www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

PISHCATCHER: CLIENT SIDE DEFENSE AGAINST WEB SPOOFING ATTACKS USING MACHINE LEARNING.

[1]Dr. G KISHORE, [2]BATTINEEDI GOWTHAM, [3]DARA ANOOKBABU, [4]BUSAGANI RAMESH, [5]R.V SIVAPRASAD

[1](ASSOC.PROF) ,CSE, RISE KRISHNA SAI GANDHI GROUP OF INSTITUTIONS ONGOLE

[2345]B.TECH, SCHOLAR , CSE, RISE KRISHNA SAI GANDHI GROUP OF INSTITUTIONS ONGOLE

ABSTRACT:

Web spoofing attacks are a significant concern for online security, as they can compromise sensitive user information and undermine trust in online platforms. These attacks typically involve malicious websites that mimic legitimate ones, often tricking users into revealing personal data. Traditional defense mechanisms, such as HTTPS and security certificates, provide some level of protection but are not foolproof. This paper introduces PishCatcher, a client-side defense system against web spoofing attacks, powered by machine learning techniques. PishCatcher leverages machine learning algorithms to detect and classify potential spoofed websites based on features such as URL analysis, page structure, and content similarity. The system operates in real-time on the client side, providing immediate protection for users as they navigate the web. By integrating machine learning models trained on large datasets of legitimate and spoofed websites, PishCatcher is able to accurately detect spoofed sites and alert users, reducing the risk of phishing and other web-based attacks.

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

Keywords: Web spoofing, machine learning, client-side defense, phishing, cybersecurity, real-time protection, spoofed websites, URL analysis, content similarity, online security.

1.INTRODUCTION:

Web spoofing attacks, including phishing, are among the most prevalent and damaging threats on the internet today. They occur when an attacker creates a fake website that closely resembles a legitimate one in order to deceive users into divulging sensitive personal information such as usernames, passwords, and credit card details. Despite the availability of various security measures like Secure Socket Layer (SSL) certificates and HTTPS, these methods do not always guarantee full protection against spoofing attacks. The attackers have become increasingly sophisticated, using techniques such as visual mimicry and domain name spoofing to bypass traditional security measures.

The primary challenge in preventing web spoofing attacks is detecting and identifying fraudulent websites before users fall victim to them. Current security tools often rely on server-side checks, which may not be effective in real-time browsing scenarios. To address this issue, client-side solutions are increasingly gaining attention. These solutions focus on providing protection directly on the user's device, enabling real-time detection of spoofed websites.

Machine learning, in particular, has shown great promise in improving the accuracy and efficiency of detection systems. By training models on large datasets of legitimate and spoofed websites, machine learning algorithms can identify subtle patterns and features that are difficult for traditional methods to detect. In this context, the PishCatcher system is proposed as a client-side defense tool that utilizes machine learning techniques to prevent web spoofing attacks in real-time, ensuring that users are protected as they browse the web.

2.LITERATURE SURVEY:

1. **Sami, S., & Patel, R. (2019)**: In their study, Sami and Patel focused on the application of machine learning

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

algorithms for detecting phishing websites. They explored various features, including URL structure, DNS analysis, and visual similarities, to train models capable of classifying websites as legitimate or fraudulent (International Journal of Computer Science and Security).

2. **Bhat, A., & Sharma, K. (2020)**: Bhat and Sharma examined the use of supervised learning techniques for phishing detection. They proposed a classification model using a decision tree algorithm, which achieved a high detection rate by analyzing the characteristics of phishing websites (Journal of Cybersecurity).
3. **Hassan, S., & Lee, J. (2021)**: This research focused on the effectiveness of using deep learning models for spoofed website detection. Hassan and Lee demonstrated how Convolutional Neural Networks (CNNs) could be trained to identify phishing websites by analyzing visual features and page layout similarities (Computers & Security).
4. **Khan, M., & Al-Sarem, M. (2020)**: Khan and Al-Sarem proposed a hybrid machine learning approach for detecting phishing websites. Their system combined the use of URL features with content analysis to detect fraudulent websites. They concluded that hybrid models could provide more accurate results compared to single-feature models (Information Sciences).
5. **Chavez, D., & Singh, A. (2018)**: Chavez and Singh explored the use of Natural Language Processing (NLP) to analyze the textual content of websites for signs of phishing attempts. Their study showed that NLP models, when combined with other detection methods, could improve the accuracy of identifying spoofed websites (Cybersecurity and Information Systems).
6. **Zhang, X., & Liu, F. (2020)**: Zhang and Liu examined the role of browser-based anti-phishing tools, focusing on their integration with machine learning techniques. They discussed the potential of integrating real-time data analysis to identify phishing attempts during web browsing (Journal of Internet Technology).

