1. A COMPARATIVE STUDY OF LIGHTWEIGHT MACHINE LEARNING

TECHNIQUES FOR CYBER-ATTACKS DETECTION IN BLOCKCHAIN-

ENABLED INDUSTRIAL SUPPLY CHAIN

**AUTHOURS: SHEREEN ISMAIL, SALAH DANDAN, DIANA W. DAWOUD** 

**PUBLISHED BY:** IEEE ACCES.

**YEAR:** 2021

**ABSTRACT** 

The integration of blockchain technology into industrial supply chains has introduced new

opportunities for enhancing transparency, traceability, and security. However, despite its inherent

resilience, blockchain systems are not entirely immune to cyber-attacks, especially at the endpoints

and interfaces with traditional IT systems. To safeguard these complex and distributed environments,

lightweight machine learning (ML) techniques are emerging as promising solutions due to their

efficiency, scalability, and suitability for edge deployment.

**ADVANTAGES:** 

Efficiency & Speed: Lightweight ML models require less computational power and memory,

enabling faster detection and real-time response.

**Edge Compatibility:** Suitable for deployment on edge devices like IoT sensors and gateways,

reducing latency and reliance on centralized systems.

**DISADVANTAGES:** 

Lower Accuracy Compared to Complex Models: May sacrifice some detection accuracy

compared to deep learning or ensemble techniques.

Limited Generalization: Might not perform well in detecting new or highly sophisticated

attack patterns without regular updates and retraining.

2. A NOVEL DEEP HIERARCHICAL MACHINE LEARNING APPROACH

FOR IDENTIFICATION OF KNOWN AND UNKNOWN MULTIPLE

SECURITY ATTACKS IN A D2D COMMUNICATIONS NETWORK

AUTHOURS: S. V. JANSI RANI, IACOVOS I. IOANNOU, SAI SHRIDHAR

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2018** 

**ABSTRACT** 

This research proposes a novel deep hierarchical machine learning (DHML) framework for

comprehensive identification and classification of multiple types of security attacks in D2D

communication networks. The approach integrates multiple layers of deep learning models—such as

stacked autoencoders, convolutional neural networks (CNN), and long short-term memory (LSTM)

networks—to create a tiered detection system capable of identifying both well-documented and

previously unseen threats. The system leverages hierarchical feature extraction and anomaly detection

at different levels of abstraction to enhance detection robustness and reduce false positives.

**ADVANTAGES:** 

Detection of Unknown Attacks (Zero-day): The hierarchical structure enables the model to

generalize well and detect new types of attacks not present in the training data.

High Accuracy and Low False Positives: Deep models excel at learning complex patterns,

leading to improved classification performance.

**DISADVANTAGES:** 

High Computational Cost: Deep hierarchical models require significant computational

resources for training and inference, which can be impractical for resource-constrained

environments.

Longer Training Time: The complexity and depth of the model can lead to prolonged

training periods, especially with large datasets.

3. AE-NET: NOVEL AUTOENCODER-BASED DEEP FEATURES FOR SQL

INJECTION ATTACK DETECTION

AUTHOURS: NISREAN THALJI, ALI RAZA, MOHAMMAD SHARIFUL ISLAM

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2023

**ABSTRACT** 

SQL Injection (SQLi) remains one of the most prevalent and dangerous web application

attacks, enabling attackers to manipulate backend databases through malicious input. Traditional

detection mechanisms often rely on signature-based or shallow machine learning approaches, which

struggle to identify obfuscated or zero-day SQLi threats.

This study introduces AE-NET, a novel deep learning architecture based on autoencoders,

designed to learn robust and high-level feature representations for the accurate detection of SQL

injection attacks. AE-NET employs an unsupervised pretraining phase to extract compressed latent

features from web traffic and query logs, which are then fine-tuned through a supervised classification

layer.

**ADVANTAGES:** 

Effective Detection of Obfuscated and Unknown Attacks: Autoencoders can learn hidden

patterns and anomalies, enabling AE-NET to detect novel or obfuscated SQL injection attacks.

Deep Feature Learning: By automatically extracting hierarchical and compressed

representations, AE-NET improves model accuracy and reduces reliance on manual feature.

**DISADVANTAGES:** 

High Training Complexity: Training autoencoder-based models requires significant

computational resources and time, especially with large datasets.

Black Box Nature: The interpretability of deep features is limited, making it difficult to explain

why a certain query is flagged as malicious.

