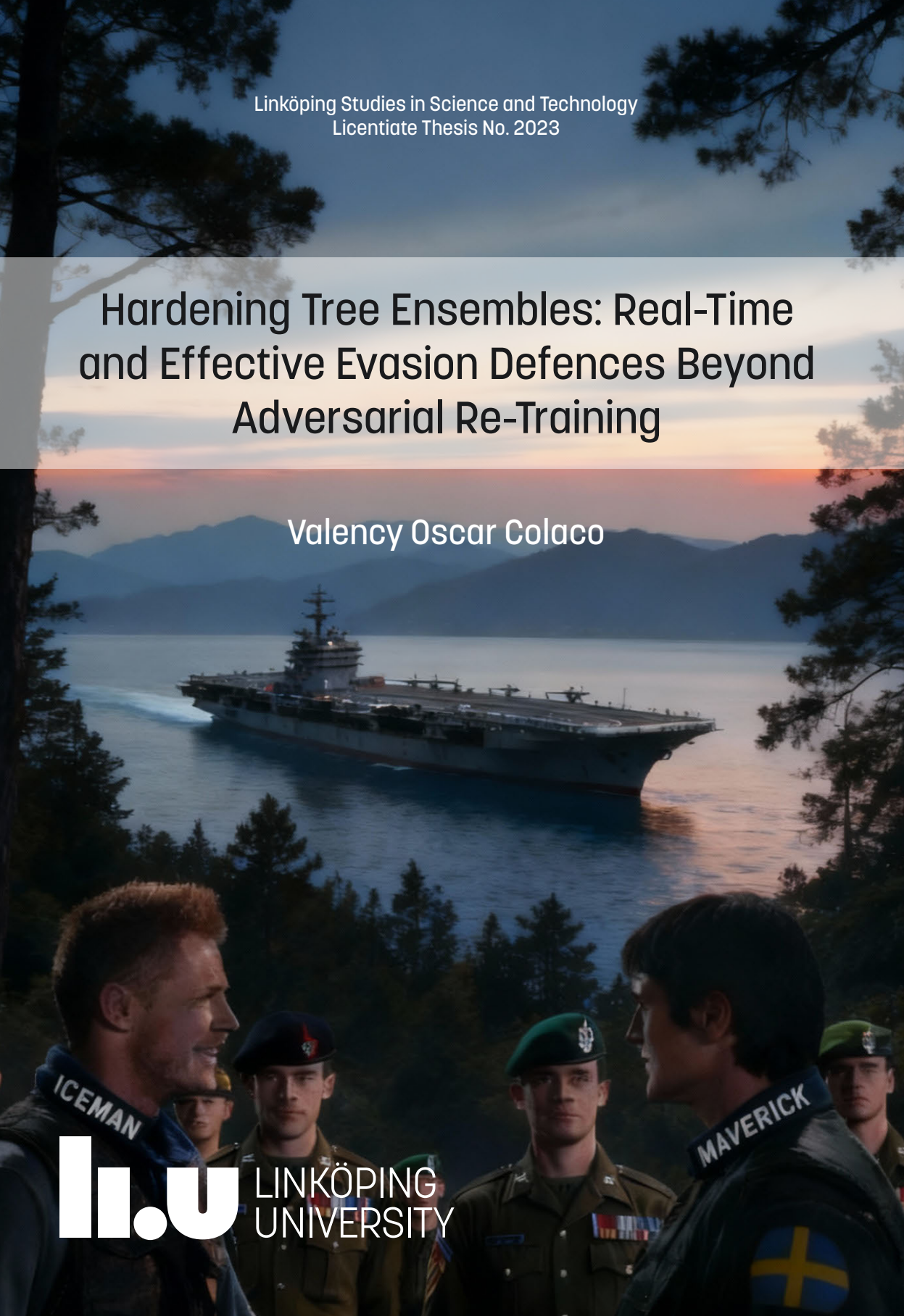
Linköping Studies in Science and Technology
Licentiate Thesis No. 2023

Hardening Tree Ensembles: Real-Time and Effective Evasion Defences Beyond Adversarial Re-Training

Valency Oscar Colaco



LINKÖPING UNIVERSITY

Linköping Studies in Science and Technology
Licentiate Thesis No. 2023

Hardening Tree Ensembles: Real-Time and Effective Evasion Defences Beyond Adversarial Re-Training

Valency Oscar Colaco

Linköping University
Department of Computer and Information Science
Division for Software and Systems (SaS)
SE-581 83 Linköping, Sweden

Linköping 2025

This is a Swedish Licentiate's Thesis

Swedish postgraduate education leads to a doctor's degree and/or a licentiate's degree.
A doctor's degree comprises 240 ECTS credits (4 years of full-time studies).
A licentiate's degree comprises 120 ECTS credits.

Except where otherwise noted, this work is licensed under a Creative Commons Attribution 4.0 International License.
https://creativecommons.org/licenses/by/4.0/

Typeset using LaTeX

The cover images for this thesis were generated using AI via replicate.com, through a combination of the ByteDance/Seedream-4 and Black-Forest-Labs/Flux-1.1-Pro models.

Printed by LiU-Tryck, Linköping 2025

Edition 1:1

© Valency Oscar Colaco, 2025
ISBN 978–91–8118–325–2 (print)
ISBN 978–91–8118–326–9 (PDF)
https://doi.org/10.3384/9789181183269
ISSN 0280-7971
Published articles have been reprinted with permission from the respective copyright holder.

POPULÄRVETENSKAPLIG SAMMANFATTNING

Maskininlärningsmodeller som trädensembler – till exempel random forests och gradient boosting-maskiner – är grupper av enkla beslutsträd som tillsammans tar bättre beslut än varje enskilt träd på egen hand. Dessa modeller presterar ofta bättre än till och med avancerade tekniker som djupa neurala nätverk i många olika uppgifter. En av deras styrkor är att beslutsprocessen kan förstås av människor, vilket gör dem särskilt användbara i säkerhetskritiska sammanhang där det är viktigt att kunna lita på systemets val. Men dessa modeller kan luras genom så kallade evasion-attacker, där angriparen använder noggrant konstruerade, små förändringar i indata för att få modellen att fatta felaktiga beslut. Denna typ av attacker har studerats ingående för neurala nätverk, men har fått betydligt mindre uppmärksamhet för trädensembler.

Denna avhandling undersöker både proaktiva och reaktiva strategier för att skydda trädensembler mot evasion-attacker, en process som kallas evasion hardening. Proaktiva metoder, som exempelvis omträning, syftar till att göra modellerna mer robusta mot angrepp. Vi visar dock att omträning ofta är ineffektivt för trädensembler och i vissa fall till och med kan göra dem mer sårbara. Detta understryker behovet av alternativa försvarsmekanismer. För att möta detta utvecklade vi två nya reaktiva försvar – *Iceman* och *Maverick* – som kan upptäcka evasion-attacker i redan tränade och driftsatta trädensembler. Dessa system kan identifiera angrepp oberoende av hur de genererats och kräver inga förändringar av de underliggande modellerna, som ofta är resultatet av omfattande ingenjörsarbete.

Iceman överträffar befintliga toppmoderna metoder för att upptäcka angrepp mot trädensembler, såsom OC-SCORE och GROOT. Nackdelen är dock att *Iceman* endast fungerar om man i förväg känner till den maximala förändring en angripare kan göra i indata, vilket begränsar dess användbarhet. För att hantera detta utvecklade vi *Maverick*, ett försvar som är särskilt utformat för situationer där denna gräns är okänd. *Maverick* uppnår markanta hastighetsförbättringar jämfört med OC-SCORE samtidigt som den behåller samma prestanda, och fungerar i samspel med *Iceman* för att ge en komplett försvarsstrategi mot evasion-attacker. Sammantaget visar våra resultat både de proaktiva metodernas begränsningar och den stora potential som reaktiva metoder har för att stärka trädensembler mot denna typ av angrepp.

ABSTRACT

Tree ensembles like random forests and gradient boosting machines are widely used machine learning (ML) models, often outperforming advanced techniques like deep neural networks on structured tabular data tasks. These models also have interpretable (human-understandable) structures that enable stakeholders to trace the decision-making process, making them particularly suitable for use in safety- and security-critical applications where trust in the model's behaviour is paramount. Despite these advantages, recent work has shown that they are highly vulnerable to adversarial examples: carefully perturbed inputs that elicit misclassifications.

These vulnerabilities are especially concerning as ML continues to permeate domains that are critical to societal functioning. Their seriousness is underscored by legislation such as the recently passed European Union Artificial Intelligence (AI) Act. This act mandates resilience against AI-specific vulnerabilities like evasion attacks caused by adversarial examples targeting ML models at inference time. Measures intended to improve resilience against such evasions, often referred to as hardening, generally involve two strategies: proactive defences, which aim to make models robust (e.g., adversarial re-training), and reactive defences, which focus on detecting and mitigating evasions at inference time. This thesis examines both strategies; it shows that proactive methods like model re-training are ineffective for tree ensembles and consequently advances the state-of-the-art in reactive defences.

In the context of re-training, doubling the training set through targeted data augmentation steps left accuracy largely unchanged. However, robustness, when quantified using formal verification techniques, dropped by 28–82% across two case studies. This indicates that model re-training alone is ineffective for tree ensembles. To address this, we leveraged formal methods to develop *Iceman*, a prototype system that uses counterexample regions which violate the robustness property to detect evasion attempts. *Iceman* can detect evasion attacks regardless of the attack generation process without modifying the underlying tree ensemble. It outperforms the current state-of-the-art methods in evasion detection, OC-Score and GROOT. Across four case studies, it improves Matthews Correlation Coefficient scores by 0.20–0.91 and achieves detection speeds 5–115x faster than OC-Score. In addition, it provides alert filtering and prioritisation capabilities with over 98% accuracy to address alert fatigue in intrusion detection systems. However, *Iceman*'s applicability is limited to scenarios with fixed attacker perturbation budgets, characterised by pre-defined constraints on the input manipulations that an attacker can apply.

To expand this applicability to unconstrained attacker perturbation budgets, we developed an additional system, called *Maverick*, designed to complement *Iceman* for a better defensive strategy. Just like *Iceman*, *Maverick* does not modify the underlying tree ensemble and can detect evasion attacks regardless of the attack generation process. We prove that *Maverick*'s core detection mechanism is mathematically equivalent to OC-Score, and present enhancements that achieve 85–563x speedups over OC-Score while maintaining identical detection performance and supporting evasion attack diagnostics with over 93% accuracy.

Acknowledgments

First and foremost, I would like to thank my supervisor, Professor Simin Nadjm-Tehrani, for challenging my ideas and teaching me to look beyond the jazz and focus on what really matters. You have fundamentally reshaped the way I think, making me question everything I once thought I knew. I would also like to thank my co-supervisor, Professor Mikael Asplund, for providing invaluable feedback throughout every stage of my work. Your breadth of knowledge across so many areas has been a true treasure, and your guidance has improved my work in immeasurable ways.

I am incredibly thankful to my parents, whose unconditional love, encouragement, and unwavering belief in me have been essential in helping me reach this point. Their support has been the steady foundation beneath every step of this journey. I am also deeply grateful to my colleagues at RTSLAB for the many discussions and valuable feedback I received during our weekly lab meetings. I truly appreciate the insightful and productive exchanges we had — whether related to my research or yours — and for all the opportunities to learn from and alongside you. To my fellow PhD students and colleagues at IDA, alongside peers at KTH, Chalmers, Lund, Umeå, Örebro, and Uppsala, our discussions and shared experiences have contributed more to this thesis than you can ever imagine.

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Valency Oscar Colaco
Linköping, December 2025

Contents

Abstract **iii**

Acknowledgments **v**

Contents **vii**

1 Introduction **1**
1.1 Aim 4
1.2 Research Questions 4
1.3 Included Papers 5
1.4 Thesis Contributions 6
1.5 Author's Contributions 6
1.6 Thesis Outline 7

2 Thesis Positioning and Terminology **9**
2.1 Preliminaries 9
2.2 Trade-Offs Between Accuracy and Robustness 16
2.3 Hardening Tree Ensembles against Evasion Attacks 17

3 Thesis Summary **19**
3.1 Summary of Results 19
3.2 Conclusion and Future Works 20

Bibliography **25**

4 Paper I **35**

5 Paper II **57**

6 Paper III **79**

Chapter One

Introduction

"Son, your ego is writing checks your body can't cash."
- Commander Stinger, Top Gun (1986)

This theatrically stark warning against Top Gun aviator Captain Pete "Maverick" Mitchell's unchecked ambition serves as an apt prelude to the field of artificial intelligence (AI), which is rapidly transforming modern society. AI now offers unprecedented opportunities across industries such as healthcare [1, 2], finance [3], defence [4], and transportation [5]. Yet, as ambition surges ahead, so too does the risk of unintended consequences when these (mostly) black-box systems behave in unforeseen ways.

Bridging the divide between what is expected of an AI-based system and what it will actually do is the very essence of AI safety [6, 7], which can be defined as ensuring that any harm to humans and the environment resulting from the deployment of AI-based systems remains below an acceptable threshold. In the AI literature, however, this definition of harm is fairly broad, encompassing aspects such as online content safety, deepfakes, and the dissemination of misinformation. To this end, this thesis defines AI safety more narrowly to focus on the harm caused by faulty inputs to AI-based safety-critical or security-critical systems.

Within the context of this narrowed definition, real-world incidents highlight the importance of AI safety. In 2018, a self-driving Uber vehicle in Arizona struck and killed a pedestrian - the first known fatality involving intelligent driving systems. Investigations by the National Transportation Safety Board (NTSB) revealed that the vehicle's AI system lacked proper fail-safes to compensate for human disengagement [8]. In 2023, a Cruise Robotaxi struck a pedestrian in San Francisco and dragged her approximately 20 feet, ultimately coming to a stop with its rear wheel resting on her legs. Investigations revealed critical limitations in the vehicle's programming: it failed to recognise that a pedestrian was about to be hit in an adjacent lane and was unable to detect that the pedestrian was being dragged, despite her legs being partially visible to one of the onboard cameras [9]. These incidents demonstrate how faulty or incomplete inputs (without any malicious intent) to AI-based systems can lead to safety violations.

Extending this notion of faulty inputs, similar incidents may also arise when adversaries intentionally manipulate inputs to cause harm. Accordingly, the definition of AI safety in this thesis also encompasses security attacks known as *adversarial examples* [10]. These are inputs deliberately crafted by malicious actors to deceive machine learning (ML) models into causing a system to take incorrect or unsafe actions. They are generated by perturbing the model's input until misclassifications occur. These perturbations may be constrained by an attacker's fixed budget (i.e., pre-defined limits placed on the attacker's input manipulation capabilities), as the attack must remain stealthy and large perturbations could be easily flagged by simple mechanisms such as input filtering or sanitisation. However, the perturbations may also be unconstrained, allowing the attacker greater freedom to manipulate the input and influence the output of AI-based systems.

Whether constrained or unconstrained, such adversarial examples can cause AI-based systems to produce unsafe or unintended outcomes. For instance, within academic settings, researchers from the University of Pennsylvania persuaded a simulated self-driving car to ignore stop signs and drive off a bridge [11]. Researchers also demonstrated successful attacks on production-grade autonomous driving software that trigger unnecessary emergency braking [12] (an action that could be life-threatening in certain scenarios, such as sudden braking on high-speed roadways). In another instance, small, innocuous-looking stickers added to a stop sign consistently fooled a neural network into misclassifying it as a 45 MPH speed-limit sign, both in controlled lab environments and real-world driving scenarios [13, 14]. These findings demonstrate how security threats (like adversarial examples) against AI-based systems can cause them to produce safety violations, highlighting the need for AI security.

