VPS Threat Monitoring Platform

Members
Member 1: Truong Dang Gia Huy 22huy.tdg@vinuni.edu.vn
Member 2: Nguyen Tuan Anh 21anh.nt@vinuni.edu.vn
Member 3: Tran Tat Hung 22hung.tt@vinuni.edu.vn

1 Introduction

1.1 Problem Statement

In today's digital landscape, Virtual Private Servers (VPS) have become increasingly popular for hosting various services, from personal websites to business applications. However, this popularity has made VPS systems prime targets for cyber attacks. Traditional network monitoring tools often provide complex, technical interfaces that are difficult for non-technical users to understand and utilize effectively. Additionally, existing solutions typically focus on either real-time monitoring or historical analysis, but rarely combine both aspects in an accessible way.

Our project addresses these challenges by developing a comprehensive network security visualization dashboard that:

- Provides real-time monitoring of network security threats
- Offers intuitive visualization of complex network data
- Combines historical analysis with current threat detection
- Makes network security monitoring accessible to non-technical users

1.2 Novelty of Solution

Our solution differs from existing network monitoring tools in several significant ways:

1.2.1 Integrated Visualization Approach

- Combines real-time monitoring with historical analysis in a single interface
- Provides interactive dashboards that make complex network data accessible
- Offers multiple visualization perspectives (temporal, geographic, network topology)

1.2.2 Advanced Analysis Features

- Implements a comprehensive threat scoring system based on multiple factors
- Provides geographic threat intelligence with country-level analysis
- Offers advanced network topology analysis with interactive filtering

1.2.3 User-Centric Design

- Designed for both technical and non-technical users
- Provides intuitive interface for security monitoring
- Offers customizable views and filtering options
- Includes automated threat detection and scoring

These innovations address the limitations of existing solutions by providing a more comprehensive, accessible, and user-friendly approach to network security monitoring.

2 Data Collection

In order to gather required dataset, we have set up a data collection pipeline, with the support of WireShark packet analysis tool. The diagram 1. We deployed a service in the Linux cloud instance, which runs the WireShark packet sniffing mode, collecting all in-bound network packet and store them in a hive of `*.pcap` files. These files are parsed with our Python parser using the `pyshark` library, exported into CSV files. The CSV files will be merged, fed into our visualization.

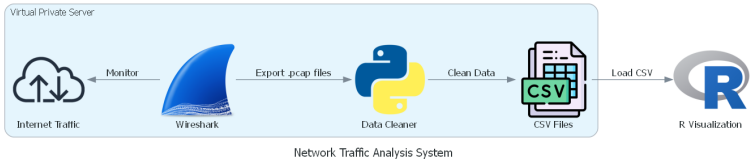

Figure 1: Data collection pipeline

The collected CSV files have the data dictionary as described in Figure 2. For the `source_country` column, we used an open-source database named `GeoLite2-Country.mmdb`, defines the IP and subnets range based on their geographical location. This approach redacted the requirement of using external IP geolocation APIs, which is costly in terms of bandwidth or cost of API keys. However, several packets have IPs that are not included in the database. Within the collected packages, 189 packages from 08 IPs can not be extracted the geolocation information. We addressed this problem by manually searching them via public service WhatIsMyIPAddress. Interestingly, all of these IPs are coming from CloudFlare tunnels - a legitimate service for routing traffics from local networks to the Internet. This reveals an interesting behavior of malicious actors.

Due to downtime from DigitalOcean, we have a quite big gaps in collecting data. However, within several days of uptime, we collected more than 22,000 packets, from more than 60 countries all around the world.

