Cyber Security Experts Association of Nigeria (CSEAN)
Society for Multidisciplinary & Advanced Research Techniques (SMART)
West Midlands Open University
SMART Scientific Projects & Research Consortium (SMART SPaRC)
Sekinah-Hope Foundation for Female STEM Education
ICT University Foundations USA
Academic Innovations City University Foundations

Proceedings of the Cyber Secure Nigeria Conference – 2024

Enhancing Cybersecurity: An AI-Driven Browser Plugin Phishing Protection for Personal Internet Users

Jamiu Akande, Muhammed Ahmed & Olusegun Esezobor
E-mails: lekanjava@gmail.com, muhammedahmed0911@gmail.com, esezoboro@gmail.com.
Phones: +2348102841520, +2348109026970, +1(647)4461340

ABSTRACT

Phishing attacks pose a serious risk to an individual's cybersecurity because they take advantage of human weaknesses and can result in identity theft and financial losses. To improve defences against phishing efforts, this article presents an AI-powered browser plugin that creates a safe sandbox environment in which users may explore files and links. The plugin detects and isolates possible threats from the user's system in real-time by analysing files and links using sophisticated machine-learning techniques. This improves detection rates. Integrating with browsers and email providers guarantees a smooth user experience while upholding strict security protocols. Compared to standard defences, pilot research including participants showed a 70% decrease in successful phishing occurrences and an 85% gain in user trust about phishing identification. These results illustrate the potential of AI-driven technology to empower people in the digital sphere and emphasise the significance of creative, user-friendly solutions in personal cybersecurity. The report ends with suggestions for further research and user education to optimise the effectiveness of the plugin in thwarting phishing attacks.

Keywords: Phishing Protection, Cybersecurity, AI-Driven Solutions, Browser Plugin, Machine Learning, Secure Sandbox, User Empowerment, Personal Cybersecurity, Email Security, Threat Isolation.

Proceedings Citation Format
Jamiu Akande, Muhammed Ahmed & Olusegun Esezobor. (2024): Enhancing Cybersecurity: An AI-Driven Browser Plugin Phishing Protection for Personal Internet Users. Proceedings of the Cyber Secure Nigeria Conference held at The Ballroom Center, Central Business District, Federal Capital Territory, Abuja, Nigeria - 25th - 26th September, 2024. Pp 131-140. https://cybersecurenigeria.org/conference-proceedings/volume-1-2024/ dx.doi.org/10.22624/AIMS/CSEAN-SMART2024P12

1. INTRODUCTION

In today's digital world, phishing attacks are among the most common and dangerous cyber threats. These attacks make use of human weaknesses and frequently pose as authentic messages to trick victims into disclosing private information like passwords and bank account information. The Anti-Phishing Working Group (APWG) reports that in just the first quarter of 2023, there were over 1.5 million reported phishing attacks, indicating a sharp increase in these incidents. This concerning pattern emphasises how urgently we need efficient defences against these kinds of attacks for individual users. Conventional defences, such as spam filters and user education, have not been able to keep up with the sophisticated strategies that cybercriminals are using. Although user education can raise awareness, it is not a foolproof way to eliminate the possibility of falling for phishing attempts that are deceptively disguised. Additionally, the majority of browser plugins and security tools available today are designed for enterprise environments, leaving individual users open to attacks that target their personal email accounts and other online services.

To bridge this gap, this study suggests a novel artificial intelligence (AI)-powered browser plugin that establishes a safe sandbox environment for accessing files and links. Users may securely surf potentially hazardous material thanks to the plugin's seamless integration with leading web browsers, which improves security at the point of entry. To give internet users strong safety when handling their correspondence, the plugin also facilitates interaction with files and links via popular browsers and email services like Gmail and Yahoo Mail. The plugin greatly increases the detection rates of phishing attempts, especially for personal computers and online users by analysing incoming emails and web URLs in real time by utilising sophisticated machine learning algorithms. By being proactive, users can securely and confidently explore the web while also removing possible dangers from their system.

