SQL INJECTION ATTACK DETECTION

PRESENTED BY

NAME: AVINASH M
REG. NO: 22011101019
CLASS: AI & DS ‘A’ (3RD YEAR)

INTRODUCTION

- SQL Injection (SQLi) is an attack that manipulates SQL queries to gain unauthorized access.

- SQL injection attacks exploit vulnerabilities in web applications that allow attackers to insert malicious SQL code into database queries

- SQL injection is a major threat. It can lead to serious data breaches and compromise sensitive information. These attacks can enable attackers to bypass authentication, steal sensitive data, modify or delete database records, and even gain control of the server.

- Objective: To Develop an AI-driven system that can automatically detect and mitigate SQLi attempts in real-time.

EXISTING SYSTEM

TRADITIONAL DETECTION MECHANISMS:

- Signature-based Intrusion Detection Systems (IDS): Monitors network traffic or system activities for malicious activity or policy violations.

- Web Application Firewalls (WAF) with predefined rule sets: Filters and monitors HTTP requests to block malicious SQLi attempts but struggles with sophisticated evasion techniques.

- Query sanitization using prepared statements: This technique separates SQL code from user-supplied data, preventing malicious input from being interpreted as executable commands.

LIMITATIONS:

- Unable to detect zero-day SQLi attacks.
- High dependency on human-defined rules.
- Computationally expensive and prone to false positives.

PROPOSED SYSTEM

(Machine Learning Based SQLi Detection)

MODEL BASED APPROACH:

- Train ML models using labelled datasets containing benign and malicious SQL queries.

- Feature extraction from SQL queries (lexical analysis, tokenization, embedding representations).

- Apply classification models such as Random Forest, Support Vector Machine (SVM), and Deep Learning models (LSTMs, CNNs).

WORKFLOW:

1. Query Preprocessing: Tokenization and feature extraction.

2. Feature Vectorization: Converting SQL queries into numerical representations.

3. Model Inference: Classify queries as benign or malicious using trained ML models.

4. Real-Time Mitigation: Flagging, blocking, or logging detected SQLi attempts.

ADVANTAGES:

- Self-learning capabilities adapt to new SQLi attack techniques.

- Reduces false positives compared to traditional methods.

- Scalable and deployable in web application firewalls (WAFs) and database security layers.

MODULES

- **Data Collection & Preprocessing:** Use publicly available SQLi datasets (e.g., CSIC 2010, Kaggle SQLi datasets). Apply NLP-based preprocessing (removing stopwords, tokenization, stemming).

- **Feature Engineering:** Extract statistical and structural features from SQL queries (query length, special character usage, token frequency). Transform queries using TF-IDF, Word Embeddings, or Graph Neural Networks.

- **Model Training & Evaluation**: Train supervised classifiers (Logistic Regression, Random Forest, CNNs, LSTMs).Use metrics like Precision, Recall, F1-score for performance evaluation.

- **Detection and Alerting:** This module analyzes incoming SQL queries using the trained model and generates alerts for suspected attacks. It also logs the detected attacks for further analysis and improvement of the system.

- **Database Integration:** This module allows the system to monitor and analyze SQL queries in real-time as they interact with the database. This ensures that the detection process is integrated seamlessly into the existing database environment.

AI-powered SQL Injection detection enhances cybersecurity by providing adaptive, scalable, and high-accuracy protection against evolving attack vectors.