4. PATTERN MINING AND DETECTION OF MALICIOUS SQL QUERIES

ON ANONYMIZATION MECHANISM

**AUTHOURS:** JIANGUO ZHENG, XINYU SHEN

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2002

**ABSTRACT** 

This research presents a hybrid framework that combines pattern mining techniques with

anonymization-aware detection to identify malicious SQL queries. The system extracts frequent and

suspicious query patterns from historical data using association rule mining and sequence analysis.

These patterns are then matched against real-time SQL queries to detect anomalies and potential

injection attempts. Special emphasis is placed on how query structures interact with anonymized

fields, ensuring that attacks targeting anonymized databases are accurately flagged.

Experimental evaluations demonstrate that the proposed method achieves high detection

accuracy with minimal false positives, especially in scenarios involving re-identification or inference

attacks. The framework offers a data privacy-conscious and context-aware solution for securing

sensitive and anonymized databases.

**ADVANTAGES:** 

Anonymization-Aware Security: Detects SQL attacks that exploit anonymized data, which

traditional systems might miss.

Improved Accuracy via Pattern Mining: Pattern mining helps uncover hidden relationships

in query behavior, improving the detection of complex and indirect attack patterns.

**DISADVANTAGES:** 

Pattern Drift Sensitivity: Attackers can change their strategies, which might render existing

mined patterns ineffective unless frequently updated.

**High Initial Setup Cost:** Mining and analyzing historical data requires time, processing power,

and proper tuning of parameters.

5. PROGESI: A PROXY GRAMMAR TO ENHANCE WEB APPLICATION

FIREWALL FOR SQL INJECTION PREVENTION

AUTHOURS: ANTONIO COSCIA, VINCENZO DENTAMARO, ANTONIO MACI

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2019** 

**ABSTRACT** 

Web Application Firewalls (WAFs) are a frontline defense against SQL Injection (SQLi)

attacks, yet traditional WAFs often rely on static signatures and pattern-matching techniques that fail

to detect obfuscated or zero-day attacks. This study introduces PROGESI (Proxy Grammar for SQL

Injection), a novel approach that augments WAF capabilities using grammar-based detection

mechanisms.

By leveraging a formal grammar model, PROGESI is able to detect deviations from normal

query structures that often indicate SQL injection attempts, regardless of their encoding or obfuscation

techniques. Unlike traditional black-box WAFs, PROGESI interprets the query logic, making it

resilient against advanced attack vectors such as tautologies, piggy-backed queries, and encoded

payloads.

**ADVANTAGES:** 

Grammar-Based Precision: Context-sensitive grammar detection ensures more accurate

identification of malicious query patterns beyond simple signature matching.

Resilience to Obfuscation: Able to detect encoded, concatenated, or semantically altered SQLi

attacks that bypass conventional WAFs.

**DISADVANTAGES:** 

Complex Grammar Design: Developing and maintaining accurate grammar models for

different SQL dialects and applications can be time-consuming.

Potential Performance Overhead: Real-time parsing and grammar validation might introduce

latency, especially under high-traffic conditions.

6. HIDS-IOMT: A DEEP LEARNING-BASED INTELLIGENT INTRUSION

DETECTION SYSTEM FOR THE INTERNET OF MEDICAL THINGS

AUTHOURS: ABDELWAHED BERGUIGA, AHLEM HARCHAY, AYMAN

**MASSAOUDI** 

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2023

**ABSTRACT** 

Web applications are increasingly vulnerable to sophisticated cyberattacks, including SQL

injection, Cross-Site Scripting (XSS), and other injection-based threats. Traditional Web Application

Firewalls (WAFs), which primarily rely on static signature-based detection, often fail to identify zero-

day attacks or obfuscated payloads. This paper proposes a hybrid Machine Learning-based Web

Application Firewall (ML-WAF) that integrates signature detection with anomaly detection,

leveraging **feature extraction techniques** to analyze web request patterns.

Experimental results using real-world traffic and benchmark datasets demonstrate that

PROGESI significantly improves SQLi detection accuracy while reducing false positives.

**ADVANTAGES:** 

Hybrid Detection Capability: Combines known attack detection (signature) with

unknown/zero-day detection (anomaly).

Adaptive and Scalable: ML models can be updated as new data becomes available, making

the WAF adaptive to new attack types.

**DISADVANTAGES:** 

Training and Maintenance Overhead: Requires high-quality, labeled data for training and

regular updates to remain effective.