AI security, in general, is defined as the process of managing the design, implementation, and operation of AI models, systems, and data throughout their lifecycle to reduce the risk of harm, whether from deliberate, unwanted, hostile or malicious acts, or from failures to act [15]. The practical importance of AI security is evident from real-world attacks on deployed systems. For instance, Slack AI was exploited via prompt injection [16], enabling attackers to trick the assistant into retrieving and leaking confidential data, such as API keys from private channels [17]. Meanwhile, Microsoft 365 Copilot suffered from a critical zero-click vulnerability known as EchoLeak (CVE-2025-32711) [18], where a crafted email containing hidden instructions could cause Copilot to exfiltrate sensitive internal data like emails, documents, or chat history without any user interaction. Finally, security researchers demonstrated a prompt injection technique called "Invitation Is All You Need" [19], where malicious Google Calendar invites injected hidden commands into Gemini's context. When the users asked Gemini to

summarise their calendar, the AI executed actions such as activating smart home devices and geo-locating the user without their consent, revealing how indirect prompt injections can trigger real-world effects. To this end, this thesis adopts a more focused definition of AI security, specifically defining it as defending AI-based systems against adversarial examples crafted to induce intentional misclassifications.

Such adversarial examples, when directed at ML models, facilitate *evasion attacks* where the attacker's goal is to compromise the integrity of the underlying system by tampering with its core functionality [13]. Recognising the seriousness of such threats, even global legislation like the recently passed European Union (EU) Artificial Intelligence (AI) Act[1] on the regulation of AI mandates the use of technical solutions to ensure cybersecurity against AI-specific vulnerabilities [20], i.e., solutions aimed at defending the models. In this context, the set of techniques used to defend ML models against evasion attacks is collectively referred to as *evasion hardening* [21].

Although a substantial body of research has explored evasion hardening for ML models such as deep neural networks (DNNs), tree ensembles have received considerably less attention. [22]. This is despite tree ensembles often outperforming DNNs on various tasks involving structured or tabular data [23, 24]. Furthermore, their interpretable (human-understandable) structures also make them particularly suitable for safety-critical applications where trust in the model's behaviour is paramount. Additionally, they are well-suited for use in cybersecurity applications [25].

For tree ensembles deployed in cybersecurity applications such as intrusion detection systems (IDSs), evasion attacks at runtime can compromise system integrity by tampering with their intended functionality (of detecting attacks). Conversely, adversaries can manipulate benign traffic to influence the IDS's output, causing it to falsely flag benign activity as malicious, dramatically increasing the number of false alarms and reducing system availability. This phenomenon can overwhelm analysts and increase incident response times, a problem commonly referred to as "alert fatigue" [26]. In both scenarios, the compromised system integrity and inflated false alarms directly erode user trust and undermine the practical usability of the system.

In summary, this thesis highlights the urgent need for AI systems, especially in safety-critical and security-critical domains, to be robust and resilient

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144

against such adversarial examples. This ensures that as ML models grow more powerful and capable, they also grow more robust and resilient, so that "writing checks" with AI doesn't force us into liabilities our systems can't cash. While robustness quantifies the degree of degradation in system functionality under input perturbations [27], resilience denotes the maximum admissible perturbation magnitude that the system can tolerate before a loss of functionality occurs [28].

1.1 Aim

This thesis aims to advance the understanding of the intertwined nature of safety and security concerns in AI systems that incorporate ML components, with a particular focus on tree ensembles. The primary goal is to develop innovative methods and tools that support safety assurances even in the presence of security threats targeting AI-based systems. The secondary goal of this thesis is also to critically challenge longstanding methodologies for achieving empirical robustness in tree ensembles against noisy inputs or adversarial examples. While this licentiate thesis does not offer complete solutions to the multifaceted challenges at the intersection of safety, security, and AI, it lays important groundwork toward achieving these objectives. Specifically, it focuses on developing practical tools to enable real-time detection of evasion attacks in tree-ensemble-based intrusion detection systems. Additionally, this thesis seeks to enhance tree-ensemble-based IDSs by incorporating advanced functionalities designed to prioritize and filter alerts while also enabling attack diagnostics through the discovery of evasion strategies.

1.2 Research Questions

This thesis is guided by two research questions that challenge longstanding methodologies for achieving robustness and propose new methods for defending tree ensembles against evasion attacks. The research questions are formulated as follows:

Q1. How does adversarial re-training impact the accuracy and robustness of tree ensembles?

In the broader literature, adversarial re-training, or enhancing the training of ML models such as deep neural networks through data augmentation, has been widely adopted as the prevalent method for improving accuracy and robustness [29, 30, 31, 32, 33, 34]. While this method has also been extensively applied to tree ensembles, at the onset of this work, no study had systematically investigated the trade-offs between accuracy and robustness resulting from the re-training process.

Q2. How can tree ensembles be defended against evasion attacks that may occur in real-time?

If evasion attacks cannot be effectively addressed through re-training (as investigated in Q1), novel methods become necessary to safeguard against them. However, this task is far from straightforward. Firstly, the literature on defensive methods is fragmented, with most approaches targeting specific evasion strategies. This presents a fundamental limitation, as a defence tailored to a particular attack strategy is ineffective against alternative strategies. Furthermore, evasion attacks can be crafted under different assumptions regarding the attacker's budget, ranging from constrained to unconstrained, depending on the threat model. Finally, such attacks may occur in real-time scenarios, where timely detection is critical to the reliable functioning of the application (e.g., automotive intrusion detection systems), or in non-real-time contexts, such as offline facial recognition (to record attendance in classrooms or workplaces), where timely detection is less critical. In this context, this research question can be further refined into three sub-questions as follows:

2a. How can tree ensemble defences be designed to remain effective regardless of any specific method used to generate evasion attacks?

2b. How can evasion defences be designed to maintain effectiveness against attackers operating under different perturbation budgets?

2c. Can evasion defences satisfy the timing requirements of real-time applications?

1.3 Included Papers

The following is a list of publications included in this thesis:

I. Valency Oscar Colaco and Simin Nadjm-Tehrani. "Formal Verification of Tree Ensembles against Real-World Composite Geometric Perturbations". In: *Workshop on Artificial Intelligence Safety 2023 (SafeAI 2023) co-located with the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI)*. CEUR-WS. 2023. URN: urn:nbn:de:0074-3381-7. URL: https://ceur-ws.org/Vol-3381/38.pdf

II. Valency Oscar Colaco and Simin Nadjm-Tehrani. "Fast Evasion Detection & Alert Management in Tree-Ensemble-Based Intrusion Detection Systems". In: *2024 IEEE 36th International Conference on Tools with Artificial Intelligence (ICTAI)*. Los Alamitos, CA, USA: IEEE Computer Society, Oct. 2024, pp. 404–412. DOI: 10.1109/ICTAI62512.2024.00065

III. Valency Oscar Colaco and Simin Nadjm-Tehrani. "Real-Time Evasion Detection in Tree Ensemble Automotive Intrusion Detection Systems". In: *2025 IEEE 16th Vehicular Networking Conference (VNC)*. 2025, pp. 1–8. DOI: 10.1109/VNC64509.2025.11054177

The publications relate to the research questions as follows: Paper I addresses Q1, while Papers II and III focus on Q2. The following subsections outline the thesis contributions along with each author's contributions in relation to the three publications above.

1.4 Thesis Contributions

The contributions of this thesis are summarised as follows:

- **Paper I** challenged the validity of model re-training via data augmentation as a means of achieving robustness in tree ensembles and presented novel insights in this regard.
- **Paper II** presented a **specialised** (designed for a specific attacker perturbation budget) tree-ensemble-based intrusion detection system called *Iceman*, which advanced the state-of-the-art in terms of evasion detection performance.
- **Paper III** presented a **generalised** (designed to operate under unknown attacker perturbation budgets) tree-ensemble-based intrusion detection system called *Maverick*, which advanced the state-of-the-art in terms of evasion detection speed.

1.5 Author's Contributions

Publication I: Formal Verification of Tree Ensembles against Real-World Composite Geometric Perturbations

While I came up with the idea that abstraction functions could be used to verify tree-ensemble-based, vision-driven systems against composite geometric perturbations - a concept that trivially extends to adversarial examples, I also wrote the majority of the paper, designed, and conducted the experiments. My co-author, meanwhile, also contributed significantly to the mathematically rigorous formulations (needed for papers with a formal verification focus) and provided key insights that strengthened the discussion section, as well as improvements to the verification algorithm and workflow presented in the paper. My overall contribution to this research was approximately 70%.

Publication II: Fast Evasion Detection & Alert Management in Tree-Ensemble-Based Intrusion Detection Systems

I proposed that counterexample regions violating the robustness property could be used to effectively detect evasion attacks within an attacker's fixed perturbation budget. Since no method existed in the literature to exhaustively enumerate these regions in high-dimensional spaces, I developed a novel approach to address this challenge for tree ensembles. I also wrote the majority of the paper and designed and conducted the experiments. My co-author contributed by restructuring and reorganizing the manuscript into a cohesive and readable format, enhancing the counterexample region generation algorithm, and improving the verification workflow and system architecture figures. My overall contribution to this research was approximately 65%.

Publication III: Real-Time Evasion Detection in Tree Ensemble Automotive Intrusion Detection Systems

From the previous publication, I identified a method in the literature that could detect evasion attacks in tree ensembles with unconstrained attacker budgets. However, their approach was extremely costly in terms of inference times, making it impractical for use in real-time applications. To address this limitation, I proposed a more efficient method and provided a proof demonstrating that my improvement was mathematically equivalent to the original method. For this publication, I proposed the method, developed the proof, wrote the majority of the paper, and designed and conducted the experiments. My co-author contributed valuable insights to the introduction and discussion sections and leveraged her extensive expertise in real-time systems to greatly enhance the paper. My overall contribution to this research was approximately 80%.

1.6 Thesis Outline

Chapter 2 provides a high-level overview of the related works and positions this thesis within that context. It also provides the background knowledge and terminology used throughout this thesis. Chapter 3 summarizes the main results, highlights how the thesis contributions address the research questions, and concludes the thesis. Finally, the included publications are appended at the end in the standardized thesis format rather than in their original two-column publication style. This change was made as the original format's two-column layout would shrink the text to an unreadable size, making it unsuitable for inclusion in this thesis.

Chapter Two

Thesis Positioning and Terminology

This chapter provides a high-level overview of related work and positions this thesis within that context. Before providing that overview, it is essential to establish a background understanding of decision trees, tree ensembles, the Verifier of Tree Ensembles (VoTE), evasion attacks, and the precise definitions of accuracy and robustness adopted in this thesis.

2.1 Preliminaries

This section outlines the theoretical foundations essential for understanding the key concepts central to this thesis.

2.1.1 Decision Trees

Decision trees are machine learning models that make predictions by learning hierarchical if-then rules through recursive partitioning of the input space. They are commonly trained using CART (Classification and Regression Trees) algorithms [38]. A decision tree implements a prediction function $t : X^n \rightarrow \mathbb{R}^m$ that maps disjoint sets of points $X_i \subset X^n$ to a single output $\overline{y_i} \in \mathbb{R}^m$, i.e.,

$$t(\overline{x}) = \begin{cases} (y_{1,1}, ..., y_{1,m}) & \text{if } \overline{x} \in X_1 \\ \vdots & \\ (y_{k,1}, ..., y_{k,m}) & \text{if } \overline{x} \in X_k \end{cases}$$

Where k is the number of disjoint sets and $X^n = \bigcup_{i=1}^{k} X_i$

The n-dimensional input domain X^n includes elements $\overline{x}$ as tuples in which each input variable x_i captures some feature of interest of the system [39]. In this thesis, we only consider binary trees in which each decision node uses a single input variable with a linear decision function, a structure called univariate hard decision trees by Irsoy et al. [40]. State-of-the-art implementations of tree ensembles like scikit-learn [41], CatBoost [42], and XGBoost [43] typically use univariate hard decision trees.

2.1.2 Tree Ensembles

Decision trees suffer from a phenomenon called overfitting, in which the models tend to memorise the training data instead of learning the patterns associated with the data to make generalised predictions. To counteract these issues, Breiman [44] proposed Random Forests.

A random forest $f : X^n \rightarrow \mathbb{R}^m$ is an ensemble of B decision trees that produces outputs by averaging the values emitted by each individual tree, i.e.,

$$f(\overline{x}) = \frac{1}{B} \sum_{b=1}^{B} t_b(\overline{x})$$

Where t_b is the prediction function of the b^{th} tree in the ensemble.

Random forests create multiple decision trees using random subsets of the data and features, then aggregate their predictions through a process known as bootstrap aggregation [45], or bagging. This approach helps reduce variance by averaging out the individual tree predictions, making the model less sensitive to noise in the training data and mitigating overfitting. In contrast, gradient-boosting machines build trees sequentially (except for XGBoost, which constructs trees in parallel), with each new tree correcting the errors of the previous trees. This process, also known as boosting [46], creates multiple weak learners and reduces bias, thereby addressing underfitting, a situation where a trained model is too simple to meaningfully (and accurately) capture the underlying patterns present in the data.

2.1.3 Classifiers

Decision trees and tree ensembles can be used as classifiers to categorise samples from an input domain into one or more classes and assign each sample a label unique to its class. Let $f(\overline{x}) = (y_1, ..., y_m)$ represent a model trained to predict the probability y_i associated with a class i within disjoint regions in the input domain, where m is the number of classes. A classifier $f_c(\overline{x})$ may then be defined as,

$$f_c(\overline{x}) = \underset{i}{argmax}\ y_i$$

In this thesis, we only consider classifiers that map each point from an input domain to exactly one class.

2.1.4 The Verifier of Tree Ensembles (VoTE)

The Verifier of Tree Ensembles (VoTE) [39] is a formal verification tool-suite specifically designed for tree ensembles, enabling the assessment of their

compliance with user-defined properties. Built upon the principles of abstract interpretation, the tool consists of two main components: the VoTE Core and the VoTE Property Checker. The VoTE Core takes as input an instance of a tree ensemble classifier (f_c) along with an input region (X^n). Simultaneously, the Property Checker receives a user-defined property (Ψ), which serves as the key criterion for the verification process. The VoTE

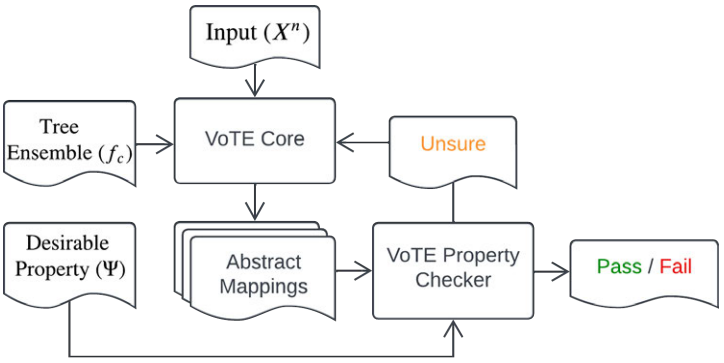


Figure 2.1: Verification workflow of the Verifier of Tree Ensembles [39, 36]

Core then partitions the input region into abstract mappings (X_i, Y_a), where $X_i \subseteq X^n$ denotes a precise input region and Y_a is a conservative approximation of the output from f_c with respect to X_i, i.e., $Y_a \supseteq \{f_c(\overline{x}) : \overline{x} \in X_i\}$.

