



Trường đại học Công nghệ  
thông tin - ĐHQG TP. Hồ  
Chí Minh



Khoa Mạng máy tính và  
Truyền thông

# Hệ thống phát hiện và phân tích tấn công SQL Injection dựa trên hành vi người dùng

Khóa luận tốt nghiệp

24/06/2025

Cán bộ hướng dẫn

ThS. Tô Nguyễn Nhật Quang  
TS. Nguyễn Ngọc Tự

Sinh viên thực hiện

Nguyễn Văn Minh Toàn  
MSSV: 21522687  
Lớp: MMCL2021

# Nội dung

## 1. Tổng quan đề tài

- Lý do chọn đề tài
- Mục đích
- Phạm vi nghiên cứu

## 2. Cơ sở lý thuyết

- SQL Injection
- Công nghệ sử dụng
- Thực nghiệm tấn công

## 3. Phân tích thiết kế hệ thống

- Kiến trúc hệ thống
- Mô hình máy học

## 4. Hiện thực hệ thống

- Ứng dụng web giám sát

## 5. Thực nghiệm, đánh giá

- Phát hiện SQL Injection
- Đánh giá hiệu suất

## 6. Kết luận, hướng phát triển

- Những kết quả đạt được
- Những điểm còn hạn chế
- Hướng phát triển

# 1. Tổng quan đề tài

## Lý do chọn đề tài

- Tấn công mạng ngày càng phổ biến, đặc biệt trên ứng dụng web.
- SQL Injection là lỗ hổng nguy hiểm, tồn tại dai dẳng.
- AI giúp phát hiện hành vi bất thường thay vì chỉ dựa vào chữ ký.

## Mục đích của đề tài

- Thiết kế hệ thống phát hiện tấn công SQL Injection.
- Mô hình (CNN) để nhận diện sớm truy vấn chứa mã độc.
- Triển khai dashboard trực quan: theo dõi IP/Port, biểu đồ tấn công, cảnh báo.

## Các hướng nghiên cứu

- Có 2 hướng tiếp cận chính:
- Công cụ quét lỗ hổng: Dùng SQLMAP, Nikto, SQLi-labs để phát hiện điểm yếu và mô phỏng tấn công.
  - Học máy (ML): Ứng dụng SVM, Random Forest, Neural Network để phân tích truy vấn và hành vi tấn công.

# 1. Tổng quan đề tài

## Các nghiên cứu ngoài nước liên quan

- Krishnan et al.: So sánh các thuật toán (CNN, SVM, Naive Bayes...). CNN đạt độ chính xác cao nhất (97%).
- Alkhatami & Alzahrani: ML trên nền tảng cloud, SVM đạt 99.42% accuracy. Mô hình phù hợp môi trường điện toán đám mây.
- Farooq: Áp dụng ensemble learning (LightGBM, XGBoost...), accuracy đạt 99.3%. Hiệu quả trong môi trường thực tế.

## Các nghiên cứu trong nước liên quan

- Trần Quang Chung (2022) – Luận văn Thạc sĩ:
  - Sử dụng các khuôn mẫu tổng quát để phát hiện và ngăn chặn SQLi.
- Đặng Thị Ngọc Tuyết (2024) – Luận văn Thạc sĩ:
  - Phân tích tự động website phát hiện cả SQLi và XSS.
- Phan Trung Hiếu & Đỗ Hữu Tú (2023) – Luận văn Cử nhân:
  - Ứng dụng học máy để phân tích câu lệnh SQL phát hiện tấn công.

# 2. Cơ sở lý thuyết

## SQL Injection

- SQL Injection là hình thức tấn công chèn mã độc SQL vào trường nhập liệu của web.
- Mục tiêu: can thiệp truy vấn cơ sở dữ liệu.
- Truy cập trái phép **đọc, thêm, xóa, sửa** dữ liệu.
- Vẫn là mối đe dọa hàng đầu đối với ứng dụng web sử dụng hệ CSDL như MySQL, Oracle, SQL Server...

## Phân loại tấn công SQL Injection

Có 9 dạng theo nghiên cứu của Waad Almadhy, Amal Alruwaili và Saloua Hendaoui năm 2022:

- First-order Injection
- Second-order Injection
- Illegal / Logically Incorrect Queries
- Tautologies
- Union Query
- Piggy-backed Queries
- Stored-Procedure Injection
- Time-based Blind