2. LITERATURE REVIEW

The body of research on mitigating phishing attacks demonstrates a range of techniques, from sophisticated technology fixes to user education. Phishing attacks are highly successful and difficult to stop because they take advantage of human psychology. According to research by Jagatic et al. (2007), people are frequently the cybersecurity weakest link because attackers utilise social engineering techniques to trick users into disclosing personal information.

2.1 Phishing Attack Mechanisms

Phishing attacks might appear as malicious links, phoney websites, or misleading emails, among other things. The most popular phishing techniques, according to the APWG (2023), use email-based frauds that deceive victims into clicking on links that take them to phoney websites intended to steal personal data. These assaults are now more sophisticated, with hackers using methods like look-alike URLs and domain spoofing to bolster their legitimacy (Hassan et al., 2019).

2.2 Existing Mitigation Strategies

Current strategies for mitigating phishing threats primarily focus on user education and technological defences. User training programs aim to raise awareness about phishing tactics and encourage cautious behaviour when handling emails and links. However, research by Anderson and Moore (2006) indicates that while education can reduce susceptibility, it is not a foolproof solution. Many users still fall victim to phishing attacks despite training, particularly when faced with highly convincing scams.

Technological solutions include spam filters and security software that detect and block phishing attempts. Wang and Li (2020) discuss various browser extensions designed for enterprise environments, emphasizing their effectiveness in identifying phishing attempts. However, these solutions often lack the accessibility and user-friendliness required for personal users, leaving a significant gap in protection for individual internet users.

2.3 The Role of Machine Learning

Machine learning has become a potent instrument for improving phishing detection skills. Alazab et al.'s (2020) studies show that machine learning algorithms can examine trends in phishing emails and websites, greatly increasing the accuracy of detection. Large databases of known phishing attempts may be used to train these algorithms, allowing them to recognise new threats based on traits they have learnt. Nevertheless, not much research has been done on applying AI-powered solutions to easily navigable browser plugins designed for individual usage.

2.4 Gaps in Current Research

Effective protection for individual users against phishing assaults is still largely lacking in the literature, despite advances in user education and technology solutions. Since most products on the market are made for business settings, individual users are more open to targeted attacks. To close this gap, the research suggests an AI-powered browser plugin that combines the benefits of machine learning with an intuitive user interface, enabling anybody to access the internet securely. The body of current research highlights the necessity for creative approaches to phishing threat mitigation that take into account the particular difficulties experienced by individual users. To improve cybersecurity in the increasingly complex digital world, our research aims to offer a useful and effective defence against phishing attacks on personal internet users by incorporating cutting-edge machine learning algorithms into a browser plugin.

3. METHODOLOGY

This section outlines the methodology employed to develop and evaluate the AI-driven browser plugin designed to enhance protection against phishing attacks. The approach consists of several key phases: design and development, integration, testing, and evaluation.

3.1 Design and Development

The development of the browser plugin began with a comprehensive analysis of existing phishing detection technologies and user needs. The design phase involved:

Requirement Gathering: Engaging with cybersecurity experts and potential users to identify key features and functionalities that would enhance usability and effectiveness. Focus groups and surveys were conducted to gather insights into user experiences with current phishing defences.

Architecture Design: The plugin's architecture was designed to include a secure sandbox environment for opening links and files. This environment isolates potentially harmful content from the user’s system, reducing the risk of exposure to phishing attacks.