These abstract mappings are then individually evaluated against the specified property. If all mappings satisfy the property, the verification outcome is reported as Pass. Conversely, if any mapping violates the property, the outcome is Fail. However, since abstract interpretation, in general, is not complete, sometimes an abstraction can be too conservative to provide a conclusive outcome. In such cases, the verification outcome is reported as Unsure, prompting the VoTE Core to refine the abstraction by further subdividing it into finer mappings before re-evaluation. The overall workflow of VoTE is illustrated in Figure 2.1.

2.1.5 Accuracy

A classifier (f_c) with respect to an input ($\overline{x}$) is considered accurate if its prediction matches the ground truth label (l), that is, if:

$$f_c(\overline{x}) = l$$

It is important to note that this notion of accuracy is defined at the level of a single input point, and in practice, accuracy level is typically measured over a test or validation set, and this definition can be readily extended to such sets, e.g., for labelled samples, $(\overline{x}, l) \in (X_{test}, y_{test})$ as:

$$\textbf{accuracy} = \frac{|\{(\overline{x}, l) \in (X_{test}, y_{test}) : f_c(\overline{x}) = l\}|}{|X_{test}|}$$

Where $|\cdot|$ is the set cardinality. When assessing classifiers over a region surrounding a given input, this pointwise definition (of accuracy) is insufficient. A stronger definition, suitable for regions, adopted in this thesis, is introduced in the following subsection.

2.1.6 Robustness

While definitions of robustness vary across the literature, in this thesis, we define robustness as the ability of a machine learning model to withstand noisy inputs or adversarial examples without a loss of its intended functionality. This definition is closely aligned with the one proposed by Törnblom et al. [39]. While their formulation primarily considers robustness in the context of noisy inputs, it can be trivially extended to include adversarial examples as well. The key distinction lies in the perturbation mechanism: noisy inputs perturb the input data randomly within a given budget, whereas adversarial examples are algorithmically crafted to intentionally elicit misclassifications. To this end, if the system is exhaustively evaluated against all noisy perturbations within a given budget, robustness against adversarial perturbations (within that same budget) is automatically covered. In essence, adversarial perturbations can be regarded as a subset of noisy perturbations, and any property that holds for noisy perturbations within a given budget will also hold for adversarial perturbations when analysed through exhaustive enumeration approaches, e.g., formal methods. The distinction between the types of perturbations (within a given budget) is illustrated in Figure 2.2.

In this context, this thesis defines robustness in terms of input–output mappings. By specifying an input region - consisting of an input point and a perturbation region definition around it, and the corresponding expected output, we can determine whether the classifier is robust (with respect to that input region). Formally, let f_c be a classifier subject to verification, $\overline{x}$ an n-dimensional input sample with label l, $\epsilon \in \mathbb{R}_{>0}$ a robustness margin, and $\Delta = \{\delta \in \mathbb{R} : -\epsilon < \delta < \epsilon\}$ the set of allowable perturbations. If $\overline{\delta}$ is an n-tuple of perturbations drawn from Δ, then the classifier is robust with respect to $\overline{x}$ and Δ if and only if:

$$\forall \overline{\delta} \in \Delta^n,\ f_c(\overline{x}) = f_c(\overline{x} + \overline{\delta}) = l \tag{2.1}$$

As before, this definition can be easily extended to compute robustness over an entire test or validation set. The ϵ can also be used to represent

the attacker's perturbation budget, transforming equation 2.1 into a definition quantifying classifier robustness against adversarial perturbations, also known as adversarial robustness.

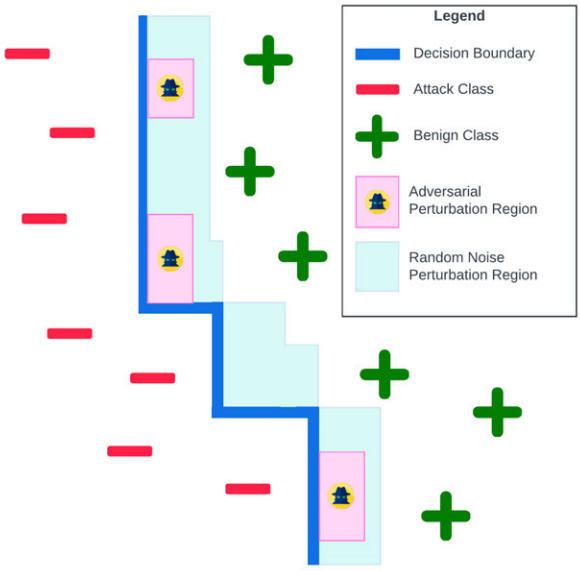
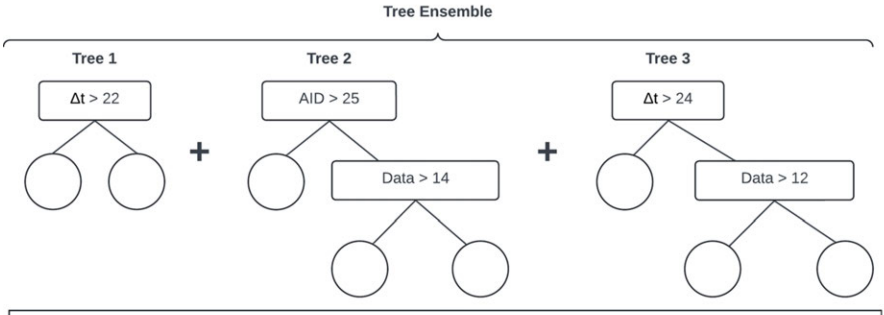


Figure 2.2: Types of Perturbations within a given budget (ϵ)

In this work, the definition in equation 2.1 is extended even further to include semantically-aware perturbations, where separate upper and lower perturbation bounds can be applied to the individual features. This extension is necessary because the thesis also addresses security applications, where perturbations become more complex. For example, in network data, modifying a feature such as a port number (e.g., changing port 22 to another value) could cause the packet to be rejected (due to blocked ports), thereby invalidating the adversarial example altogether. In this context, allowing separate ϵ values per feature enables the modelling of perturbations that are more likely to be realistic and semantically valid.

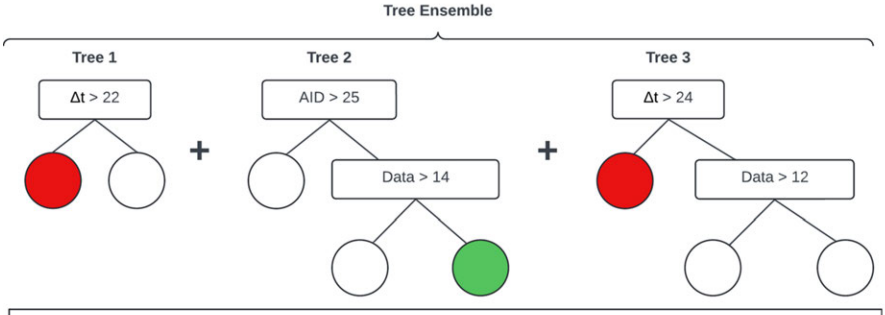
To illustrate what such perturbations look like in a practical setting, consider a running example from an in-vehicular network intrusion detection use case. Note that this example is provided for explanatory purposes only. Specifically, Figure 2.3 shows a tree ensemble model trained on CAN (Controller Area Network) data.

When the sample CAN message is passed to the tree ensemble, the model output is classified as an attack, as shown in Figure 2.4.



Sample CAN Message				
Inter-arrival time (Δt)	**Arbitration ID (AID)**	**Data**	**Ground Truth**	**Tree Ensemble Prediction**
20 µs	30	15	Attack	

Figure 2.3: Tree Ensemble Model trained on CAN Data



Sample CAN Message				
Inter-arrival time (Δt)	**Arbitration ID (AID)**	**Data**	**Ground Truth**	**Tree Ensemble Prediction**
20 µs	30	15	Attack	Majority (●, ●, ●) = ● (Attack)

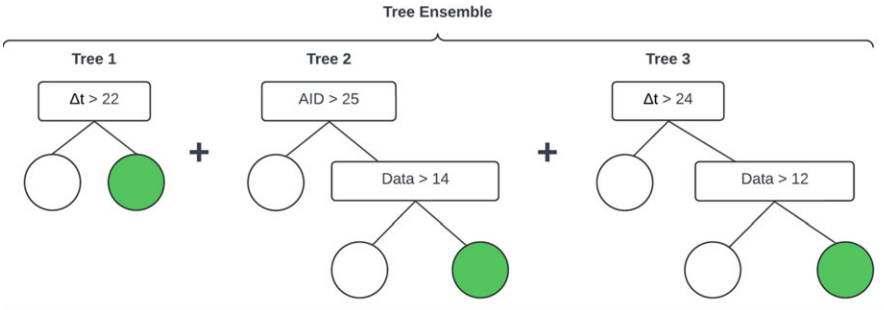
Figure 2.4: Tree Ensemble Model Prediction Mechanism

However, if one of the features, namely, the inter-arrival time, is manipulated (for instance, by simply delaying the message), the prediction from the tree ensemble completely flips, as shown in Figure 2.5.

This kind of stealthy manipulation exemplifies how adversarial examples are constructed in practical settings. Note that modifying other features - for instance, the data field - would likely cause the CAN checksum mechanism to fail and the message to be rejected, which would negate the intended purpose of the adversarial example.

2.1.7 Evasion Attacks

An evasion attack is an intentional manipulation of input samples at inference time with the goal of causing a trained classifier f_c to misclassify them,



Sample CAN Message				
Inter-arrival time (Δt)	**Arbitration ID (AID)**	**Data**	**Ground Truth**	**Tree Ensemble Prediction**
20 µs + 5 µs (delay)	30	15	Attack	Majority (●, ●, ●) = ● (Benign)

Figure 2.5: Tree Ensemble Model Prediction Flip due to Stealthy Manipulations

without altering the classifier itself. Formally, given a classifer f_c and an input $\overline{x}$ with ground truth label l, an adversary seeks to perturb the input into $\overline{x}' = \overline{x} + \overline{\delta}$ such that:

$$f_c(\overline{x}) = l \wedge f_c(\overline{x}') \neq l$$

Where $\overline{x}'$ is known as an adversarial example, and $\overline{\delta}$ is a tuple of perturbations crafted under constraints (like ϵ) that limit its magnitude to maintain plausibility and stealth. Building on the CAN example from the previous subsection, Figure 2.6 illustrates how evasion attacks, enabled by adversarial examples, can bypass detection mechanisms while retaining their underlying malicious logic. In this example, the perturbation budget (ϵ) for the `Data` and `AID` fields is set to 0, while the ϵ values for Δt lies in the range $0 \leqslant \epsilon \leqslant 5$, indicating that packets can be delayed by up to 5µs.

Sample CAN Message				
Inter-arrival time (Δt)	**Arbitration ID (AID)**	**Data**	**Ground Truth**	**Tree Ensemble Prediction**
20 µs	30	15	Attack	Majority (●, ●, ●) = ● (Attack)

Normal Sample

Sample CAN Message				
Inter-arrival time (Δt)	**Arbitration ID (AID)**	**Data**	**Ground Truth**	**Tree Ensemble Prediction**
20 µs + 5 µs (delay)	30	15	Attack	Majority (●, ●, ●) = ● (Benign)

Adversarial Example

Stealthy Manipulation — **Evasion of Detection Mechanism**

Figure 2.6: Evasion Attacks facilitated by Adversarial Examples

Evasion attacks can be either targeted, wherein the adversary aims to induce

the classifier to output a specific desired class label, or untargeted, where the objective is merely to cause a misclassification regardless of the output label. Hardening machine learning models against evasion attacks generally involves two strategies. Reactive defences focus on detecting and mitigating evasions, while proactive defences aim to design models inherently robust to evasions through techniques such as adversarial re-training. This thesis examines both approaches. In particular, it provides key insights into why certain proactive methods, such as adversarial re-training, may be unsuitable for defending tree ensembles. It also advances the state-of-the-art in reactive defences by proposing novel methods to detect evasion attacks more effectively in diverse scenarios. The upcoming sections place these contributions in the context of the wider literature and discuss them in greater detail.

2.2 Trade-Offs Between Accuracy and Robustness

Proactive defences, which aim to make machine learning models robust against adversarial examples - such as adversarial re-training, have long been regarded as a key strategy for hardening these models against evasion attacks. While extensively studied and empirically shown to enhance robustness in neural networks, as reflected by numerous publications [29, 30, 31, 32, 33, 34] in top conferences and forums - it can come at the cost of reduced performance on clean, unperturbed samples [47], potentially undermining the very purpose of the trained model. Since this thesis deals with tree ensembles, which do not possess gradients, numerous methods commonly used for neural networks are not directly applicable. Instead, the following adversarial re-training techniques have been proposed for tree ensembles.

Kantchelian et al. [21] present a technique called *adversarial boosting* that hardens a tree-based model by augmenting the training set of each boosting round with evading instances. Chaitou et al. [48] proposed a method to improve adversarial training by directing it to focus on the best evasion candidates (samples whose adversarial neighbourhood included a sample from the original dataset). Ji et al. [49] presented a method called DEEPARMOUR based on feature reconstruction, weighted voting and adversarial re-training for robust malware classification. Vitorino et al. [50] pursued an adversarial realism paradigm (in terms of perturbations) and proposed the use of adversarial training for robust tree-based models. Li et al. [51] proposed a robust learning framework called RAD based on iterative re-training with adversarial examples to boost evasion robustness of learning algorithms against evasion attacks. Zenden et al. [52] provided a method based on adversarial re-training for robust intrusion detection in in-vehicular networks. While these approaches provided empirical improvements in robustness, they did

not investigate the implications of the altered decision boundaries resulting from adversarial re-training. This aspect is particularly important for tree ensembles that could potentially be deployed in safety-critical applications, which demand formal (mathematically rigorous) arguments for operational safety. To address this knowledge gap, Paper I in this thesis leveraged formal verification, which here refers to mathematically checking whether tree ensembles satisfied specified properties. Using this approach, the paper presented novel insights into the trade-offs between accuracy and robustness in the presence of such modified decision boundaries.

2.3 Hardening Tree Ensembles against Evasion Attacks

Apart from adversarial re-training, several defences for hardening tree ensembles against evasion attacks have been proposed in the literature.

Model Enhancements: Chen et al. [53] proposed a booster-fixer training framework for enforcing security properties in classifiers. Their approach followed a counterexample-guided inductive synthesis (CEGIS) [54] paradigm in which counterexamples generated during the formal verification process were used to optimise the models until the robustness property was eventually satisfied. Vos et al. [55] proposed a method called robust re-labelling (based on finding the minimum vertex cover of bipartite graphs) that optimally changed the prediction labels of decision tree leaves to maximize adversarial robustness. Pashamokhtari et al. [56] proposed the use of patching (adding an additional check on the upper bound of the feature associated with a leaf node) for improved resilience of decision trees against adversarial volumetric attacks. Yang et al. [57] proposed a weighted random forest method based on randomization to improve robustness of tree-based models against evasion attacks. Also, a considerable body of work exists on training robust decision trees and tree ensembles against evasion attacks [58, 59, 60, 61, 62, 21].

