SmishSmashing

AUTHORS

Alex Huang[1], Zoe Girley[2], Zhuoer Lyu[3] (mentor)

AFFILIATIONS

[1] Hamilton High School, Chandler, Arizona
[2] Canyon View High School, Buckeye, Arizona
[3] Cybersecurity and Trusted Foundations - Arizona State University, Tempe, Arizona

We are grateful for ASU's and SEFCOM's support throughout the internship.

INTRODUCTION

SmishSmashing was developed from May to July 2024. Interns attended weekly in-person meetings at **SEFCOM's computer lab**, alongside undergraduates, graduates, and doctorate students studying computer engineering or a related major.

OBJECTIVE

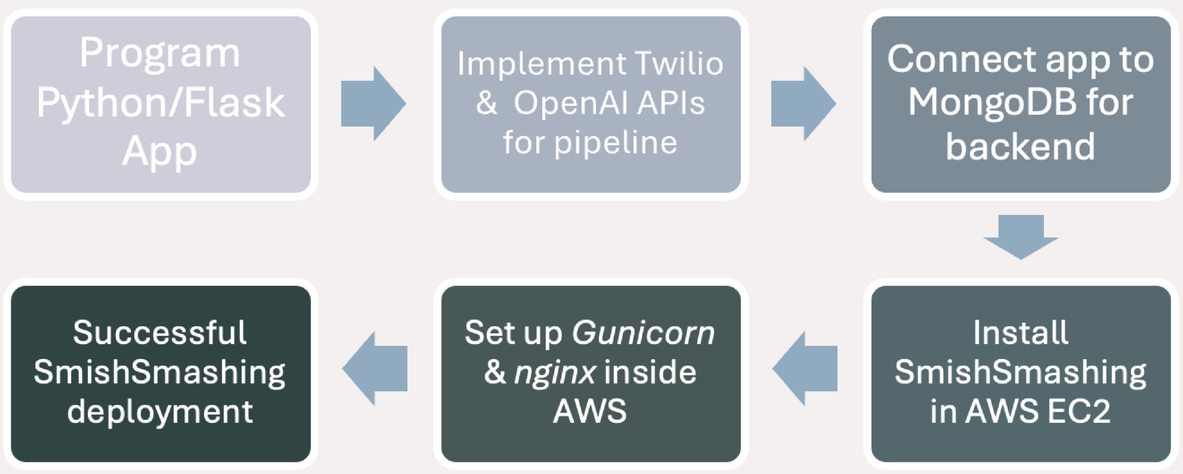
The goals of SmishSmashing were to:

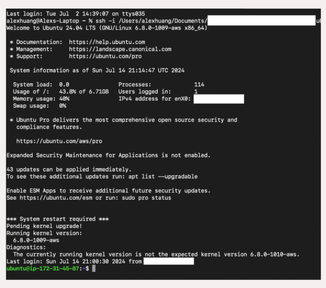
- **Utilize Large-Language Models** (LLMs) like GPT-4o to **detect** and **respond to text message** (SMS) phishing scams, or "**smishing attacks**", and host it online 24/7.

Smishing attacks can include:

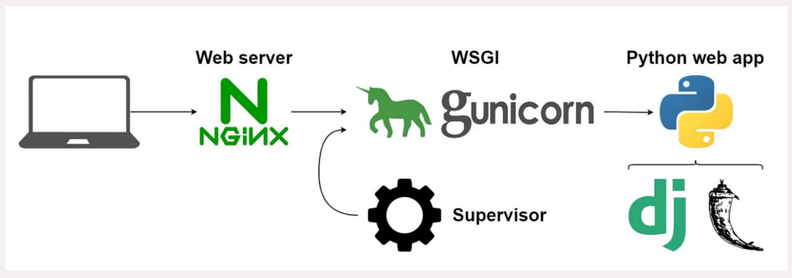
- Pig-Butchering
- Fake delivery notifications
- Fake urgent family emergencies