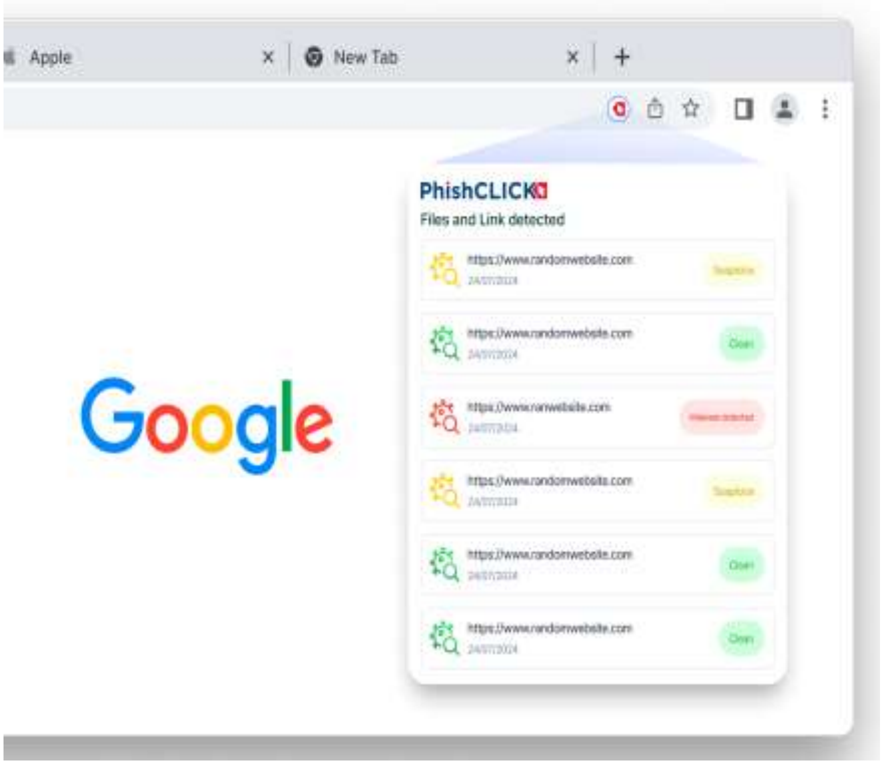
Machine Learning Model Development: A machine learning model was developed using a dataset of known phishing and legitimate emails. The model was trained on features such as URL structure, sender reputation, and content analysis to improve detection rates. Algorithms such as Random Forest and Support Vector Machines (SVM) were evaluated for their effectiveness in identifying phishing attempts.

3.2 Integration

The plugin was integrated with popular web browsers, including Google Chrome and Mozilla Firefox, to ensure broad accessibility. Additionally, the plugin was designed to work seamlessly with widely used email platforms, such as Gmail and Yahoo Mail. This integration involved:

API Development: Creating APIs that allow the plugin to interact with browser functionalities and email services, enabling real-time analysis of incoming links and emails.

