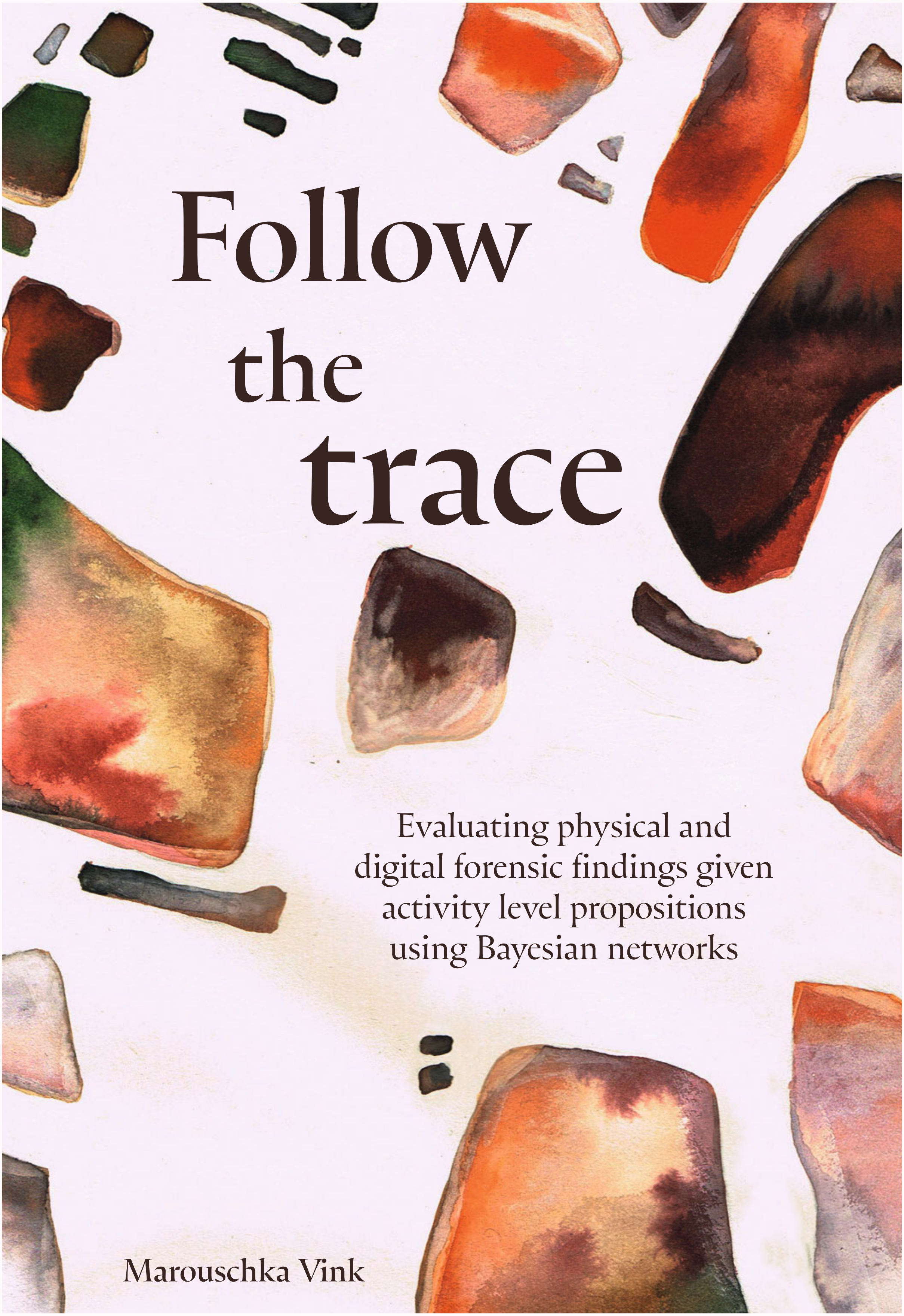
Follow the trace

Evaluating physical and digital forensic findings given activity level propositions using Bayesian networks

Vink, M.

Publication date
2025



Follow the trace

Evaluating physical and digital forensic findings given activity level propositions using Bayesian networks

Marouschka Vink

Follow the trace

Evaluating physical and digital forensic findings given activity level propositions using Bayesian networks

M. Vink

ISBN: 978-94-6522-693-4
Author: Marouschka Vink
Cover design by Jessy Zuiderwijk, @jessydoestattoos
Printed and bound by Ridderprint, www.ridderprint.nl

The printing of this thesis was financially supported by the Co van Ledden Hulsebosch Center, The Netherlands Center for Forensic Science and Medicine, and the Netherlands Forensic Institute.

