CYBER SECURITY & ETHICAL HACKING

Group-5

INSTRUCTOR : ABDUR REHMAN

PROJECT:
SMART EMAIL SPOOFING DETECTION AND ANALYSIS DASHBOARD

GROUP MEMBERS:

SHEEZA ALAM KHAN

HOORIYA EHTESHAM

HAREEM MALLICK

PhishShield: Smart Email Spoofing Detection and Analysis Dashboard

Introduction

We developed a **smart email spoofing detection system** designed to identify suspicious and spoofed emails through **email header analysis**. The system helps users and cybersecurity professionals verify the authenticity of emails and prevent phishing attacks. It provides a **live dashboard** that visualizes real-time detection results, authentication checks, and analysis summaries.

Overview

Our project focuses on detecting spoofed emails using the metadata embedded within **email headers**. These headers contain authentication data such as **SPF**, **DKIM**, and **DMARC**, along with information about the **sender domain, return path, and routing servers**.

We built a **Flask-based web application** that allows users to upload or paste raw email headers. The system automatically breaks down the header into structured data, performs authentication checks, and then classifies the email as *Legitimate*, *Possibly Spoofed*, or *Suspicious*.

The project also includes a **smart dashboard** that provides an analytical view of the results displaying success rates, total analyses, verdict counts, and recent reports.

Motivation

Email spoofing remains one of the most common entry points for **phishing attacks** and **social engineering campaigns**. Attackers often forge sender addresses to trick users into sharing sensitive information or clicking malicious links.

Our motivation was to build a **user-friendly yet technically deep** system that demonstrates how cybersecurity tools like **MXToolbox** perform header analysis. This project helped us understand how spoofing can be detected programmatically by examining authentication mechanisms and header anomalies.

Methodologies

1. Email Header Parsing

When a user submits an email header, our system extracts and analyzes important fields such as:

- **From**
- **Return-Path**
- **Received**
- **Authentication-Results**
- **SPF**, **DKIM**, and **DMARC** indicators

We implemented a Python-based parser that reads the raw text, identifies relevant lines, and separates key-value pairs for further analysis.

2. Authentication Analysis

The parsed data is passed through logic that verifies:

- **SPF (Sender Policy Framework):** Checks if the sending IP is authorized for the sender's domain.
- **DKIM (DomainKeys Identified Mail):** Verifies whether the message content was signed and not altered in transit.
- **DMARC (Domain-based Message Authentication, Reporting, and Conformance):** Confirms domain alignment and policy compliance.
- **Return-Path Validation:** Ensures that the return path matches the sender domain.
- **Keyword Scanning:** Detects phishing-related terms such as "urgent," "update," "verify," etc.
- **Attachment Inspection:** Flags risky file types (.exe, .scr, etc.).

Each test outputs a pass, fail, or neutral result. The system calculates a **confidence score** based on these outcomes.

3. Verdict Generation

A scoring model assigns weights to the checks:

- High scores (80–100%) → **Legitimate**
- Medium scores (50–79%) → **Possibly Spoofed**
- Low scores (0–49%) → **Suspicious**

The verdict is displayed along with the timestamp and stored in the database for historical review.

4. Smart Dashboard

We created a **dashboard interface** that visualizes key metrics:

- **Total Emails Analyzed**
- **SPF, DKIM, and DMARC Success Rates**
- **Authentication Success Breakdown**
- **Verdict Distribution**
- **Recent Analyses Table**

Each record includes the email subject, verdict, confidence score, and analysis time. A **View Report** button lets users open detailed reports for specific analyses.

5. Detailed Report View

When viewing an individual report, users can see:

- Subject and timestamp
- Overall verdict and score
- Recommended action (e.g., “Be careful — verify sender”)
- Detailed check results for SPF, DKIM, DMARC, Return-Path, Keywords, Attachments
- Explanations for each failed or uncertain check

This section helps users understand *why* an email is classified as spoofed or legitimate.

End Result

The final system provides:

- A fully functional **Flask web interface** for analyzing email headers.
- A **real-time dashboard** showing all analysis results and verdict statistics.
- **Detailed individual reports** that explain every authentication check.
- A secure and educational environment for understanding **email spoofing detection**.

The dashboard successfully analyzes multiple email samples, calculating average success rates such as:

- **SPF Success:** 65.4%
- **DKIM Success:** 65.4%
- **DMARC Success:** 53.8%

Recent analyses demonstrate the system's ability to correctly identify spoofed and legitimate emails with clarity and precision.