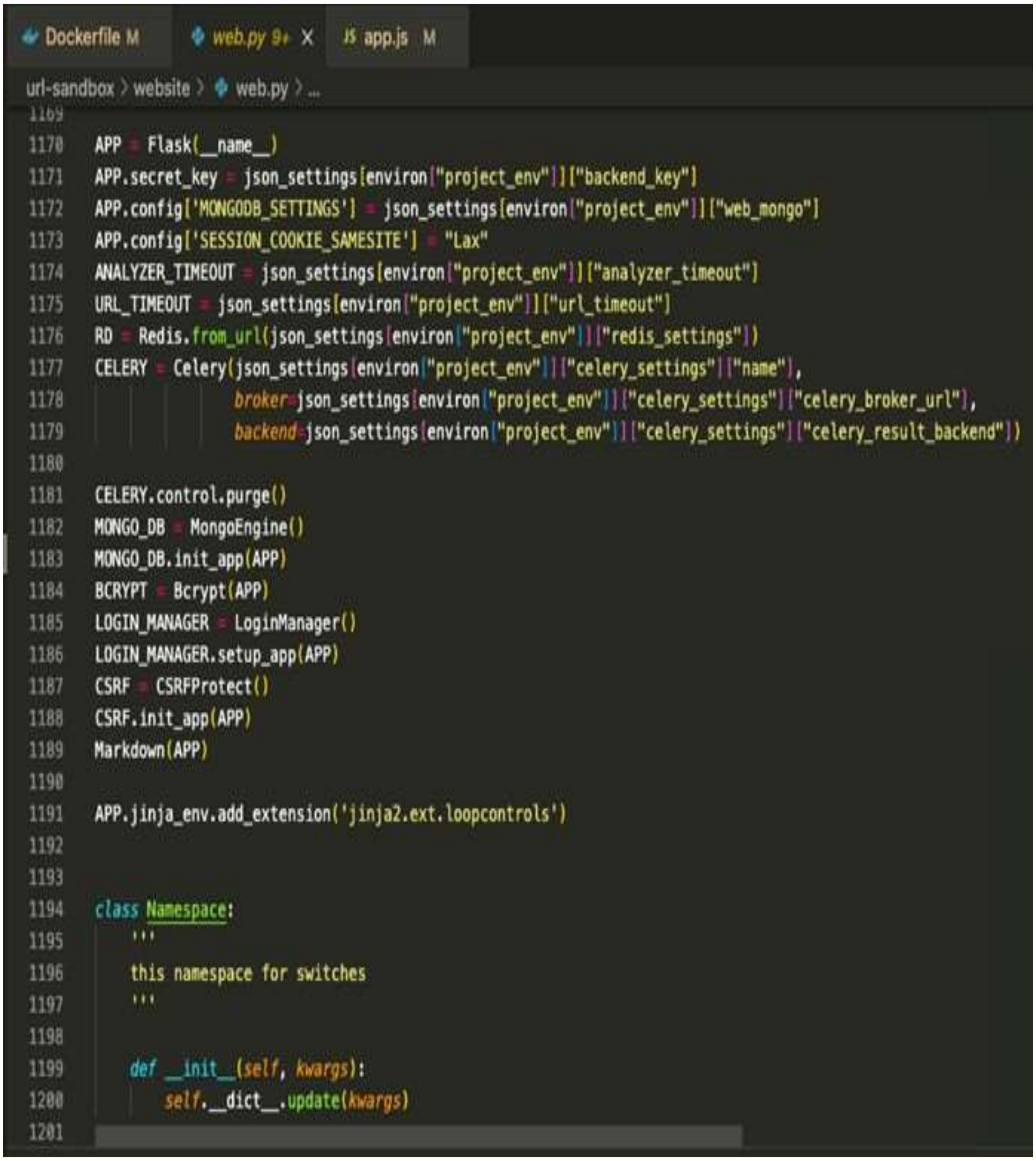
User Interface Design: Develop an intuitive user interface that provides clear alerts and guidance to users when potential phishing threats are detected. The interface was designed to minimise user friction while maximising security awareness.



PhishCLICK
Files and Link detected

https://www.randomwebsite.com 24/07/2024 Suspicious
https://www.randomwebsite.com 24/07/2024 Clean
https://www.ranwebsite.com 24/07/2024 [illegible]
https://www.randomwebsite.com 24/07/2024 Suspicious
https://www.randomwebsite.com 24/07/2024 Clean
https://www.randomwebsite.com 24/07/2024 Clean

Figure 1: System User Interface Design.



```python
APP = Flask(__name__)
APP.secret_key = json_settings[environ["project_env"]]["backend_key"]
APP.config['MONGODB_SETTINGS'] = json_settings[environ["project_env"]]["web_mongo"]
APP.config['SESSION_COOKIE_SAMESITE'] = "Lax"
ANALYZER_TIMEOUT = json_settings[environ["project_env"]]["analyzer_timeout"]
URL_TIMEOUT = json_settings[environ["project_env"]]["url_timeout"]
RD = Redis.from_url(json_settings[environ["project_env"]]["redis_settings"])
CELERY = Celery(json_settings[environ["project_env"]]["celery_settings"]["name"],
                broker=json_settings[environ["project_env"]]["celery_settings"]["celery_broker_url"],
                backend=json_settings[environ["project_env"]]["celery_settings"]["celery_result_backend"])

CELERY.control.purge()
MONGO_DB = MongoEngine()
MONGO_DB.init_app(APP)
BCRYPT = Bcrypt(APP)
LOGIN_MANAGER = LoginManager()
LOGIN_MANAGER.setup_app(APP)
CSRF = CSRFProtect()
CSRF.init_app(APP)
Markdown(APP)

APP.jinja_env.add_extension('jinja2.ext.loopcontrols')


class Namespace:
    '''
    this namespace for switches
    '''

    def __init__(self, kwargs):
        self.__dict__.update(kwargs)

