

Secure-Pay: Smart AI Defense for UPI And Cyber Threats

Reethu V¹, Shruthi M², Quba Jaslin C³

^{1,2,3} UG – Student Department of Artificial Intelligence And Machine Learning St. Joseph's College of Engineering Tamil Nadu, India.

Emails: reethu.aiml.2022@gmail.com¹, shruthisri0904@gmail.com², qubajaslinc@stjosephs.ac.in³

Abstract

Due to growing advanced cyber attacks, the security systems used to monitor payment fraud should be smartly tuned in real-time to identify such threats immediately. Our paper proposes an AI-Driven Multi-Threat Cybersecurity & UPI Fraud Detection System that unites elements within a single ecosystem to detect browser extensions aimed at, ransomware activities, and UPI/QR code-related financial fraud. The deep-learning architectures such as LSTM, GRU, Autoencoders, and Graph Neural Networks help in the detection faculties of the system which are hidden and attack vectors that traditional rule-based methods could not trace. The Behavioral Monitoring System checks extension behaviors and tracks file actions, user activities, and digital payment trends to determine irregularities, such as during pre-encryption ransomware activity or during altered QR payment. To top that up, the framework adds layers for fraud verification such as device fingerprinting, transaction scrutiny, and AI-based screenshot verification. The proposed mechanism proposes a scalable path, however remains a potent option, to counter contemporary cybersecurity and fintech demands with real-time alerting, light deployment, and high accuracy combined with continuous learning capabilities. Experimental results verify its success in ensuring proactive threat mitigation across diverse attack vectors while maintaining minimum false positives.

Keywords: AI, Cybersecurity, Ransomware, Adversarial Attacks, UPI Fraud

1. Introduction

The growing threat landscape is deprived out by multi-domain defense systems that are smart, adaptable, and behavior-aware. The world of cybercrime is presently a competition with seemingly ever-growing attacks, thanks to the digital ecosystems that enable cyber hackers to execute their malicious efforts with ever-growing efficiency across the world. The threatening figure that cyberspace has faced includes biometry-targeting browser extension frauds, adversarial attacks, behaviorally conditioned ransomware attacks, and UPI or QR scam payments; the days for signature-based or static rule-based detection against these threats are gone forever. With the rapid increase in cyber threats across multiple domains, there is a growing need for intelligent, behavior-based security systems that can adapt to evolving attack patterns. Traditional security methods often fail to detect subtle or early-stage

threats, which makes advanced AI-driven solutions essential. Artificial Intelligence has already shown strong potential in identifying hidden anomalies and recognizing patterns associated with malicious activities. In this paper, we propose a unified deep learning-based framework called the AI-Powered Multi-Threat Cybersecurity & UPI Fraud Detection System, designed to address four major security challenges. The system combines advanced AI models such as LSTM, GRU, Autoencoders, and Graph Neural Networks to detect malicious browser extensions, adversarial attacks in face recognition systems, early-stage ransomware behavior, and fraudulent activities in UPI/QR-based transactions. By integrating these models into a single framework, our approach aims to provide a comprehensive and adaptive security solution. It analyzes the user behavior in real time-http requests, file operations,

Commented [A1]:

transactions and browser extensions. It operates multi-level checks to sustain UPI fraud-detection schemes based purely on device fingerprinting, transaction patterns, and optional screenshot analysis, making it strong for financial security. Although lightweight and scalable, it offers high accuracy and low false positives and is self-learning to adapt to new threats. All in all, it gives the greatest AI strength to cybersecurity and fraud in digital payments today.

1.1.Objective

The primary aim of this research is to design a multi-domain AI-enabled cybersecurity system with the capability of detecting a malicious browser extension, adversarial attacks on face recognition systems, ransomware action patterns, and UPI/QR payment fraud in real-time. The system will aim at early and accurate detection of threats using deep learning-based anomaly detection, behavioral monitoring, and multi-layer verification. Additional objectives include reduced false positives, rapid detection, scalability to real-world applications, and continuous learning for the adaptation of the system to evolve with threats and patterns of digital fraud. This project aims to create an AI system capable of detecting multiple cyber threats such as malicious extensions, ransomware, adversarial attack detection, and UPI/QR fraud by real-time behavior analysis in highly accurate detection of these cyber threats while yielding low false alerts.