Complex Feature Engineering: Designing and tuning the feature set is time-consuming and

may require domain expertise.

7. EMPIRICAL EVALUATION OF ATTACKS AGAINST IEEE 802.11

ENTERPRISE NETWORKS: THE AWID3 DATASET

**AUTHOURS:** EFSTRATIOS CHATZOGLOU, GEORGIOS KAMBOURAKIS

**PUBLISHED BY:** IEEE ACCES.

**YEAR:** 2016

**ABSTRACT** 

This work serves two key objectives. First, it markedly supplements and extends the well

AWID corpus by capturing and studying traces of a wide variety of attacks hurled in the IEEE 802.1X

Extensible Authentication Protocol (EAP) environment. Second, given that all the 802.11-oriented

attacks have been carried out when the defenses introduced by Protected Management Frames (PMF)

were operative, it offers the first to our knowledge full-fledged empirical study regarding the

robustness of the IEEE 802.11w amendment, which is mandatory for WPA3 certified devices. Under

both the aforementioned settings, the dataset and study at hand are novel and are anticipated to be of

significant aid towards designing and evaluating intrusion detection systems.

**ADVANTAGES:** 

Realistic Testbed: The dataset was collected using a realistic testbed setup, ensuring that the

captured data reflects real-world scenarios and device behaviors.

Support for Intrusion Detection Research: By providing detailed documentation and a

variety of attack scenarios, AWID3 serves as a valuable resource for designing and evaluating

intrusion detection systems (IDS).

**DISADVANTAGES:** 

Limited Scope on Physical Layer Attacks: The dataset does not cover physical (PHY) layer

attacks, focusing instead on MAC layer and above. This limits research on PHY-specific.

Potential for Imbalanced Data: As with many intrusion detection datasets, there may be an

imbalance between normal and attack traffic, which can affect the training and evaluation of

machine learning.

8. SURVEY: INTRUSION DETECTION SYSTEM IN SOFTWARE-DEFINED

**NETWORKING** 

**AUTHOURS:** AHMED H. JANABI, TRIANTAFYLLOS KANAKIS

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2023

**ABSTRACT** 

Software-Defined Networking (SDN) has emerged as a transformative networking paradigm

that decouples the control plane from the data plane, offering centralized control and programmability.

This architectural shift presents both opportunities and challenges for network security. Intrusion

Detection Systems (IDS) play a crucial role in identifying malicious activities and potential threats

within SDN environments. This survey reviews the current landscape of IDS in SDN, analyzing

various detection techniques, architectures, and their applicability in dynamic and programmable

networks. It also examines the integration of machine learning, deep learning, and hybrid approaches

for enhanced threat detection. Furthermore, the paper discusses performance evaluation metrics, real-

world deployments, and outlines future research directions to address existing limitations.

**ADVANTAGES:** 

Centralized Monitoring: SDN's centralized control enables global network visibility,

allowing IDS to analyze traffic patterns more effectively.

Programmability: IDS can be dynamically updated or reconfigured without manual

intervention across multiple devices.

**DISADVANTAGES:** 

Single Point of Failure: The centralized controller is a critical component; if compromised, the

entire network's security is at risk.

**High Latency:** Centralized analysis might introduce latency, especially during high-volume

traffic or complex processing.

9. RESEARCH INTO THE SECURITY THREAT OF WEB APPLICATION

**AUTHOURS:** Yanling Zhang, Ting Zhang

**PUBLISHED BY:** IEEE ACCES.

**YEAR:** 2022

**ABSTRACT** 

Web applications have become a critical component of modern digital infrastructure,

enabling seamless interaction and service delivery across various sectors. However, their growing

complexity and constant connectivity make them prime targets for cyber threats. This research

explores the security vulnerabilities inherent in web applications, including common attack vectors

such as SQL injection, cross-site scripting (XSS), and cross-site request forgery (CSRF). It analyzes

real-world case studies, threat modeling approaches, and current defensive mechanisms. The study

also evaluates the effectiveness of modern security frameworks, penetration testing, and secure coding

practices. By identifying prevalent threats and evaluating countermeasures, this research aims to

enhance the security posture of web applications and inform best practices in development and

deployment.

**ADVANTAGES:** 

Improved Risk Awareness: Helps developers and organizations understand the most pressing

security threats.

Proactive Defense Strategies: Encourages implementation of preventive measures like input

validation and secure session handling.

**DISADVANTAGES:** 

Complexity in Implementation: Securing web applications can require significant changes to

design and development processes.