## 2. Cơ sở lý thuyết

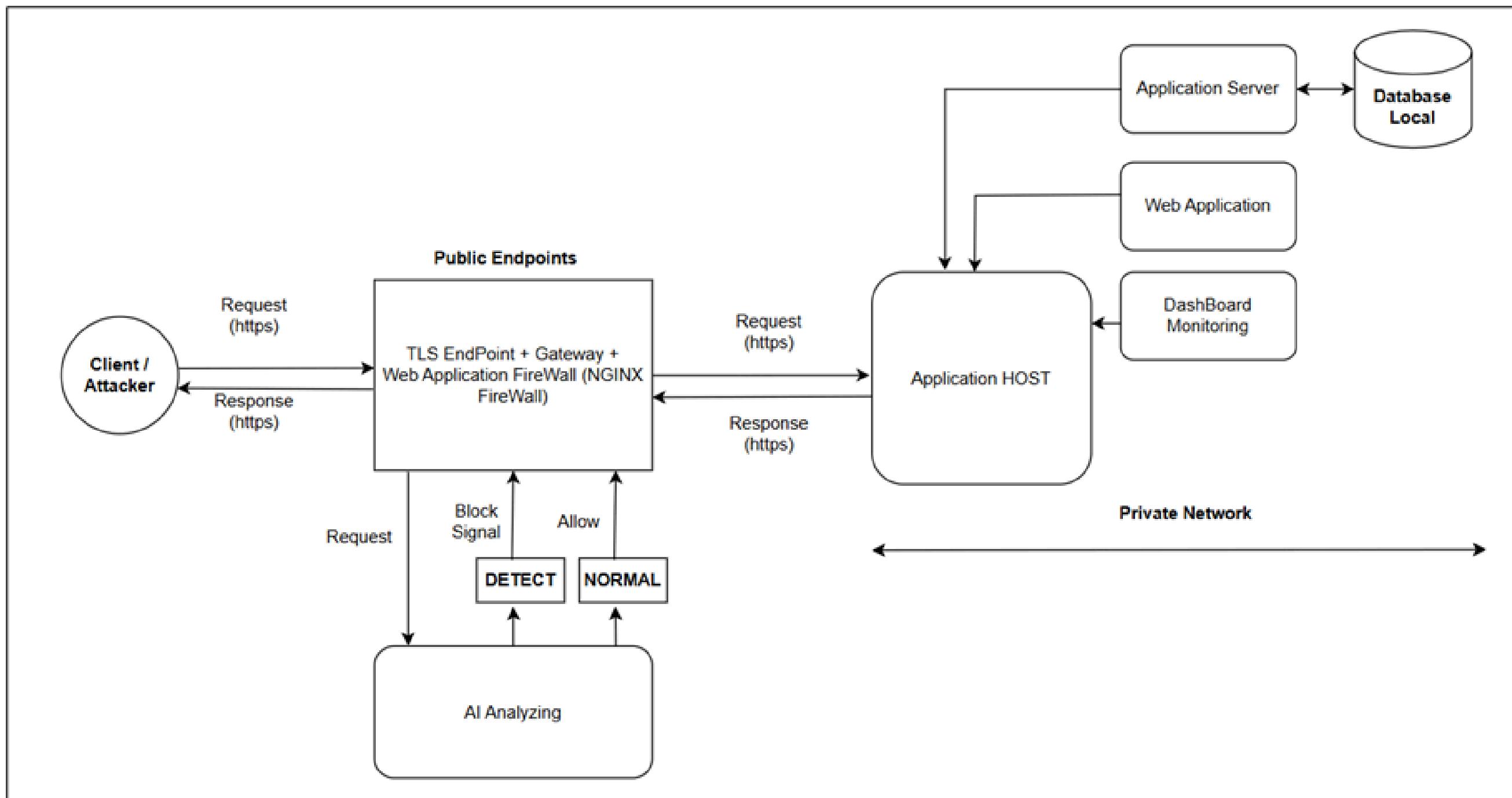
**Các công nghệ đã sử dụng trong luận văn:**

- JavaScript:
  - Triển khai mô hình MERN Stack.
- Python:
  - Machine Learning.
- MongoDB
  - Lưu trữ dữ liệu.
- Docker
  - Đóng gói và chạy ứng dụng độc lập.
- NGINX
  - Triển khai Gateway, TLS/SSL, WAF.

**Thử nghiệm tấn công SQL Injection**

- ZAP
  - Tấn công trên OWASP Juice Shop
- Tấn công với SQLmap (trên DVWA)

# 3. Phân tích thiết kế hệ thống



# 3. Phân tích thiết kế hệ thống

## Client:

Là người dùng cuối hoặc kẻ tấn công.

- Gửi yêu cầu đến hệ thống, có thể chứa mã độc như SQL Injection.
- Dù sử dụng HTTPS, dữ liệu vẫn có thể bị chèn mã độc → cần thêm lớp bảo vệ phía sau.

## Public Endpoint

- TLS Endpoint: Mã hóa dữ liệu, đảm bảo an toàn trong truyền tải.
- Web Application Firewall (WAF): Lọc và kiểm tra yêu cầu, ngăn chặn các tấn công đã biết hoặc bất thường.
- AI Analyzing: Phân tích hành vi truy cập để quyết định chặn hoặc cho phép.