The research described in this thesis received financial support from the Netherlands Forensic Institute as part of a multidisciplinary project called *Solide Activiteitniveau door Kruisbestuiving* (personal translation: *Solid Activity Level through Cross-pollination*).

CLHC
NETHERLANDS CENTER FOR FORENSIC SCIENCE AND MEDICINE

Netherlands Forensic Institute
Ministry of Justice and Security

Follow the trace
Evaluating physical and digital forensic findings given activity level propositions using Bayesian networks

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 11 november 2025, te 13.00 uur

door Marouschka Vink
geboren te Delfzijl

Promotiecommissie

Promotor:	prof. dr. M.J. Sjerps	Universiteit van Amsterdam
Copromotor:	dr. E. Musta	Universiteit van Amsterdam
Overige leden:	prof. dr. J.M. Mooij	Universiteit van Amsterdam
	prof. dr. A.C. van Asten	Universiteit van Amsterdam
	dr. M.J. Blom	Universiteit van Amsterdam
	prof. dr. H.B. Verheij	Rijksuniversiteit Groningen
	dr. D. Taylor	Forensic Science South Australia
	prof. dr. ir. H.M.A. van Beek	Open Universiteit

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Voor papa,

mijn allergrootste fan.

Personal note

During my physics master's program, I lost the connection between what I studied and the real world—even though understanding the world is fundamentally what physics is about. After transitioning from physics to the forensic science master's program, I rediscovered that connection. I have a strong preference for structure and enjoy making complex topics accessible to a broader audience. When I discovered reasoning under uncertainty in forensic casework—particularly during evidence evaluation given activity level propositions—I finally found a profession that aligned with my natural tendencies.

I will never forget the day our class was first introduced to Bayesian networks as a tool for modeling complex case scenarios. Modeling these networks resulted in structured chaos and it just clicked with me. And after almost five years of having the software Hugin opened on my laptop, I am still not done with it.

While this thesis marks the end of my PhD, it also serves as a starting point for my career and further exploration. This journey has contributed significantly to my development as a person and as a scientist. Through this thesis, I hope to give something meaningful back to this beautiful discipline.

Summary

Problem statement

In the Likelihood Ratio (LR) framework, forensic scientists evaluate their findings against (at least) two competing propositions. Forensic scientists are increasingly asked to evaluate their findings given activity level propositions (Evett *et al.*, 2002; Prinz *et al.*, 2024). These propositions address issues such as which activity, performed by whom and when, led to the traces, rather than focusing on the source and/or donor of these traces. "Activity level evaluations" can be complex due to the volume of available results, the amount and nature of information required, and the varying framework of case circumstances. Such complex casework benefits from structured reasoning, and Bayesian networks prove to be valuable tools for this purpose (e.g., Aitken *et al.* (2003); Dawid & Evett (1997); Garbolino & Taroni (2002); Taroni *et al.* (2014)). After examining the existing research on Bayesian networks for modeling activity level evaluations in forensic casework, we identified three research directions.

First, modeling Bayesian networks is a complex task that requires considerable time investment from forensic experts. Such time is not always available, and experts can benefit from more generic models, called "template Bayesian networks," that can be applied across various cases with different circumstances. A template model for evaluating DNA trace evidence given activity level propositions exists (Taylor *et al.*, 2018). This template is very practical and graphically intuitive as it is somewhat narrative-based, mapping possible pathways from alleged activities to traces found on items of interest (i.e., "Follow the trace"). Such a narrative template model is lacking for other forensic disciplines.

Second, forensic cases involving activity level evaluations often contain forensic results from multiple disciplines. Combining these results in a logical, balanced, robust, and transparent manner (Willis *et al.*, 2015) presents challenges. Key challenges include combining LR conclusions from different disciplines that address varying sets of (sub)propositions and determining whether multiple findings are conditionally independent given the case propositions. This interdisciplinary casework could also potentially benefit from a template Bayesian network.

Third, we observe a difference in evidence interpretation practices between digital forensic science and physical forensic disciplines, with the LR framework being used less in digital forensic science despite it currently being one of the most prominent disciplines in forensic casework. The question is whether and how methods and template Bayesian networks developed for physical forensic findings can also be applied to this domain.

The overarching research objective can be summarized as "making activity level evaluations and Bayesian networks more accessible and uniform across forensic disciplines—including digital forensic science."