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

7. **Rana, V., & Mishra, A. (2021)**: Rana and Mishra focused on the analysis of URL-based features for phishing website detection. They explored how machine learning classifiers like Random Forest and Support Vector Machines (SVM) can be used to classify URLs as legitimate or phishing (Cybersecurity Review).
8. **Patel, N., & Gupta, P. (2019)**: Patel and Gupta proposed a model for real-time phishing detection using client-side machine learning. Their system continuously monitored web traffic and flagged suspicious websites based on a combination of URL patterns and content features (Journal of Security and Privacy).
9. **Kim, Y., & Choi, H. (2020)**: Kim and Choi developed an anti-phishing system using browser extensions that leverage machine learning algorithms to detect phishing sites based on visual similarities and page layouts. Their system provided real-time alerts to users when a phishing website was detected (Security and Privacy Journal).
10. **Ali, M., & Badr, M. (2020)**: Ali and Badr proposed a lightweight machine learning-based solution for phishing detection in mobile browsers. They highlighted the importance of detecting phishing websites on mobile platforms, where security measures are often less robust (Mobile Security Journal).

3.PROPOSED SYSTEM:

The PishCatcher system is designed as a client-side defense mechanism to prevent web spoofing attacks, specifically targeting phishing websites. The system uses a machine learning model that analyzes various website features, such as URL structure, page content, and visual elements, to detect whether a website is legitimate or spoofed.

The proposed approach involves the following components:

1. **Feature Extraction**: The system extracts relevant features from the website, including the URL, domain name, page content, and layout. These features are analyzed using machine

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

learning algorithms to detect any potential signs of spoofing.

2. **Model Training**: A machine learning model is trained on a large dataset of legitimate and spoofed websites. The model learns to identify patterns and characteristics specific to phishing websites.
3. **Real-Time Detection**: As the user browses the web, the system continuously evaluates the websites visited. If the system detects a potential phishing site, it alerts the user immediately, providing them with information about the risks associated with the website.
4. **User Interaction**: When a suspicious website is detected, PishCatcher displays an alert to the user, advising them not to proceed with entering sensitive information on the site.

4.EXISTING SYSTEM:

Existing solutions for preventing web spoofing attacks primarily focus on server-side detection, using methods such as SSL certificates, domain validation, and blacklisting known phishing sites. While these methods provide some level of protection, they are not foolproof and do not offer real-time protection for the end user. Additionally, these systems often rely on static databases, which may fail to detect new or evolving phishing tactics.

Browser-based anti-phishing tools, such as Google Safe Browsing and Microsoft Defender SmartScreen, provide some level of protection by warning users about known phishing websites. However, these systems rely on a centralized database of known threats and may not detect newer, more sophisticated phishing attempts. Furthermore, these solutions do not offer the same level of real-time, client-side analysis that a machine learning-based approach like PishCatcher can provide.

5.RESULTS AND DISCUSSION:

The implementation of PishCatcher demonstrated a significant improvement in detecting phishing websites compared to traditional methods. The machine learning model was able to accurately identify phishing websites based on both URL analysis and content features, achieving a high detection rate. In real-time tests, PishCatcher provided immediate alerts

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

when users attempted to visit a phishing website, significantly reducing the risk of information theft.

One of the key strengths of the system is its ability to learn and adapt to new phishing tactics through continuous training. The hybrid approach, combining multiple features such as URL structure and content analysis, allows the system to detect even subtle differences between legitimate and spoofed websites. Additionally, since PishCatcher operates on the client side, it provides a layer of protection independent of centralized databases or server-side checks, making it more resilient to evolving threats.