Adversarial Example Detection: Yang et al. [63] proposed a framework called ML-LOO to detect adversarial examples through thresholding a scale estimate of feature attribution scores. Smutz et al. [64] proposed an approach based on the mutual agreement analysis of the individual votes within an ensemble of learners to detect evasion attacks. Lecuyer et al. [65] proposed a certified defence based on differential privacy to detect adversarial examples. Devos et al. [66] proposed a method called OC-SCORE to detect evasion attacks in tree ensembles by analysing the set of leaves activated by the adversarial example in the ensemble's constituent trees.

Ensembling and Other Methods: Apruzzese et al. [67] presented a method

called APPCON, which is an ensemble of application-specific detectors to mitigate evasion attacks against ML-based intrusion detectors. In another paper [68], they presented a method based on defensive distillation to harden random forest cyber detectors against adversarial attacks. Similarly, in [69], they proposed a countermeasure against evasion attacks called feature removal which reduced the attack severity, but also affected the detector performance in scenarios that were not subject to evasion attacks. Izmailov et al. [70] presented a method based on combinatorial diversification of classifiers along with fusion randomization for defending against evasion attacks. Nguyen et al. [71] presented a robust scheme for intrusion detection based on an enhanced robust decision tree and a tabular deep learning model. Mikhail et al. [72] proposed an approach called PROCMONML, which generated behavioral host-based analytics designed to enhance resilience against adversary evasion, thereby raising the evasion cost for attackers.

While the aforementioned approaches provided incremental improvements in evasion hardening for tree ensembles, they were superseded by two methods: GROOT [58] and OC-SCORE [66], which represent the state-of-the-art in proactive and reactive defences, respectively. In this context, papers II and III sought to advance these methods by improving both their evasion detection performance and speed. GROOT was fast since it only required executing a tree ensemble. However, its robust splitting criteria would sometimes overenlarge class regions depending on the proximity of training samples to the decision boundaries, leading to an increased number of misclassifications on the test samples. Moreover, it was limited to scenarios with fixed (known) attacker perturbation budgets. In contrast, OC-SCORE did not modify the tree ensemble decision boundaries and could operate under unconstrained or unknown perturbation budgets but suffered from the critical limitation of high prediction times. These issues rendered both GROOT and OC-SCORE unsuitable for use in time-sensitive critical applications where attackers could use a wide range of perturbation strategies. To address this, Papers II and III improve and extend upon these methods, making them applicable in the aforementioned scenarios while also providing additional capabilities to strengthen the evasion attack analysis.

Chapter Three

Thesis Summary

This chapter provides a summary of the research conducted in this thesis. It begins by presenting the main findings and concludes by addressing the research questions. Additionally, it reflects on future directions while highlighting the challenges that remain in advancing this line of research.

3.1 Summary of Results

While detailed results from each paper are provided in the appended publications, this section highlights the key findings that address the research questions in the context of this thesis.

Paper I: We observed a significant discrepancy between the accuracy and robustness (quantified through formal verification) of tree ensembles, specifically random forests and gradient boosting machines. In both case studies, despite doubling the size of the training set using composite geometric perturbations, perturbations that were expected to be encountered during the verification stage, the reductions in robustness were substantial. While data augmentation had little impact on model accuracy, the robustness drops were staggering: for the Handwritten Digit Recognition [73] case study, random forests experienced decreases of 37–82%, and gradient boosting machines saw reductions of 36–71%. Similarly, in the Handwritten Character Recognition [74] case study, robustness decreased by 28–51% for random forests and 35–71% for gradient boosting machines. These findings indicated that data augmentation alone was insufficient for enhancing the robustness of tree ensembles, highlighting the need for alternative approaches.

Paper II: We presented our specialised defence, called *Iceman* and compared it with two existing approaches, OC-SCORE [66] and GROOT [58]. Our results demonstrated that *Iceman* effectively defended tree ensembles against evasion attacks while maintaining a balanced classification performance, measured using the Matthews Correlation Coefficient (MCC) [75]. Across four case studies related to safety and security, *Iceman* improved MCC scores by 0.20–0.91 compared to GROOT and by 0.02–0.79 compared to OC-SCORE. In terms of prediction time, while GROOT was the fastest method, requiring only the execution of a tree ensemble, *Iceman* was 5–115× faster than OC-

SCORE. Additionally, *Iceman* offered capabilities beyond evasion detection, including alert filtering and prioritisation in intrusion detection systems, achieving an accuracy exceeding 98%. While these results advanced the state-of-the-art, *Iceman* was limited to scenarios with a fixed attacker perturbation budget. This limitation highlighted the need for alternate approaches that could complement *Iceman* in scenarios where the attacker's perturbation budget was unknown.

Paper III: To address the need for an approach to complement *Iceman* (from Paper II) in scenarios where the attacker's perturbation budget was unknown, we improved upon the OC-SCORE method. In this context, we presented a generalised defence called *Maverick* that maintained an identical evasion detection performance to OC-SCORE while achieving speedups of 85–563x, making it suitable for time-sensitive, real-time applications such as automotive intrusion detection systems. Moreover, *Maverick* could also contribute to the evasion attack diagnostic process, achieving an accuracy exceeding 93%.

3.2 Conclusion and Future Works

As outlined previously, this thesis seeks to critically challenge the validity of model re-training for ensuring robustness in tree ensembles against noisy inputs or adversarial examples, while proposing novel defences against evasion attacks that advance the state-of-the-art. In this context, the **answers to the research questions** are as follows:

Q1. How does adversarial re-training impact the accuracy and robustness of tree ensembles?

Model re-training via data augmentation provided modest improvements in accuracy. However, robustness decreased in most experiments, suggesting that **re-training may inadvertently make the construction of adversarial examples easier for attackers**. This occurs because the re-training process creates jagged (overly complex) decision boundaries, so augmented samples are classified correctly, but the majority of original samples become vulnerable to misclassifications through perturbations. This explains the substantial drops in robustness and shows that model re-training does not achieve its intended defensive purpose.

An instance of this phenomenon is illustrated in Figure 3.1. Consider a tree ensemble model with a simple decision boundary (before re-training) and two sets of inputs. Initially, the model achieves 100% accuracy. However, when model robustness is quantified for

the attack class by defining margins around the attack samples, the measured robustness is about 70%. Including samples (from both attack and benign classes) that violate the robustness property in the augmentation process makes the tree ensemble's decision boundaries more jagged (after re-training), as the model adjusts to fit these new points. While the accuracy still remains at 100%, the new decision boundaries now encroach into the robustness margins of the original attack samples. Consequently, the model robustness for the attack class decreases from 70% to 30%. This reduction in the robustness margins lowers the perturbation magnitude (required to cross the decision boundaries), thereby decreasing the practical effort needed to craft adversarial examples.

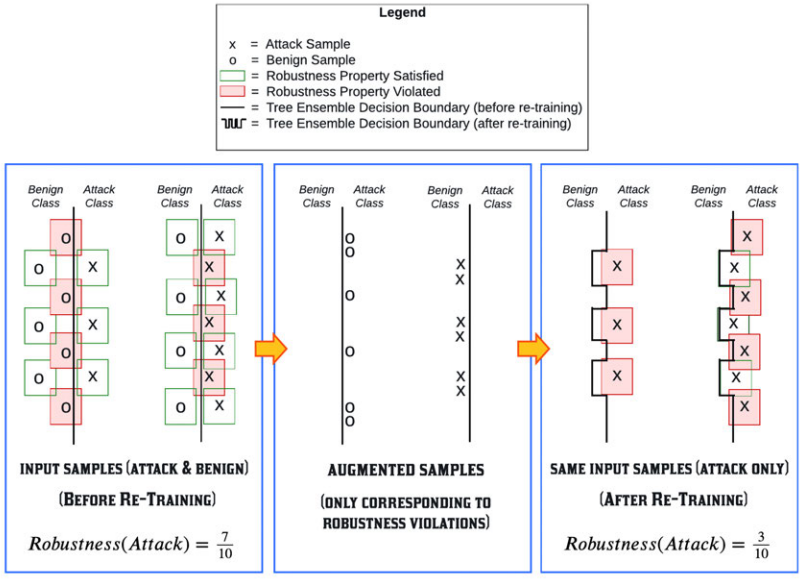


Figure 3.1: Robustness Impacts due to Model Re-training

Q2a. How can tree ensemble defences be designed to remain effective regardless of any specific method used to generate evasion attacks?

Tree ensemble evasion defences can be designed to remain independent of specific attack generation strategies through two approaches:

– **Identification of counterexample regions:** This approach involves determining regions in the model's input space that are vulnerable to evasion attacks, which requires prior knowledge of the attacker's budget. Once the budget is known, formal verification tools such as VoTE, based on abstract interpretation, which operate on regions rather than individual points (like SMT/MILP), can

be used to enumerate these counterexample regions that violate the robustness property. By exhaustively enumerating all such regions, the resulting defence becomes independent of specific evasion strategies. This is because any evasion attempt (crafted using the same budget as the one used to enumerate the counterexample regions) capable of flipping a classification will likely fall within, or close to, these counterexample regions.

– **Analysis of reached leaf nodes:** Adversarial examples, irrespective of their generation method, tend to traverse unusual paths in tree ensembles compared to correctly classified normal samples. This occurs because normal samples typically reach leaves that are strongly predictive of a single class, whereas small perturbations in the input can redirect them to rarely visited leaves. By quantifying the discrepancy between the leaf nodes reached by normal samples and those reached by adversarial examples, evasion attempts can be detected. This approach requires no knowledge of the attacker's perturbation budget.

Q2b. How can evasion defences be designed to maintain effectiveness against attackers operating under different perturbation budgets?

Evasion defences for tree ensembles can maintain effectiveness across different attacker budgets through **a combination of specialised and generalised defensive strategies**. Specialised defences (like *Iceman* from Paper II), tailored to a specific perturbation budget, tend to perform well under those conditions but degrade significantly when the attacker exceeds that budget. Generalised defences (like *Maverick* from Paper III), on the other hand, make no assumptions about the attacker's budget and therefore offer broader protection across multiple perturbation budgets, though typically with lower effectiveness for any single budget. A hybrid approach leverages both specialised defences to deliver strong protection when the attacker's budget is known or can be estimated with a fair degree of certainty, while generalised defences safeguard in scenarios where the attacker's budget is unknown, thereby compensating for each other's limitations.

Q2c. Can evasion defences satisfy the timing requirements of real-time applications?

Evasion defences for tree ensembles can satisfy the timing requirements of real-time applications like CAN-based automotive intrusion detection systems (as shown in Papers II and III) through a combination of **simple distance measurements** (between a counterexample

region and an incoming vector) and **ML-based inference**. These techniques effectively replace costly, time-intensive prediction mechanisms, enabling real-time evasion detection without compromising detection accuracy.

Regarding future works, a fundamental question remains: why are tree ensembles even vulnerable to adversarial examples? The answer likely lies in the way these models are constructed, where feature values are used to iteratively split the input space until the decision trees in the ensemble are formed. This process does not incorporate context-specific information about the intricate and complex relationships among the input features. By modelling such contextual information, it may be possible to augment tree ensemble models with the capability to understand and reason about the input before making predictions. This could also enable models to abstain from making uncertain predictions. Insights from cognitive science, particularly studies on how human brains extract invariant features while ignoring noise [76, 77], could guide the development of such capabilities. A simple example would be augmenting models with the reasoning behind how humans can correctly recognise perturbed digits, whereas tree ensembles often make very confident but incorrect predictions under similar perturbations. This could potentially advance the safe and secure, real-world deployment of tree ensembles in safety-critical and security-critical applications.

Bibliography

[1] Bradley Erickson, Panagiotis Korfiatis, Zeynettin Akkus, and Timothy Kline. "Machine learning for medical imaging". In: *Radiographics* 37.2 (2017), pp. 505–515.

[2] Robert Ledley and Lee Lusted. "Reasoning foundations of medical diagnosis: symbolic logic, probability, and value theory aid our understanding of how physicians reason". In: *Science* 130.3366 (1959), pp. 9–21.

[3] Longbing Cao. "AI in finance: challenges, techniques, and opportunities". In: *ACM Computing Surveys (CSUR)* 55.3 (2022), pp. 1–38.

[4] Yoesoep Edhie Rachmad. *Artificial Intelligence, Autonomous Weapons, and the Future of War*. The United Nations and The Education Training Centre, 2011.

[5] Dean Pomerleau. "Alvinn: An autonomous land vehicle in a neural network". In: *Advances in Neural Information Processing Systems* 1 (1988).

[6] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané. "Concrete problems in AI safety". In: *arXiv preprint arXiv:1606.06565* (2016).

[7] Yoshua Bengio, Sören Mindermann, Daniel Privitera, Tamay Besiroglu, Rishi Bommasani, Stephen Casper, Yejin Choi, Philip Fox, Ben Garfinkel, Danielle Goldfarb, et al. "International AI Safety Report". In: *arXiv preprint arXiv:2501.17805* (2025).

[8] National Transportation Safety Board. "Collision between vehicle controlled by developmental automated driving system and pedestrian". In: *Transportation Safety Board, Washington, DC, USA* (2019).

[9] Philip Koopman. "Lessons from the cruise robotaxi pedestrian dragging mishap". In: *IEEE Reliability Magazine* 1.3 (2024), pp. 54–61.

[10] Ian Goodfellow, Jonathon Shlens, and Christian Szegedy. "Explaining and harnessing adversarial examples". In: *arXiv preprint arXiv:1412.6572* (2014).

[11] Alexander Robey, Zachary Ravichandran, Vijay Kumar, Hamed Hassani, and George J Pappas. "Jailbreaking llm-controlled robots". In: *arXiv preprint arXiv:2410.13691* (2024).

[12] Saurabh Jha, Shengkun Cui, Subho Banerjee, James Cyriac, Timothy Tsai, Zbigniew Kalbarczyk, and Ravishankar K Iyer. "Ml-driven malware that targets av safety". In: *2020 50th annual IEEE/IFIP international conference on dependable systems and networks (DSN)*. IEEE. 2020, pp. 113–124.

[13] Battista Biggio, Igino Corona, Davide Maiorca, Blaine Nelson, Nedim Šrndić, Pavel Laskov, Giorgio Giacinto, and Fabio Roli. "Evasion attacks against machine learning at test time". In: *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2013, Prague, Czech Republic, September 23-27, 2013, Proceedings, Part III 13*. Springer. 2013.

[14] Kevin Eykholt, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. "Robust physical-world attacks on deep learning visual classification". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 1625–1634.

[15] Rosamund Powell, Sam Stockwell, Nalanda Sharadjaya, and Hugh Boyes. *Towards Secure AI: How far can international standards take us?* 2024. URL: https://cetas.turing.ac.uk/publications/towards-secure-ai.

[16] Yi Liu, Gelei Deng, Yuekang Li, Kailong Wang, Zihao Wang, Xiaofeng Wang, Tianwei Zhang, Yepang Liu, Haoyu Wang, Yan Zheng, et al. "Prompt injection attack against llm-integrated applications". In: *arXiv preprint arXiv:2306.05499* (2023).

[17] PromptArmor. *Data exfiltration from Slack AI via indirect prompt injection*. 2024. URL: https://promptarmor.substack.com/p/data-exfiltration-from-slack-ai-via.

[18] Carson Ezell, Xavier Roberts-Gaal, and Alan Chan. "Incident Analysis for AI Agents". In: *arXiv preprint arXiv:2508.14231* (2025).

[19] Ben Nassi, Stav Cohen, and Or Yair. "Invitation Is All You Need! Promptware Attacks Against LLM-Powered Assistants in Production Are Practical and Dangerous". In: *arXiv preprint arXiv:2508.12175* (2025).