Resource Intensive: Security testing, audits, and ongoing monitoring require time, expertise,

and financial resources.

10. A SYSTEMATIC LITERATURE REVIEW ON THE CHARACTERISTICS

AND EFFECTIVENESS OF WEB APPLICATION VULNERABILITY

**SCANNERS** 

**AUTHOURS: SULIMAN ALAZMI, DANIEL CONTE DE LEON** 

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2012

**ABSTRACT** 

Web Application Vulnerability Scanners (WAVS) are essential tools in identifying and

mitigating security flaws in web-based systems. As cyber threats become increasingly sophisticated,

the reliability and accuracy of these scanners are critical for ensuring robust security. This systematic

literature review examines the characteristics, detection capabilities, and limitations of popular

WAVS tools. The study categorizes scanners based on scanning techniques, such as static analysis,

dynamic analysis, and hybrid approaches. It evaluates their effectiveness against common

vulnerabilities listed in the OWASP Top 10, considering metrics like detection rate, false positives,

and scanning depth. The review also explores usability factors, integration capabilities, and

adaptability to evolving web technologies. Ultimately.

**ADVANTAGES:** 

Automated Vulnerability Detection: Scanners provide rapid, automated identification of

common web vulnerabilities, saving time and effort.

Improved Security Posture: Regular scanning helps maintain secure web applications by

identifying risks before exploitation.

**DISADVANTAGES:** 

False Positives and Negatives: Scanners may report incorrect results, requiring manual

verification and increasing workload.

Limited Context Awareness: Scanners may struggle with dynamic content, JavaScript-heavy

pages, or complex logic flows.

11. ONLINE BANKING USER AUTHENTICATION METHODS:

SYSTEMATIC LITERATURE REVIEW

**AUTHOURS:** NADER ABDEL KARIM, HASAN KANAKER

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2019** 

**ABSTRACT** 

The security of online banking systems heavily relies on robust user authentication

mechanisms to protect against fraud, identity theft, and unauthorized access. This systematic literature

review examines the evolution, classification, and effectiveness of user authentication methods in

online banking environments. The study explores various authentication categories, including

knowledge-based (passwords, PINs), possession-based (tokens, smart cards), and biometric-based

(fingerprints, facial recognition) techniques. Additionally, it evaluates emerging methods such as

behavioral biometrics and multi-factor authentication (MFA). The review assesses each approach

based on usability, security strength, implementation complexity, and resistance to common attack

vectors. By synthesizing insights from academic and industry sources, this review identifies trends.

**ADVANTAGES:** 

Enhanced Security: Multi-factor and biometric authentication methods provide stronger

protection against unauthorized access.

User Convenience: Biometric and mobile-based authentication can offer seamless and quick

login experiences.

**DISADVANTAGES:** 

Usability Issues: Complex or multi-step authentication can frustrate users and lead to poor user

experience.

**Privacy Concerns:** Biometric data, once compromised, cannot be changed and raises serious

privacy implications.

12. ACROSS THE SPECTRUM IN-DEPTH REVIEW AI-BASED MODELS

FOR PHISHING DETECTION

**AUTHOURS:** SHAKEEL AHMAD, RAHIEL AHMAD, ISMAIL ERGEN

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2023

**ABSTRACT** 

Phishing attacks remain one of the most prevalent and damaging forms of cybercrime,

targeting individuals and organizations through deceptive communication techniques. With the

growing sophistication of phishing tactics, traditional detection methods often fall short. This in-depth

review explores the landscape of Artificial Intelligence (AI)-based models developed for phishing

detection, covering a wide spectrum of machine learning (ML), deep learning (DL), and hybrid

approaches. The paper categorizes models based on input features such as URL characteristics, email

metadata, website content, and behavioral patterns. It evaluates model performance using metrics like

accuracy, precision, recall, and false positive rates. Furthermore, the review examines the strengths

and limitations of various AI techniques including decision trees, random forests, support vector

machines, neural networks, and ensemble methods.

**ADVANTAGES:** 

**High Detection Accuracy**: AI models can identify subtle patterns in phishing content that

traditional methods might miss.

Real-Time Detection: AI enables fast, automated decision-making, crucial for early-stage

phishing attack.

**DISADVANTAGES:** 

Complexity in Setup: Defining valid query patterns for complex systems can be time-

consuming.

Performance Overhead: Real-time parsing and analysis may slightly degrade system

performance, especially under high traffic.