## Public Endpoint

- Tiếp nhận các yêu cầu đã được lọc.
- Gồm các thành phần: Application Server, Web App, Dashboard Monitoring, Database Local.
- Thực hiện xử lý nghiệp vụ và lưu trữ nội dung một cách an toàn.

# 3. Phân tích thiết kế hệ thống

## Kiến trúc CNN được đề xuất:

- Sử dụng Convolutional Neural Network (CNN) để phân loại truy vấn SQL là benign hay malicious.
- Kiến trúc gồm:
  - 3 lớp Conv2D với số filters tăng dần:  $64 \rightarrow 128 \rightarrow 256$ , dùng kernel  $3 \times 3$  + ReLU + MaxPooling( $2 \times 2$ ).
  - 4 lớp Dense:  $256 \rightarrow 128 \rightarrow 64 \rightarrow 1$  với ReLU và sigmoid (đầu ra xác suất).

# 4. Hiện thực hệ thống

## Ứng dụng web giám sát – Injex Watch

- Injex Watch là ứng dụng giám sát trung tâm của hệ thống phát hiện tấn công SQLi.

### Dashboard

- Hiển thị danh sách thiết bị đang được giám sát
- Các thông tin: tên thiết bị, IP, port, trạng thái (màu sắc), thời gian cập nhật
- Chức năng: thêm xóa sửa thiết bị.

The screenshot shows a dark-themed web application interface titled "GIÁM SÁT" (Monitoring). At the top, there is a search bar and a user profile icon labeled "Toàn 21522687". Below the title, it says "Danh sách các thiết bị" (List of devices) and has a "DĂNG KÝ" (Register) button. The main area is a table with the following data:

Tên thiết bị	Địa chỉ IP	Port	Trạng thái	Thêm ngày	Cập nhật cuối	Thao tác	
Device 1	10.10.10.1	4545	● inactive	20:50:51 20/06/2025	20:50:51 20/06/2025		
Device 2	192.168.48.128	5011	● inactive	20:50:31 20/06/2025	20:50:31 20/06/2025		
Device 3	172.27.112.1	5050	● active	20:49:58 20/06/2025	20:49:58 20/06/2025		
Device 4	172.27.222.3	5555	● inactive	20:51:23 20/06/2025	20:51:23 20/06/2025		

# 4. Hiện thực hệ thống

## Danh sách đen:

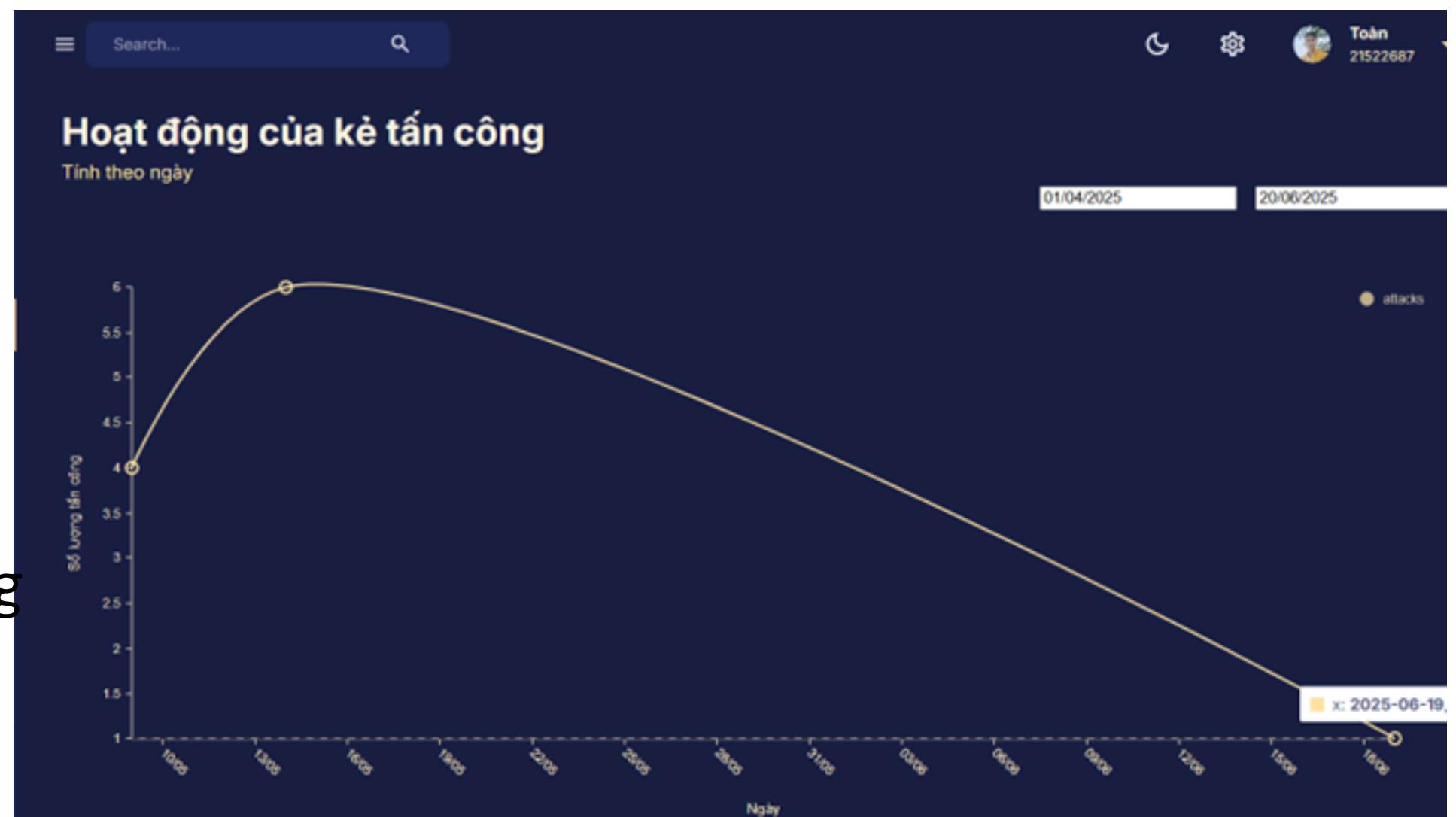
- Lưu các IP bị phát hiện tấn công (qua ML)
- Hiển thị: ID, IP, vị trí địa lý, thời gian tấn công gần nhất
- Cho phép xem vị trí IP trên bản đồ (Leaflet)
- Hỗ trợ: chặn IP, cảnh báo, cập nhật tường lửa

