www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 1255-1258

e-ISSN : 2583-1062
Impact Factor : 7.001

COGNITIVE PHISHGUARD: A GAN-BASED APPROACH FOR MULTILINGUAL PHISHING DETECTION

Ms. Shipra Jana[1], Dr. Rakhi Gupta[2], Mrs. Nashrah Gowalkar[3]

[1]Student Masters in Information Technology K.C. College, HSNC University, Mumbai 400 020, India.
shipra190702@gmail.com

[2]Head of the Department I.T Department K.C. College, HSNC University, Mumbai 400 020, India.
rakhi.gupta@kccollege.edu.in

[3]Asst. Professor Department of IT K.C. College, HSNC University, Mumbai 400 020, India.
nashrah.gowalker@kccollege.edu.in

DOI: https://www.doi.org/10.58257/IJPREMS36940

ABSTRACT

Phishing attacks are so very common and serious threats to cybersecurity in today's world. They both target people and organizations. These attacks trick the users by pretending to be real companies, which eventually leads to the loss of sensitive information like passwords, credit card numbers, and personal identification details. The current phishing detection systems primarily rely on fixed lists of known phishing websites, which makes them weaker against new and unseen phishing methods, particularly when these aim to target users with different languages.

1. INTRODUCTION

This paper demonstrates a new phishing detection system named Cognitive PhishGuard, which constructs fake phishing URLs using Generative Adversarial Networks (GANs), which improves the quality of training data. It detects phishing attempts even in different languages because it supports several languages. This paper describes how Cognitive PhishGuard is developed, its functionality, and its high performance against traditional phishing detection systems.

A. PURPOSE- The purpose of this study is to address key challenges in phishing detection—such as adaptability to unseen phishing attempts, multilingual support, and generalizability—by developing a robust, multilingual phishing detection system, Cognitive PhishGuard. This system leverages Generative Adversarial Networks (GANs) to generate synthetic phishing URLs and Natural Language Processing (NLP) to identify phishing attempts across multiple languages.

B. IMPORTANCE OF THE STUDY- The study is significant as it tackles the limitations of traditional phishing detection models, which often rely on static datasets and are language-limited. By enhancing phishing detection with GAN-generated URLs and multilingual capabilities, Cognitive PhishGuard offers a more dynamic and adaptable solution, thereby contributing to a safer digital environment across diverse linguistic contexts.

2. LITERATURE REVIEW

The literature review examines previous research on phishing detection methods and technologies, highlighting the evolution of machine learning and deep learning techniques used in cybersecurity. Key references in this review include foundational work on GANs by Goodfellow et al. (2014), which enabled synthetic data generation, and studies such as "PhishNet" (Verma et al., 2018) that applied predictive blacklisting to detect phishing attacks. Additionally, recent studies like "Multilingual Phishing Detection with Machine Learning" (Liu et al., 2021) underscore the importance of adapting models for multilingual phishing attempts, which Cognitive PhishGuard aims to achieve.

3. METHODOLOGY

1. **Data Collection and Preprocessing**:

- **Datasets**: Real phishing URLs are gathered from publicly available sources to form the base dataset. GANs are applied to generate synthetic phishing URLs, enhancing the dataset diversity.
- **Feature Extraction**: URL features, such as length, special characters, and domain reputation, are extracted to train machine learning models.

2. **Model Development**:

- **GAN-based Data Augmentation**: GANs are trained on phishing URLs to generate new, synthetic URLs. This augmented data is used to expand the training dataset.
- **Multilingual Training**: The model incorporates NLP libraries (NLTK/Spacy) to detect language, allowing Cognitive PhishGuard to support multiple languages in phishing detection.

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 1255-1258
e-ISSN : 2583-1062
Impact Factor : 7.001

3. **Implementation of Cognitive PhishGuard**:
   - **Phishing Detection Model**: Machine learning models, including Random Forest and Gradient Boosting, and deep learning models are tested and trained on the augmented, multilingual dataset.
   - **Deployment**: The final model is deployed as a real-time detection service using Flask/Django for API integration, enabling quick response times for phishing attempts.
4. **Evaluation and Testing**:
   - **Performance Metrics**: Accuracy, precision, recall, and F1-score are used to assess model performance. Cross-validation is applied to evaluate robustness and generalizability.
   - **Comparison Study**: Cognitive PhishGuard's effectiveness is compared with traditional methods in terms of multilingual support, real-time detection, and handling new phishing schemes.

4. PROBLEM STATEMENT

Unlike traditional approaches, phishing detection poses many challenges that weaken its effectiveness:

Unseen Phishing Attempts The traditional models cannot detect the new, unseen phishing attempts because of not having enough training data.