13. PHISHCATCHER: CLIENT-SIDE DEFENSE AGAINST WEB SPOOFING

ATTACKS USING MACHINE LEARNING

AUTHOURS: MUZAMMIL AHMED, AAKASH AHMAD, WILAYAT KHAN

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2021** 

**ABSTRACT** 

Phishing attacks continue to pose a major threat to internet users, exploiting web spoofing

techniques to deceive individuals into divulging sensitive information. "PhishCatcher" introduces a

client-side defense mechanism that leverages machine learning to detect and prevent phishing attacks

in real-time. Unlike traditional blacklist-based or heuristic systems, PhishCatcher uses a trained

classifier to analyze various features of web pages—such as URL structure, DOM elements, visual

similarities, and SSL certificate anomalies—to identify potentially malicious sites. The model runs

locally in the user's browser or as a lightweight extension, offering proactive protection without the

need for constant server communication. Through extensive testing on large datasets of phishing and

legitimate websites, Phish Catcher demonstrates high accuracy and low false positive rates, making

it a practical and efficient solution for everyday users.

**ADVANTAGES:** 

**Real-Time Protection:** Detects phishing attempts instantly at the client-side without relying on

server-side lookups.

**Privacy-Preserving:** Keeps user data on the client, avoiding privacy concerns related to

sending browsing data to external servers.

**DISADVANTAGES:** 

Model Drift: As phishing tactics evolve, the machine learning model may need frequent

retraining to stay effective.

Resource Usage: On-device models may increase CPU/memory usage, especially on lower-

end systems.

14. SPARQ: A CYBER-RESILIENT VOLTAGE REGULATION USING SOFT

Q-LEARNING APPROACH FOR AUTONOMOUS GRID OPERATIONS

**AUTHOURS: MOHAMED MASSAOUDI** 

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2022

**ABSTRACT** 

The increasing integration of distributed energy resources (DERs) and the rising threat of

cyberattacks demand robust, intelligent, and adaptive solutions for power grid voltage regulation.

SPARQ (Soft Q-learning-based Autonomous Regulation for Quality voltage) introduces a novel,

cyber-resilient voltage control framework leveraging Soft Q-Learning (SQL), a reinforcement

learning technique known for its stability and robustness under uncertainty. SPARQ enables

decentralized, autonomous decision-making among smart grid components, allowing for adaptive

voltage regulation even in adversarial or fault-prone conditions. By learning optimal control strategies

through interaction with the environment, SPARQ effectively mitigates cyber-physical risks while

improving grid reliability, responsiveness, and operational efficiency.

**ADVANTAGES:** 

Cyber-Resilience: SPARQ is designed to detect and respond to cyber threats, enhancing the

grid's ability to maintain operations during attacks.

Autonomous Decision-Making: The use of Soft Q-Learning allows components to make

intelligent, decentralized decisions without relying on constant communication with a central

controller.

**DISADVANTAGES:** 

Computational Complexity: Training SQL models can be resource-intensive and time-

consuming, especially in large-scale systems.

Implementation Overhead: Requires retrofitting or updating existing grid infrastructure with

smart components capable of learning and decision-making.

15. REAL-TIME HEALTHCARE RECOMMENDATION SYSTEM FOR

SOCIAL MEDIA PLATFORMS

**AUTHOURS:** E. MARUTHAVANI, S. P. SHANTHARAJAH

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2013

**ABSTRACT** 

With the rise in health-related discussions on social media platforms, there is an opportunity

to harness user-generated content for delivering personalized healthcare recommendations in real

time. This paper proposes a Real-Time Healthcare Recommendation System (RTHRS) integrated

with social media platforms to monitor, analyze, and interpret users' health-related posts using natural

language processing (NLP), sentiment analysis, and medical knowledge graphs. The system provides

users with instant, context-aware suggestions, such as lifestyle tips, early warning signs, and when to

seek professional care. The platform leverages deep learning models for user profiling and

collaborative filtering to tailor recommendations while maintaining privacy and ethical guidelines.

The system aims to bridge the gap between informal health discussions and actionable healthcare

guidance, promoting early intervention and public health awareness.

**ADVANTAGES:** 

Real-Time Response: Provides instant recommendations based on current user behavior and

posts.

**User Engagement:** Meets users where they are — on social media — increasing the reach and

impact of health advice.

**DISADVANTAGES:** 

Privacy and Ethical Concerns: Analyzing user content may raise issues around consent, data

misuse, and surveillance.

