of a DSS. Therefore, implementation of a shared intelligence database is a significant step forward for this problem, as shown in Figure 2.

![Figure 2](image-url)  
*Figure 2. Technical setup of the system by [26] shows that each state police database automatically feeds into the shared intelligence database.*
Figure 3 shows the summary of the different roles of drug profiling. This also involves a component from DSS, which is knowledge-base that usually employs intelligence methods.

![Diagram of Drug Profiling Process](image)

**Figure 3. Summary of the different roles of drug profiling presented by [7].**

5. Applying the Bayesian Method for Forensic Science in Crime Investigation

The Bayesian method is a mathematical model and computerized method commonly applied in computer science (CS) and forensic science (FS) fields. Bayesian networks (BN) known to be useful for forensic DNA analysis in FS [27] and has been used to assess model adequacy. This involves the analysis of regression model before inferences. Comparisons of using the Bayesian method has been compared by [28] with fuzzy and neural networks. In computational forensics [29], which is an applied field in CS and FS, fuzzy Bayesian is applied [30].

As shown in Figure 4 and Figure 5, one of the proposed Bayesian networks (BN) can be used to evaluate cases involving transfer that include:

i) Random match probability of compared characteristics among database of a related population;

ii) Possibility that the stain has been left by a suspect (regardless of guilt or innocence), and;

iii) The relevance of a crime stain for the case (probabilistic in nature). The models shown in Figure 4 and Figure 5 are applicable for DNA evidence and evaluation using BN.

The nodes in Figure 4 are defined as follows: A: ‘X’ committed the murder; B: ‘Y’ committed the murder; E: eyewitness evidence of arrow between ‘X’, ‘Y’ and the victim some time before the crime was committed; F: fibres from a jacket similar to one found in the possession of ‘X’ are found at the scene of the crime; H: ‘Y’ drives the car of ‘X’ regularly; T: ‘Y’ picks up fibres from the jacket of ‘X’.
Figure 4. Bayesian network proposed by [31]

The definition of the variables in Figure 5 is as follows; $H$: the suspect is the offender; $G$: the crime stain came from the offender; $F$: the crime stain came from the suspect; $E$: the suspect’s blood sample matches the crime stain.

Figure 5. Bayesian network for ‘one-trace’ cases involving transfer from the offender [32]

The simplest form is a two-node network fragment that describes the state of variables observed, such as $E$ (short for ‘DNA match’ as in Figure 5) to come out with an inference on propositions of the source, which is a determination of whether the ‘matching’ suspect is or is not the source of the crime stain. Thus, if the latter propositions are represented by $F$, then the corresponding network fragment is $F \rightarrow E$. Currently there are several research studies focusing on the application of BN to inference problems related to the results of forensic DNA analysis. The study [33] shows how graphical structures for BN can be derived from initial pedigree presentations of forensic DNA identification. Table 2 shows the related works using the Bayesian approach for forensic science in crime investigation.
Table 2. Related works of the Bayesian approach for forensic science in crime investigation

<table>
<thead>
<tr>
<th>Topic (Author)</th>
<th>Main points</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning about Bayesian networks for forensic interpretation: An example based on ‘the problem of multiple propositions’ [34].</td>
<td>Bayesian networks and probabilistic evaluation. A gap with respect to writings that focus on foundational aspects and on how these may be acquired by interested scientists new to these topics. The presentation uses an example that relies on a casework scenario drawn from published literature, involving a questioned signature.</td>
<td>A complicating aspect of that case study - proposed to students in a teaching scenario - is due to the need to consider multiple competing propositions, which is an outset that may not readily be approached within a likelihood ratio based framework without drawing attention to some additional technical issues.</td>
</tr>
<tr>
<td>Signature verification using Bayesian approach [35].</td>
<td>Authors describe signature verification using a non-parametric Bayesian approach. The fully Bayesian approach has been shown to be powerful in machine learning.</td>
<td>Given sample(s) of Genuine signatures of an individual, the task of signature verification is a problem of classifying a questioned signature as Genuine or Forgery.</td>
</tr>
<tr>
<td>Constructing Bayesian networks for criminal profiling from limited data [36].</td>
<td>Authors develop an approach to obtain a Bayesian network model of criminal behaviour based on certain specific criteria in order to predict the unknown offender, given evidence at the crime scene.</td>
<td>Authors present a methodology for obtaining a Bayesian network (BN) model of offender behaviour from a database of clear homicides.</td>
</tr>
</tbody>
</table>

The court is looking into evaluating the ratio of the probability that the suspect left the crime stain over the probability that the suspect did not leave the crime stain [37]. There could also be a compromise of evidence collection procedures that may lead to a wrong conviction and interpretation.