ID	IP Address	Location	Last Attack
6855620b04282350cd539dfc	196.110.41.59	Lat: 10.822, Long: 106.6257	20:28:43 20/06/2025
6825a9ccb95e89d1284b49a5	153.156.176	Lat: 10.822, Long: 106.6257	15:46:04 15/05/2025
6825a8c9b95e89d1284b4995	153.156.176	Lat: 10.822, Long: 106.6257	15:41:45 15/05/2025
68257266b20bc13ea1fabc5e	153.156.176	Lat: 10.822, Long: 106.6257	11:49:42 15/05/2025
682572288e41863d81d73f15	153.156.176	Lat: 10.822, Long: 106.6257	11:48:40 15/05/2025
682571a6e8baa5b113f7bfa2	153.156.176	Lat: 0, Long: 0	11:46:30 15/05/2025
68255d510734d152955e39ad	153.156.176	Lat: 10.822, Long: 106.6257	10:19:45 15/05/2025
681ea2b88caa845aae6bac88	14.161.3.129	Lat: 10.822, Long: 106.6257	07:50:00 10/05/2025

# 4. Hiện thực hệ thống

## Thống kê:

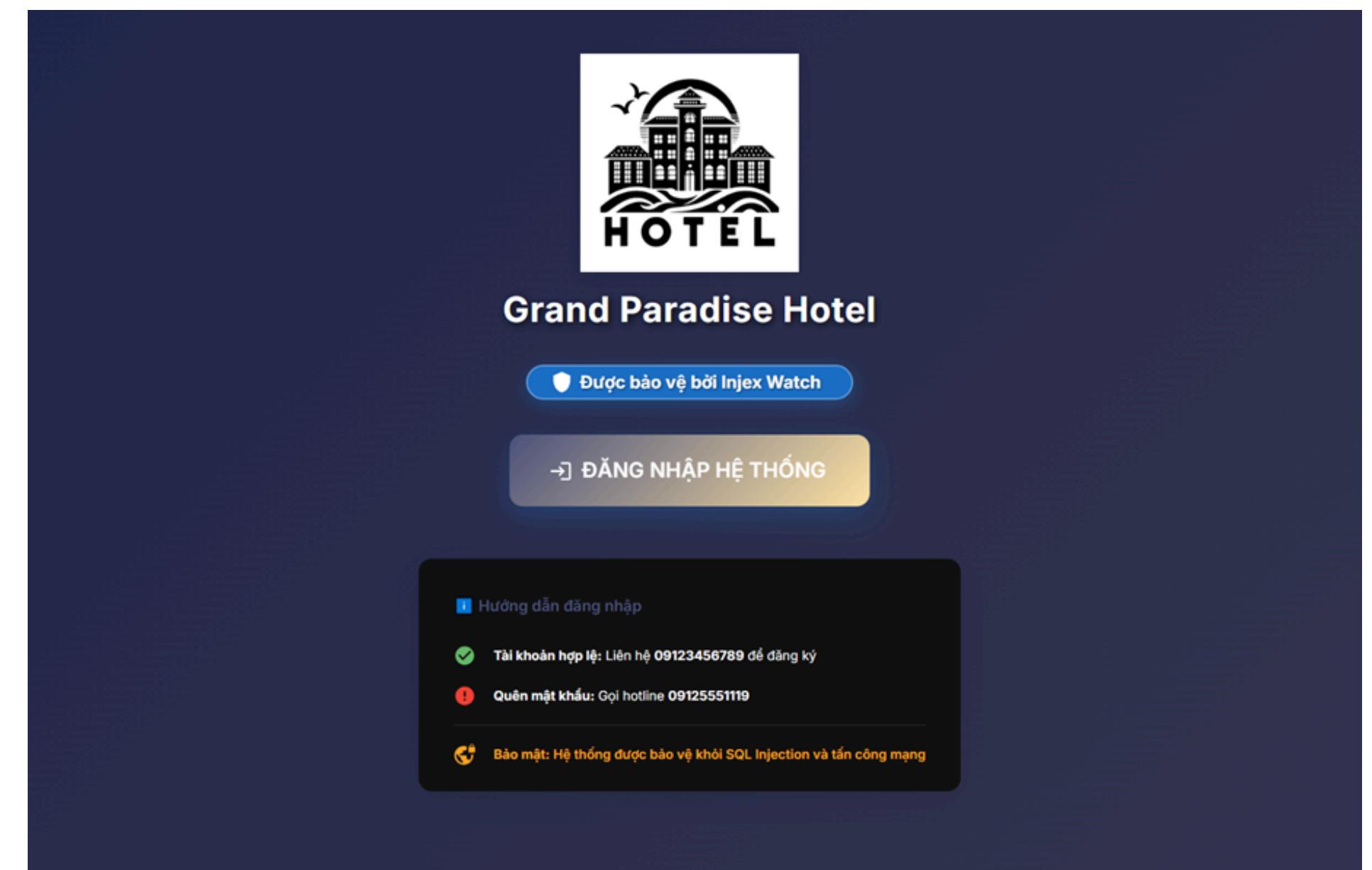
- Hiển thị biểu đồ số lượng tấn công theo ngày
- Dạng biểu đồ đường (line chart), dễ quan sát
- Có thể chọn thời gian thống kê linh hoạt
- Giúp đánh giá xu hướng và hiệu quả phòng thủ



# 5. Thực nghiệm đánh giá, bàn luận

## Triển khai phần mềm bảo mật trên trang web thực tế

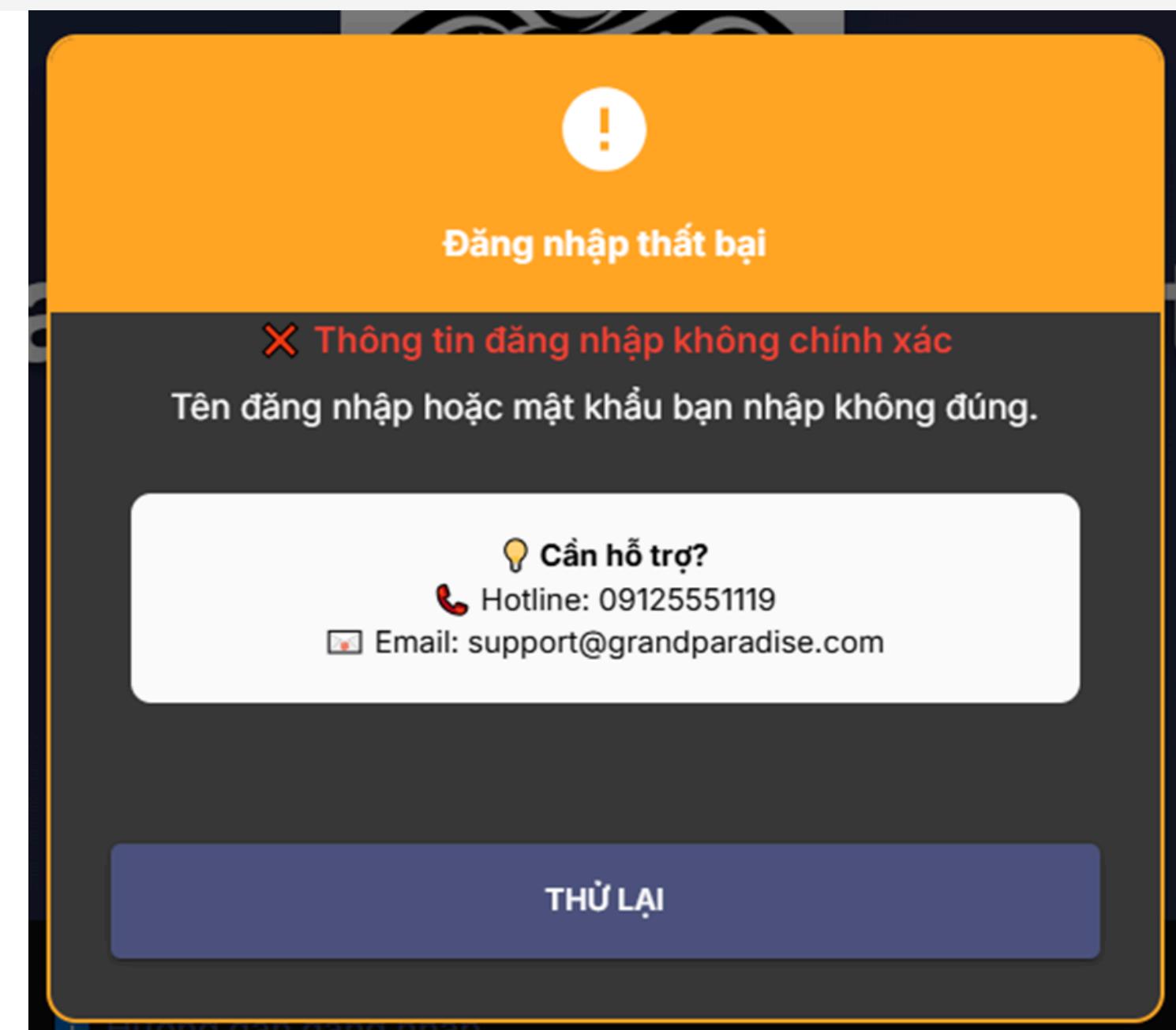
- Trang web thử nghiệm: Grand Paradise Hotel – có giao diện đăng nhập (username, password).