Misinformation Risk: System might misinterpret sarcasm, humor, or figurative language

common on social media.

16. IDENTIFICATION OF SQL INJECTION SECURITY VULNERABILITIES

IN WEB APPLICATIONS BASED ON BINARY CODE SIMILARITY

**AUTHOURS: JIANHUA WANG** 

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2020** 

**ABSTRACT** 

SQL injection remains one of the most critical security vulnerabilities in web applications,

enabling attackers to gain unauthorized access to databases by manipulating user inputs. Traditional

detection approaches often rely on static or dynamic analysis of source code, which may be

unavailable or obfuscated in real-world scenarios. This paper presents a novel method for identifying

SQL injection vulnerabilities by analyzing binary code similarity. By comparing compiled binaries

of target applications with known vulnerable code patterns, the system can detect potential injection

points without access to source code. Using graph-based models and machine learning techniques,

the approach extracts semantic features from binary code, enabling precise and scalable vulnerability

detection.

**ADVANTAGES:** 

Works Without Source Code: Ideal for auditing proprietary or legacy applications where

source code is not available.

**Resilient to Obfuscation:** Binary-level analysis can uncover vulnerabilities even in heavily

obfuscated or packed code.

**DISADVANTAGES:** 

High Computational Cost: Binary analysis, especially involving similarity matching, can be

resource-intensive and slow.

False Positives/Negatives: Similar code structures may not always imply the same

vulnerabilities, potentially leading to misclassification.

17. SYNTHESIS OF ALLOWLISTS FOR RUNTIME PROTECTION AGAINST

**SQLI** 

**AUTHOURS:** Neel Gandhi, Jaykumar Patel, Rajdeepsinh Sisodiya, Nishant Doshi

**PUBLISHED BY: IEEE ACCES.** 

**YEAR:** 2023

**ABSTRACT** 

SQL injection (SQLi) continues to be a critical threat to the security of web applications, often

resulting in data breaches and system compromise. Traditional input validation and filtering

techniques are prone to bypass by skilled attackers. This paper presents a novel approach for runtime

protection against SQLi attacks through the automatic synthesis of allowlists — predefined sets of

permissible SQL query structures derived from legitimate application behavior. By analyzing normal

query patterns during application execution, the system builds allowlists representing safe and

expected query templates. At runtime, any deviation from these templates triggers alerts or blocks the

query. This lightweight, language-agnostic technique enhances security without requiring source code

modification. Evaluations on real-world web applications demonstrate.

**ADVANTAGES:** 

Strong Runtime Protection: Immediately blocks unexpected or malicious SQL queries that

don't match the known-good patterns.

Language-Independent: Can work across applications built in different programming

languages, as it monitors SQL queries rather than code.

**DISADVANTAGES:** 

Cold Start Problem: Initially, the allowlist may be incomplete, possibly blocking legitimate

but unobserved query patterns.

Maintenance Overhead: Applications that frequently change database logic may require

retraining or manual updates to the allowlist.

18. EFFECTIVE FILTER FOR COMMON INJECTION ATTACKS IN

ONLINE WEB APPLICATIONS

**AUTHOURS: SANTIAGO IBARRA-FIALLOS** 

**PUBLISHED BY: IEEE ACCES.** 

**YEAR: 2017** 

**ABSTRACT** 

Injection attacks, such as SQL injection, cross-site scripting (XSS), and command injection,

remain among the most prevalent security threats targeting online web applications. These attacks

exploit insufficient input validation and sanitization to manipulate application behavior or

compromise data integrity. This paper presents an Effective Filtering System that provides a unified

defense against multiple types of injection attacks. The system employs a hybrid approach combining

pattern-based detection, context-aware input sanitization, and machine learning classification to

identify and block malicious inputs in real time. It dynamically adapts to application-specific contexts

such as query construction, HTML rendering, or system command execution, improving accuracy

while minimizing false positives. Tested across various real-world web environments, the proposed

filter demonstrates high detection rates.

**ADVANTAGES:** 

Multi-Injection Protection: Guards against a wide range of injection attacks, including SQLi,

XSS, and command injection.

Context-Aware Filtering: Adapts input validation based on where the data is used (e.g.,

database, HTML, shell), reducing false positives.

**DISADVANTAGES:** 

False Positives in Edge Cases: May incorrectly block legitimate requests if user input closely

resembles known attack patterns.

Maintenance Burden: Requires regular updates to detection patterns and retraining of models

to stay current with new attack vectors.