Variable	Class	Description
timestamp	integer	UNIX timestamp of the trafifc
source_ip	string	IPv4 address of Source
source_country	string	The country where **source_ip** comes from, based on geolocation
destination_port	integer	The destination port that **source_ip** connects to.
protocol	string	Protocol used by the packet
length	integer	Length in bytes of the packet

Figure 2: Data dictionary of the CSV files

3 Data Visualization

3.1 Architecture

Our network security visualization dashboard employs a modular architecture built on the R/Shiny framework, designed to transform complex network traffic data into actionable security insights. The system architecture consists of three primary layers:

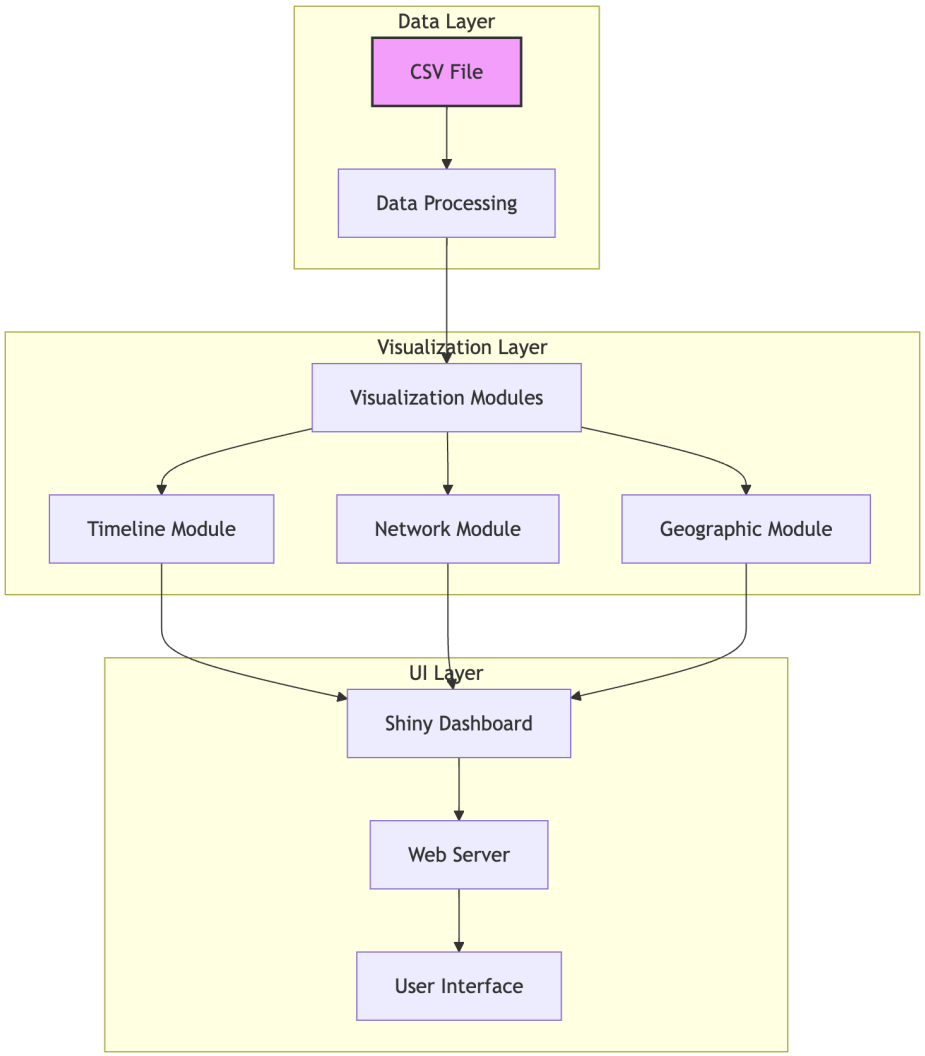

Figure 3: Architecture of the Data Visualization Part

3.2 Technology Stack

3.2.1 Core Framework

- **R**: Statistical computing language providing the foundation
- **Shiny**: Reactive web framework for interactive applications
- **shinydashboard**: Dashboard layout and component framework

3.2.2 Data Processing Libraries

- **data.table**: Data manipulation
  - Chosen for its superior performance with large datasets
  - Automatically optimising operations internally and very effectively by knowing precisely the data required for each operation, leading to very fast and memory-efficient code.
- **dplyr**: Intuitive data transformation pipelines
  - Used for data transformations