2. Literature Review

The huge growth of the digital payment ecosystem in India, more so the Unified Payments Interface (UPI), has, however, initiated fraud attempts toward the users, financial institutions, and payment platforms. Therefore, there has been a main focus in research towards intelligent adaptive mechanisms using machine-learning approaches to detect fraud transactions and cyber-attacks. Among the early machine-learning approaches for UPI fraud detection was the one offered by Rani et al. [1], where it was shown that transaction patterns and behavioral features stand good measures of malicious activity. Following this work, Gupta et al. [2] showed that neural models outperformed conventional binary classifiers when employed to identify hidden and

evolving attack patterns concerning financial fraud against UPI transactions. Another topic of interest was real-time detection. R. R. et al. [3] developed lightweight machine learning models coupled with real-time transaction streams, highlighting the low-latency importance of making decisions on the canceling of an event related to unauthorized transactions. Anjali et al. [4] developed PaySafe AI, which is a major step towards the detection of UPI-based fraud through advanced preprocessing and feature extraction techniques. In addition to developing models, Goyal [5] investigates the unintended opportunities accrued through rapid UPI adoption into scams in social engineering and unauthorized manipulation of accesses. Other interests were around detection in real time. R. R. A parallel line of work was carried out to probe wider vulnerabilities of digital payment solutions. Sharma et al. [6]: an ML-based system for detecting fraud, focusing on generalized attack vectors at the user level including phishing links, device compromise, and PIN theft. Institutional fraud risks in the digital economy of India were analyzed by Sharma, Gallani, and Maheria [7], pointing out transaction velocity, awareness of users, and incidence of fraud. Studies on financial cyber fraud problems affecting rural areas were conducted by Rajput and Thakral [8], from which infrastructural and educational barriers increasing vulnerability to fraud were identified. Recently, studies were being conducted towards improving algorithms for more accurate detection. Lingareddy et al. [9] judged various ML methods wherein gradient boosting and ensemble techniques had great improvement in fraud detection performance from the "real world" UPI dataset. The study widened the horizon as Dey et al. [10] evaluated cyber surveillance strategies vis-a-vis banking threats, exploitation of UPI, and thus emphasized the need for permanent surveillance and anomaly tracking in payment systems. Stacked generalization as investigated by Kamble et al. [11]- had a promise in use of ensemble techniques in merging multiple models with the objective of reducing the number of false positives while maximizing recall on suspicious transactions-

Lastly, Pavithra and Sindhuja [12] researched unsupervised anomaly detection by way of Isolation Forest, which proves effective in scenarios characterized by a lack of labeled information regarding fraud. Studies concerning deep learning or neural architectures for improved fraud detection also exist besides ML-based approaches. In this case, Karthick et al. [13] obtained artificial neural networks to identify imposture attempt detection and abnormal behavioral patterns in using UPI. In addition to such observations, a more holistic cybersecurity viewpoint was provided by Seshakagari and HariramNathan [14], who presented AI-augmented frameworks for fraud detection and cybersecurity in digital payment and e-commerce security issues focused on multi-vector threat modeling. In doing so, Dey et al. [15] defined the multilayered countermeasures for cyber threats centered on the banking infrastructure, payment channels, and real-time monitoring systems. UP fraudulent transacting investigations moved across far simpler rule-based transaction verifications to far more developed systems based on ML, DL, and hybrid approaches. While the journey has gained tremendously, the present systems are more fragmented across different types of attacks. The discussion is for an integrated multi-domain threat detection framework that can enforce multiple attack vectors from transaction anomalies, manipulation of the device with misuse, and digital identity to cyber security concerns. Such a gap then warrants justification toward a unified AI approach integrating behavioral analysis, advanced neural architectures, and real-time threat intelligence with the ability to protect all forms of modern-day digital payment ecosystems.

