Tool Name: Homoglyph Domain Inspector

By: Soham Kadu
Intern ID: *235*

Overview

Homoglyph-based attacks are a common trick in phishing, where fake domains replace normal characters with visually similar ones from other languages (Unicode scripts). This tool is a Python utility that scans text files for URLs and flags those that may be using homoglyph characters to impersonate real domains.

It is designed for security analysts, incident responders, and awareness trainers who need a **fast, automated way** to find dangerous links hidden in plain sight.

Purpose of the Tool

- Detect domains that visually look legitimate but contain hidden Unicode characters.
- Help identify phishing attempts before they can cause harm.
- Speed up security checks in large datasets like chat logs, email dumps, or reports.
- Serve as a teaching resource for security awareness programs.

Core Features

- **File Scanning:** Reads plain text files and extracts all URLs.
- **Unicode Analysis:** Detects mixed-script usage (Latin, Greek, Cyrillic, etc.).
- **Suspicion Marking:** Flags domains where the ASCII conversion differs from the original.
- **Multi-platform Support:** Works on Windows, Linux, and macOS.
- **Low Resource Usage:** Requires only Python and the homoglyphs library.
- **Portable:** No heavy setup; can be added to investigation scripts or CI/CD pipelines.

Practical Benefits

- Finds dangerous URLs in **seconds** instead of manually reviewing them.
- Reduces human error when scanning long lists of links.
- Makes phishing simulations more realistic by generating real-world lookalike domains.
- Can be used during incident triage, digital forensics, or awareness workshops.

Modules in the Tool

1. **Command-line Interface (CLI):** Accepts file names for batch scanning.
2. **Detection Engine:** Uses the homoglyphs library to identify Unicode lookalikes.
3. **Report Generator:** Produces a clear output showing normal and suspicious links separately.

Command to Run:

bash

python homoglyph_checker.py testlinks.txt

Sample Output:

bash

Links found:

  https://google.com

  http://google.com/fake

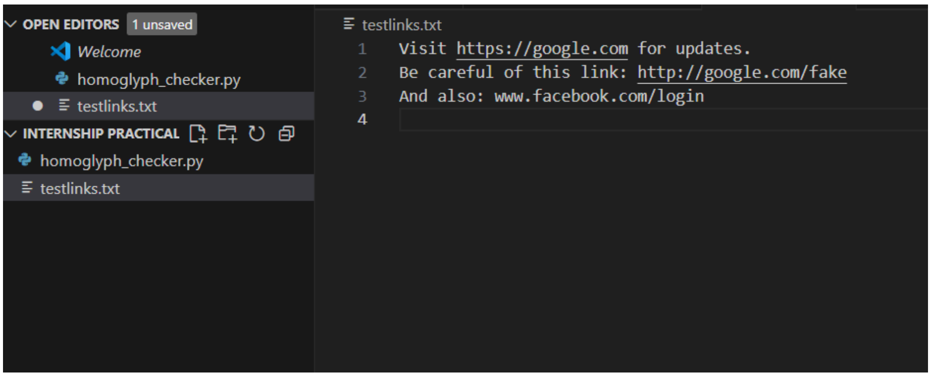
  www.facebook.com/scam

Suspicious links:

  http://google.com/fake

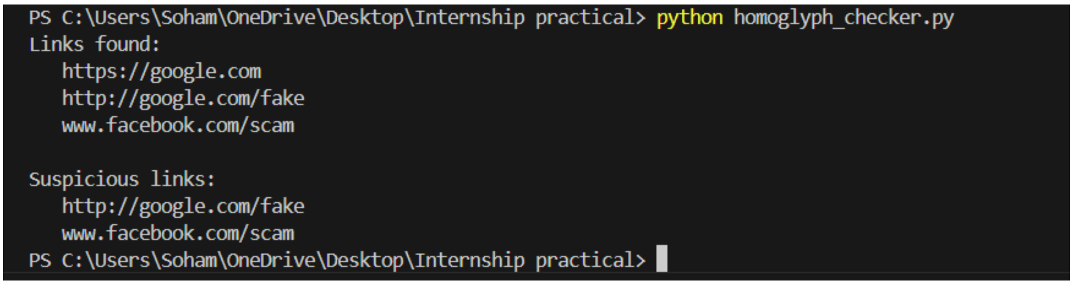
  www.facebook.com/scam

📷 Screenshot – Input File:



```
Visit https://google.com for updates.
Be careful of this link: http://google.com/fake
And also: www.facebook.com/login
```

📷 Screenshot – Script And Output



```
PS C:\Users\Soham\OneDrive\Desktop\Internship practical> python homoglyph_checker.py
Links found:
   https://google.com
   http://google.com/fake
   www.facebook.com/scam

Suspicious links:
   http://google.com/fake
   www.facebook.com/scam
PS C:\Users\Soham\OneDrive\Desktop\Internship practical> 
```

Operational Summary

- Reads the provided file.
- Uses regex to collect all URLs.
- Compares each URL's ASCII-normalized version to the original.
- If they differ, flags the link as suspicious.
- Displays results in an easy-to-read format.

Best Use Scenarios

- Checking suspicious links in incident reports.
- Auditing public-facing documents before publishing.
- Verifying URLs in phishing awareness campaigns.
- As a quick security gate in CI/CD pipelines.

Who Should Use This Tool

- SOC Analysts
- Threat Intelligence Teams
- Cybersecurity Trainers
- Red Team Members

Skills Required

- Basic Python and CLI familiarity.
- Understanding of phishing attack patterns.
- Awareness of Unicode and homoglyph risks (optional but useful).

Limitations and Future Improvements

Limitation	Suggested Enhancement
Works only on plain text files	Add HTML and email file support
No scoring system	Add risk score based on link similarity
Unicode database updates required	Automate update checks

Limitation	**Suggested Enhancement**
Lacks integration with threat intel feeds	Connect to domain reputation services

---

Why This Tool is Valuable

- Saves analyst time.
- Increases accuracy in phishing detection.
- Lightweight and portable.
- Easy to integrate into existing workflows.

Conclusion

Homoglyph-based phishing is a subtle but dangerous technique. The Homoglyph Domain Inspector offers an efficient way to spot and flag these malicious links, making it a valuable addition to any security toolkit.