6. Applying the Likelihood Ratio (LR) Method for FS in Crime Investigation

Obtaining the correct likelihood ratio (LR) for the hypotheses based on available evidences can help in building Bayesian networks (BN) to represent the described problems. Courts can use the weight of the evidence provided as an instrument to judge from a result of a Bayesian analysis. Scientists can combine their domain knowledge and data in one model, apply the technique into generating hypotheses, and solve problems of uncertain information systematically. BN and LR are sometimes used together depending on the problem domains.

There are many variables that can influence LR, and some are estimated directly while others are very subjective and controversial. Further systematic study is required to identify and prove hypotheses using LR or other alternative techniques. Different authors have measured LR in a variety of ways. Nevertheless, it is useless to develop techniques that have no impact towards the crime investigation process. The formula for LR is as follows: -Since a prosecution’s hypothesis that the suspect committed the crime is based on the family members, it can be simplified with \( H_0 = 1 \). The general equation of this formula is:

\[
LR = \frac{P(E|H_0,L)}{P(E|H_1,L)} = \frac{1}{P(E|H_1,L)}
\]  

\( H_0 = \) The source of DNA is from the family members of the suspect  
\( H_1 = \) The source of DNA is from an unknown person

In relation to the LR method, the mathematical model used to calculate allele frequencies of population statistics in a population database follow the concept of Bayes’ theorem [38, 39].

Taken together, the discussed studies have many methods and concepts that are currently used to solve scientific problems with computer software. The challenge will be to ensure that proper methodology is used going forward. Some of the dataset will be taken from domain experts in forensics,
while the rest is obtained from published data. It is first important to understand the complexity of LR because there will be multiple probability types of data to be analysed. There are current related research studies in this problem domain that discuss how to estimate empirical precision using LR [40], how to evaluate and compare data using LR [41], and the validity and reliability using LR [42]. Table 3 shows some of the related works using LR for forensic science in crime investigation.

Table 3. Related work on the likelihood ratio for forensic science in crime investigation

<table>
<thead>
<tr>
<th>Topic (Author)</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>An extended likelihood ratio framework for interpreting evidence [43].</td>
<td>Reviews some current methods including likelihood ratio-based approach and Bayesian approach to interpret evidence and discusses previously identified shortcomings.</td>
</tr>
<tr>
<td>The use of DNA statistics in criminal trials [37].</td>
<td>Determines the likelihood ratio from sample data using assumptions regarding the general population.</td>
</tr>
<tr>
<td>Evaluation of DNA match probability in criminal case [44].</td>
<td>Describes the weight as the likelihood ratio, which is the ratio of the probability for the DNA evidence assuming that the evidence sample came from the suspect to the corresponding probability assuming that the evidence sample came from another person.</td>
</tr>
<tr>
<td>A likelihood ratio approach to familial searching of large DNA databases [45].</td>
<td>Developed a familial searching approach that uses a combination of filtering by the number of shared alleles and ordering by likelihood ratio.</td>
</tr>
<tr>
<td>Likelihood ratio tests for triply multivariate data with structured correlation on spatial repeated measurements [46].</td>
<td>Studies the problems of tests of hypotheses on the Kronecker product structured covariance matrices with multiple q-variate observations over u sites and over p time / spatial points under the assumption of multivariate normality.</td>
</tr>
</tbody>
</table>

7. Applying Image Processing Methods for FS in Crime Investigation

Image processing methods can be applied in forensic science. A review on this issue has been conducted by [47]. Face recognition [48] can be implemented for crime investigation, as well as forensic sketch matching [49]. There are many issues in digital image forensics [50-53]. Fuzzy can also be applied in image forensics [54, 55]. Fuzzy matching can also be used in genome classification [56]. Recent works in fuzzy has expanded from fundamental concepts into database [57] and other applications [28, 58, 59]. Forensic image retrieval such as that described in [60] also involves matching like [49, 52]. Even analysis of a fingerprint [61] or closed-circuit television (CCTV) [62] involves - image processing. Table 4 lists the related works of image processing for forensic science in crime investigation.