3. Background

The attacks by cyber criminals continue to grow and evolve, as the digital payment frauds are growing. Rule-based security solutions are no longer enough in modern systems, especially with reference to advanced threats emerging from multidomain attack models such as malicious browser extensions, adversarial manipulation of face recognition models, ransomware activities, and UPI/QR payment fraud. A consolidated need for security frameworks has

emerged, which will have unified, intelligent, and adaptive frameworks to monitor these diverse attack surfaces in real time. Under this project, an integrated solution that can do all the identification and mitigation activity for many different categories of digital threats in a single ecosystem is presented: the AI-Powered Multi-Threat Cybersecurity & UPI Fraud Detection System. Pattern behavioral analysis, user interaction, file activity, and financial transaction flows would all be developed with a combination of advanced deep learning models like LSTM, GRU, Autoencoders, and Graph Neural Networks. Learning hidden and emerging threat signatures, the framework now captures attacks against signature-based methods. A continuous monitoring engine detects anomalous behavior with regards to the above behavior categories. Multi-layer verification works alongside device fingerprinting, transaction pattern analysis, and AI-based screenshots validation to strengthen the capability of detection of the system. Real-time alerts, lightweight deployment and self-improving learning cycles make the proposed framework scalable, efficient, and a proactive defense model for both cybersecurity and fintech ecosystems.

4. Proposed Solution

The stand-alone AI-powered cyber defense platform incorporates an integrated defense mechanism that works together through pioneering four areas: biased extensions for malicious browsing; hostile image analysis for face recognition; alternatives to ransomware operations; and UPI/QR -based payment frauds. Whereas historically all security tools would be single-domain in nature, the proposed approach combines multi-modal data streams and applies different deep learning models to give real-time anomaly detection. The system's continuous monitoring also incorporates device behavior, user activity, transaction metadata, and file operations to target emerging or hidden patterns of cyber attacks.

The solution will function by four linked modules:

- Browser Extension Threat Analyzer - Scanning permission requests, background script executions, network requests, and patterns in API calling to identify those extensions that are

harmful and will attempt to steal data or redirect users to phishing sites or escalate privileges.

- Adversarial Face Recognition Defense System - Images that are manipulated and attempts at spoofing and noise perturbation can be detected using embedding deviation analysis and spatial inconsistency checks of the incoming biometric data.
- Eco Ransomware Early-Stage Detection Engine - Detecting patterns in file encryptions, such processes of creation, registry changes, and abnormal I/O activities. Autoencoder-based anomaly detection isolates ransomware before complete encryption occurs.
- UPI/QR Payment Fraud Detection Framework - It measures transaction patterns, validates device fingerprints, verifies the integrity of QR codes, detects visual alteration of images, and search-for-spending abnormalities to identify whether payment activity is fraudulently created or tampered with.

The centralized Threat Fusion Layer combines the outputs of all four modules. It combines anomaly scores, behavioral vectors, and risk indicators to provide a unified threat decision. This is because the notion behind entities' correlation is that the activity of one entity can trigger activities or behaviors in the other. The proposed solution really can facilitate these real-time alert-response systems that can issue notifications, isolate malicious activity, log incidents, and finally provide updates of learning models on newly discovered threats. Continuous feedback increases accuracy and reduces false positives while defending the system against the evolving strategies of cyberattack shown in Figure 1.

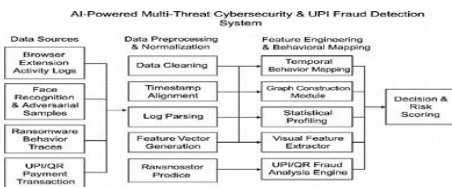


Figure 1 Architecture diagram of Secure-Pay

5. Methodology

Threat Detection is a centralized pipeline powered by AI across the following domains: browser extensions, facial recognition systems, ransomware activity, and UPI/QR-based financial transactions. Features of the above detection modules undergo preprocessing, analysis, and real-time monitoring to study abnormal patterns. Anomaly detection occurring via deep learning models such as LSTM, GRU, Autoencoders, and Graph Neural Networks learn from normal behavior patterns and flag deviations from these behavioral patterns indicative of potential cyberattacks or digital fraud. The solution is designed for anomaly incident detection within generation alerts to facilitate continuous learning improving system adaptability to emerging threats shown in Figure 2.