# 5. Thực nghiệm đánh giá, bàn luận

## Tình huống kiểm thử

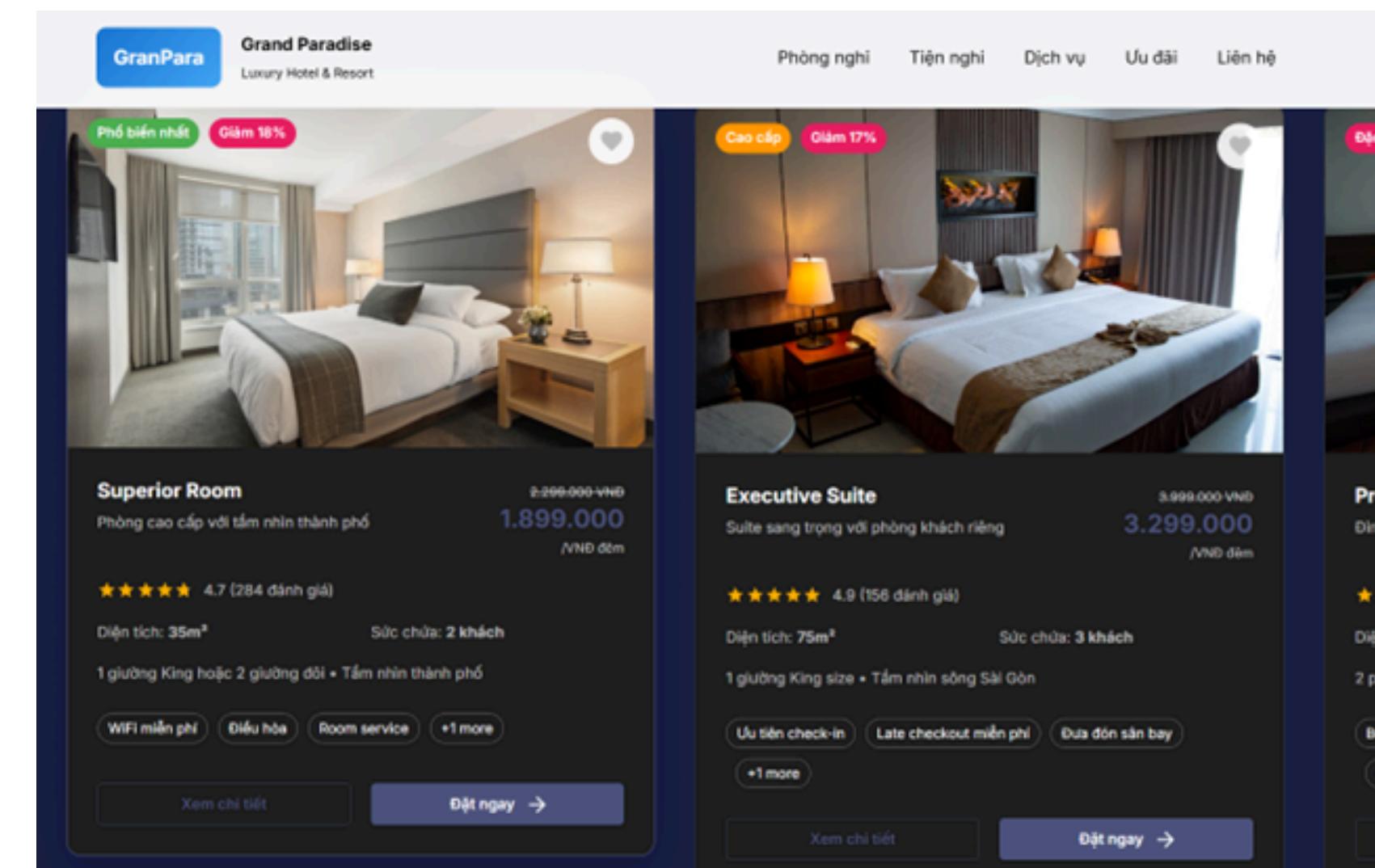
- Đăng nhập bình thường
  - Nếu sai username/password → hiển thị thông báo lỗi, vẫn kiểm tra qua AI.
  - Nếu đúng thông tin → đăng nhập thành công, xác thực qua cookie bảo mật (Secure cookie).