Table 4. Related works of image processing for forensic science in crime investigation

<table>
<thead>
<tr>
<th>Topic (Author)</th>
<th>Main points</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature verification using Bayesian approach [35].</td>
<td>Authors describe signature verification using a non-parametric Bayesian approach. The fully Bayesian approach has been shown to be powerful in machine learning.</td>
<td>Given sample(s) of Genuine signatures of an individual, the task of signature verification is a problem of classifying a questioned signature as genuine or forgery.</td>
</tr>
<tr>
<td>3D Processing &amp; Visualization of Scanned Forensic Data [63].</td>
<td>Authors mention that in forensic pathology, the documentation of surface injuries is the major task.</td>
<td>In this contribution, a semi-automatic approach is presented for the processing of data from 3D photogrammetry for the visualization of the body surface.</td>
</tr>
</tbody>
</table>
### Topic (Author)

| Computerized matching of Shoeprints based on Sole Pattern [64]. | The computerized matching is done using concept Zernike moments | Shoeprints are common clues left at crime scenes that provide valuable evidence for identifying criminals. A shoeprint matching method based on power spectral density (PSD) and Zernike moment have been investigated. |
| Person Identification based on Barefoot 3D Sole Shape [65]. | Data normalization is performed and person identification is carried out using principal component analysis (PCA) and linear discriminant analysis (LDA). | PCA and LDA are in conjunction with 1-NN classification, and support vector machines (SVM) in conjunction with output value thresholding. |
| Fuzzy methods for forensic data analysis [66]. | Authors describe a methodology and an automatic procedure for inferring accurate and easily understandable expert-system-like rules from forensic data. | This methodology is based on the fuzzy set theory. The algorithms they used are described in detail, and were tested on forensic data sets. |

### 8. Conclusion

For Nowadays, research in the area of science and criminology are very important for social and economic development. An increase in crime imposes a direct impact on these areas in our country. As an example, forensic DNA analysis has been shown to be a reliable and accurate technique in identifying suspects related to crimes since the 1980s. Moreover, crime prevention remains a critical issue for future studies, and models for criminal incident prediction [67] will be very helpful to law enforcement.

Information technology and human resources utilized in the field of forensic science and computer science will be able to help use computerized analysis for forensic science data in crime investigation. Once successfully developed, the solution will be very interesting and worthwhile, because this knowledge can also be implemented in other specialized scientific fields. These studies will be important for improving the development of decision support systems (DSS) for use in forensic science during crime investigation and its research strategies, especially in Malaysia.

The main contribution of this paper is serve as a comprehensive reference to other DSS researchers. The survey scope is within the DSS methods that can be applied to field forensic science in crime investigation. In this paper, numerous different including integrated methods applied in DSS have been reviewed. Anyway, compared to other fields like water management, forest fire management, tourism advise, or even clinical DSS, application of DSS into forensic science in crime investigation is rare. This makes this field that are being explored are comparatively can provide more potential theories and methods to be applied for future work.

### 9. Acknowledgement

The authors wish to thank the Royal Malaysian Police (RMP), and especially Dato’ Seri Mohd Bakri Zinin, Director of the Criminal Investigation Department, Royal Malaysian Police (PDRM). This work was supported in part by a Fundamental Research Grant Scheme (FRGS), Ministry of Higher Education (MOHE), vot number 55077, and grant eScienceFund from the Ministry of Science and Technology Innovation (MOSTI), vot number 52056, managed by Research and Management Innovation Centre (RMIC) at Universiti Malaysia Terengganu (UMT).
10. References