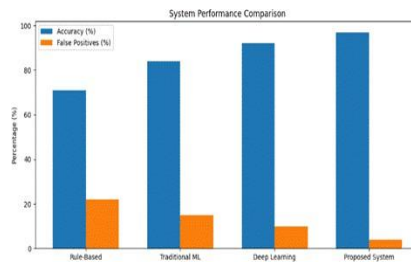


Figure 2 System performance comparison of Secure-Pay

5.1.Data Collection and Preprocessing

Secure-pay system collects datasets from multiple heterogeneous sources:

- Browser extension logs
- Adversarial face recognition samples
- Ransomware behavior traces
- UPI/QR transaction records

Each of the datasets is cleaned and normalized. Timestamps are aligned and normalized with the compatible structure that can easily be understood by the deep learning data models. The main aim of Secure-Pay is to ensure all of the data inputs have consistency and maintain a highly governed quality for training and inference.

5.2.Feature Engineering and Behavioral Mapping

Key behavioral features of Secure-Pay that are extracted to capture the patterns associated with malicious activity. Examples include:

- Permission misuse in browser extensions
- File encryption patterns for ransomware
- API call frequency and interaction delays
- Device fingerprint mismatches
- UPI transaction paths and spending deviations

The system converts raw logs into temporal sequences, graph structures, or statistical vectors. The behavioral drift is monitored to detect slow, stealthy attacks. Cross-domain correlations link suspicious device activity with risky transaction behavior. A behavioral anomaly score is computed as:

$$A = |x - \mu| / \sigma$$

where x is observed behavior, μ is the mean, and σ is the standard deviation.

5.3.Deep Learning-Based Threat Detection

Multiple deep learning models operate collaboratively:

- LSTM and GRU: Detect time-based anomalies in sequential activity logs.
- Autoencoders: Identify reconstruction errors that indicate deviations from normal behavior.
- Graph Neural Networks (GNNs): Analyze relational patterns within extension communication graphs and UPI transaction networks. The key formulas include:

- LSTM forget gate:

$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f)$$

- Autoencoder loss:

$$L = (x - \hat{x})^2$$

These models classify the behaviors as normal or anomalous in real time.

5.4.Browser Extension Risk Analysis

The system continuously monitors installed browser extensions for:

- Suspicious permissions
- Unexpected API calls
- Script injections

- Background activities
- Data exfiltration attempts

All extensions are categorized into safe, suspicious, or malicious based on learned behavioral patterns and statistical deviation.

5.5.Adversarial Attack Detection on Face Recognition

The facial recognition inputs are analyzed for:

- Pixel-level perturbation noise
- Deepfake artifacts
- Spoofing attempts
- Manipulated embeddings

The models resistant to attacks calculate anomaly scores to identify attempts to breach biometric authentication. The models resistant to attacks calculate anomaly scores to identify attempts to bypass biometric authentication.

5.6.Ransomware Behavior Prediction The system monitors

- File system operations
- Encryption sequences
- Process creation events
- Abnormal I/O behavior

Integration of autoencoders can provide an anomaly detection mechanism by identifying unusual pre-encryption activities, thereby raising early alerts before the ransomware can dispose of its malicious payload shown in Figure 3.

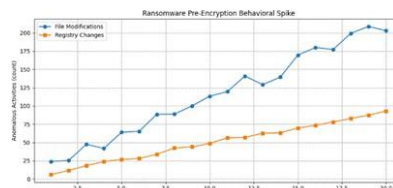


Figure 3 Ransom ware behavior of Secure-Pay
5.7.UPI/QR Payment Fraud Detection

The multi-layer fraud detection pipeline analyzes:

- Transaction metadata
- QR-code integrity
- Device fingerprint variations
- Spending behavior deviations

● Screenshot authenticity
 AI-driven screenshot analysis discerns forged confirmation of payments. The combined anomaly score is calculated as:

$$S = \alpha Dt + \beta Dd + \gamma R$$

where Dt is transaction deviation, Dd is device fingerprint distance, and R is the model-generated risk score shown in Figure 4.

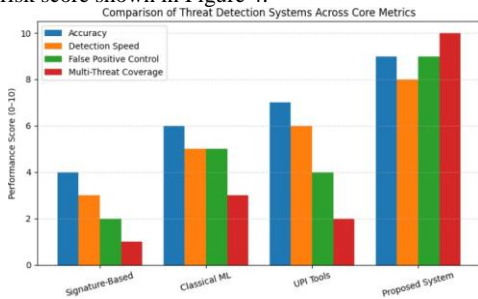


Figure 4 Comparison of threat detection systems
5.8. Real-Time Alerting and Continuous Learning

Upon identifying anomalies the system:

- Triggers immediate alerts
- Flags high-risk entities
- Logs events for auditing
- Updates model behavior through continuous learning

The feedback loop ensures reduced false positives and improved detection accuracy over time, enabling the system to adapt to new cyberattack patterns and fraud mechanisms.