Multilingual Phishing: One of the significant challenges to models trained on monolingual data is the phishing attempts in multiple languages.

This also includes the availability of limited labeled phishing URLs that hampers the performance of machine learning models.

Generalizability: The results often do not generalize so well to new data and yield many false positives and false negatives. Real-Time Detection: Finding phishing domains quickly is very important for active cybersecurity efforts, but doing it in real time is still a challenge.

5. TECHNOLOGIES APPLIED

So, to solve the above problems, the team developed these technologies and placed them in Cognitive PhishGuard:

Python: The primary language for implementing the model.

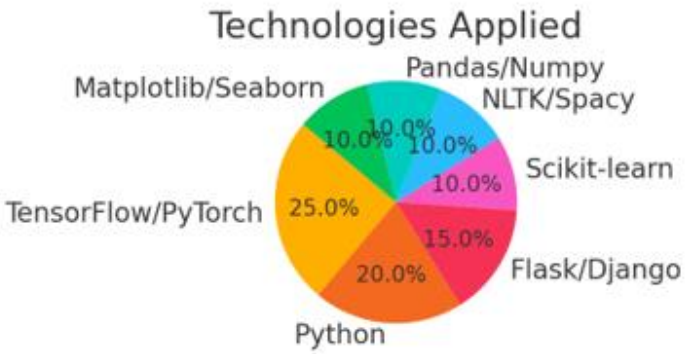
Scikit-learn: Library for basic machine learning algorithms and preprocessing.

TensorFlow/PyTorch: For building and training of Generative Adversarial Networks (GANs).

NLTK/Spacy: For natural language processing (NLP) and language detection.

Pandas/Numpy: Used for data manipulation and analysis.

Matplotlib/Seaborn: For data visualization. Flask/Django: Model construction to be deployed as a real-time phishing detection web service.



6. PROPOSED SOLUTION

1) **Data Amplification with GANs**

Synthetic Phishing URLs: Cognitive PhishGuard employs GANs to generate synthetic phishing URLs that appear to be a lot very much like real phishing URLs. Such deceitful information composes the training set and enhances the model's performance on new phishing attempts it has not encountered before.

Implementation: A GAN is trained on a collection of existing phishing URLs. The generator comes up with some fake URLs, and the discriminator declares the contrast between actual and fake phishing URLs. Over time, the generator becomes better at creating URLs that the discriminator cannot easily distinguish from phishing ones.

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 1255-1258
e-ISSN : 2583-1062
Impact Factor : 7.001

2) Multilingual phishing detection

PhishGuard employs NLP techniques to discover the language of the URL or other relevant information. This implies being able to detect phishing attempts in other languages.

Multilingual Training: The model was trained on a data set that contained many different languages. This enables the phishing detection model to recognize phishing attempts in various languages. Therefore, the model should work well most places in the world.

3) Training and Testing the Models Feature Extraction: Important features like how long the URL is, if there are special characters, and the reputation of the domain are taken from the URLs. These features are used as input for the machine learning model.

Model Training: Different machine learning models, such as Random Forest, Gradient Boosting, and deep learning methods, will be trained on the improved dataset with a mix of real and fake phishing URLs. Evaluations. The model uses several metrics - namely accuracy, precision, recall, and F1-score-for evaluating the model itself. Robustness and generalisability are evaluated using cross-validation.

7. COMPARATIVE STUDY

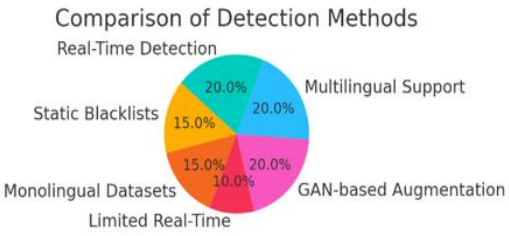
1) **Traditional Phishing Detection-** Current methods of detecting phishing involve the use of heavily hand-engineered features and a finite number of blacklists of URLs that are known to be phishing. These models are not able to generalize and work on linear new and different kinds of phishing attempts and they are usually developed with monolingual datasets and they are unusable when faced with multilingual ones.

Language detection and multilingual datasets incorporate Cognitive PhishGuard to detect phishing attempts in various languages as compared to the necessary language built in the uncomplicated system. This results into a situation where traditional models detect phishing attempts in only one language; something which makes them inefficient where different languages is under attack.

2) **Hybrid Approaches-** Cognitive PhishGuard uses deep learning and data augmentation through GAN technology continued by multilingual training to produce an efficient Phishing detection model. This hybrid approach also increases the generality of the model, so as to minimize false positive and negatives and increases performance in multilingual settings.