Part II

Starting small: from idioms to case models

Following the Introduction in Part I, Part II explores the idiom-based approach (Neil *et al.*, 2000) to modeling Bayesian networks to reduce construction time and complexity. Idioms are building blocks that represent commonly recurring reasoning patterns. These patterns can be translated into small, primarily qualitative Bayesian networks and combined to build larger case models. Collections of idioms already exist for several scientific disciplines. In Chapter 2, we create one specifically for modeling the evaluation of forensic findings given activity level propositions. Using this collection, we build a generic yet sufficiently specific template Bayesian network that evaluates various types of transfer evidence given activity level propositions about disputed actors and/or activities (Chapter 3). This template builds upon the ideas of Taylor *et al.* (2018) and Kokshoorn *et al.* (2017).

Part III

Bayesian networks and interdisciplinary casework

Part III studies how to combine the evaluation of findings from various disciplines in interdisciplinary forensic casework using Bayesian networks. We use a hypothetical strangulation case where DNA traces and fibers are recovered from a sweater found near the crime scene. The case focuses on two disputed elements: the sweater's connection to the incident and its relation to the suspect. In Chapter 4, we build a case model for this example using the template Bayesian network from Part II. The case model includes a set of association propositions about whether or not the sweater was worn by the offender during the incident. These propositions make it possible to combine the evaluations of DNA traces and fiber traces on the sweater given the relevant propositions in the case ("who strangled the victim?"). In Chapter 5, we generalize this case-specific network into a template Bayesian network that evaluates multiple observations on an item with an uncertain relation to the activities, in addition to a disputed actor and/or activity.

Part IV

Exploring digital forensic science

Part IV studies how to advance logical, balanced, robust and transparent reasoning in digital forensic science to better align it with physical forensic disciplines. Chapter 6 shows that the LR framework can be used in relatively simple digital forensic casework with a single finding: distance data from the iPhone Health app. We build a method for the interpretation of this data in cases where an individual's walking distance is disputed. Chapters 7 and 8 address the use of Bayesian networks in more complex casework with multiple findings: Trojan horse defense cases. Because well-formulated propositions are essential to using the LR framework and to modeling Bayesian networks, we split this research into two parts. Chapter 7 demonstrates how to formulate propositions for various types of disputes in Trojan horse defense cases. One of these disputes covers the intent of an individual, which translates into propositions including the phrase "(un)knowingly." We argue that, although these are offense level propositions, the digital forensic expert can—with the use of appropriate expert knowledge—evaluate forensic findings given such propositions. In Chapter 8, we present the step-by-step modeling process of a case-specific Bayesian network, applying the template model from Parts II and III, and found no difficulties using it for Trojan horse defense cases.

Conclusion

Our overall aim is to improve probabilistic reasoning in mono- and interdisciplinary forensic casework addressing activity level propositions. By using concrete case examples throughout our work, we demonstrate the benefits of the LR framework and Bayesian networks for evaluating forensic findings given activity level propositions across various disciplines—including digital forensic science. The research resulted in a collection of idioms, two generally applicable template Bayesian networks, and laid the foundation for two use cases of the LR framework in the digital forensic domain: a method for the interpretation of iPhone Health app data and a Bayesian network for the evaluation of digital findings in Trojan horse defense cases.

Contents

List of Figures

Clarification of wording

During the research for this thesis, I discovered that certain terminology is debated within the forensic science community. Below, I present the most prominent words or phrases in my thesis and the different ways I named them throughout the work. I decided to keep the original wording throughout the text while adding this clarification about what I mean—or do not mean—by these terms.

1. *(Forensic) evidence, findings, observations*
   These words are used interchangeably for the results obtained by the forensic scientists. During the course of this thesis, I learned that "evidence" may not be the preferred term because of its legal connotation.

2. *Activity level evaluations, evidence evaluations at activity level, evaluating evidence given activity level propositions, addressing activity level issues*
   With these phrasings I do not intend to say that the forensic scientist evaluates propositions; rather, the forensic scientist evaluates their findings *given* at least two propositions (which can be *at* activity level).

3. *Hypotheses, propositions*
   Both hypotheses and propositions are workable formulations of scenarios for forensic scientists and are denoted as "H" in the likelihood ratio, prior odds, and posterior odds.

4. *Forensic practitioner, scientist, expert*

These terms refer to the same profession and are used interchangeably throughout this thesis.

5. *Physical forensic science, traditional forensic science*
   Both are defined as forensic science disciplines dealing with physical (trace) evidence (e.g., fibers, glass, DNA).