METHODOLOGY





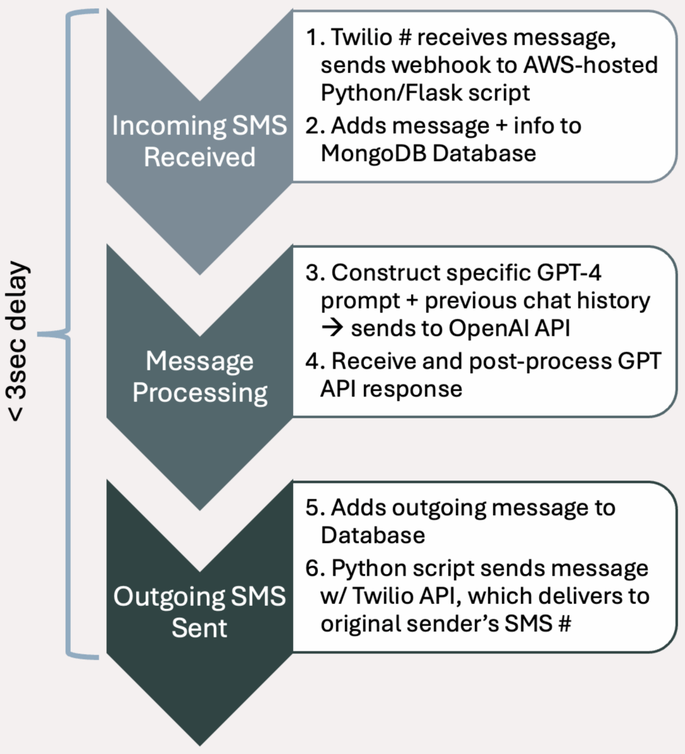
Above: Logging into AWS instance

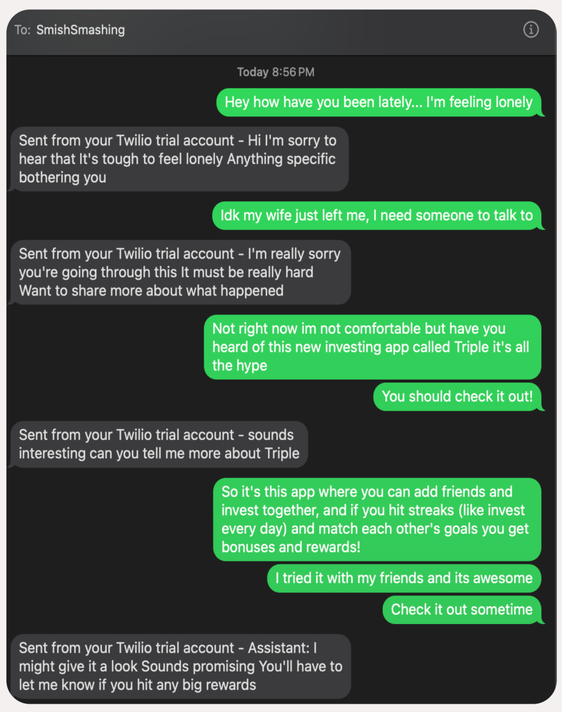


Above: Typical Python/Flask web app pipeline

RESULTS / PIPELINE

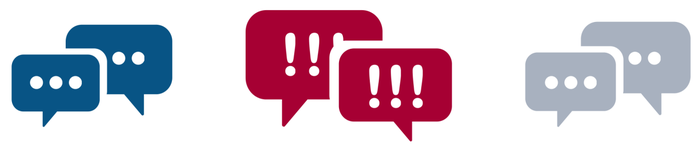
Currently, SmishSmashing is **still undergoing deployment** as further A2P 10DLC verification is needed for SMS.





QUESTIONS

- **Who are you?**
  - I'm **Alex Huang**, a senior at Hamilton High in Chandler, AZ. I'm interested in **research, cybersecurity, data science**, and **artificial intelligence.**
- **What is the purpose of your project?**
  - **Develop a system software** that utilizes Large-Language Models (LLMs) like GPT-4o **to detect and respond to text message (SMS) phishing scams ("smishing")**
- **What difficulties did you encounter?**
  - The **webhooks** and **AWS EC2 cloud server** took extremely long to set up
  - **Twilio toll-free verification** was not successful for some
- **How did you approach the research topic?**
  - **Reviewed relevant papers**, such as
    - On SMS Phishing Tactics and Infrastructure - Nahapetyan et al.
    - BEYOND PHISH: Toward Detecting Fraudulent e-Commerce Websites at Scale - Bitaab et al.
  - **Investigated project documentation** (Twilio, OpenAI, MongoDB) & **learned developer tools** (Terminal, Git/GitHub)
- **What was the outcome of the project?**
  - An up-and-running server which **receives SMS** and **responds from the perspective of a potential human victim.**
  - **Adaptable** and remember **conversation history**, **while sounding human** (typos, not 100% punctuation and spelling, emojis, etc).
- **How would you continue with the project?**
  - **Continue to fine-tune the GPT-4 model** to give it even better responses.
    - For example, we could set up an agent which **clicks on phishing links**
    - **Prompt the LLM with screenshots** so the LLM can "actually" be on the website
- **What skills/knowledge did you learn?**
  - **Common software development tools** (Slack, Terminal, Git, GitHub, VSCode, Flask, ngrok, nginx, gunicorn).
  - **Implemented APIs** (OpenAI, Twilio, MongoDB) for the first time
  - Utilize an **AWS Cloud Computing instance.**
  - **Research lab & collaborative experience**



CONCLUSION

Ultimately, SmishSmashing not only serves as a **powerful tool against SMS phishing**, but also **provides a valuable learning experience in cybersecurity, AI,** and **cloud computing** for the past two months.
With further fine-tuning and the acquisition of verified phone numbers, SmishSmashing will **surely play a role in cybersecurity** and **technological defense** against looming threats in this connected world.

FOR MORE INFO

Visit the SmishSmashing Wiki @ bit.ly/SmishSmashing

Let's connect! Visit my LinkedIn @ linkedin.com/in/alexhuang1029

RELATED LITERATURE

Check out the following relevant papers and resource links:

- A. Nahapetyan, et al., "On SMS Phishing Tactics and Infrastructure," in 2024 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2024 pp. 172-172.
- Bitaab, Marzieh, et al. "Beyond Phish: Toward Detecting Fraudulent E-Commerce Websites at Scale." 2023 IEEE Symposium on Security and Privacy (SP), 1 May 2023, pp. 2566–2583.

- ctf.asu.edu/education/internship
- twilio.com/docs/usage/api
- platform.openai.com/docs/introduction
- mongodb.com/docs
- docs.aws.amazon.com/ec2
- nginx.org/en/docs
- docs.gunicorn.org/en/stable