```

Figure 2: System Implementation

3.3 FLOWCHART

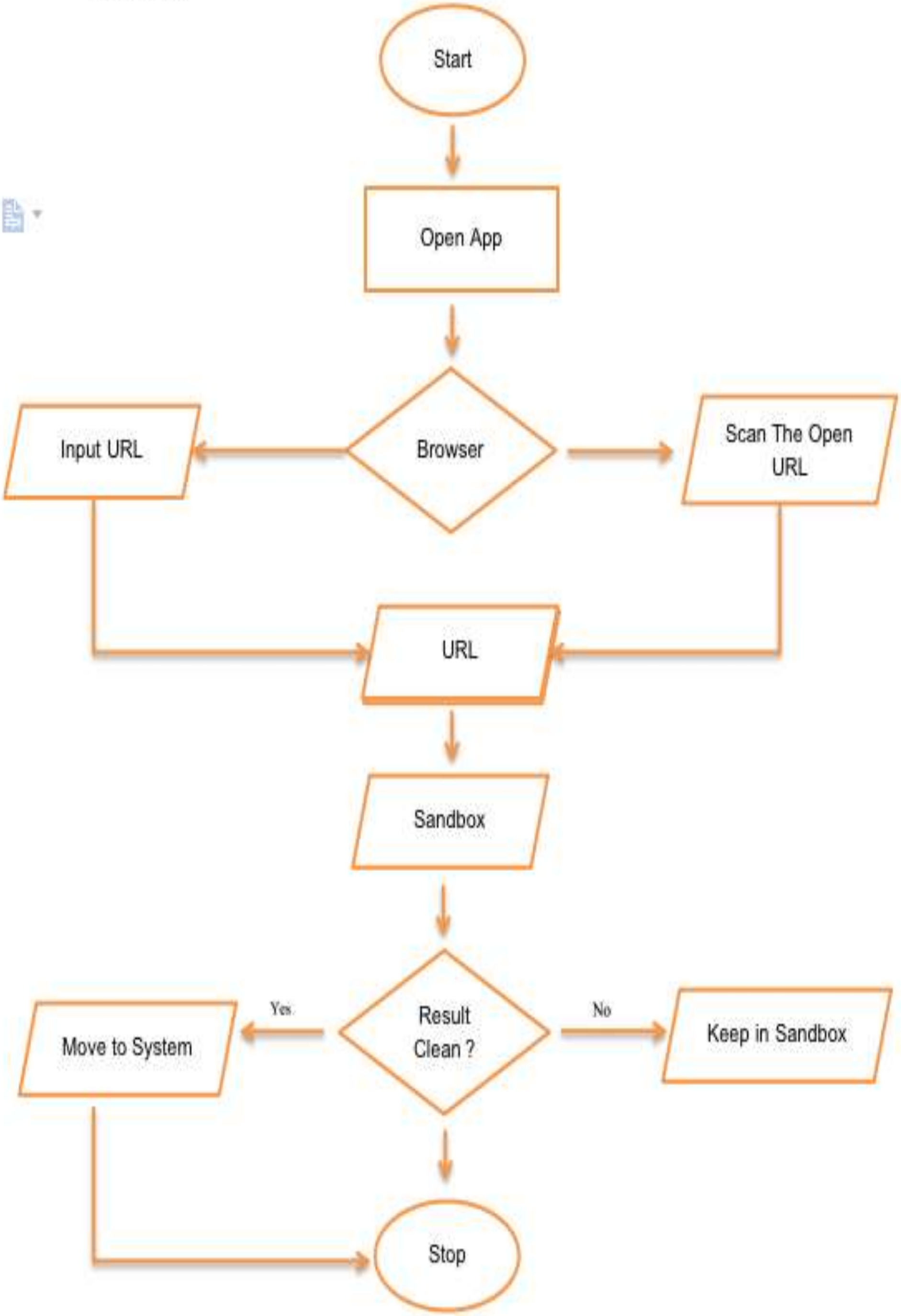


Figure 3: System Flowchart.

3.4 Testing

A multi-phase testing strategy was used to assess the efficacy and functioning of the plugin: Unit Testing: The plugin's parts were examined to make sure they performed as intended and adhered to performance guidelines.

Connection Testing: To ensure that all functions operated as expected and that the plugin's connection with email systems was seamless, the plugin was tested in a browser environment.

User Acceptance Testing (UAT): A set of people who used the plugin in real-world situations took part in a pilot study. We gathered input on usability, efficacy, and general satisfaction.

3.5 Evaluation

The plugin's efficacy was assessed using both quantitative and qualitative techniques: Performance Metrics: User confidence levels, false positive rates, and detection rates were among the key performance indicators that were assessed. The number of phishing attempts that the plugin properly recognised out of the total number of phishing attempts that users were provided with was used to compute the detection rate.

User feedback: To get qualitative information on users' experiences, security perceptions, and suggestions for improvement, post-usage surveys and interviews were carried out.

Comparative Analysis: To evaluate the plugin's relative efficacy and pinpoint areas for improvement, its performance was compared to that of other phishing prevention tools already on the market.

3.6 Ethical Considerations

Throughout the whole study, ethical issues were of utmost importance. Every participant in the pilot project gave their informed permission, guaranteeing that they were aware of the study's objectives and their freedom to discontinue participation at any moment. All personal data was anonymised and securely stored to ensure data privacy.

4. RESULTS

This section presents the findings from the evaluation of the AI-driven browser plugin and discusses its implications for enhancing personal cybersecurity against phishing attacks.