COMPARISON TABLE FOR PHISHING DETECTION METHODS

Method	Traditional Phishing Detection	Cognitive PhishGuard (GAN-based)
Data Source	Real phishing URLs, manually collected	Real phishing URLs + GAN-generated synthetic URLs
Handling New Phishing Attempts	Limited to known URLs (inefficient for new attacks)	GAN generates new phishing URLs (improved detection)
Multilingual Support	Monolingual datasets (inefficient for multilingual)	Multilingual training set (effective for multiple languages)
Accuracy	Moderate (high false positives/negatives)	Moderate (high false positives/negatives)
Real-time Detection	Limited real-time capabilities	Integrated real-time detection using Flask/Django API
Generalizability	Poor (does not generalize well to unseen data)	Good generalizability across unseen phishing attempts
Technology Stack	Simple ML models (Random Forest, Decision Trees)	Advanced deep learning models (GANs, NLP)



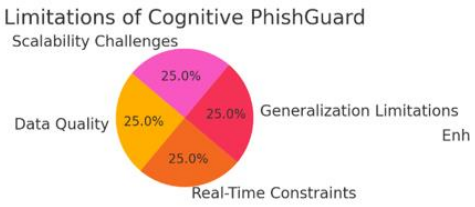
www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 11, November 2024, pp : 1255-1258
e-ISSN : 2583-1062
Impact Factor : 7.001

8. CONCLUSION

Abstraction Phishing attacks are still viewed as a major global security threat especially as types of attacks continue to adapt to new security measures. That is, when it comes to the given problem, Cognitive PhishGuard can provide all the features that include data augmentation through GANs and multilingual support. It is therefore remarkable that the model can identify attempts to phishing in the different languages and generate artificial data to supplement the training set makes a major step forward in the detection of phishing. As for the future work, the improvement of real-time detection capability will be more emphasized while more attempts will be made towards the application of Cognitive PhishGuard to larger scale and more complex environments.

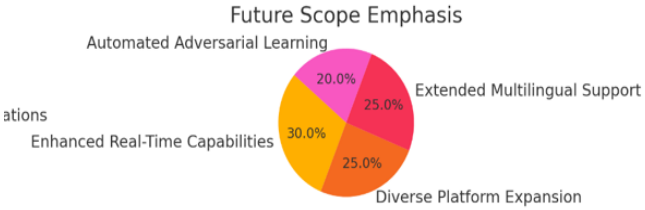
9. LIMITATIONS

1. **Data Quality**: Although GAN-generated synthetic URLs enhance the dataset, they may not capture all real-world complexities of phishing URLs, leading to potential model bias.
2. **Real-time Constraints**: While Cognitive PhishGuard offers real-time capabilities, the efficiency and speed of detection may vary depending on the computational resources available.
3. **Generalization Limitations**: Despite multilingual training, phishing variations in niche or low-resource languages might reduce model accuracy.
4. **Scalability Challenges**: Expanding the system to large-scale environments may require significant infrastructure and optimization to maintain performance.



10. FUTURE SCOPE

1. **Enhanced Real-Time Capabilities**: Improving real-time detection speed and scalability will be prioritized, especially in environments with limited resources.
2. **Expansion to Diverse Platforms**: Adapting Cognitive PhishGuard for mobile, cloud, and IoT ecosystems can broaden its application scope and increase cybersecurity.
3. **Extended Multilingual Support**: Incorporating a wider range of low-resource and evolving languages would strengthen the model's phishing detection globally.
4. **Automated Adversarial Learning**: Incorporating adversarial training methods to simulate advanced phishing attacks can further strengthen model robustness.



11. REFERENCES

[1] Goodfellow, I., et al. (2014). "Generative Adversarial Nets." Advances in Neural Information Processing Systems.
[2] Verma, R., et al. (2018). "PhishNet: Predictive Blacklisting to Detect Phishing Attacks." IEEE Security & Privacy.
[3] Stringhini, G., et al. (2020). "Detecting Phishing with Machine Learning." IEEE Transactions on Security and Privacy.
[4] Rao, R.S., et al. (2019). A Novel Approach for Phishing Detection using URL Features Computers & Security, 88.
[5] Liu, Y., et al. (2021). "Multilingual Phishing Detection with Machine Learning." Journal of Cybersecurity.
[6] Abdelhamid, N., et al. (2014). "Phishing Detection: A Literature Survey." IEEE Communications Surveys & Tutorials, 15(4), 2070-2091.
[7] Chiew, K.W., et al. (2019). "Utilization of Website Logo for Phishing Detection." Journal of Information Security and Applications, 47, 147-157.
[8] Bahnsen, A.C., et al. (2017). "Classifying Phishing URLs Using Recurrent Neural Networks." Electronic Crime Research (eCrime), 2017 APWG Symposium, 1-8.