3.2.3 Visualization Libraries

Interactive Time Series

- **dygraphs**:
  - For time series visualization
  - Built-in zooming and panning capabilities
  - Data updates support

Network Graphs

- **visNetwork**:
  - Based on vis.js library
  - Physics-based layout algorithms
  - Interactive node/edge manipulation
  - Customizable styling and behavior

Interactive Plots

- **plotly**:
  - WebGL rendering for performance
  - Rich interactivity out-of-the-box
  - Extensive chart type support

Static Visualizations

- **ggplot2**:
  - Grammar of Graphics implementation
  - Extensive theming capabilities

3.2.4 Supporting Libraries

- **scales**: Data transformation and formatting
- **viridis**: Perceptually uniform color scales
- **DT**: Interactive data tables with search/sort/filter

3.2.5 Core Features

The visualization layer consists of four specialized modules, each addressing different aspects of network security monitoring:

1. **Dashboard Overview Module**: Provides at-a-glance security status with six key metrics (total attacks, unique attackers, average threat level, top origin, most targeted port, data volume), recent attack trends visualization, and a detailed top attackers table.

Figure 4: Dashboard

2. **Timeline Analysis Module**: Implements interactive time series visualization using dygraphs, enabling temporal pattern detection with multiple aggregation levels and metrics.

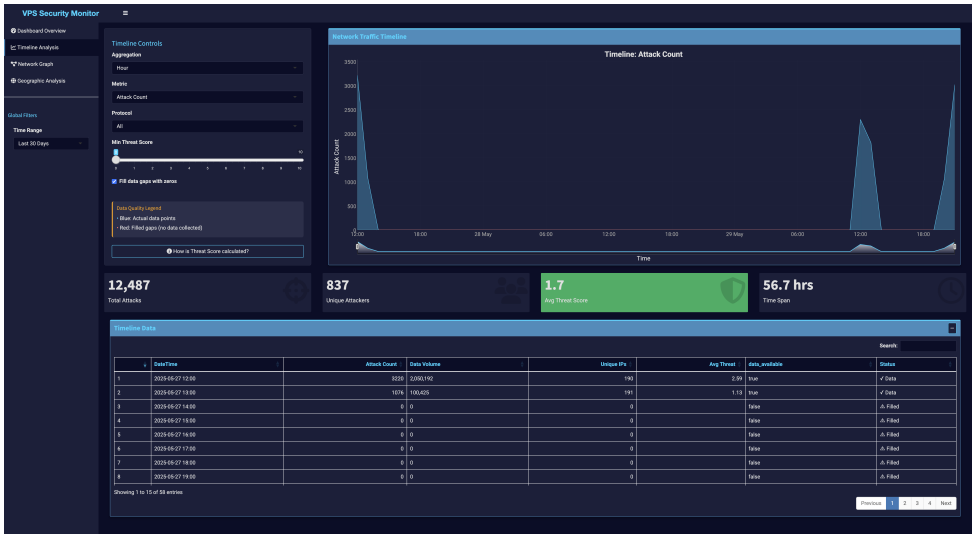

Figure 5: Timeline Analysis

3. **Network Graph Module**: Creates force-directed network topology visualizations using visNetwork, revealing connection patterns between attackers and targeted services.

Figure 6: Network Graph

4. **Geographic Analysis Module**: Generates choropleth world maps using Plotly, providing geographic threat intelligence and country-level attack statistics.