What We Have Learned

Through this project, we learned:

- How email headers work and how spoofing is performed.
- The role of **SPF, DKIM, and DMARC** in preventing email forgery.
- How to parse complex text-based headers into structured Python data.
- How to build **Flask applications** with templates and dashboards.
- How to interpret and visualize **email authentication results**.

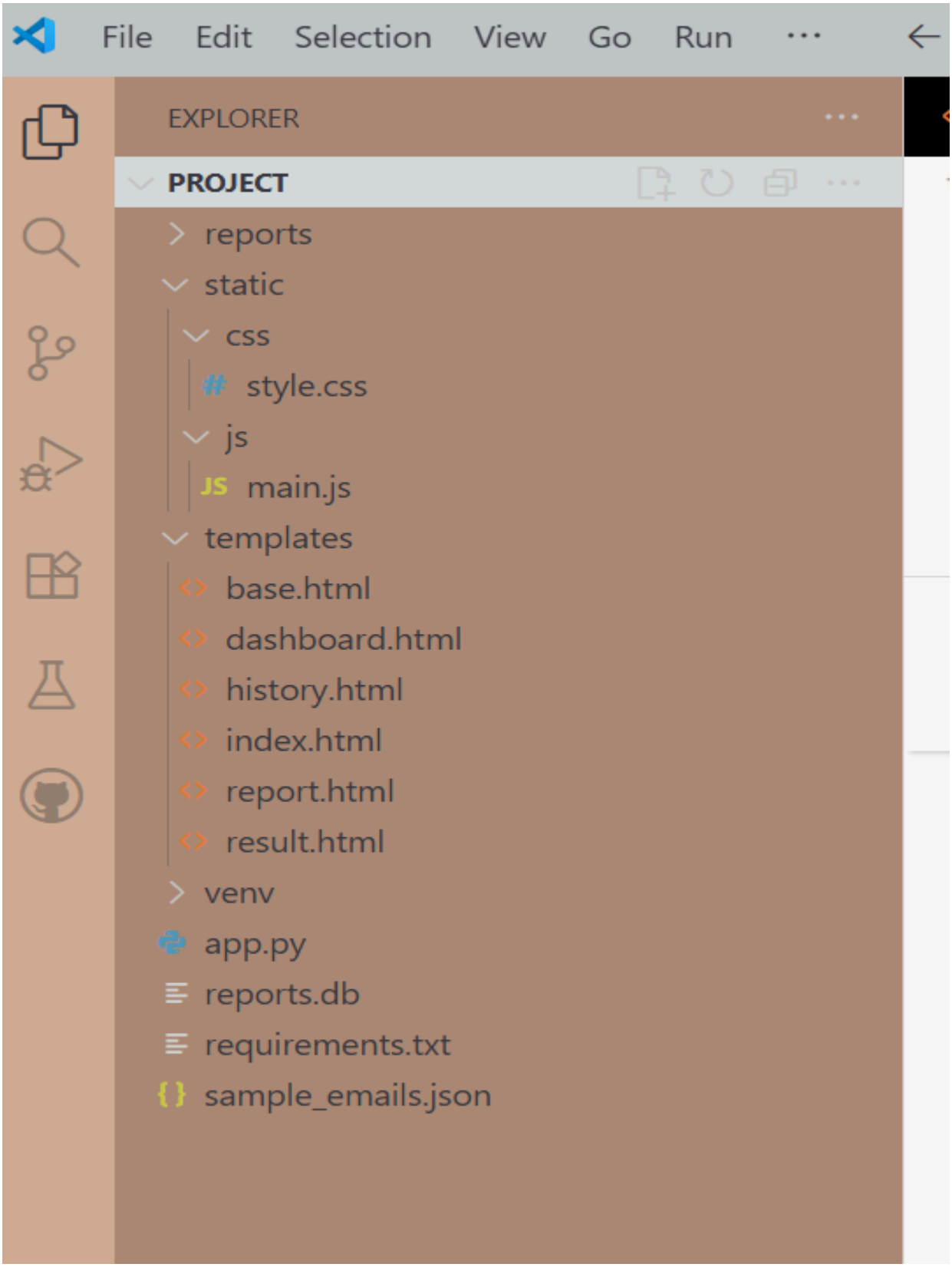
We also learned to design intuitive user interfaces for cybersecurity tools that make technical analysis accessible to both experts and non-technical users.