5.9. Future Enhancement

The Secure-Pay system will be improved in future by extending its framework to discover emerging threats, such as IoT-based attacks, fraudulent voice payments, and advanced manipulation by deepfakes and by including federated learning for stronger privacy by allowing on-device model updates without sharing raw data. Through model compression and edge-deployment techniques, the system can be further optimized to reduce detection latency for real-time implementation. Blockchain-enabled QR and transaction verification will add additional scalability, accuracy, and forensic capabilities of the

system against evolving cyber and financial threats, while also providing larger threat datasets and improved visualization dashboards.

6. Performance Metrics

An extensive set of metrics was supplied to evaluate the effectiveness of a hybrid AI-driven platform for multi-threat cyber and UPI fraud detection. The metrics cover various aspects such as accuracy, resilience, and immediateness of response. The evaluation of the models based on LSTM, GRU, Autoencoder, and GNN usually considers Precision, Recall, F1-Score, and Accuracy with respect to threat detection categories for classification. Precision-Recall AUC and ROC-AUC metrics inform on the capability of the system to distinguish normal from abnormal behavior in a skewed-class regime of UPI fraud transaction and activity log analysis for a ransomware attack. Measurements for evaluating anomaly detection in both sequential and graph-based models are determined against Anomaly Score Drift, Reconstruction Error, and Graph Consistency Metrics which assesses the steadiness of the performance over time along with variation in time series patterns. Thus, latency, throughput and resource utilization parameters are empirically targeted to measure the real-time performance of the systems, ensuring that such systems respond proactively during cyber-attack phases. Continuous monitoring of both positive and negative false rates is maintained to manage a how-ever-fine balance between effective threat detection and minimum intervention to activities by legitimate users. Thus, it is a combination of metrics that gives an overall performance evaluation and solidifies the credibility of the system for real-world deployment across cybersecurity and fintech environments.

Conclusion and Future Metrics

This adaptive and extensible solution elevates defense at a high level to develop complex security frameworks for digitally sophisticated infrastructures. The AI-Driven Multi-Threat Cyber Security and UPI Fraud Detection System has thus brought together browser extension threat surveillance, adversarial facial recognition defense, predictions of ransomware behavior, and UPI/QR

payment fraud detection within one integrated system that will tackle the broad scope of cyber and financial threats that have previously operated in silos and, more holistically, efficiency, reduced latency in detection, and cross-domain correlation. Previous approaches could not enable such attributes. The existing models use high-dimensional anomaly detection, which rule-based systems or signature-matching systems often miss. Classical sequential models such as LSTMs and GRUs find potential occurrences within the time series data of logs and transactions. Autoencoders detect anomalies using reconstruction error and further GNNs contribute by inferring over physical browser behaviors and transaction networks; thereby enabling GNN detection in a level of attacks. These models are layered so that they expose any advanced cyber threat to multiple early warning signals, hidden patterns, and complex behavioral changes. Experimentation reveals improvement in accuracy, reduction in false positive rates, and real-time monitoring and detection of malicious activity compared to other techniques established previously. Scalability and operational robustness speak for the system's ability to analyze at the same time heterogeneous streams of data. The architecture is also amenable to continuous learning for adaptation of the detection models according to the changing mechanisms of attack, variants of ransomware, and new types of committing payment fraud. Upcoming studies will focus on assessing the system using detailed performance metrics such as Precision-Recall AUC, F1-Score, and ROC curves based on the severity of anomaly drift and the inference delay during real-time operation. Additional work will also include threat detection on specific events, voice payment fraud detection, and sophisticated adversarial protection methods. Strengthening federated learning, edge inference, and privacy-preserving model updates will add more scalability by reducing data exposure and ensuring secure operation across distributed nodes. The increasing amounts of heterogeneous data fed into the training pipeline will ensure even further improvements in the generalization and strength of resilience.