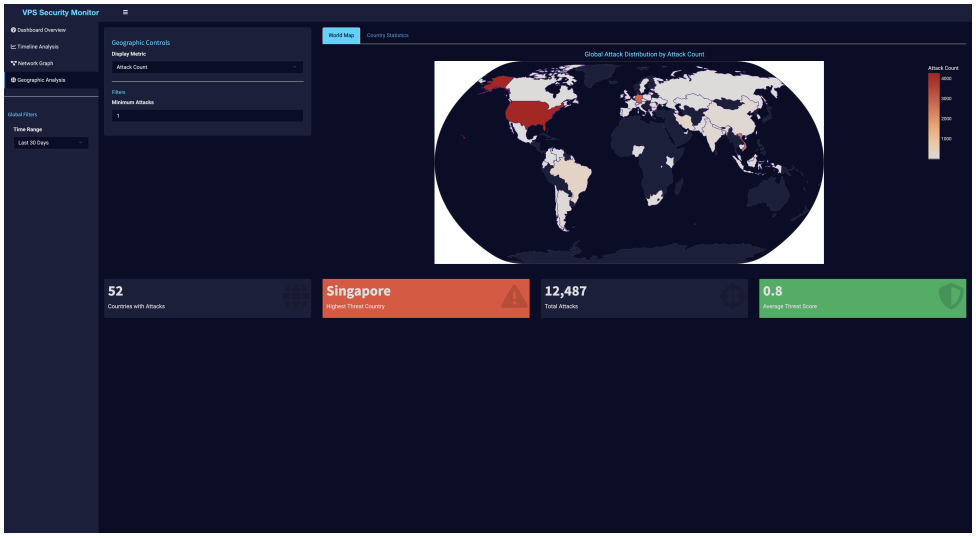

Figure 7: Geographic Analysis

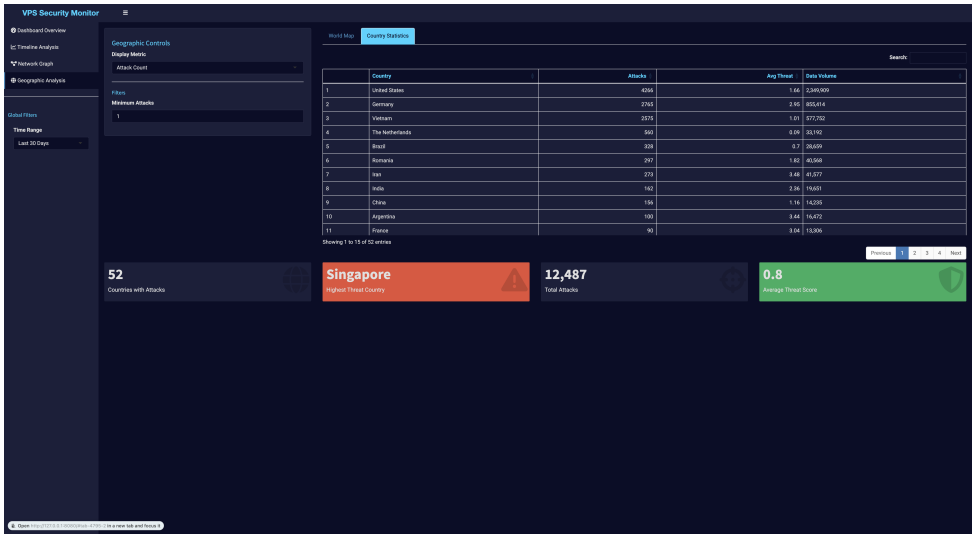

Figure 8: Country Statistics

3.2.6 Other Interactive Features

- **Dynamic Threat Scoring**: Real-time calculation based on multiple risk factors
- **Smart Search Functionality**: Quickly locate specific IPs or ports across visualizations
- **Customizable Views**: Users can filter by threat level, protocol, country, and time range
- **Data Gap Visualization**: Clearly indicates periods of missing data to maintain analytical integrity

3.3 User Manual

3.3.1 Set Up:

Follow instructions in README.md file on Github

3.3.2 Dashboard Overview

The main dashboard provides immediate security status through:

- **Key Metrics**: Six value boxes display critical security indicators
- **Attack Trends**: Area chart shows attack frequency over time
- **Top Attackers Table**: Interactive table with sorting and searching capabilities

3.3.3 Timeline Analysis

To analyze temporal patterns:

1. Select aggregation level: Minute (detailed), Hour (balanced), or Day (overview)
2. Choose metric: Attack Count, Data Volume, Unique IPs, or Threat Score
3. Apply filters: Protocol type and minimum threat score
4. Enable "Fill Gaps" to visualize missing data periods in red
5. Interact with the chart: Drag to zoom, double-click to reset

3.3.4 Network Graph

To explore network topology:

1. Select port display count (5, 10, or 20) to manage visualization complexity
2. Use filters to focus on specific node types, threat levels, or countries
3. Search for specific IPs or ports using the search bar
4. Interact with nodes: Click for details, drag to rearrange
5. Navigate: Mouse wheel to zoom, drag background to pan

3.3.5 Geographic Analysis

To examine global threat distribution:

1. Choose display metric: Attack Count, Threat Score, or Data Volume
2. Set minimum attack threshold to filter insignificant sources
3. Hover over countries for detailed statistics
4. Switch to "Country Statistics" tab for tabular view

3.3.6 Global Controls

- **Time Range Filter**: Located in sidebar, affects all modules
- **Accessibility Toggle**: Enable high contrast or large font modes
- **Data Export**: Download filtered data from any module

3.4 Discussion of Results

Our visualization dashboard has demonstrated significant effectiveness in identifying and analyzing network security threats. Through extensive testing with real VPS traffic data, we observed several key insights:

3.4.1 Pattern Recognition

The timeline analysis revealed distinct attack patterns, with peaks during specific hours suggesting automated bot activities. The ability to aggregate data at different time scales proved crucial for identifying both short-term attack bursts and long-term trends.

3.4.2 Geographic Insights

Geographic visualization exposed concentrated attack origins from specific regions, enabling targeted security measures. Countries with high attack volumes often correlated with known cybercrime hotspots.

3.4.3 Network Topology Analysis

The network graph effectively revealed targeted attack campaigns, where multiple IPs from the same geographic region targeted specific service ports. This visualization proved particularly valuable for identifying coordinated attacks that traditional log analysis might miss.

3.5 Limitations

While our solution provides comprehensive security visualization, several limitations should be acknowledged:

3.5.1 Performance Constraints

- **Network Graph Scalability**: Limited to displaying top 5, 10, or 20 ports to maintain readability and performance. Attempting to visualize all connections would result in an illegible graph and browser performance issues.
- **Data Volume**: Large datasets (>1 million records) may experience rendering delays, particularly in the network graph module.

3.5.2 Scope Limitations

- **Inbound Traffic Only**: The system currently processes only inbound network signals, missing potential threats from compromised internal systems communicating outward.
- **Static Data Processing**: Real-time streaming is not yet implemented, requiring periodic data updates.
- **Limited Protocol Analysis**: Deep packet inspection is not performed, focusing only on metadata analysis.

3.6 Future Directions

The network security visualization dashboard presents numerous opportunities for enhancement and expanded applications:

3.6.1 Real-time Capabilities

- **WebSocket Integration**: Implement live data streaming for immediate threat detection
- **Alert System**: Automated notifications for high-threat activities
- **Predictive Analytics**: Machine learning models to forecast attack patterns

3.6.2 Enhanced Analysis Features

- **Behavioral Profiling**: Identify attacker patterns and create threat profiles
- **Correlation Analysis**: Detect relationships between different attack campaigns
- **Outbound Traffic Monitoring**: Extend analysis to include egress traffic for comprehensive security coverage

These enhancements would transform our dashboard from a monitoring tool into a comprehensive security intelligence platform, capable of proactive threat prevention rather than reactive analysis.

4 Conclusions

This project successfully developed a comprehensive network security visualization dashboard aimed at making complex VPS network monitoring more accessible to both technical and non-technical users. The solution differentiates itself by integrating real-time monitoring with historical analysis, offering intuitive visualizations, advanced threat scoring, and geographic intelligence. Through testing with real VPS traffic data, the dashboard demonstrated significant effectiveness in identifying and analyzing network security threats, revealing distinct attack patterns, concentrated geographic origins, and targeted attack campaigns. This user-centric approach provides a valuable tool for enhancing network security monitoring.