Conclusion

Our **PhishShield** system effectively demonstrates the process of **email spoofing detection** using header-based analysis. By breaking down authentication data, scanning for phishing indicators, and presenting visual insights, we created a tool that bridges cybersecurity education and practical application.

This project enhanced our understanding of **email security, authentication protocols, and real-time web visualization**, marking a significant step forward in our cybersecurity learning journey.

Project Files (VS CODE):



New Analysis | Dashboard | History

Email Spoof Detector

Paste raw headers or full .eml content

```
Delivered-To: shizaalam1789077@gmail.com
Received: by 2002:a05:7108:5526:b0:4d5:3783:bbf9 with SMTP id ay38csp1237207gdb;
        Fri, 24 Oct 2025 00:27:37 -0700 (PDT)
X-Google-Smtp-Source: AGHT+IHF/YBaZ5U4sOU9coQ0488plzr/AtNssIQXPLeRpgAjcJuzN1DkpSZybAMbU6GT2ACVGKAa
X-Received: by 2002:a05:6000:2586:b0:3eb:d906:e553 with SMTP id ffacd0b85a97d-42704dab707mr17558379f8f.55.1761290856814;
        Fri, 24 Oct 2025 00:27:36 -0700 (PDT)
ARC-Seal: i=1; a=rsa-sha256; t=1761290856; cv=none;
        d=google.com; s=arc-20240605;
        b=edqYNzbWHvtjhkQy0AiJO3U5112TYkOkkgMn8BdQxVL2qb7NcgHaxmWKB3an1ciCqF
         7rWC1/H4X1z2gsE4zcr4XxBubevBx83hJfhFH7mH96e31ivuTKEIKWyPdd3pWHh1wEHO
         nlHJKsoovnb4NDazdmfSXEBzndhu/tM6LGWF+ZzEO4rax+8hM1MZmPCiogSFUBM5zdcS
```

Choose File No file chosen Analyze

Demo Samples

Legitimate Email

Spoofed Email

Ambiguous Email

New Analysis | Dashboard | History

Analysis Result

Safe

Confidence Score: **85.0%**

Action for you

Looks OK — verify links.

Top reasons flagged

- Return-Path domain (kjhgfghjk.uytrftyureds.checkss.it.com) ≠ From domain (bse3hnqc.checkss.it.com>)
- Duplicate headers detected

Quick definitions (for non-technical)

SPF	Who is allowed to send mail for this domain.
DKIM	Signature proving message integrity.
DMARC	Policy telling receivers how to treat failing mail.

View History

Email Spoofing Detector

Home Dashboard

Email Analysis History

ID	Timestamp (UTC)	Subject	Verdict	Score (%)	Summary	Actions
27	2025-11-07T20:26:33.400343Z	shizaalam1789077 - You have won a 170 Piece Stanley Tool Set 🎁 #ID53663	Legitimate	85.0	Return-Path domain (kjhgfghjk.uytrftyureds.checkss.it.com) ≠ From domain (bse3hnqc.checkss.it.com>); Duplicate headers detected	View
26	2025-11-07T19:37:03.081235Z	Weekly update	Possibly spoofed	55.0	Phishing keywords: update	View
25	2025-11-07T19:34:53.257222Z	shizaalam1789077 - You have won a 170 Piece Stanley Tool Set 🎁 #ID53663	Legitimate	85.0	Return-Path domain (kjhgfghjk.uytrftyureds.checkss.it.com) ≠ From domain (bse3hnqc.checkss.it.com>); Duplicate headers detected	View
24	2025-11-07T19:26:57.210639Z	Weekly update	Possibly spoofed	55.0	Phishing keywords: update	View
23	2025-11-07T19:26:03.709562Z	micro1 interview invite	Legitimate	85.0	Return-Path domain (em8057.micro1.ai) ≠ From domain (micro1.ai>); Duplicate headers detected	View
22	2025-11-07T19:23:44.880772Z	Application submitted	Possibly spoofed	75.0	Return-Path domain (mailing.nobelhub.com) ≠ From domain (nobelhub.com); Duplicate headers detected	View
21	2025-11-07T19:21:48.657372Z	Hello	Legitimate	100.0	-	View
20	2025-11-07T19:21:35.236535Z	Application submitted	Possibly spoofed	75.0	Return-Path domain (mailing.nobelhub.com) ≠ From domain (nobelhub.com); Duplicate headers detected	View
19	2025-11-07T18:18:13.851688Z	Weekly update	Possibly spoofed	55.0	Phishing keywords in subject: update	View
18	2025-11-07T18:15:06.897254Z	Hello	Legitimate	100.0	-	View
17	2025-11-07T18:08:39.021325Z	Urgent! Verify your account	Suspicious	25.0	SPF failed; DKIM failed; DMARC failed	View
16	2025-11-07T18:08:19.628149Z	Urgent! Verify your account	Suspicious	25.0	SPF failed; DKIM failed; DMARC failed	View