[20] Matthias Wagner, Markus Borg, and Per Runeson. "Navigating the Upcoming European Union AI Act". In: *IEEE Software* 41.1 (2023).

[21] Alex Kantchelian, J Doug Tygar, and Anthony Joseph. "Evasion and hardening of tree ensemble classifiers". In: *International Conference on Machine Learning*. PMLR. 2016.

[22] Ke He, Dan Dongseong Kim, and Muhammad Rizwan Asghar. "Adversarial machine learning for network intrusion detection systems: A comprehensive survey". In: *IEEE Communications Surveys & Tutorials* 25.1 (2023).

[23] Léo Grinsztajn, Edouard Oyallon, and Gaël Varoquaux. "Why do tree-based models still outperform deep learning on typical tabular data?" In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 507–520.

[24] Ravid Shwartz-Ziv and Amitai Armon. "Tabular data: Deep learning is not all you need". In: *Information Fusion* 81 (2022), pp. 84–90.

[25] Tommaso Zoppi, Andrea Ceccarelli, Tommaso Puccetti, and Andrea Bondavalli. "Which algorithm can detect unknown attacks? Comparison of supervised, unsupervised and meta-learning algorithms for intrusion detection". In: *Computers & Security* 127 (2023), p. 103107.

[26] Wajih Ul Hassan, Shengjian Guo, Ding Li, Zhengzhang Chen, Kangkook Jee, Zhichun Li, and Adam Bates. "Nodoze: Combatting threat alert fatigue with automated provenance triage". In: *network and distributed systems security symposium*. 2019.

[27] Thomas Dietterich. "Steps toward robust artificial intelligence". In: *AI Magazine, AAAI* 38.3 (2017), pp. 3–24.

[28] Chih-Hong Cheng, Georg Nührenberg, and Harald Ruess. "Maximum resilience of artificial neural networks". In: *International symposium on automated technology for verification and analysis*. Springer. 2017, pp. 251–268.

[29] Maksym Andriushchenko and Matthias Hein. "Provably robust boosted decision stumps and trees against adversarial attacks". In: *Advances in Neural Information Processing Systems* 32 (2019).

[30] Stefano Calzavara, Claudio Lucchese, and Gabriele Tolomei. "Adversarial training of gradient-boosted decision trees". In: *Proceedings of the 28th ACM international conference on information and knowledge management*. 2019, pp. 2429–2432.

[31] Giovanni Apruzzese, Mauro Andreolini, Mirco Marchetti, Andrea Venturi, and Michele Colajanni. "Deep reinforcement adversarial learning against botnet evasion attacks". In: *IEEE Transactions on Network and Service Management* 17.4 (2020), pp. 1975–1987.

[32] Tong Zhao, Yozen Liu, Leonardo Neves, Oliver Woodford, Meng Jiang, and Neil Shah. "Data augmentation for graph neural networks". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 12. 2021, pp. 11015–11023.

[33] Alex Hernández-García and Peter König. “Further advantages of data augmentation on convolutional neural networks”. In: *International Conference on Artificial Neural Networks*. Springer. 2018, pp. 95–103.

[34] Rasool Fakoor, Jonas Mueller, Nick Erickson, Pratik Chaudhari, and Alexander Smola. “Fast, accurate, and simple models for tabular data via augmented distillation”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 8671–8681.

[35] Valency Oscar Colaco and Simin Nadjm-Tehrani. “Formal Verification of Tree Ensembles against Real-World Composite Geometric Perturbations”. In: *Workshop on Artificial Intelligence Safety 2023 (SafeAI 2023) co-located with the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI)*. CEUR-WS. 2023. URN: urn:nbn:de:0074-3381-7. URL: https://ceur-ws.org/Vol-3381/38.pdf.

[36] Valency Oscar Colaco and Simin Nadjm-Tehrani. “Fast Evasion Detection & Alert Management in Tree-Ensemble-Based Intrusion Detection Systems”. In: *2024 IEEE 36th International Conference on Tools with Artificial Intelligence (ICTAI)*. Los Alamitos, CA, USA: IEEE Computer Society, Oct. 2024, pp. 404–412. DOI: 10.1109/ICTAI62512.2024.00065.

[37] Valency Oscar Colaco and Simin Nadjm-Tehrani. “Real-Time Evasion Detection in Tree Ensemble Automotive Intrusion Detection Systems”. In: *2025 IEEE 16th Vehicular Networking Conference (VNC)*. 2025, pp. 1–8. DOI: 10.1109/VNC64509.2025.11054177.

[38] Leo Breiman, Jerome Friedman, Richard A Olshen, and Charles J Stone. *Classification and regression trees*. Routledge, 2017.

[39] John Törnblom and Simin Nadjm-Tehrani. “An abstraction-refinement approach to formal verification of tree ensembles”. In: *Computer Safety, Reliability, and Security: SAFECOMP Workshops, WAISE, Proceedings 38*. Springer. 2019.

[40] Ozan Irsoy, Olcay Taner Yıldız, and Ethem Alpaydın. “Soft decision trees”. In: *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*. IEEE. 2012, pp. 1819–1822.

[41] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. “Scikit-learn: Machine learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[42] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. “CatBoost: unbiased boosting with categorical features”. In: *Advances in Neural Information Processing Systems* 31 (2018).

[43] Tianqi Chen and Carlos Guestrin. “Xgboost: A scalable tree boosting system”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016.

[44] Leo Breiman. “Random forests”. In: *Machine learning* 45 (2001).

[45] Leo Breiman. “Bagging predictors”. In: *Machine Learning* 24.2 (1996), pp. 123–140.

[46] Yoav Freund and Robert Schapire. “A decision-theoretic generalization of on-line learning and an application to boosting”. In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139.

[47] Meysam Sadeghi and Erik Larsson. “Physical adversarial attacks against end-to-end autoencoder communication systems”. In: *IEEE Communications Letters* 23.5 (2019).

[48] Hassan Chaitou, Thomas Robert, Jean Leneutre, and Laurent Pautet. “Intrinsic weaknesses of IDSs to malicious adversarial attacks and their mitigation”. In: *International Conference on Smart Business Technologies*. Springer. 2022.

[49] Yuede Ji, Benjamin Bowman, and H Howie Huang. “Securing malware cognitive systems against adversarial attacks”. In: *2019 IEEE international conference on cognitive computing (ICCC)*. IEEE. 2019.

[50] João Vitorino, Isabel Praça, and Eva Maia. “Towards adversarial realism and robust learning for IOT intrusion detection and classification”. In: *Annals of Telecommunications* 78.7 (2023).

[51] Bo Li and Yevgeniy Vorobeychik. “Evasion-robust classification on binary domains”. In: *ACM Transactions on Knowledge Discovery from Data (TKDD)* 12.4 (2018).

[52] Ivo Zenden, Han Wang, Alfonso Iacovazzi, Arash Vahidi, Rolf Blom, and Shahid Raza. “On the Resilience of Machine Learning-Based IDS for Automotive Networks”. In: *IEEE Vehicular Networking Conference (VNC)*. IEEE. 2023.

[53] Yizheng Chen, Shiqi Wang, Yue Qin, Xiaojing Liao, Suman Jana, and David Wagner. “Learning security classifiers with verified global robustness properties”. In: *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*. 2021.

[54] Armando Solar-Lezama, Liviu Tancau, Rastislav Bodik, Sanjit Seshia, and Vijay Saraswat. “Combinatorial sketching for finite programs”. In: *Proceedings of the 12th international conference on Architectural support for programming languages and operating systems*. 2006.

[55] Daniël Vos and Sicco Verwer. “Adversarially robust decision tree relabeling”. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer. 2022.

[56] Arman Pashamokhtari, Gustavo Batista, and Hassan Habibi Gharakheili. "Adiotack: Quantifying and refining resilience of decision tree ensemble inference models against adversarial volumetric attacks on IoT networks". In: *Computers & Security* 120 (2022).

[57] Fan Yang, Zhiyuan Chen, and Aryya Gangopadhyay. "Using randomness to improve robustness of tree-based models against evasion attacks". In: *Proceedings of the ACM International Workshop on Security and Privacy Analytics*. 2019.

[58] Daniël Vos and Sicco Verwer. "Efficient training of robust decision trees against adversarial examples". In: *International Conference on Machine Learning*. PMLR. 2021.

[59] Stefano Calzavara, Lorenzo Cazzaro, Giulio Ermanno Pibiri, and Nicola Prezza. "Verifiable learning for robust tree ensembles". In: *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*. 2023.

[60] Yizheng Chen, Shiqi Wang, Weifan Jiang, Asaf Cidon, and Suman Jana. "{Cost-Aware} robust tree ensembles for security applications". In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021.

[61] Daniël Vos and Sicco Verwer. "Robust optimal classification trees against adversarial examples". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. 8. 2022.

[62] Stefano Calzavara, Claudio Lucchese, Gabriele Tolomei, Seyum Assefa Abebe, and Salvatore Orlando. "Treant: training evasion-aware decision trees". In: *Data Mining and Knowledge Discovery* 34.5 (2020).

[63] Puyudi Yang, Jianbo Chen, Cho-Jui Hsieh, Jane-Ling Wang, and Michael Jordan. "ML-LOO: Detecting adversarial examples with feature attribution". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 04. 2020.

[64] Charles Smutz and Angelos Stavrou. "When a Tree Falls: Using Diversity in Ensemble Classifiers to Identify Evasion in Malware Detectors." In: *NDSS*. 2016.

[65] Mathias Lecuyer, Vaggelis Atlidakis, Roxana Geambasu, Daniel Hsu, and Suman Jana. "Certified robustness to adversarial examples with differential privacy". In: *2019 IEEE symposium on security and privacy (SP)*. IEEE. 2019, pp. 656–672.

[66] Laurens Devos, Lorenzo Perini, Wannes Meert, and Jesse Davis. "Detecting Evasion Attacks in Deployed Tree Ensembles". In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer. 2023.

[67] Giovanni Apruzzese, Mauro Andreolini, Mirco Marchetti, Vincenzo Giuseppe Colacino, and Giacomo Russo. "AppCon: Mitigating evasion attacks to ML cyber detectors". In: *Symmetry* 12.4 (2020), p. 653.

[68] Giovanni Apruzzese, Mauro Andreolini, Michele Colajanni, and Mirco Marchetti. "Hardening random forest cyber detectors against adversarial attacks". In: *IEEE Transactions on Emerging Topics in Computational Intelligence* 4.4 (2020).

[69] Giovanni Apruzzese, Michele Colajanni, Luca Ferretti, and Mirco Marchetti. "Addressing adversarial attacks against security systems based on machine learning". In: *2019 11th international conference on cyber conflict (CyCon)*. Vol. 900. IEEE. 2019, pp. 1–18.

[70] Rauf Izmailov, Peter Lin, Sridhar Venkatesan, and Shridatt Sugrim. "Combinatorial Boosting of Ensembles of Diversified Classifiers for Defense Against Evasion Attacks". In: *MILCOM 2021-2021 IEEE Military Communications Conference (MILCOM)*. IEEE. 2021.

[71] Dat-Thinh Nguyen and Kim-Hung Le. "The robust scheme for intrusion detection system in internet of things". In: *Internet of Things* 24 (2023).

[72] Joseph Mikhail, Jamie Williams, and George Roelke. "procmonML: Generating evasion resilient host-based behavioral analytics from tree ensembles". In: *Computers & Security* 98 (2020).

[73] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11 (2002), pp. 2278–2324.

[74] Gregory Cohen, Saeed Afshar, Jonathan Tapson, and Andre Van Schaik. "EMNIST: Extending MNIST to handwritten letters". In: *2017 international joint conference on neural networks (IJCNN)*. IEEE. 2017, pp. 2921–2926.

[75] Davide Chicco and Giuseppe Jurman. "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification". In: *BioData Mining* 16.1 (2023).

[76] Nancy Kanwisher, Josh McDermott, and Marvin Chun. "The fusiform face area: a module in human extrastriate cortex specialized for face perception". In: *Journal of Neuroscience* 17.11 (1997), pp. 4302–4311.

[77] Mortimer Mishkin and Leslie Ungerleider. "Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys". In: *Behavioural Brain Research* 6.1 (1982), pp. 57–77.

[78] Battista Biggio, Blaine Nelson, and Pavel Laskov. "Poisoning attacks against support vector machines". In: *arXiv preprint arXiv:1206.6389* (2012).

Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

https://doi.org/10.3384/9789181183269

Department of Computer and Information Science
Linköpings universitet

Licentiate Theses

Linköpings Studies in Science and Technology
Faculty of Arts and Sciences

No 17 **Vojin Plavsic:** Interleaved Processing of Non-Numerical Data Stored on a Cyclic Memory. (Available at: FOA, Box 1165, S-581 11 Linköping, Sweden. FOA Report B30062E)

No 28 **Arne Jönsson, Mikael Patel:** An Interactive Flowcharting Technique for Communicating and Realizing Algorithms, 1984.

No 29 **Johnny Eckerland:** Retargeting of an Incremental Code Generator, 1984.

No 48 **Henrik Nordin:** On the Use of Typical Cases for Knowledge-Based Consultation and Teaching, 1985.

No 52 **Zebo Peng:** Steps Towards the Formalization of Designing VLSI Systems, 1985.

No 60 **Johan Fagerström:** Simulation and Evaluation of Architecture based on Asynchronous Processes, 1985.

No 71 **Jalal Maleki:** ICONStraint, A Dependency Directed Constraint Maintenance System, 1987.

No 72 **Tony Larsson:** On the Specification and Verification of VLSI Systems, 1986.

No 73 **Ola Strömfors:** A Structure Editor for Documents and Programs, 1986.

No 74 **Christos Levcopoulos:** New Results about the Approximation Behavior of the Greedy Triangulation, 1986.

No 104 **Shamsul I. Chowdhury: S**tatistical Expert Systems - a Special Application Area for Knowledge-Based Computer Methodology, 1987.

No 108 **Rober Bilos:** Incremental Scanning and Token-Based Editing, 1987.

No 111 **Hans Block:** SPORT-SORT Sorting Algorithms and Sport Tournaments, 1987.

No 113 **Ralph Rönnquist:** Network and Lattice Based Approaches to the Representation of Knowledge, 1987.

No 118 **Mariam Kamkar, Nahid Shahmehri:** Affect-Chaining in Program Flow Analysis Applied to Queries of Programs, 1987.

No 126 **Dan Strömberg:** Transfer and Distribution of Application Programs, 1987.

No 127 **Kristian Sandahl:** Case Studies in Knowledge Acquisition, Migration and User Acceptance of Expert Systems, 1987.

No 139 **Christer Bäckström:** Reasoning about Interdependent Actions, 1988.

No 140 **Mats Wirén:** On Control Strategies and Incrementality in Unification-Based Chart Parsing, 1988.

No 146 **Johan Hultman:** A Software System for Defining and Controlling Actions in a Mechanical System, 1988.

No 150 **Tim Hansen:** Diagnosing Faults using Knowledge about Malfunctioning Behavior, 1988.

No 165 **Jonas Löwgren:** Supporting Design and Management of Expert System User Interfaces, 1989.

No 166 **Ola Petersson:** On Adaptive Sorting in Sequential and Parallel Models, 1989.

No 174 **Yngve Larsson:** Dynamic Configuration in a Distributed Environment, 1989.