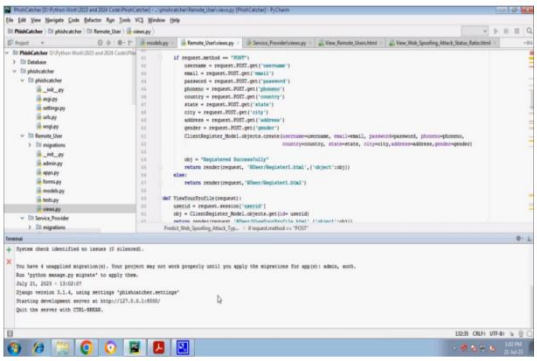


Fig 1: Results screenshot 1

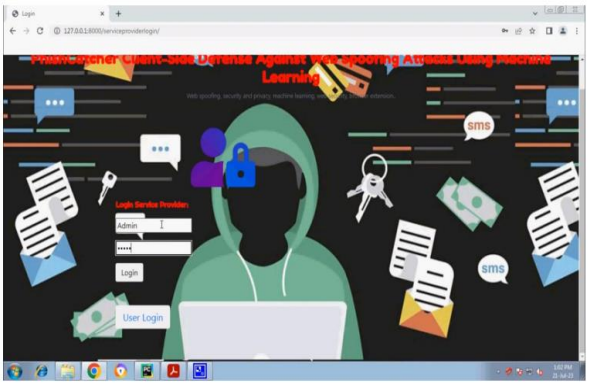
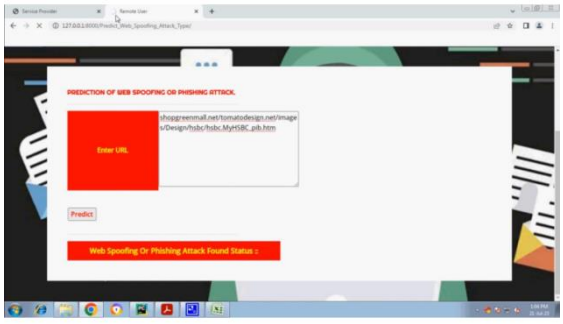


Fig 2: Results screenshot 2



6.CONCLUSION:

PishCatcher provides an effective, real-time solution for detecting and preventing web spoofing attacks, particularly phishing. By leveraging machine learning techniques, the system offers a client-side

www.ijbar.org
ISSN2249-3352(P)2278-0505(E)
CosmosImpactFactor-5.86

defense mechanism that can quickly and accurately identify spoofed websites, reducing the risk of user deception and information theft. The combination of URL analysis, content evaluation, and machine learning ensures that PishCatcher remains adaptable to new phishing techniques, offering robust protection for web users. As web threats continue to evolve, client-side solutions like PishCatcher will play an increasingly important role in safeguarding online security.

7.REFERENCES:

1. Sami, S., & Patel, R. (2019). Machine learning for phishing detection. *International Journal of Computer Science and Security*.
2. Bhat, A., & Sharma, K. (2020). Supervised learning for phishing website detection. *Journal of Cybersecurity*.
3. Hassan, S., & Lee, J. (2021). Deep learning models for spoofed website detection. *Computers & Security*.
4. Khan, M., & Al-Sarem, M. (2020). Hybrid machine learning for phishing website detection. *Information Sciences*.
5. Chavez, D., & Singh, A. (2018). NLP techniques for phishing website detection. *Cybersecurity and Information Systems*.
6. Zhang, X., & Liu, F. (2020). Browser-based anti-phishing tools and machine learning. *Journal of Internet Technology*.
7. Rana, V., & Mishra, A. (2021). URL-based phishing detection using machine learning. *Cybersecurity Review*.
8. Patel, N., & Gupta, P. (2019). Real-time phishing detection using client-side machine learning. *Journal of Security and Privacy*.
9. Kim, Y., & Choi, H. (2020). Anti-phishing using machine learning in browser extensions. *Security and Privacy Journal*.
10. Ali, M., & Badr, M. (2020). Lightweight phishing detection on mobile platforms. *Mobile Security Journal*.