Email Spoofing Detector

Home Dashboard

Email Spoofing Report

Subject: shizaalam1789077 - You have won a 170 Piece Stanley Tool Set 🎁 #ID53663 Legitimate

Timestamp (UTC): 2025-11-07T20:26:33.400343Z

Confidence Score: 85.0%

Recommended Action: Looks OK — verify links.

Detailed Checks

SPF	pass
DKIM	pass
DMARC	pass
RETURNPATH	fail
KEYWORDS	pass
ATTACHMENTS	pass
HEADER_ANOMALIES	duplicate

Potential Issues:

- Return-Path domain (kjhgfghjk.uytrftyureds.checkss.it.com) ≠ From domain (bse3hnqc.checkss.it.com>)
- Duplicate headers detected

Explanations

- **SPF:** SPF passed successfully.
- **DKIM:** DKIM passed successfully.
- **DMARC:** DMARC passed successfully.
- **RETURNPATH:** Mismatch often indicates spoofing.
- **KEYWORDS:** No phishing words in subject.
- **ATTACHMENTS:** No executable attachments found.
- **HEADER_ANOMALIES:** Duplicate header fields may indicate tampering.

← Back to History

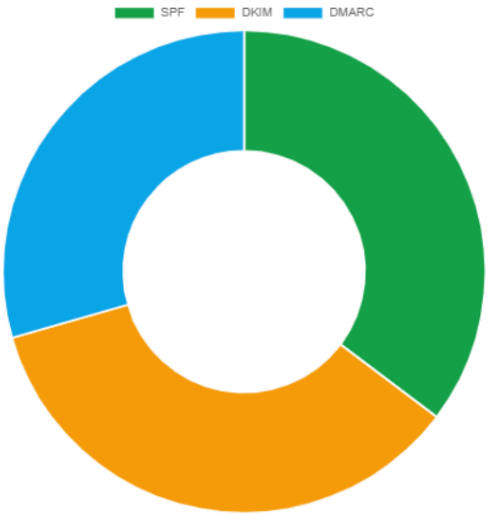
New Analysis | History

Smart Dashboard

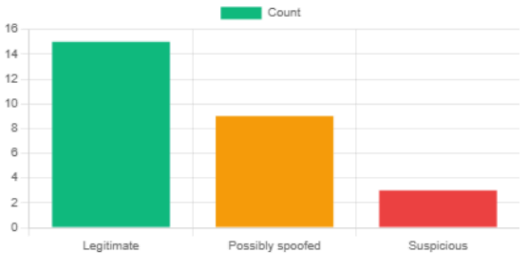
Live overview — updates automatically.

27	66.7%	66.7%	55.6%
Total Emails	SPF Success	DKIM Success	DMARC Success

Auth Success Breakdown



Verdict Counts



Recent Analyses

ID	Subject	Verdict	Score	Time
-	shizaalam1789077 - You have won a 170 Piece Stanley Tool Set 🎁 #ID53663	Legitimate	85%	2025-11-07T20:26:33.400343Z
-	Weekly update	Possibly spoofed	55%	2025-11-07T19:37:03.081235Z
-	shizaalam1789077 - You have won a 170 Piece Stanley Tool Set 🎁 #ID53663	Legitimate	85%	2025-11-07T19:34:53.257222Z
-	Weekly update	Possibly spoofed	55%	2025-11-07T19:26:57.210639Z
-	micro1 interview invite	Legitimate	85%	2025-11-07T19:26:03.709562Z
-	Application submitted	Possibly spoofed	75%	2025-11-07T19:23:44.880772Z
-	Hello	Legitimate	100%	2025-11-07T19:21:48.657372Z
-	Application submitted	Possibly spoofed	75%	2025-11-07T19:21:35.236535Z
-	Weekly update	Possibly spoofed	55%	2025-11-07T18:18:13.851688Z
-	Hello	Legitimate	100%	2025-11-07T18:15:06.897254Z