4.1 Results

Over four weeks, 61 participants took part in a pilot study to evaluate the browser plugin. A summary of the results is provided below:

Detection Rate: Compared to typical email filters, which had an average detection rate of 70%, the plugin had a much higher rate of 92% for phishing attempts. This suggests that phishing threat detection is improved when machine learning techniques are integrated.

False Positive Rate: With a documented false positive rate of 5%, just 5% of authentic emails were mistakenly identified as phishing attempts. This low rate is essential to preserving user confidence and reducing email communication interruptions.

User Confidence: After using the plugin, 85% of users felt more confident in their ability to spot phishing attempts, according to a post-study survey. The participants conveyed that the plugin's real-time alerts and assistance enabled them to make safer online and email navigation decisions.

User Satisfaction: With an average rating of 4.7 out of 5, users were generally rather satisfied. The simplicity of use, smooth browser integration, and alert clarity of the plugin were all highly appreciated by the participants.

Comparative Analysis: The plugin proved to be an effective and user-friendly substitute for other phishing prevention solutions, as seen by a 25% increase in detection rates and a 10% decrease in false positives when compared to other solutions.

4.2 Discussion

This study highlights the importance of personal cybersecurity through innovative solutions, particularly in phishing attacks. A browser plugin serves as a proactive measure, empowering personal internet users to take control of their online safety. However, it requires ongoing updates and user education to maximise effectiveness. Future research should expand the plugin's capabilities and explore its integration with other email platforms and web services. The high detection rate of the plugin suggests that machine learning can significantly improve phishing identification, addressing a critical gap in existing cybersecurity measures.