No 177 **Peter Åberg:** Design of a Multiple View Presentation and Interaction Manager, 1989.

No 181 **Henrik Eriksson:** A Study in Domain-Oriented Tool Support for Knowledge Acquisition, 1989.

No 184 **Ivan Rankin:** The Deep Generation of Text in Expert Critiquing Systems, 1989.

No 187 **Simin Nadjm-Tehrani:** Contributions to the Declarative Approach to Debugging Prolog Programs, 1989.

No 189 **Magnus Merkel:** Temporal Information in Natural Language, 1989.

No 196 **Ulf Nilsson:** A Systematic Approach to Abstract Interpretation of Logic Programs, 1989.

No 197 **Staffan Bonnier:** Horn Clause Logic with External Procedures: Towards a Theoretical Framework, 1989.

No 203 **Christer Hansson:** A Prototype System for Logical Reasoning about Time and Action, 1990.

No 212 **Björn Fjellborg:** An Approach to Extraction of Pipeline Structures for VLSI High-Level Synthesis, 1990.

No 230 **Patrick Doherty:** A Three-Valued Approach to Non-Monotonic Reasoning, 1990.

No 237 **Tomas Sokolnicki:** Coaching Partial Plans: An Approach to Knowledge-Based Tutoring, 1990.

No 250 **Lars Strömberg:** Postmortem Debugging of Distributed Systems, 1990.

No 253 **Torbjörn Näslund:** SLDFA-Resolution - Computing Answers for Negative Queries, 1990.

No 260 **Peter D. Holmes:** Using Connectivity Graphs to Support Map-Related Reasoning, 1991.

No 283 **Olof Johansson:** Improving Implementation of Graphical User Interfaces for Object-Oriented Knowledge- Bases, 1991.

No 298 **Rolf G Larsson:** Aktivitetsbaserad kalkylering i ett nytt ekonomisystem, 1991.

No 318 **Lena Strömbäck:** Studies in Extended Unification-Based Formalism for Linguistic Description: An Algorithm for Feature Structures with Disjunction and a Proposal for Flexible Systems, 1992.

No 319 **Mikael Pettersson:** DML-A Language and System for the Generation of Efficient Compilers from Denotational Specification, 1992.

No 326 **Andreas Kågedal:** Logic Programming with External Procedures: an Implementation, 1992.

No 328 **Patrick Lambrix:** Aspects of Version Management of Composite Objects, 1992.

No 333 **Xinli Gu:** Testability Analysis and Improvement in High-Level Synthesis Systems, 1992.

No 335 **Torbjörn Näslund:** On the Role of Evaluations in Iterative Development of Managerial Support Systems, 1992.

No 348 **Ulf Cederling:** Industrial Software Development - a Case Study, 1992.

No 352 **Magnus Morin:** Predictable Cyclic Computations in Autonomous Systems: A Computational Model and Implementation, 1992.

No 371 **Mehran Noghabai:** Evaluation of Strategic Investments in Information Technology, 1993.
No 378 **Mats Larsson:** A Transformational Approach to Formal Digital System Design, 1993.
No 380 **Johan Ringström:** Compiler Generation for Parallel Languages from Denotational Specifications, 1993.
No 381 **Michael Jansson:** Propagation of Change in an Intelligent Information System, 1993.
No 383 **Jonni Harrius:** An Architecture and a Knowledge Representation Model for Expert Critiquing Systems, 1993.
No 386 **Per Österling:** Symbolic Modelling of the Dynamic Environments of Autonomous Agents, 1993.
No 398 **Johan Boye:** Dependency-based Groudness Analysis of Functional Logic Programs, 1993.
No 402 **Lars Degerstedt:** Tabulated Resolution for Well Founded Semantics, 1993.
No 406 **Anna Moberg:** Satellitkontor - en studie av kommunikationsmönster vid arbete på distans, 1993.
No 414 **Peter Carlsson:** Separation av företagsledning och finansiering - fallstudier av företagsledarutköp ur ett agent-teoretiskt perspektiv, 1994.
No 417 **Camilla Sjöström:** Revision och lagreglering - ett historiskt perspektiv, 1994.
No 436 **Cecilia Sjöberg:** Voices in Design: Argumentation in Participatory Development, 1994.
No 437 **Lars Viklund:** Contributions to a High-level Programming Environment for a Scientific Computing, 1994.
No 440 **Peter Loborg:** Error Recovery Support in Manufacturing Control Systems, 1994.
FHS 3/94 **Owen Eriksson:** Informationssystem med verksamhetskvalitet - utvärdering baserat på ett verksamhetsinriktat och samskapande perspektiv, 1994.
FHS 4/94 **Karin Pettersson:** Informationssystemstrukturering, ansvarsfördelning och användarinflytande - En komparativ studie med utgångspunkt i två informationssystemstrategier, 1994.
No 441 **Lars Poignant:** Informationsteknologi och företagsetablering - Effekter på produktivitet och region, 1994.
No 446 **Gustav Fahl:** Object Views of Relational Data in Multidatabase Systems, 1994.
No 450 **Henrik Nilsson:** A Declarative Approach to Debugging for Lazy Functional Languages, 1994.
No 451 **Jonas Lind:** Creditor - Firm Relations: an Interdisciplinary Analysis, 1994.
No 452 **Martin Sköld:** Active Rules based on Object Relational Queries - Efficient Change Monitoring Techniques, 1994.
No 455 **Pär Carlshamre:** A Collaborative Approach to Usability Engineering: Technical Communicators and System Developers in Usability-Oriented Systems Development, 1994.
FHS 5/94 **Stefan Cronholm:** Varför CASE-verktyg i systemutveckling? - En motiv- och konsekvensstudie avseende arbetssätt och arbetsformer, 1994.
No 462 **Mikael Lindvall:** A Study of Traceability in Object-Oriented Systems Development, 1994.
No 463 **Fredrik Nilsson:** Strategi och ekonomisk styrning - En studie av Sandviks förvärv av Bahco Verktyg, 1994.
No 464 **Hans Olsén:** Collage Induction: Proving Properties of Logic Programs by Program Synthesis, 1994.
No 469 **Lars Karlsson:** Specification and Synthesis of Plans Using the Features and Fluents Framework, 1995.
No 473 **Ulf Söderman:** On Conceptual Modelling of Mode Switching Systems, 1995.
No 475 **Choong-ho Yi:** Reasoning about Concurrent Actions in the Trajectory Semantics, 1995.
No 476 **Bo Lagerström:** Successiv resultatavräkning av pågående arbeten. - Fallstudier i tre byggföretag, 1995.
No 478 **Peter Jonsson:** Complexity of State-Variable Planning under Structural Restrictions, 1995.
FHS 7/95 **Anders Avdic:** Arbetsintegrerad systemutveckling med kalkylprogram, 1995.
No 482 **Eva L Ragnemalm:** Towards Student Modelling through Collaborative Dialogue with a Learning Companion, 1995.
No 488 **Eva Toller:** Contributions to Parallel Multiparadigm Languages: Combining Object-Oriented and Rule-Based Programming, 1995.
No 489 **Erik Stoy:** A Petri Net Based Unified Representation for Hardware/Software Co-Design, 1995.
No 497 **Johan Herber:** Environment Support for Building Structured Mathematical Models, 1995.
No 498 **Stefan Svenberg:** Structure-Driven Derivation of Inter-Lingual Functor-Argument Trees for Multi-Lingual Generation, 1995.
No 503 **Hee-Cheol Kim:** Prediction and Postdiction under Uncertainty, 1995.
FHS 8/95 **Dan Fristedt:** Metoder i användning - mot förbättring av systemutveckling genom situationell metodkunskap och metodanalys, 1995.
FHS 9/95 **Malin Bergvall:** Systemförvaltning i praktiken - en kvalitativ studie avseende centrala begrepp, aktiviteter och ansvarsroller, 1995.
No 513 **Joachim Karlsson:** Towards a Strategy for Software Requirements Selection, 1995.
No 517 **Jakob Axelsson:** Schedulability-Driven Partitioning of Heterogeneous Real-Time Systems, 1995.
No 518 **Göran Forslund:** Toward Cooperative Advice-Giving Systems: The Expert Systems Experience, 1995.
No 522 **Jörgen Andersson:** Bilder av småföretagares ekonomistyrning, 1995.
No 538 **Staffan Flodin:** Efficient Management of Object-Oriented Queries with Late Binding, 1996.
No 545 **Vadim Engelson:** An Approach to Automatic Construction of Graphical User Interfaces for Applications in Scientific Computing, 1996.
No 546 **Magnus Werner :** Multidatabase Integration using Polymorphic Queries and Views, 1996.
FiF-a 1/96 **Mikael Lind:** Affärsprocessinriktad förändringsanalys - utveckling och tillämpning av synsätt och metod, 1996.
No 549 **Jonas Hallberg:** High-Level Synthesis under Local Timing Constraints, 1996.
No 550 **Kristina Larsen:** Förutsättningar och begränsningar för arbete på distans - erfarenheter från fyra svenska företag. 1996.

No 557 **Mikael Johansson:** Quality Functions for Requirements Engineering Methods, 1996.
No 558 **Patrik Nordling:** The Simulation of Rolling Bearing Dynamics on Parallel Computers, 1996.
No 561 **Anders Ekman:** Exploration of Polygonal Environments, 1996.
No 563 **Niclas Andersson:** Compilation of Mathematical Models to Parallel Code, 1996.
No 567 **Johan Jenvald:** Simulation and Data Collection in Battle Training, 1996.
No 575 **Niclas Ohlsson:** Software Quality Engineering by Early Identification of Fault-Prone Modules, 1996.
No 576 **Mikael Ericsson:** Commenting Systems as Design Support—A Wizard-of-Oz Study, 1996.
No 587 **Jörgen Lindström:** Chefers användning av kommunikationsteknik, 1996.
No 589 **Esa Falkenroth:** Data Management in Control Applications - A Proposal Based on Active Database Systems, 1996.
No 591 **Niclas Wahllöf:** A Default Extension to Description Logics and its Applications, 1996.
No 595 **Annika Larsson:** Ekonomisk Styrning och Organisatorisk Passion - ett interaktivt perspektiv, 1997.
No 597 **Ling Lin:** A Value-based Indexing Technique for Time Sequences, 1997.
No 598 **Rego Granlund:** C^3Fire - A Microworld Supporting Emergency Management Training, 1997.
No 599 **Peter Ingels:** A Robust Text Processing Technique Applied to Lexical Error Recovery, 1997.
No 607 **Per-Arne Persson:** Toward a Grounded Theory for Support of Command and Control in Military Coalitions, 1997.
No 609 **Jonas S Karlsson:** A Scalable Data Structure for a Parallel Data Server, 1997.
FiF-a 4 **Carita Åbom:** Videomötesteknik i olika affärssituationer - möjligheter och hinder, 1997.
FiF-a 6 **Tommy Wedlund**: Att skapa en företagsanpassad systemutvecklingsmodell - genom rekonstruktion, värdering och vidareutveckling i T50-bolag inom ABB, 1997.
No 615 **Silvia Coradeschi**: A Decision-Mechanism for Reactive and Coordinated Agents, 1997.
No 623 **Jan Ollinen:** Det flexibla kontorets utveckling på Digital - Ett stöd för multiflex? 1997.
No 626 **David Byers:** Towards Estimating Software Testability Using Static Analysis, 1997.
No 627 **Fredrik Eklund:** Declarative Error Diagnosis of GAPLog Programs, 1997.
No 629 **Gunilla Ivefors:** Krigsspel och Informationsteknik inför en oförutsägbar framtid, 1997**.**
No 631 **Jens-Olof Lindh:** Analysing Traffic Safety from a Case-Based Reasoning Perspective, 1997
No 639 **Jukka Mäki-Turja:**. Smalltalk - a suitable Real-Time Language, 1997.
No 640 **Juha Takkinen**: CAFE: Towards a Conceptual Model for Information Management in Electronic Mail, 1997.
No 643 **Man Lin**: Formal Analysis of Reactive Rule-based Programs, 1997.
No 653 **Mats Gustafsson**: Bringing Role-Based Access Control to Distributed Systems, 1997.
FiF-a 13 **Boris Karlsson:** Metodanalys för förståelse och utveckling av systemutvecklingsverksamhet. Analys och värdering av systemutvecklingsmodeller och dess användning, 1997.
No 674 **Marcus Bjäreland:** Two Aspects of Automating Logics of Action and Change - Regression and Tractability, 1998.
No 676 **Jan Håkegård**: Hierarchical Test Architecture and Board-Level Test Controller Synthesis, 1998.
No 668 **Per-Ove Zetterlund**: Normering av svensk redovisning - En studie av tillkomsten av Redovisningsrådets rekommendation om koncernredovisning (RR01:91), 1998.
No 675 **Jimmy Tjäder**: Projektledaren & planen - en studie av projektledning i tre installations- och systemutvecklingsprojekt, 1998.
FiF-a 14 **Ulf Melin**: Informationssystem vid ökad affärs- och processorientering - egenskaper, strategier och utveckling, 1998.
No 695 **Tim Heyer**: COMPASS: Introduction of Formal Methods in Code Development and Inspection, 1998.
No 700 **Patrik Hägglund:** Programming Languages for Computer Algebra, 1998.
FiF-a 16 **Marie-Therese Christiansson:** Inter-organisatorisk verksamhetsutveckling - metoder som stöd vid utveckling av partnerskap och informationssystem, 1998.
No 712 **Christina Wennestam:** Information om immateriella resurser. Investeringar i forskning och utveckling samt i personal inom skogsindustrin, 1998.
No 719 **Joakim Gustafsson:** Extending Temporal Action Logic for Ramification and Concurrency, 1998.
No 723 **Henrik André-Jönsson:** Indexing time-series data using text indexing methods, 1999.
No 725 **Erik Larsson:** High-Level Testability Analysis and Enhancement Techniques, 1998.
No 730 **Carl-Johan Westin:** Informationsförsörjning: en fråga om ansvar - aktiviteter och uppdrag i fem stora svenska organisationers operativa informationsförsörjning, 1998.
No 731 **Åse Jansson:** Miljöhänsyn - en del i företags styrning, 1998.
No 733 **Thomas Padron-McCarthy:** Performance-Polymorphic Declarative Queries, 1998.
No 734 **Anders Bäckström:** Värdeskapande kreditgivning - Kreditriskhantering ur ett agentteoretiskt perspektiv, 1998.
FiF-a 21 **Ulf Seigerroth:** Integration av förändringsmetoder - en modell för välgrundad metodintegration, 1999.
FiF-a 22 **Fredrik Öberg:** Object-Oriented Frameworks - A New Strategy for Case Tool Development, 1998.
No 737 **Jonas Mellin:** Predictable Event Monitoring, 1998.
No 738 **Joakim Eriksson:** Specifying and Managing Rules in an Active Real-Time Database System, 1998.
FiF-a 25 **Bengt E W Andersson:** Samverkande informationssystem mellan aktörer i offentliga åtaganden - En teori om aktörsarenor i samverkan om utbyte av information, 1998.

No 742 **Pawel Pietrzak:** Static Incorrectness Diagnosis of CLP (FD), 1999.

No 748 **Tobias Ritzau:** Real-Time Reference Counting in RT-Java, 1999.

No 751 **Anders Ferntoft:** Elektronisk affärskommunikation - kontaktkostnader och kontaktprocesser mellan kunder och leverantörer på producentmarknader, 1999.

No 752 **Jo Skåmedal:** Arbete på distans och arbetsformens påverkan på resor och resmönster, 1999.