# 5. Thực nghiệm đánh giá, bàn luận

## Tình huống kiểm thử

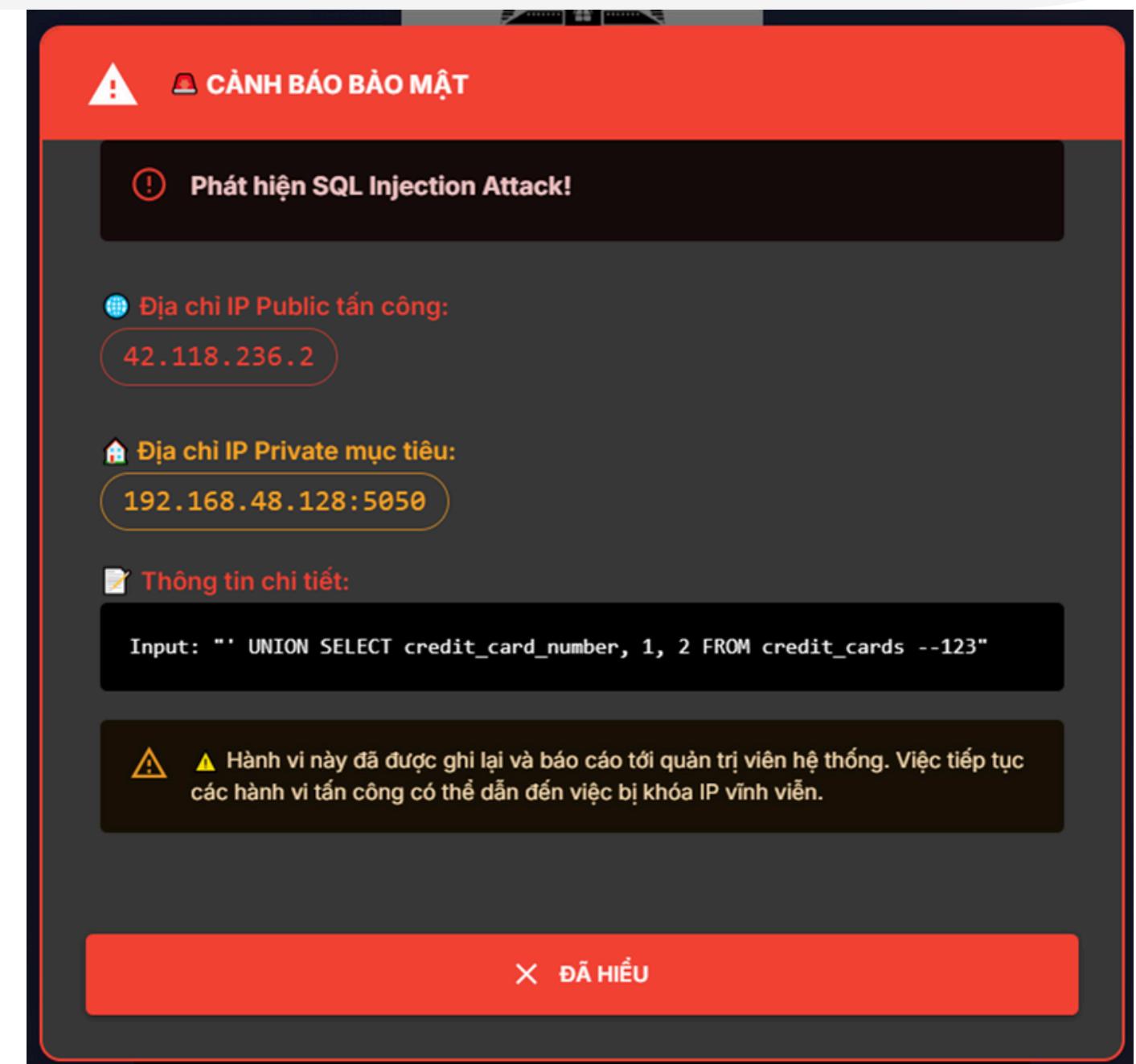
- Đăng nhập bình thường
  - Nếu đúng thông tin → đăng nhập thành công, xác thực qua cookie bảo mật (Secure cookie).