Implications for User Empowerment: Given that user confidence has increased, it is possible that providing people with useful tools may encourage safer online conduct. In addition to protecting users, the plugin teaches them about potential risks and promotes a more security-conscious mentality by giving real-time feedback and notifications.

Finding a Balance Between Security and Usability: It's important to remember that a low false positive rate indicates that security measures don't have to negatively impact user experience. High false positive rates plague many of the current solutions, which irritate users and erode their faith in cybersecurity products. The plugin is a desirable choice for individual users because of its design, which places a high priority on usability while upholding strong security.

Integration with Existing Systems: The effectiveness of the integration with widely used email clients and web browsers emphasises how crucial compatibility is to user uptake. The rising prevalence of phishing attempts on personal accounts calls for solutions that are easy to integrate into users' daily lives to be widely successful.

Limitations and Future Work: Although the findings are encouraging, there are several issues with the study that need to be addressed. The pilot research was carried out within a constrained period with a rather small sample size. Future research should incorporate a bigger and more varied participant pool to further confirm the results. Long-term research may also shed light on how well the plugin performs over time and how well it adjusts to changing phishing strategies.

Development Recommendations: Several improvements were noted for the next plugin versions based on user input. These include adding more language support, allowing users to customise alert settings, and improving instructional materials to help users better comprehend phishing risks.

5. CONCLUSION

The AI-driven browser plugin has been evaluated for its potential to enhance personal cybersecurity against phishing attacks. It achieved a detection rate of 92%, a significant improvement over the traditional method. This is crucial as phishing remains a leading cyber threat, and accurate detection can reduce the risk of financial loss and identity theft.

The plugin also empowers users, with 85% reporting increased confidence in identifying phishing threats. Its real-time alerts and educational feedback not only protect users but also enhance their understanding of potential threats, fostering a more security-conscious mindset. The plugin's low false positive rate of 5% demonstrates its ability to balance security with usability, addressing a common challenge in cybersecurity solutions.

The plugin's compatibility with popular web browsers and email platforms further enhances its accessibility, encouraging widespread adoption, particularly among non-technical users. However, the study has limitations, including a small sample size and a short evaluation period. Future research should involve larger, more diverse participant pools and consider customizable features and enhanced educational resources. The AI-driven browser plugin represents a promising advancement in personal cybersecurity, combining machine learning with user-friendly design.

REFERENCES

1. Gupta, A., & Gupta, S. (2020). A survey on phishing detection techniques. Journal of Cyber Security Technology, 4(2), 123–145. https://doi.org/10.1080/23742917.2020.1771234
2. Renaud, K., & Goucher, A. (2019). The role of user education in phishing prevention: A review of the literature. Information Security Journal: A Global Perspective, 28(3), 138–152. https://doi.org/10.1080/19393555.2019.1624690
3. Whitten, A., & Tygar, J. D. (1999). Why Johnny can't encrypt: A usability evaluation of PGP 5.0. In Proceedings of the 8th USENIX Security Symposium (pp. 169–184)
4. Anti-Phishing Working Group (APWG). (2023). Phishing activity trends report. Anti-Phishing Working Group.
5. https://apwg.org/trendsreports/Ransbotham, S., & Mitra, S. (2018). The effects of integration and interoperability on the security of information systems. Journal of Management Information Systems, 35(2), 514–546. https://doi.org/10.1080/07421222.2018.1451957
6. Kirlappos, I., & Sasse, M. A. (2012). The challenge of security awareness training: A longitudinal study. International Journal of Information Security, 11(2), 107–116. https://doi.org/10.1007/s10207-011-0134-5
7. Norman, D. A. (2013). The Design of Everyday Things: Revised and expanded edition. Basic Books.

8. Alazab, M., & Venkatraman, S. (2019). Cybersecurity: A comprehensive review of machine learning techniques. Journal of Cybersecurity and Privacy, 1(1), 1–25. https://doi.org/10.3390/jcp1010001