No 753 **Johan Alvehus:** Mötets metaforer. En studie av berättelser om möten, 1999.

No 754 **Magnus Lindahl:** Bankens villkor i låneavtal vid kreditgivning till högt belånade företagsförvärv: En studie ur ett agentteoretiskt perspektiv, 2000.

No 766 **Martin V. Howard:** Designing dynamic visualizations of temporal data, 1999.

No 769 **Jesper Andersson:** Towards Reactive Software Architectures, 1999.

No 775 **Anders Henriksson:** Unique kernel diagnosis, 1999.

FiF-a 30 **Pär J. Ågerfalk:** Pragmatization of Information Systems - A Theoretical and Methodological Outline, 1999.

No 787 **Charlotte Björkegren:** Learning for the next project - Bearers and barriers in knowledge transfer within an organisation, 1999.

No 788 **Håkan Nilsson:** Informationsteknik som drivkraft i granskningsprocessen - En studie av fyra revisionsbyråer, 2000.

No 790 **Erik Berglund:** Use-Oriented Documentation in Software Development, 1999.

No 791 **Klas Gäre:** Verksamhetsförändringar i samband med IS-införande, 1999.

No 800 **Anders Subotic:** Software Quality Inspection, 1999.

No 807 **Svein Bergum**: Managerial communication in telework, 2000.

No 809 **Flavius Gruian:** Energy-Aware Design of Digital Systems, 2000.

FiF-a 32 **Karin Hedström:** Kunskapsanvändning och kunskapsutveckling hos verksamhetskonsulter - Erfarenheter från ett FOU-samarbete, 2000.

No 808 **Linda Askenäs:** Affärssystemet - En studie om teknikens aktiva och passiva roll i en organisation, 2000.

No 820 **Jean Paul Meynard:** Control of industrial robots through high-level task programming, 2000.

No 823 **Lars Hult:** Publika Gränsytor - ett designexempel, 2000.

No 832 **Paul Pop:** Scheduling and Communication Synthesis for Distributed Real-Time Systems, 2000.

FiF-a 34 **Göran Hultgren:** Nätverksinriktad Förändringsanalys - perspektiv och metoder som stöd för förståelse och utveckling av affärsrelationer och informationssystem, 2000.

No 842 **Magnus Kald:** The role of management control systems in strategic business units, 2000.

No 844 **Mikael Cäker:** Vad kostar kunden? Modeller för intern redovisning, 2000.

FiF-a 37 **Ewa Braf**: Organisationers kunskapsverksamheter - en kritisk studie av ”knowledge management”, 2000.

FiF-a 40 **Henrik Lindberg:** Webbaserade affärsprocesser - Möjligheter och begränsningar, 2000.

FiF-a 41 **Benneth Christiansson:** Att komponentbasera informationssystem - Vad säger teori och praktik?, 2000.

No. 854 **Ola Pettersson:** Deliberation in a Mobile Robot, 2000.

No 863 **Dan Lawesson**: Towards Behavioral Model Fault Isolation for Object Oriented Control Systems, 2000.

No 881 **Johan Moe:** Execution Tracing of Large Distributed Systems, 2001.

No 882 **Yuxiao Zhao:** XML-based Frameworks for Internet Commerce and an Implementation of B2B e-procurement, 2001.

No 890 **Annika Flycht-Eriksson:** Domain Knowledge Management in Information-providing Dialogue systems, 2001.

FiF-a 47 **Per-Arne Segerkvis**t: Webbaserade imaginära organisationers samverkansformer: Informationssystemarkitektur och aktörssamverkan som förutsättningar för affärsprocesser, 2001.

No 894 **Stefan Svarén:** Styrning av investeringar i divisionaliserade företag - Ett koncernperspektiv, 2001.

No 906 **Lin Han**: Secure and Scalable E-Service Software Delivery, 2001.

No 917 **Emma Hansson:** Optionsprogram för anställda - en studie av svenska börsföretag, 2001.

No 916 **Susanne Odar:** IT som stöd för strategiska beslut, en studie av datorimplementerade modeller av verksamhet som stöd för beslut om anskaffning av JAS 1982, 2002.

FiF-a-49 **Stefan Holgersson:** IT-system och filtrering av verksamhetskunskap - kvalitetsproblem vid analyser och beslutsfattande som bygger på uppgifter hämtade från polisens IT-system, 2001.

FiF-a-51 **Per Oscarsson:** Informationssäkerhet i verksamheter - begrepp och modeller som stöd för förståelse av informationssäkerhet och dess hantering, 2001.

No 919 **Luis Alejandro Cortes:** A Petri Net Based Modeling and Verification Technique for Real-Time Embedded Systems, 2001.

No 915 **Niklas Sandell:** Redovisning i skuggan av en bankkris - Värdering av fastigheter. 2001.

No 931 **Fredrik Elg:** Ett dynamiskt perspektiv på individuella skillnader av heuristisk kompetens, intelligens, mentala modeller, mål och konfidens i kontroll av mikrovärlden Moro, 2002.

No 933 **Peter Aronsson:** Automatic Parallelization of Simulation Code from Equation Based Simulation Languages, 2002.

No 938 **Bourhane Kadmiry**: Fuzzy Control of Unmanned Helicopter, 2002.

No 942 **Patrik Haslum**: Prediction as a Knowledge Representation Problem: A Case Study in Model Design, 2002.

No 956 **Robert Sevenius:** On the instruments of governance - A law & economics study of capital instruments in limited liability companies, 2002.

FiF-a 58 **Johan Petersson:** Lokala elektroniska marknadsplatser - informationssystem för platsbundna affärer, 2002.

No 964 **Peter Bunus:** Debugging and Structural Analysis of Declarative Equation-Based Languages, 2002.

No 973 **Gert Jervan:** High-Level Test Generation and Built-In Self-Test Techniques for Digital Systems, 2002.
No 958 **Fredrika Berglund:** Management Control and Strategy - a Case Study of Pharmaceutical Drug Development, 2002.
FiF-a 61 **Fredrik Karlsson:** Meta-Method for Method Configuration - A Rational Unified Process Case, 2002.
No 985 **Sorin Manolache:** Schedulability Analysis of Real-Time Systems with Stochastic Task Execution Times, 2002.
No 982 **Diana Szentiványi:** Performance and Availability Trade-offs in Fault-Tolerant Middleware, 2002.
No 989 **Iakov Nakhimovski:** Modeling and Simulation of Contacting Flexible Bodies in Multibody Systems, 2002.
No 990 **Levon Saldamli:** PDEModelica - Towards a High-Level Language for Modeling with Partial Differential Equations, 2002.
No 991 **Almut Herzog:** Secure Execution Environment for Java Electronic Services, 2002.
No 999 **Jon Edvardsson:** Contributions to Program- and Specification-based Test Data Generation, 2002.
No 1000 **Anders Arpteg:** Adaptive Semi-structured Information Extraction, 2002.
No 1001 **Andrzej Bednarski:** A Dynamic Programming Approach to Optimal Retargetable Code Generation for Irregular Architectures, 2002.
No 988 **Mattias Arvola:** Good to use! : Use quality of multi-user applications in the home, 2003.
FiF-a 62 **Lennart Ljung:** Utveckling av en projektivitetsmodell - om organisationers förmåga att tillämpa projektarbetsformen, 2003.
No 1003 **Pernilla Qvarfordt:** User experience of spoken feedback in multimodal interaction, 2003.
No 1005 **Alexander Siemers:** Visualization of Dynamic Multibody Simulation With Special Reference to Contacts, 2003.
No 1008 **Jens Gustavsson:** Towards Unanticipated Runtime Software Evolution, 2003.
No 1010 **Calin Curescu:** Adaptive QoS-aware Resource Allocation for Wireless Networks, 2003.
No 1015 **Anna Andersson:** Management Information Systems in Process-oriented Healthcare Organisations, 2003.
No 1018 **Björn Johansson:** Feedforward Control in Dynamic Situations, 2003.
No 1022 **Traian Pop:** Scheduling and Optimisation of Heterogeneous Time/Event-Triggered Distributed Embedded Systems, 2003.
FiF-a 65 **Britt-Marie Johansson:** Kundkommunikation på distans - en studie om kommunikationsmediets betydelse i affärstransaktioner, 2003.
No 1024 **Aleksandra Tešanovic:** Towards Aspectual Component-Based Real-Time System Development, 2003.
No 1034 **Arja Vainio-Larsson**: Designing for Use in a Future Context - Five Case Studies in Retrospect, 2003.
No 1033 **Peter Nilsson:** Svenska bankers redovisningsval vid reservering för befarade kreditförluster - En studie vid införandet av nya redovisningsregler, 2003.
FiF-a 69 **Fredrik Ericsson:** Information Technology for Learning and Acquiring of Work Knowledge, 2003.
No 1049 **Marcus Comstedt:** Towards Fine-Grained Binary Composition through Link Time Weaving, 2003.
No 1052 **Åsa Hedenskog:** Increasing the Automation of Radio Network Control, 2003.
No 1054 **Claudiu Duma:** Security and Efficiency Tradeoffs in Multicast Group Key Management, 2003.
FiF-a 71 **Emma Eliason:** Effektanalys av IT-systems handlingsutrymme, 2003.
No 1055 **Carl Cederberg:** Experiments in Indirect Fault Injection with Open Source and Industrial Software, 2003.
No 1058 **Daniel Karlsson:** Towards Formal Verification in a Component-based Reuse Methodology, 2003.
FiF-a 73 **Anders Hjalmarsson:** Att etablera och vidmakthålla förbättringsverksamhet - behovet av koordination och interaktion vid förändring av systemutvecklingsverksamheter, 2004.
No 1079 **Pontus Johansson:** Design and Development of Recommender Dialogue Systems, 2004.
No 1084 **Charlotte Stoltz:** Calling for Call Centres - A Study of Call Centre Locations in a Swedish Rural Region, 2004.
FiF-a 74 **Björn Johansson:** Deciding on Using Application Service Provision in SMEs, 2004.
No 1094 **Genevieve Gorrell:** Language Modelling and Error Handling in Spoken Dialogue Systems, 2004.
No 1095 **Ulf Johansson:** Rule Extraction - the Key to Accurate and Comprehensible Data Mining Models, 2004.
No 1099 **Sonia Sangari:** Computational Models of Some Communicative Head Movements, 2004.
No 1110 **Hans Nässla:** Intra-Family Information Flow and Prospects for Communication Systems, 2004.
No 1116 **Henrik Sällberg:** On the value of customer loyalty programs - A study of point programs and switching costs, 2004.
FiF-a 77 **Ulf Larsson:** Designarbete i dialog - karaktärisering av interaktionen mellan användare och utvecklare i en systemutvecklingsprocess, 2004.
No 1126 **Andreas Borg:** Contribution to Management and Validation of Non-Functional Requirements, 2004.
No 1127 **Per-Ola Kristensson:** Large Vocabulary Shorthand Writing on Stylus Keyboard, 2004.
No 1132 **Pär-Anders Albinsson:** Interacting with Command and Control Systems: Tools for Operators and Designers, 2004.
No 1130 **Ioan Chisalita:** Safety-Oriented Communication in Mobile Networks for Vehicles, 2004.
No 1138 **Thomas Gustafsson:** Maintaining Data Consistency in Embedded Databases for Vehicular Systems, 2004.
No 1149 **Vaida Jakoniené:** A Study in Integrating Multiple Biological Data Sources, 2005.
No 1156 **Abdil Rashid Mohamed**: High-Level Techniques for Built-In Self-Test Resources Optimization, 2005.
No 1162 **Adrian Pop:** Contributions to Meta-Modeling Tools and Methods, 2005.
No 1165 **Fidel Vascós Palacios:** On the information exchange between physicians and social insurance officers in the sick leave process: an Activity Theoretical perspective, 2005.
FiF-a 84 **Jenny Lagsten:** Verksamhetsutvecklande utvärdering i informationssystemprojekt, 2005.
No 1166 **Emma Larsdotter Nilsson:** Modeling, Simulation, and Visualization of Metabolic Pathways Using Modelica, 2005.

No 1167 **Christina Keller:** Virtual Learning Environments in higher education. A study of students' acceptance of educational technology, 2005.
No 1168 **Cécile Åberg:** Integration of organizational workflows and the Semantic Web, 2005.
FiF-a 85 **Anders Forsman:** Standardisering som grund för informationssamverkan och IT-tjänster - En fallstudie baserad på trafikinformationstjänsten RDS-TMC, 2005.
No 1171 **Yu-Hsing Huang:** A systemic traffic accident model, 2005.
FiF-a 86 **Jan Olausson:** Att modellera uppdrag - grunder för förståelse av processinriktade informationssystem i transaktionsintensiva verksamheter, 2005.
No 1172 **Petter Ahlström**: Affärsstrategier för seniorbostadsmarknaden, 2005.
No 1183 **Mathias Cöster**: Beyond IT and Productivity - How Digitization Transformed the Graphic Industry, 2005.
No 1184 **Åsa Horzella**: Beyond IT and Productivity - Effects of Digitized Information Flows in Grocery Distribution, 2005.
No 1185 **Maria Kollberg**: Beyond IT and Productivity - Effects of Digitized Information Flows in the Logging Industry, 2005.
No 1190 **David Dinka**: Role and Identity - Experience of technology in professional settings, 2005.
No 1191 **Andreas Hansson**: Increasing the Storage Capacity of Recursive Auto-associative Memory by Segmenting Data, 2005.
No 1192 **Nicklas Bergfeldt:** Towards Detached Communication for Robot Cooperation, 2005.
No 1194 **Dennis Maciuszek:** Towards Dependable Virtual Companions for Later Life, 2005.
No 1204 **Beatrice Alenljung:** Decision-making in the Requirements Engineering Process: A Human-centered Approach, 2005.
No 1206 **Anders Larsson:** System-on-Chip Test Scheduling and Test Infrastructure Design, 2005.
No 1207 **John Wilander:** Policy and Implementation Assurance for Software Security, 2005.
No 1209 **Andreas Käll:** Översättningar av en managementmodell - En studie av införandet av Balanced Scorecard i ett landsting, 2005.
No 1225 **He Tan:** Aligning and Merging Biomedical Ontologies, 2006.
No 1228 **Artur Wilk:** Descriptive Types for XML Query Language Xcerpt, 2006.
No 1229 **Per Olof Pettersson:** Sampling-based Path Planning for an Autonomous Helicopter, 2006.
No 1231 **Kalle Burbeck:** Adaptive Real-time Anomaly Detection for Safeguarding Critical Networks, 2006.
No 1233 **Daniela Mihailescu:** Implementation Methodology in Action: A Study of an Enterprise Systems Implementation Methodology, 2006.
No 1244 **Jörgen Skågeby:** Public and Non-public gifting on the Internet, 2006.
No 1248 **Karolina Eliasson:** The Use of Case-Based Reasoning in a Human-Robot Dialog System, 2006.
No 1263 **Misook Park-Westman:** Managing Competence Development Programs in a Cross-Cultural Organisation - What are the Barriers and Enablers, 2006.
FiF-a 90 **Amra Halilovic:** Ett praktikperspektiv på hantering av mjukvarukomponenter, 2006.
No 1272 **Raquel Flodström**: A Framework for the Strategic Management of Information Technology, 2006.
No 1277 **Viacheslav Izosimov:** Scheduling and Optimization of Fault-Tolerant Embedded Systems, 2006.
No 1283 **Håkan Hasewinkel:** A Blueprint for Using Commercial Games off the Shelf in Defence Training, Education and Research Simulations, 2006.
FiF-a 91 **Hanna Broberg:** Verksamhetsanpassade IT-stöd - Designteori och metod, 2006.
No 1286 **Robert Kaminski:** Towards an XML Document Restructuring Framework, 2006.
No 1293 **Jiri Trnka:** Prerequisites for data sharing in emergency management, 2007.
No 1302 **Björn Hägglund:** A Framework for Designing Constraint Stores, 2007.
No 1303 **Daniel Andreasson**: Slack-Time Aware Dynamic Routing Schemes for On-Chip Networks, 2007.
No 1305 **Magnus Ingmarsson:** Modelling User Tasks and Intentions for Service Discovery in Ubiquitous Computing, 2007.
No 1306 **Gustaf Svedjemo**: Ontology as Conceptual Schema when Modelling Historical Maps for Database Storage, 2007.
No 1307 **Gianpaolo Conte**: Navigation Functionalities for an Autonomous UAV Helicopter, 2007.
No 1309 **Ola Leifler**: User-Centric Critiquing in Command and Control: The DKExpert and ComPlan Approaches, 2007.
No 1312 **Henrik Svensson**: Embodied simulation as off-line representation, 2007.
No 1313 **Zhiyuan He:** System-on-Chip Test Scheduling with Defect-Probability and Temperature Considerations, 2007.
No 1317 **Jonas Elmqvist:** Components, Safety Interfaces and Compositional Analysis, 2007.
No 1320 **Håkan Sundblad:** Question Classification in Question Answering Systems, 2007.
No 1323 **Magnus Lundqvist**: Information Demand and Use: Improving Information Flow within Small-scale Business Contexts, 2007.
No 1329 **Martin Magnusson:** Deductive Planning and Composite Actions in Temporal Action Logic, 2007.
No 1331 **Mikael Asplund:** Restoring Consistency after Network Partitions, 2007.
No 1332 **Martin Fransson:** Towards Individualized Drug Dosage - General Methods and Case Studies, 2007.
No 1333 **Karin Camara:** A Visual Query Language Served by a Multi-sensor Environment, 2007.
No 1337 **David Broman:** Safety, Security, and Semantic Aspects of Equation-Based Object-Oriented Languages and Environments, 2007.
No 1339 **Mikhail Chalabine:** Invasive Interactive Parallelization, 2007.
No 1351 **Susanna Nilsson:** A Holistic Approach to Usability Evaluations of Mixed Reality Systems, 2008.
No 1353 **Shanai Ardi:** A Model and Implementation of a Security Plug-in for the Software Life Cycle, 2008.
No 1356 **Erik Kuiper:** Mobility and Routing in a Delay-tolerant Network of Unmanned Aerial Vehicles, 2008.
No 1359 **Jana Rambusch**: Situated Play, 2008.
No 1361 **Martin Karresand:** Completing the Picture - Fragments and Back Again, 2008.