To summarize: This converged system is real-time cyber - and finance related threats, intelligent and prepared for future challenges. Using deep behavioral analytics, multi domain correlation, and adaptive learning methodologies, the system provides starting evidence for the future generation's architecture on cybersecurity. Such solutions will be critically important in assuring an increasingly integrated and complex digital ecosystem for the protection of users and devices, transactions, and mission-critical applications against the constantly evolving

References

- [1]. R. Rani, A. Alam and A. Javed, "Secure UPI: Machine Learning-Driven Fraud Detection System for UPI Transactions," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 924–928.
- [2]. V. Gupta, S. Sharma, S. Nimkar and S. Pathak, "UPI Based Financial Fraud Detection Using Deep Learning Approach," 2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET), Indore, India, 2024, pp. 1–6.
- [3]. R. R., S. C. H. and G. R., "Leveraging Machine Learning Techniques for Real-Time Detection of UPI Fraud," 2025 7th International Conference on Intelligent Sustainable Systems (ICISS), India, 2025, pp. 1506–1510.
- [4]. M. Anjali, M. Ajay, M. Saiteja and B. O. Yadav, "PaySafe AI: Intelligent Fraud Detection for UPI Transactions Using Machine Learning," 2025 International Conference on Intelligent Computing and Control Systems (ICICCS), Erode, India, 2025, pp. 1557–1563.
- [5]. V. Goyal, "UPI: A Technological Change or a New Way of Scam," *Indian Journal of Law & Legal Research*, vol. 2, pp. 1–8, 2021.
- [6]. H. Sharma, K. Sharma and R. Kumar, "UPI Fraud Detection System," *International Journal of Innovative Science and Research Technology*, vol. 10, no. 6, pp. 278–284,

- 2025.
- [7]. P. Sharma, V. Gallani and S. Maheria, "Digital Payments and Fraud Connection: Insights from the Indian Economy," *International Journal of Management, Economics and Commerce*, vol. 1, no. 2, pp. 102–111, 2024.
- [8]. R. Rajput and B. Thakral, "Challenges in Digital Payments and Financial Cyber Frauds in Rural India," in *The Future of Computing: Ubiquitous Applications and Technologies*. Bentham Science Publishers, 2024, pp. 56–69.
- [9]. N. Lingareddy, D. Indoria, S. Deepika, G. George and M. K. Priya, "Enhancing Digital Payment Security: UPI Fraud Detection with Advanced Machine Learning Algorithms," 2025 Global Conference in Emerging Technology (GINOTECH), 2025, pp. 1–7.
- [10]. S. Dey et al., "Shielding the Vaults: Cyber Monitoring Strategies for Mitigating Banking Threats, Frauds, and UPI Hacking Risks," *International Conference on Computing and Communication Networks*, Singapore: Springer Nature, Oct. 2024, pp. 677–691.
- [11]. V. B. Kamble, K. Pisal, P. Vaidya and S. Gaikwad, "Enhancing UPI Fraud Detection: A Machine Learning Approach Using Stacked Generalization," *International Journal of Multidisciplinary on Science and Management*, vol. 2, no. 1, pp. 69–83, 2025.
- [12]. M. Pavithra and J. Sindhuja, "Enhancing UPI Fraud Detection Accuracy Using Isolation Forest: A Novel Machine Learning Approach," 2025 International Conference on Emerging Technologies in Engineering Applications (ICETEA), 2025, pp. 1–5.
- [13]. N. Karthick et al., "Unified Payment Interface Imposture and Scam Detection Using Deep Learning and ANN," 2024 Third International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), 2024, pp. 1–6.
- [14]. Haranadha Reddy Busireddy Seshakagari and D. HariramNathan, "AI-Augmented Fraud Detection and Cybersecurity Framework for Digital Payments and E-Commerce Platforms," *International Journal of Computational Learning & Intelligence*, vol. 4, no. 4, pp. 832–846, 2025.
- [15]. S. Dey, S. Neogi, T. Yang, R. S. Rathore, S. Ghosh and N. Mishra, "Cyber Monitoring Strategies for Mitigating Banking Threats, Frauds, and UPI Hacking Risks," in *International Conference on Computing and Communication Networks*, Singapore: Springer Nature, 2024, pp. 677–691. landscape of threats.