# 5. Thực nghiệm đánh giá, bàn luận

## Tấn công SQLi bị chặn

- Phát hiện qua CNN → Chặn truy vấn.
  - Hiển thị cảnh báo:
  - IP public của attacker
  - IP nội bộ máy bị tấn công
  - Mã SQL độc hại
  - Thông báo nguy cơ bị khóa IP



# 5. Thực nghiệm đánh giá, bàn luận

## Trên giao diện Injex Watch

- Thiết bị chuyển sang trạng thái: Under Attack (SQL Injection).
- Hiển thị:
  - Thiết bị: GranPara
  - IP: 192.168.56.1
  - Port: 5050
- Ghi nhận:
  - IP tấn công: 42.118.236.2
  - Tọa độ: Lat: 10.822, Long: 106.6257
  - Thời gian tấn công: 03:34:23 – 23/06/2025

Tên thiết bị	Địa chỉ IP	Port	Trạng thái
GranPara	192.168.56.1	5050	Online
Device 2	192.168.253.241	5051	Offline

IP Address	Location
42.118.236.2	Lat: 10.822, Long: 106.6257
42.118.236.2	Lat: 10.823, Long: 106.63
171.242.199.66	Lat: 9.9384, Long: 106.3455
171.242.199.66	Lat: 9.9384, Long: 106.3455

# 5. Thực nghiệm đánh giá, bàn luận

## Đánh giá hiệu suất mô hình

Công cụ đánh giá: Ma trận Confusion

- TP: Dự đoán đúng truy vấn độc hại
- TN: Dự đoán đúng truy vấn hợp lệ
- FP: Cảnh báo sai – truy vấn hợp lệ bị nhầm là độc hại
- FN: Bỏ sót tấn công – truy vấn độc hại bị nhầm là hợp lệ

Các chỉ số hiệu suất:

- Accuracy: 93.93% – tổng thể mô hình hoạt động tốt
- Precision: 85.50% – cảnh báo đúng nhiều
- Recall: 93.40% – ít bỏ sót tấn công
- F1-Score: 89.29% – cân bằng giữa cảnh báo sai và bỏ sót

# 6. Kết luận và hướng phát triển

## So sánh với các mô hình khác

CNN vượt trội nhờ:

- Tự động học đặc trưng qua các lớp Conv
- Xử lý tốt dữ liệu có cấu trúc như truy vấn SQL
- Hiệu suất cao và ổn định trên dữ liệu mới

## Những kết quả đạt được

- Hoàn thành hệ thống phát hiện SQLi đa lớp, tích hợp Injex Watch và mô hình CNN
- Giao diện trực quan, dễ giám sát
- Mô hình CNN đạt độ chính xác cao, tổng quát tốt
- Tích hợp thành công các công nghệ: Docker, NGINX, MERN Stack, TLS, Reverse Proxy
- Thử nghiệm thực tế cho kết quả khả quan

# 6. Kết luận và hướng phát triển

## Hạn chế

- Dữ liệu huấn luyện chưa đa dạng (thiếu Blind SQLi, obfuscated payloads)
- Chưa có cơ chế học liên tục, cập nhật mô hình
- Thiếu phản ứng chủ động: tự chặn IP.
- Chưa thử nghiệm hiệu năng quy mô lớn
- Chưa hỗ trợ phát hiện tấn công chéo nền tảng (cross-db)

## Hướng phát triển

- Mở rộng tập dữ liệu & tăng độ phức tạp các mẫu tấn công
- Áp dụng mô hình nâng cao: Transformer, BiLSTM, Semi-supervised
- Tăng cường phòng thủ chủ động: Tự động chặn IP, gửi cảnh báo, kết nối với SIEM
- Kiểm thử hiệu suất hệ thống lớn, hỗ trợ phát hiện đa dạng tấn công (XSS, Command Injection...)
- Cải tiến Dashboard & phát triển app mobile giám sát

Xin chân thành  
cảm ơn thầy cô đã  
theo dõi