No 1363 **Per Nyblom:** Dynamic Abstraction for Interleaved Task Planning and Execution, 2008.
No 1371 **Fredrik Lantz:** Terrain Object Recognition and Context Fusion for Decision Support, 2008.
No 1373 **Martin Östlund:** Assistance Plus: 3D-mediated Advice-giving on Pharmaceutical Products, 2008.
No 1381 **Håkan Lundvall:** Automatic Parallelization using Pipelining for Equation-Based Simulation Languages, 2008.
No 1386 **Mirko Thorstensson:** Using Observers for Model Based Data Collection in Distributed Tactical Operations, 2008.
No 1387 **Bahlol Rahimi:** Implementation of Health Information Systems, 2008.
No 1392 **Maria Holmqvist:** Word Alignment by Re-using Parallel Phrases, 2008.
No 1393 **Mattias Eriksson**: Integrated Software Pipelining, 2009.
No 1401 **Annika Öhgren:** Towards an Ontology Development Methodology for Small and Medium-sized Enterprises, 2009.
No 1410 **Rickard Holsmark:** Deadlock Free Routing in Mesh Networks on Chip with Regions, 2009.
No 1421 **Sara Stymne:** Compound Processing for Phrase-Based Statistical Machine Translation, 2009.
No 1427 **Tommy Ellqvist**: Supporting Scientific Collaboration through Workflows and Provenance, 2009.
No 1450 **Fabian Segelström:** Visualisations in Service Design, 2010.
No 1459 **Min Bao**: System Level Techniques for Temperature-Aware Energy Optimization, 2010.
No 1466 **Mohammad Saifullah**: Exploring Biologically Inspired Interactive Networks for Object Recognition, 2011
No 1468 **Qiang Liu**: Dealing with Missing Mappings and Structure in a Network of Ontologies, 2011.
No 1469 **Ruxandra Pop**: Mapping Concurrent Applications to Multiprocessor Systems with Multithreaded Processors and Network on Chip-Based Interconnections, 2011.
No 1476 **Per-Magnus Olsson**: Positioning Algorithms for Surveillance Using Unmanned Aerial Vehicles, 2011.
No 1481 **Anna Vapen**: Contributions to Web Authentication for Untrusted Computers, 2011.
No 1485 **Loove Broms:** Sustainable Interactions: Studies in the Design of Energy Awareness Artefacts, 2011.
FiF-a 101 **Johan Blomkvist:** Conceptualising Prototypes in Service Design, 2011.
No 1490 **Håkan Warnquist:** Computer-Assisted Troubleshooting for Efficient Off-board Diagnosis, 2011.
No 1503 **Jakob Rosén:** Predictable Real-Time Applications on Multiprocessor Systems-on-Chip, 2011.
No 1504 **Usman Dastgeer:** Skeleton Programming for Heterogeneous GPU-based Systems, 2011.
No 1506 **David Landén:** Complex Task Allocation for Delegation: From Theory to Practice, 2011.
No 1507 **Kristian Stavåker**: Contributions to Parallel Simulation of Equation-Based Models on Graphics Processing Units, 2011.
No 1509 **Mariusz Wzorek:** Selected Aspects of Navigation and Path Planning in Unmanned Aircraft Systems, 2011.
No 1510 **Piotr Rudol:** Increasing Autonomy of Unmanned Aircraft Systems Through the Use of Imaging Sensors, 2011.
No 1513 **Anders Carstensen:** The Evolution of the Connector View Concept: Enterprise Models for Interoperability Solutions in the Extended Enterprise, 2011.
No 1523 **Jody Foo:** Computational Terminology: Exploring Bilingual and Monolingual Term Extraction, 2012.
No 1550 **Anders Fröberg:** Models and Tools for Distributed User Interface Development, 2012.
No 1558 **Dimitar Nikolov:** Optimizing Fault Tolerance for Real-Time Systems, 2012.
No 1582 **Dennis Andersson:** Mission Experience: How to Model and Capture it to Enable Vicarious Learning, 2013.
No 1586 **Massimiliano Raciti:** Anomaly Detection and its Adaptation: Studies on Cyber-physical Systems, 2013.
No 1588 **Banafsheh Khademhosseinieh:** Towards an Approach for Efficiency Evaluation of Enterprise Modeling Methods, 2013.
No 1589 **Amy Rankin:** Resilience in High Risk Work: Analysing Adaptive Performance, 2013.
No 1592 **Martin Sjölund:** Tools for Understanding, Debugging, and Simulation Performance Improvement of Equation-Based Models, 2013.
No 1606 **Karl Hammar:** Towards an Ontology Design Pattern Quality Model, 2013.
No 1624 **Maria Vasilevskaya:** Designing Security-enhanced Embedded Systems: Bridging Two Islands of Expertise, 2013.
No 1627 **Ekhiotz Vergara:** Exploiting Energy Awareness in Mobile Communication, 2013.
No 1644 **Valentina Ivanova:** Integration of Ontology Alignment and Ontology Debugging for Taxonomy Networks, 2014.
No 1647 **Dag Sonntag:** A Study of Chain Graph Interpretations, 2014.
No 1657 **Kiril Kiryazov:** Grounding Emotion Appraisal in Autonomous Humanoids, 2014.
No 1683 **Zlatan Dragisic:** Completing the Is-a Structure in Description Logics Ontologies, 2014.
No 1688 **Erik Hansson:** Code Generation and Global Optimization Techniques for a Reconfigurable PRAM-NUMA Multicore Architecture, 2014.
No 1715 **Nicolas Melot:** Energy-Efficient Computing over Streams with Massively Parallel Architectures, 2015.
No 1716 **Mahder Gebremedhin:** Automatic and Explicit Parallelization Approaches for Mathematical Simulation Models, 2015.
No 1722 **Mikael Nilsson:** Efficient Temporal Reasoning with Uncertainty, 2015.
No 1732 **Vladislavs Jahundovics:** Automatic Verification of Parameterized Sytems by Over-Approximation, 2015.
FiF 118 **Camilla Kirkegaard:** Adding Challenge to a Teachable Agent in a Virtual Learning Environment, 2016.
No 1758 **Vengatanathan Krishnamoorthi:** Efficient and Scalable Content Delivery of Linear and Interactive Branched Videos, 2016.
No 1771 **Andreas Löfwenmark:** Timing Predictability in Future Multi-Core Avionics Systems, 2017.
No 1777 **Anders Andersson:** Extensions for Distributed Moving Base Driving Simulators, 2017.
No 1780 **Olov Andersson:** Methods for Scalable and Safe Robot Learning, 2017.
No 1782 **Robin Keskisärkkä**: Towards Semantically Enabled Complex Event Processing, 2017.
No 1783 **Daniel de Leng**: Spatio-Temporal Stream Reasoning with Adaptive State Stream Generation, 2017.

No 1827 **Johan Falkenjack**: Towards a Model of General Text Complexity for Swedish, 2018.

No 1836 **Magdalena Granåsen**: Exploring C2 Capability and Effectiveness in Challenging Environments: Interorganizational Crisis Management, Military Operations and Cyber Defence, 2019.

No 1848 **Alachew Mengist**: Methods and Tools for Efficient Model-Based Development of Cyber-Physical Systems with Emphasis on Model and Tool Integration, 2019.

No 1871 **Klervie Toczé**: Latency-aware Resource Management at the Edge, 2020.

No 1881 **Chih-Yuan Lin**: A Timing Approach to Network-based Anomaly Detection for SCADA Systems, 2020.

No 1886 **August Ernstsson**: Designing a Modern Skeleton Programming Framework for Parallel and Heterogeneous Systems, 2020.

No 1892 **John Törnblom**: Formal Verification of Tree Ensembles in Safety-Critical Applications, 2020.

FiF 128 **Amanda Olmin**: On Uncertainty Quantification in Neural Networks: Ensemble Distillation and Weak Supervision, 2022.

No 1937 **John Tinnerholm**: A Composable and Extensible Environment for Equation-based Modeling and Simulation of Variable Structured Systems in Modelica, 2022.

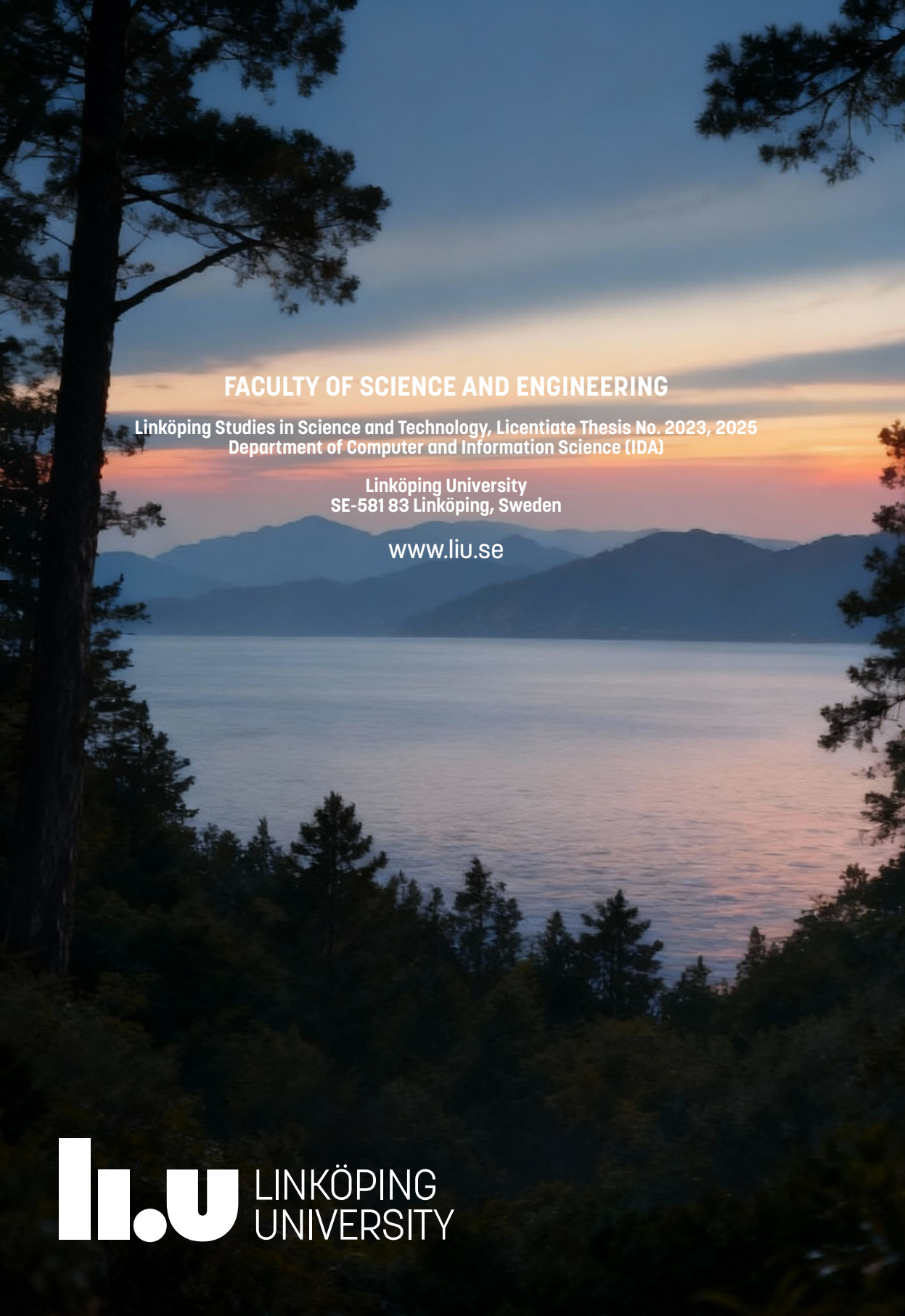
FiF 134 **Woodrow Hao Chi Kiang**: Uncertainty Estimation in Models of Multivariate Trait Evolution on Given Phylogenies, 2024.

No 2011 **Johannes Wilson**: Towards Practical Formal Verification of Cryptographic Protocol Specifications and Implementations, 2025.

No 2012 **Abdelazim G. Hussien**: Contributions to Metaheuristic Algorithms for Real-World Engineering Problems, 2025.

No 2019 **Ahmad B. Usman**: Trustworthy Software States through Attestation and Secure Updates, 2025.

No 2023 **Valency Colaco**: Hardening Tree Ensembles: Real-Time and Effective Evasion Defences Beyond Adversarial R-Training, 2025.



FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology, Licentiate Thesis No. 2023, 2025
Department of Computer and Information Science (IDA)

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se

LINKÖPING UNIVERSITY